AIR TRAVEL DEMAND AND AIRLINE SEAT INVENTORY MANAGEMENT

by

PETER PAUL BELOBABA

SM, Massachusetts Institute of Technology, 1982.

BES, University of Waterloo, Canada, 1980

Doctoral Dissertation in Flight Transportation

submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1987

Signature of Author: ______________________________

Department of Aeronautics and Astronautics, May 1987

Certified by: ______________________________

Dr. Robert W. Simpson, Thesis Supervisor
Professor of Aeronautics and Astronautics
Director, Flight Transportation Laboratory

Certified by: ______________________________

Dr. Amedeo R. Odoni
Professor of Aeronautics and Astronautics and Civil
Engineering; Co-Director, Operations Research Center

Certified by: ______________________________

Dr. Nigel H. M. Wilson,
Professor of Civil Engineering

Accepted by: ______________________________

Dr. Harold Y. Wachman,
Professor of Aeronautics and Astronautics
Chairman, Departmental Graduate Committee

Air Travel Demand and Airline Seat Inventory Management

by Peter Paul Belobaba

Submitted on May 1, 1987 in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy in Flight Transportation
at the Massachusetts Institute of Technology

ABSTRACT

Many airlines practice differential pricing of fare products that share a common inventory of available seats on an aircraft. Seat inventory management is the process of limiting the number of seats made available to each fare class. The objective of both strategies is to maximize the total revenues generated by the mix of fare products sold for a flight.

This dissertation first examines the evolution of airline marketing and seat inventory management practices. A demand segmentation model is developed to help explain current airline fare structures. A conceptual model of the consumer choice process for air travel is then presented, and extended to describe the airline reservations process and the probabilistic elements that can affect seat inventory control.

A survey of current airline practice in this area revealed that seat inventory control is an *ad-hoc* process which depends heavily on human judgement. Past work on quantitative approaches has focused on large-scale optimization models that solve simple representations of the problem. A primary objective of this research was the development of a quantitative approach based on the practical constraints faced by airlines.

The Expected Marginal Seat Revenue (EMSR) model developed in this thesis is a decision framework for maximizing flight leg revenues which can be applied to multiple nested fare class inventories. It is applied to a dynamic process of booking limit revision for future flight departures, and overbooking factors as well as fare class upgrade probabilities are incorporated. Examples of EMSR model results are presented, and a critical analysis of the demand assumptions and sensitivity of the model is performed.

The EMSR model was implemented as part of an automated seat inventory control system at Western Airlines and tested on a sample of actual flights. Compared to flights managed by existing manual methods, flights for which fare class booking limits were set and revised automatically on the basis of the EMSR decision model carried more passengers at a lower yield, and generated higher total revenues.

Thesis Committee:

Dr. Robert W. Simpson, Professor of Aeronautics and Astronautics;
Director, Flight Transportation Laboratory; Committee Chairman.

Dr. Amedeo R. Odoni, Professor of Aeronautics and Astronautics and
Civil Engineering; Co-Director, Operations Research Center.

Dr. Nigel H. M. Wilson, Professor of Civil Engineering.

Acknowledgements

I was extremely fortunate to have a Doctoral Committee comprised of three members who contributed substantially and in different respects to the format and contents of this dissertation. Professor Robert Simpson was involved most closely in the various research efforts that contributed to this work, guiding the empirical analysis and the development of revenue maximization approaches. More recently, he has provided detailed feedback on the sections that deal with airline economics and operations, insisting that the semantics be precise so that this work might become a useful reference document.

Professor Amedeo Odoni contributed to the quantitative sections of this thesis, helping me to explain the mathematical formulations as clearly as possible. Professor Nigel Wilson provided much needed assistance with respect to the logic and consistency of the discussion from a systems analysis perspective, ensuring that the concepts presented can be understood by those not intimately familiar with airline operations. I am grateful to my committee members not only for their contribution to this thesis, but also for the stimulating courses that I have taken under their instruction at M.I.T. All three possess a dedication to teaching that seems to be increasingly rare in research-oriented academic environments.

A doctoral program that takes almost four years to complete requires a substantial amount of financial support, and I am privileged to acknowledge funding from a number of sources. I would like to thank the National Social Sciences and Humanities Research Council (NSSHRC) of Canada for the Doctoral Fellowship support. Additional support during the 1985-86 academic year was provided by the United Parcel Service Foundation Doctoral Fellowship, administered by the Center for Transportation Studies at M.I.T.

Research funding and technical assistance was also provided by the M.I.T./Industry Cooperative Research Program at the Flight Transportation Laboratory. As a member of the Program, Trans World Airlines furnished reservations data for empirical analysis early in this research. McDonnell-Douglas representative John Stroup assisted in the initial development of computer code for seat allocation models.

A research agreement between the Flight Transportation Laboratory and Western Airlines allowed me to develop and implement my ideas in an actual airline environment.

Many thanks to Cal Rader and Dennis Fitzpatrick, who initiated what proved to be an extremely rewarding cooperative effort. Those involved in the project at Western were very helpful, and deserve special mention: Judy Whitaker, Rainer Siegert, Diane McGinty, Sherry Jackson, Lynne Graham, and John Sefton. I share their sadness over the passing of Western Airlines from the industry. After making more than 20 trips to the West Coast, I must admit that Western was "The only way to fly!"

Personal thanks go to my closest friends: Mike Meyer, for being the first to encourage me to pursue a Ph.D.; Brad Martin, for providing last-minute help with tables and figures; Carole Marks, for her genuine interest in my progress and feigned interest in my research; and Eric and Liane Schreffler, for their hospitality in Los Angeles.

This dissertation is dedicated to my parents, who always wanted me to become a doctor.

Contents

List of Figures

List of Tables

INTRODUCTION

Having seen a newspaper advertisement for air travel between Boston and Los Angeles for $ 198 round-trip, a potential traveler calls the airline and is told that the lowest available fare for the specific dates and flights requested will be $ 358. After agreeing to accept this higher fare, the traveler ends up sitting next to another traveler who has paid over $ 500 for one-way passage. This scenario has become increasingly familiar to air travelers since the deregulation of the U.S. airline industry in 1978. Many airlines offer a wide range of fares for travel in a single city-pair market as they attempt to fill empty seats and realize as much revenue as possible from each seat sold.

Seat inventory management is the process of balancing the number of seats sold at each of the fare levels offered so as to maximize total passenger revenues. This practice enables airlines to influence their total revenues on a flight-by-flight basis, within a given price structure. Seat inventory control and pricing are two distinct strategies that comprise what can be referred to as airline revenue management. This dissertation provides a comprehensive discussion of the airline revenue management problem, with an emphasis on seat inventory control.

The evolution of the theory and practice of seat inventory management is reviewed first in Part One. The economic principles behind airline pricing and market demand segmentation practices are used to explain the airline fare structures currently in place. A model of how consumers make air travel decisions is developed, and the implications for the airline reservations process are discussed. The seat inventory control problem faced by airlines is then defined and a survey of current practice allows the practical constraints for solution approaches to be identified.

Part Two is devoted to quantitative methods for seat inventory control, beginning with a review of previous work in the area. The shortcomings of applying past approaches to the practical problems faced by airlines lead to the development of a quantitative decision framework for revenue maximization that can be applied directly to existing fare structures and reservations systems. The Expected Marginal Seat Revenue (EMSR) model is applied to the problem of revising booking limits on different fare types dynamically, and extended to account for overbooking and passenger upgrades when reservations requests are denied. The EMSR model is examined critically with respect to its demand assumptions and sensitivity to input variables, and its application to origin-destination seat inventory control is considered.

The final part of this dissertation describes the application of the EMSR framework at an airline, as the development of an automated seat inventory control system at Western Airlines is discussed. A test of the automated system relative to existing manual seat inventory control methods was performed, and the results demonstrate the potential benefits of employing quantitative decision models in the process. The dissertation concludes with a discussion of future seat inventory control system development, and directions for further theoretical and empirical research.

PART ONE
The Evolution of Seat Inventory Management: Theory and Practice

Airlines market a "product" in the form of a seat on an aircraft scheduled to fly from one point to another at a specified time in the future. Consumers purchase the transportation service provided by the airline product, along with the meals, drinks and entertainment that might be included. Although what airlines market is primarily a service, the limited number of seats that may be sold for a future flight departure represent a fixed product "inventory" with several unique properties. Identical units in this inventory (seats) may be associated with different purchase conditions and service amenities, allowing them to be priced and marketed as distinct service options. Furthermore, the unsold product (i.e., an empty seat) actually increases in value to some consumers over time (i.e., as the desired departure date and time approach), but only up to the point at which the aircraft departs. Having reached its highest value just prior to departure, the unsold airline seat cannot be sold at any price after the aircraft departs.

In coping with these unique properties of their product "inventory", airlines have implemented pricing and marketing schemes designed to minimize the "spoilage" of unsold seats. For many years, carriers pursued incremental demand with "discounted" fares and services while striving to retain their original full-fare traffic bases. With the increased freedoms of deregulation in the United States, this effort to stimulate demand and fill empty seats has intensified. A wide variety of differentiated service/price combinations are offered in most markets. With the emergence of differential pricing of seats that can share a common aircraft cabin, and therefore the same seat "inventory", the problem of managing these seat inventories so as to maximize revenue and operating profits has also emerged.

Many airlines have therefore become extremely interested in what is commonly referred to as "yield management". *Yield* is the revenue per passenger-mile of traffic carried by an airline. An airline's yield will be a function of both the prices it charges for its differentiated service options and the number of seats actually sold at each price. *Yield management* thus involves two major components — pricing and seat inventory control. The emphasis of this dissertation is on the latter. *Seat inventory control* is the process of limiting the number of seats to be made available to service options with different price levels on a future flight departure.

Part One of this thesis is devoted to the theory and practice of airline seat inventory control. It introduces the basics of the seat inventory control problem and outlines its implications from the perspective of both the airline wishing to manage its seat inventory and the consumer faced with a series of travel decisions given a choice of service options. The current state of seat inventory control practices is also reviewed.

Chapter One reviews the development of airline practices and the inherent characteristics of airline operations that led to the need for more sophisticated seat inventory control. The economic concepts of differential pricing and market segmentation are presented in the airline context as the underlying rationale for the evolution of current marketing practices in the airline industry. Included is an overview of the wide range of fare levels, classes of service, and ticket/travel conditions found in the U.S. airline marketplace.

Chapter Two opens with a model of the decision process for air travel consumers that has developed in response to changing airline marketing and pricing strategies. Given a range of price and service options, and an ever-increasing choice of carriers and routing alternatives, consumers have changed the way in which they make travel decisions. This model of consumer behavior is extended to illustrate the reservations process as viewed from the airline's perspective. Greater choice on the part of the consumer has meant even less predictable demand for the individual airline flight than was the case a decade ago.

The complexity of the seat inventory control problem for large airlines is explored in Chapter Three. The characteristics of airline operations and reservations procedures that contribute to the size and definition of the problem are discussed. An overview of current seat inventory control practices is then presented. Based on a survey of nine large North American airlines, this discussion of current practice highlights both the range of sophistication found in the industry and the apparent need for better mathematical and modeling techniques for seat inventory control.

Chapter 1

Airline Economics and Pricing Practices

Determining the price a potential passenger would pay to fly between two cities used to be a relatively simple exercise. The biggest choice facing the passenger was generally between the first class and regular economy or coach fares, which were identical on all carriers and were derived on the basis of mileage-based fare structures imposed by the Civil Aeronautics Board, or IATA for international services. In certain markets, price discounts might have been available to those meeting certain age requirements (children, senior citizens), traveling as a family or group, belonging to specific organizations (military, travel industry), or flying on selected "night coach" flights. Apart from "night coach" fares, these special fares represented true ***discounts*** in price for a standard service. Potential passengers had only to choose between the two or three carriers serving the city-pair market desired on the basis of reputation, service and schedule convenience.

Needless to say, the airline pricing situation has changed dramatically. The practice of discounting prices on standard service options has given way to the creation of differentiated service options offered at prices lower than the standard coach fare. This chapter describes the evolution of airline marketing strategies that has led to the wide range of price and service options offered in the current airline marketplace. The theoretical economic concepts on which these strategies have been based are presented first, and the development of a need to control seat inventories more effectively is outlined. The second section of this chapter explains how airlines have applied these theoretical concepts and adapted them to the practical constraints of their reservations and ticketing processes. Finally, a description of the price and service options available in most domestic airline markets at the present time provides an up-to-date illustration of the extent to which the simple pricing structure of the 1960's has changed.

1.1 Empty Seats and Capacity-Controlled Fares

The introduction of advance purchase excursion (APEX) fares in international markets enabled any passenger, regardless of age or group affiliation, to purchase air travel at a reduced price. These fares required advance purchase of a round-trip ticket and a minimum length of stay at the destination point. The reduced price was therefore available only for a service option that was different in its conditions of use from the standard coach fare, even though the flights, seats, and on-board service amenities were generally identical. This type of fare spread to U.S. domestic markets in 1975 when American Airlines introduced its first "super-saver" fares.

Restrictions on the use of these lower fares were imposed by the airlines to limit their use to leisure or vacation travelers. The original objective of these fares was to stimulate new demand in the discretionary travel market segment, which would in turn fill otherwise empty seats and generate additional revenues for the airline. By applying advance purchase, length of stay and round-trip itinerary requirements, airlines hoped to keep full-fare business travelers from making use of these lower fares, and thereby introduced a new type of air travel service option to the market.

Research done by the Boeing Aircraft Company into the nature of empty seats on the larger wide-body aircraft that entered service in the 1970's encouraged airlines to apply such restrictions to reduced fares to prevent the *diversion* of existing demand from higher fare levels [1]. Boeing argued that, due to an inherent variability in demand for air travel and an inability on the part of the airline to provide exactly the number of seats required to satisfy all full-fare demand, empty seats are inevitable. The inevitability of having unsold seats and the potential for recovering additional revenues from them led to Boeing's development of an approach for "surplus seat management" [2], which involved forecasting the number of seats that would otherwise remain empty.

While inevitable, the existence of empty seats in an airline's operation is not entirely undesirable. Given the flight-to-flight variation in full-fare demand, some empty seats are necessary to provide a "buffer" for accommodating unexpectedly high demand and to maintain a high quality of service in terms of seat "availability" for full-fare passengers [3]. This buffer ensures a reasonable probability of the airline being able to accommodate last-minute passengers with changes to their travel plans. Beyond the number of seats required to maintain a reasonable "availability" buffer, additional empty seats may also contribute to a feeling of spaciousness for those on board. On the other hand, too many empty seats (overcapacity) contribute nothing in revenue and represent a sunk cost to the airline.

Under ideal conditions, airlines can increase their revenues by filling up some or all of these vacant seats with incremental demand without a corresponding increase in operating costs. Because of the large proportion of airline operating costs that can be considered fixed in the very short-run (given a commitment to operate a flight schedule over a month or longer), the marginal cost of carrying an additional passenger in an otherwise empty seat is very low. The incremental costs involved in reservations, ticketing, baggage handling and meal service for an additional passenger typically amount to much less than $25 [4]. As a result, prices much lower than full-fare can be charged to attract incremental demand. As long as the incremental revenue per passenger exceeds the marginal cost of carrying each one, a contribution will be made to the fixed costs of the flight. It can thus be entirely rational, in the short run, for an airline to offer seats that are sure to remain unsold at prices as low as, say $29, for a scheduled transcontinental flight.

In the case of air transportation, determining the relevant marginal cost for pricing purposes is difficult. The marginal unit of production for an airline in the short-run is a flight departure with a fixed seating capacity, whereas the marginal unit sold to the consumer is a single seat on that flight. For a scheduled future flight departure, fixed aircraft operating costs per passenger decrease as the number of passengers actually carried increases, up to aircraft capacity. Assuming a single class of service, variable costs per passenger are constant, independent of the number carried on a particular flight. Total costs per passenger are thus minimized when the fixed costs of operating the flight are spread over a full planeload of passengers.

Given these conditions, a single price charged for all seats based strictly on marginal variable costs might not cover the total operating costs of the flight. For that matter, depending on the cost structure of the airline and the level of demand, there might not be any one price that will cover total operating costs of the flight. It might be possible, however, for firms in an oligopolistic industry to segregate the purchases of individual consumers and charge each consumer as much as he/she is willing to pay. In theory, each incremental purchaser would be charged less and less as more output is sold, up to the point at which price equals marginal cost. In practice, a firm might be able to identify distinct groups of consumers and charge each group a different price for a homogeneous product.

Figure 1.1 illustrates how ***differential pricing*** can enable the firm to cover total costs with total revenues, whereas strict marginal cost pricing would not. For an isolated flight scheduled to operate with a given aircraft and fixed number of identical seats, the marginal variable costs are constant per passenger carried. The fixed aircraft operating costs per passenger decrease with the number of seats filled, meaning total costs per passenger also decrease, as shown. In Figure 1.1, the demand curve was drawn below the total cost per passenger curve to illustrate a case in which no one price can generate total revenues to cover total flight operating costs.

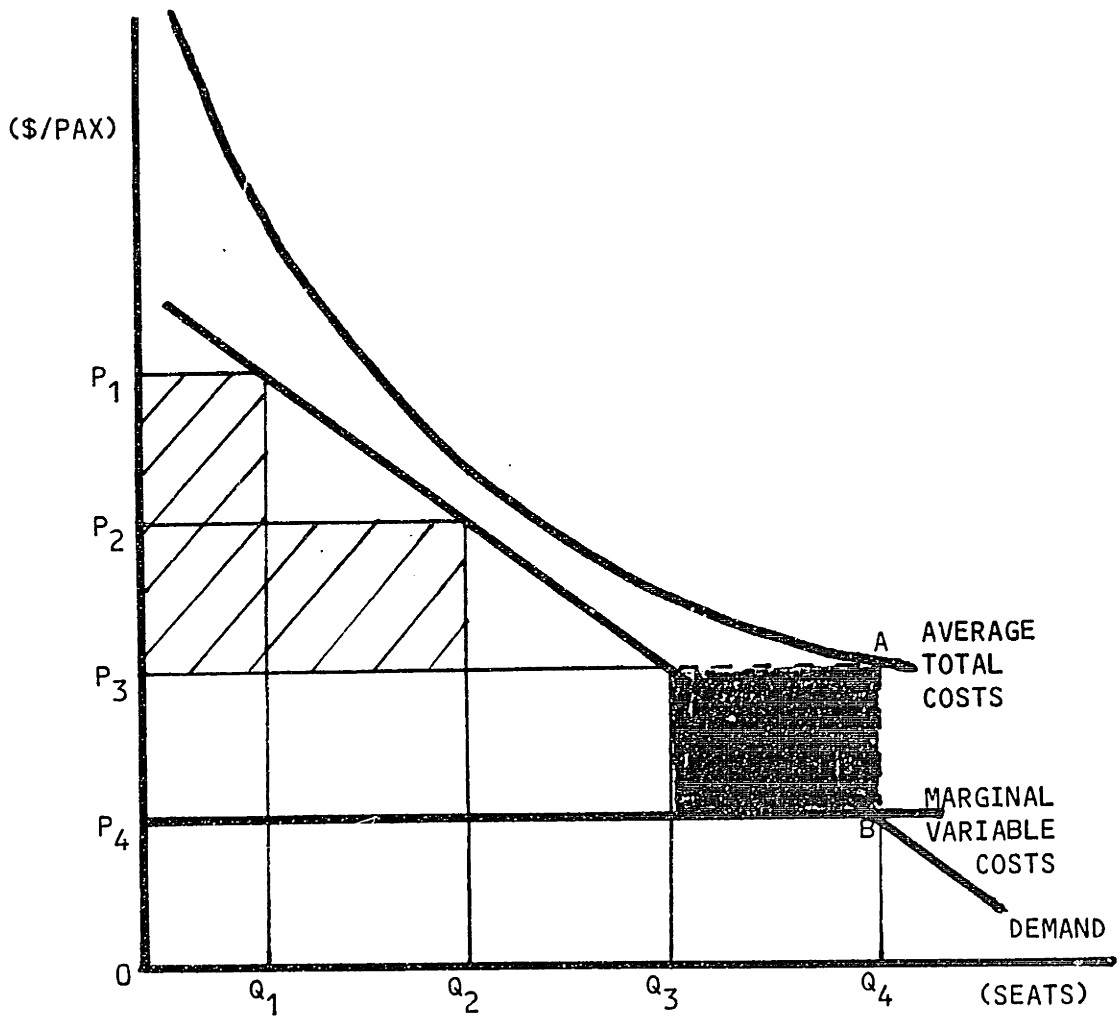
If price is set equal to marginal variable costs and applied to all passengers, Q_4 passengers would be carried at a total revenue equal to the area of $0P_4BQ_4$. Total operating costs would be equal to the area $0P_3AQ_4$, resulting in an operating loss equal to P_4P_3AB. With the differential pricing strategy shown, Q_1 passengers would pay P_1, $Q_2 - Q_1$ passengers would pay P_2, and so on. Under this strategy, total revenues exceed total operating costs of the flight as long as the striped area is greater than the shaded area.

Pricing in airline markets, which are oligopolistic in most cases, is well-suited to this concept of differential pricing. Such a pricing strategy is economically desirable if [5]:

1. the relevant costs of taking on incremental passengers are less than the average total costs of the flight;

2. lower prices are required to stimulate incremental demand;

Figure 1.1: Differential Pricing of Airline Seats: Single Flight Example



3. the incremental demand is sufficiently elastic that reduced fares increase total revenues by more than the increase in total costs.

Differential pricing of identical seats on a flight departure can enable the airline to generate the additional revenue required to cover the total costs of the flight from those passengers using the reduced fare options. Airline efficiency can be enhanced, as better use of existing equipment with its sunk costs can be realized by filling otherwise empty seats. It is also possible for the airline to reduce cyclical variations in demand over time through selective time-dependent price reductions and effective seat inventory control, as will be discussed.

From the consumer's perspective, the practice of differential pricing by airlines can in theory benefit *all* passengers, resulting in a *pareto optimal* market situation. As long as the full-fare passengers do not pay more than they would have in the absence of differential pricing, they are no worse off. Full-fare passengers can in fact benefit from the practice if low-fare passengers contribute to common costs and help to keep full fares lower than they otherwise would be. And, as long as seat inventory control is practiced effectively, full-fare passengers need not be denied seats due to high demand for low fares. In the longer run, full-fare passengers stand to gain from the potential increase in flight frequencies required to accommodate increased total market demand.

Low-fare passengers benefit from the availability of lower-priced air travel options, either by saving money on a planned trip or by being able to take a trip that would not otherwise have been taken. The only theoretical problem associated with the practice of differential pricing occurs when the lowest fares become too low. If the low-fare option, which was offered only to fill "surplus" seats, is priced below its marginal variable costs, a situation of cross-subsidization between passenger groups exists, compromising both economic efficiency and consumer satisfaction [6]. It can be argued, however, that in the theoretically ideal situation in which the seats made available at lower fares are identifiable as surplus and full-fare passengers are not able to take advantage of the low-fare options, it would take an almost ridiculously low discount fare to raise valid concerns about cross-subsidization.

Application of the differential pricing concept in actual airline markets requires the identification of distinct segments of the total demand for air travel. Market demand segmentation involves dividing the total demand for a product into separate groups of potential buyers who might require or prefer different product characteristics. If differentiated air travel options can be offered to each demand segment, then price differentials can be applied more effectively. In the Boeing approach to surplus seat management, demand segmentation through the application of "restrictions" to the lower-priced service options is essential to maximize stimulated demand and to minimize the diversion of full-fare passengers to the lower fare options.

The development of an effective approach to seat inventory management is thus dependent on the attributes of the different service options and the market demand for each. Airlines have attempted to differentiate their reduced-fare options from their full-fare services to keep the cross-elasticities of demand between options low, thereby reducing the potential for diversion. The concept of ***product differentiation*** has been applied by airlines to their service options. Each option is in essence a ***fare product*** that can be defined by the restrictions on its purchase and use, differences in service amenities, and a price level. Under ideal conditions, each fare product should be associated with restrictions, amenities and a price that will make it attractive to a particular market demand segment.

Airlines traditionally have identified demand segments under the assumption, supported by empirical evidence, that there exist substantial differences in demand elasticities between business and leisure or vacation travelers, and little or no cross-elasticity [7]. While there are differences in the literature with respect to the specifics of airline market demand segmentation, there is a consensus that price and service elasticities of demand are the strongest determinants of demand segments [8, 9, 10, 11]. Total market demand is a function of both the fare levels offered and the levels of service provided by the airline and its competitors. Different segments of that demand will place different weights on the importance of low fares versus the reduced restrictions and/or increased service amenities associated with higher-priced fare products.

It is important at this point to emphasize the distinction between the airline's output, usually considered to be an available seat flown from one airport to another, and the airline product as purchased by the consumer. The demand for air travel is a derived demand, meaning the consumer does not purchase a quantity of available seat-miles as if they were a commodity. The value of air travel is related to being at a specific place over a certain time period. A ticket purchased for air travel is associated with a specific origin, departure time, destination and arrival time, which comprise an air travel "package". Included in this package are the attributes of the fare products being considered, such as ticket purchase and itinerary limitations, refundability, as well as amenities such as on-board space and service.

For the purposes of demand segmentation, airlines have recognized that differentiated fare products can in fact be offered at different price levels because of the price and service elasticities of air travel demand. Travelers very sensitive to price will base air travel decisions almost exclusively on a lowest fare criterion and will be willing to contend with the restrictions, reduced service amenities, and perhaps less convenient flight times or routings associated with the lowest-priced fare products. Conversely, passengers with time constraints and service-sensitive travelers will value level of service factors such as schedule convenience, travel flexibility and inflight amenities, to the point that price might not be a decision factor in selecting air travel. Of course, there exists a continuum between these two extremes, along which the majority of air travelers are likely to fall.

Given this continuum between extreme price sensitivity and extreme service level sensitivity, it can be difficult to divide the total demand in an airline market into well-defined segments. Most airlines have in the past attempted to use various socio-economic and travel characteristics of air travelers to identify distinct travel groups. A distinction was made between the business and leisure segments of total market demand, under the assumption that persons traveling on business are far more likely to be sensitive to time or level of service factors and relatively price-insensitive. At the same time, leisure travelers have been assumed to be more concerned about price, and perhaps less sensitive to certain service-related characteristics.

Airlines attempted to ensure that only the leisure travel segment purchases low-priced fare products through the imposition of ticketing and travel restrictions designed to "fence out" the full-fare business segment. The most commonly imposed fences were conditions requiring advance ticket purchase and round-trip travel with a minimum length of stay at the destination, defined so as to preclude those on business trips from making use of the discounted fares.

The use of "super-saver"-type fare products spread industry-wide in the United States in the mid-1970's, and airlines found that properly fenced low fares could in fact generate incremental leisure traffic. The problem that airlines faced, however, was that the seats being sold to leisure passengers at lower fares were not necessarily those that would have gone empty. Low-fare passengers wanted to choose the same peak period flights that were already popular with business travelers. Even worse, because advance purchase conditions required reduced fare passengers to book earlier than most full-fare passengers, passengers paying lower fares in fact displaced full-fare passengers from the most desirable flights. Airlines experienced flight loads that were not significantly higher, total revenues that were actually lower on peak period flights, and a loss of goodwill by business travelers frustrated by their inability to obtain seats on desired flights.

This experience led to the introduction of *capacity-controlled* reduced fares. Airlines recognized the need to limit the number of seats made available to their low-priced fare products and, furthermore, to distinguish between low-fare seat availability on peak and off-peak flights. More sophisticated statistical analysis was required to predict the number of full-fare passengers that could be expected for each flight departure, to establish an adequate availability buffer of empty seats above this expected full-fare demand and, then, to allocate any remaining seats to lower-priced fare products. The number of low-fare seats could thus vary by day of week and time of day for flights in a given market.

A methodology for managing capacity controlled fare products was presented by Boeing in a package it called the "Surplus Seat System" [12]. Boeing's approach used historical full-fare demand data by flight, day of week and season of the year to derive an estimate of expected full fare demand, also by flight. A growth factor was applied, and an optimal buffer determined to reduce full-fare reservations denials to a targeted level. Remaining seats were deemed to be "surplus" and made available to low-fare passengers.

1.2 Changes to the Surplus Seat Concept

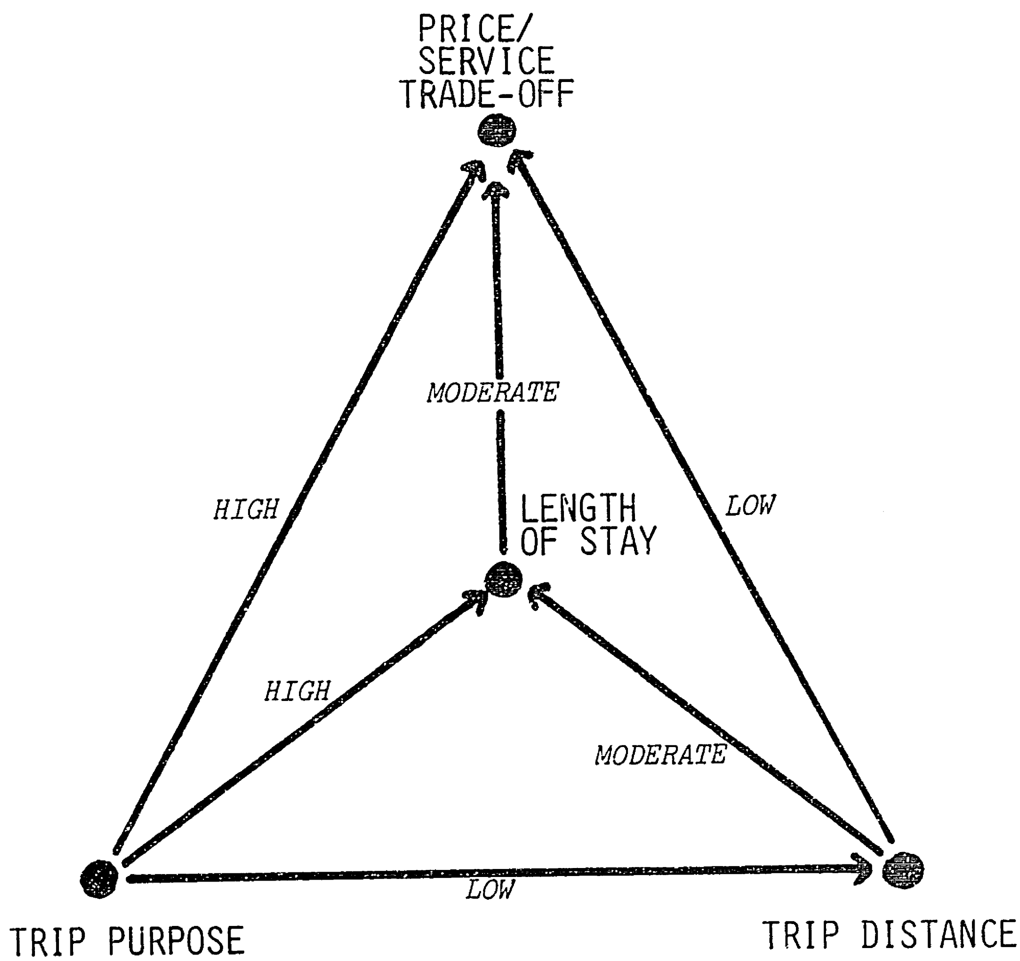
The basic principles of surplus seat management and capacity-controlled fares remain valid and continue to be used by airlines today. There have been changes, however, to the simple two-class demand segmentation and differential pricing model proposed originally by Boeing. Passenger behavior and demand characteristics have evolved, making the identification of distinct demand segments more difficult. Furthermore, the competitive environment of the airline industry has changed dramatically. With flexible pricing and competition from low-cost new entrants, it has become more difficult for any airline to use differential pricing to its fullest potential. And, given that less than 10 percent of airline passengers traveled at the full coach fare in 1986 [13], the notion that reduced fares represent a "surplus seat" product for airlines is no longer valid. Each of these changes is discussed briefly below.

First, market demand segments have become more difficult for airlines to identify. The business versus leisure classification of passengers adopted initially by airlines offering differentiated fare products overlooked all other types of non-business travel. Leisure (vacation) travel is but one component of the non-business or personal travel segment, which includes travel to visit friends or relatives ("VFR"), and personal emergency travel. Increases in the proportion of personal relative to business air travel have made the identification of price- versus service-sensitive demand segments more complicated. A survey of passengers conducted by the Canadian Transport Commission (CTC) in 1979 found that up to 45 percent of business and mixed business-pleasure trips in domestic markets could be considered price-sensitive, while 39 percent of non-business trips were in fact non-discretionary and thus service-sensitive [14].

Findings such as these indicate that market demand segmentation based solely on trip purpose might not be the most appropriate. The characteristics of price-sensitive versus on-demand travel suggest that, in most cases, a trade-off is made by the passenger between the price of an airline ticket and the need to have travel flexibility and service amenities, or even the need to travel at all. The CTC has undertaken numerous passenger surveys in an attempt to identify more clearly the main determinants of the price versus service trade-off, examining both trip and traveler characteristics. While traveler characteristics had "very little impact on price sensitivity independent of the type of trip being taken", trip characteristics showed a strong relationship with the price versus service trade-off [15].

Trip purpose, trip distance and length of stay were all examined for their influence on price and service sensitivity, and the results of the various CTC surveys can best be summarized in diagram form (see Figure 1.2). Trip purpose exhibited the strongest relationship with the price-service trade-off, although considerable overlap by price-sensitive business travelers and service-sensitive non-business travelers was evident in the results, as mentioned earlier. Trip purpose also showed a strong relationship with length of stay,

Figure 1.2: Trip Characteristics and the Price vs. Service Trade-off



Note: The arrows indicate the direction of influence, if any exists.

which in turn had only a moderate influence on the price-service trade-off. Trip distance showed little relationship with the other factors, apart from a moderate influence on length of stay.

Given that trip purpose seems to be the most important determinant of where a passenger lies along the price-service trade-off continuum, it is not surprising that airlines have traditionally focused on the business versus non-business demand segment definition. Many of the restrictions applied to low-priced fare products are designed to prevent the diversion of business travelers from higher fares. The overlap between the traditional market demand segments implies, however, that trip purpose as the sole demand segmentation criterion does not distinguish among the traveler groups that have emerged with the evolution of multiple fare product offerings. Just as there are business travelers willing to give up some service amenities in exchange for a lower fare, there are non-business travelers willing to pay more for better service and travel flexibility.

The concept of a price-service trade-off continuum, as described by the CTC studies, allows demand segments to be defined independently of trip purpose. Its disadvantage is that the continuum is one-dimensional — the consumer is assumed to be either price-sensitive or service-sensitive, but not both. Furthermore, the term "service-sensitive" is ambiguous, in that it refers to both the need for on-demand travel flexibility and a preference for other service amenities. A consumer's need to have travel flexibility reflects a sensitivity to the time elements of air travel (i.e., departure time, enroute time, and return time), which could well be independent of the many other "service" dimensions of a fare product.

Schweiterman [16] presented another approach to air travel demand segmentation, identifying three segments on the basis of a somewhat arbitrary breakdown of both the time- and price-sensitivity continua, as follows:

1. *Highly discretionary consumers,* traveling for personal reasons, willing to alter their travel plans over a large time "window" to obtain the lowest fares, and able to meet advance purchase and length of stay restrictions;

2. *Moderately discretionary consumers,* who will travel only with a discount from the full coach fare and are able to meet less stringent advance purchase requirements (up to 7 days in advance), but are unable to meet minimum stay requirements

of a week or more, and are not willing to accept significant deviations from their desired flight times;

3. *Non-discretionary consumers,* such as employees traveling for business purposes, not willing to compromise schedule convenience, and unable to meet any advance purchase or minimum stay conditions.

Schweiterman's three-segment model makes reference to trip purpose, but in fact relies implicitly on the notion of trip "value" to consumers in order to distinguish between discretionary and non-discretionary trips. The "value" of a trip to the consumer in Schweiterman's model is determined by the presence or absence of time constraints on being at a location at a desired time. The value of a trip may decrease, to zero in some cases, if these time constraints cannot be met.

We can develop a demand segmentation model that incorporates explicitly the notions of both the price-service trade-off and the value of a trip as determined by the consumer's time-sensitivity with respect to a given trip. Both concepts are important in the identification of demand segments in airline markets where a wide variety of fare products are offered. By separating time-sensitivity from price-sensitivity and its implied trade-off between price and service attributes other than time, we can characterize consumer groups without reference to trip purpose and at the same time avoid the overlap of the above demand segment definitions. Furthermore, we can define consumer groups independently of the fare products available at any particular time.

The availability of differentiated fare products at various price levels requires consumers to make a trade-off between the higher-priced fare products associated with high levels of service amenities and fewer restrictions, and the lower-priced fare products with greater restrictions and, in some cases, reduced service amenities. Given a range of fare product options and a budget constraint, each consumer will choose the fare product that maximizes some measure of "value", subject to the budget constraint. Alternatively, the consumer might choose the fare product that minimizes dollar costs, subject to a minimum "value" threshold. In either case, a choice is made between "value" and cost.

This choice is related to the price-sensitivity of the consumer for a given trip, and is not independent of the time-sensitivity of the trip being considered. The time-sensitivity of a trip is determined by the length of the "time window" over which a trip may be taken and still provide the consumer with a certain "value" of being at the desired location. A totally non-discretionary trip is one that must be taken at a specified time, meaning the acceptable time window for the trip is very short. On the other hand, a totally discretionary trip is one for which the acceptable time window for travel is extremely long. For the purposes of market demand segmentation, then, the time-sensitivity continuum incorporates the notions of discretionary and non-discretionary travel, without reference to trip purpose.

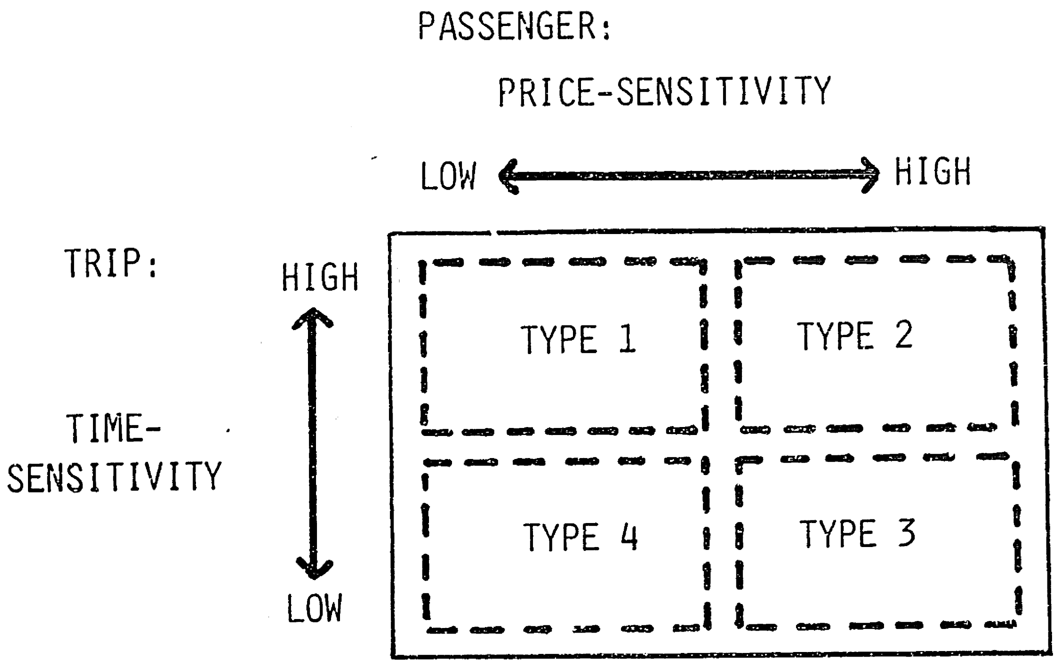
A consumer's location along both the price-sensitivity and time-sensitivity continua can differ from one planned trip to the next. An individual can be insensitive to price for an extremely time-sensitive trip, be it for business or personal purposes. The same individual can be extremely price-sensitive for a trip that has no time-sensitivity associated with it, again regardless of trip purpose.

Given that any one trip will lie somewhere along the time-sensitivity continuum, and given that the potential traveler must fall somewhere along the price-sensitivity continuum for each trip being considered, these two scales provide the basis for a demand segmentation model that includes all possible trip/consumer characteristics. We can define four generalized demand segments by dividing each continuum into two sections, as shown in Figure 1.3.

The market demand segments defined by Figure 1.3 are relevant only to the extent that differentiated fare products are offered to each segment. If only one fare product is available on all flights in a market, all four segments shown would be forced to purchase that fare product, or not travel at all. Given that a range of differentiated fare products is offered in most markets, we can identify the prominent characteristics of consumers in each quadrant of this segmentation model. The dividing lines between the segments are by no means exact, and each segment still encompasses a wide range of consumer characteristics and trip itinerary requirements. As such, no single definition of the typical consumer in each segment will apply to all consumers in that group. Nevertheless, the characteristics of most of the travelers falling into each segment can be described as follows:

1. *Type 1: Time-sensitive and insensitive to price.* These consumers prefer to travel on flights that meet their schedule requirements, and are willing to purchase the

Figure 1.3: Market Demand Segmentation Model



highest-priced fare products to do so. They might even be willing to pay a premium price for the extra amenities of a business or first class service. Travel flexibility and last-minute seat availability are extremely important to this group.

2. *Type 2: Time-sensitive but price-sensitive.* A large proportion of air travelers likely belong to this demand segment. These consumers must make a trip but are willing to be somewhat flexible in order to secure a reduced fare. They cannot book far enough in advance to obtain the lowest fares, although they might be willing to rearrange a trip to meet less stringent requirements if the savings are great enough. This group is willing to make stops or connections en route, and will accept less convenient flight times.

3. *Type 3: Price-sensitive and insensitive to time constraints.* These consumers are willing to change their time and day of travel, and even destination airports, to find a seat at the lowest possible fare. Some will be induced to travel by the availability of an extremely low fare. This group is willing to stop or make connections and can meet virtually any travel or ticketing conditions associated with a low fare.

4. *Type 4: Insensitive to both price and time constraints.* This group includes the relatively few consumers that have little or no time constraints for travel, yet are willing to pay for high levels of service, including the flexibility of flying with little advance notice on preferred flights.

For demand segmentation purposes, Type 4 consumers can be combined with the Type 1 group, since both are willing to purchase a high-priced fare product to secure a high level of service and/or travel flexibility, regardless of their trip purpose.

This market demand segmentation model is still broad in its definitions, and many airlines currently have in place fare structures with more than three types of product, as will be described. Nevertheless, the model includes all consumer and trip characteristics and as such can be used to classify smaller demand segments more precisely, if necessary.

Airlines abandoned the original two-segment market demand model due in part to a realization that the characteristics of consumer demand for air travel had changed. At the same time that defining demand segments was becoming more difficult, airlines recognized that differential pricing strategies targeted at multiple demand segments offered the potential of even greater increases in total flight revenues. The greater the number of segments that can be identified and differentiated fare products that can be offered for sale, the closer airlines can come to the idealistic model of a different price for each seat sold. In spite of the practical problems of identifying distinct demand segments, most airlines instituted multiple-tier fare structures.

Ironically, this movement to multi-tier pricing, when combined with increased price competition, actually made it more difficult for airlines to use the differential pricing concept effectively. In fact, it has been argued that the ability of airlines to charge different demand segments significantly different fares has actually decreased since deregulation [17]. With the development of hub connecting complexes and the entry of low-fare airlines, many airline markets receive at least connecting service from several carriers and exhibit substantial price competition. Under these conditions, carriers maintain a multi-level fare structure in name, but are often forced to make a large proportion of their seats available at lower fares without imposing the restrictions intended to differentiate the low-priced fare products and prevent diversion, rather than risking a loss of market share. The strategy of offering only one low-priced fare product, adopted by some new entrant carriers, has further undermined the differential pricing objectives of the major airlines.

The problem of identifying market demand segments and the obstacles to effective differential pricing practices, as discussed, lead to a third way in which the original surplus seat concept has changed. With an almost universal availability and usage of low-priced fare products, the term "surplus seats" may no longer be appropriate in the context of seat inventory control. The original surplus seat concept considered low-fare seats to be a by-product of the airline's operations, which had traditionally been devoted to serving the full-fare passenger.

The high proportion of passengers now carried at reduced fares indicates that the original surplus seat concept has been eroded severely. Airlines appear to be providing capacity for low-fare as well as full-fare passengers, in which case the low-fare products are a part of the airline's primary output. If this is in fact the case, economic efficiency would require the low-fare products to contribute in a substantial amount to airline capacity (fixed) costs, in addition to covering their marginal variable costs. The inability of airlines to segment market demand precisely or to charge radically different fares for different products in many markets suggests that such a contribution is being made.

While the differential pricing concepts and demand analysis techniques of the original surplus seat management model are still valid, their application in the current airline marketplace has become substantially more complex. Rather than estimating the demand for a primary product and allocating all remaining capacity to a marginal product, airlines must recognize that each of the several products they offer on a flight makes a revenue contribution to fixed costs. A seat inventory management system in today's airline environment must estimate the demand for multiple fare products and then allocate seats to each on the basis of the relative revenue contribution of each and the likelihood that seats will ultimately go empty. This approach underlies the mathematical models presented in Part Two of this dissertation.

1.3 Current Airline Fare Structures

A description of the fare products being offered in a particular market at any point in time could well be obsolete within 24 hours. In addition to the complexity of price levels, service amenities and restrictions, airline fare product offerings are characterized by the speed with which they can change. Given complete pricing freedom in domestic markets, U.S. airlines are able to make changes to their fare structures by origin-destination market instantaneously within their own computer reservations systems, and overnight through the Air Tariff Publishing Company (ATPCO). ATPCO routinely handles thousands of individual market changes to fare levels and/or rules in its electronic database each day.

It is therefore impossible to present a comprehensive description of current airline fare products that would apply to all markets and carriers, or which would remain valid for any length of time. It is possible, however, to describe the general structure of price levels, service amenities, and purchase/travel conditions that characterize the fare products offered by most major carriers in most U.S. domestic markets. Apart from markets in which price competition is unusually intense due to the presence of a low-cost competitor, and apart from fluctuations in price levels during seasonal system-wide "fare wars", the range of fare products available in most markets remains relatively stable. The discussion that follows focuses on the stable components of current fare structures, and explores them in the context of the differential pricing and market demand segmentation practices described earlier.

Airlines wishing to segment the total market demand for air travel to their revenue-maximizing advantage must design a range of fare "products" that will appeal to, and be used exclusively by, each of the demand segments. An ideal fare structure will minimize any "seepage" between segments, particularly the diversion of those willing to pay higher fares to the lower-priced fare products. Airlines use restrictions on reduced fares as disincentives to prevent downward diversion of higher fare passengers, and increased service amenities as an incentive for consumers to purchase higher-priced fare products. In fact, the *absence* of purchase or travel restrictions has come to be regarded as a service amenity of full fare products.

Advance purchase requirements, minimum stay conditions, and round-trip travel requirements have been applied to the lowest fares almost universally since the introduction of super-saver fares in the mid-1970's. While the nature of these restrictions has not changed, the levels of each restriction imposed on particular fare types change often, both with the level of price reduction involved and the degree to which airlines feel they must match competitors in both price levels and fare rules. Airlines have also constrained the use of certain reduced fares with routing and/or flight limitations, day-of-week/time-of-day restrictions and, more recently, cancellation/change penalties.

These restrictions on the purchase and use of low-priced fare products are often accompanied by capacity controls or limits on the number of seats available to particular fares or types of fares. The specific limits on low-fare seat availability are not known to consumers, and as such are not explicit attributes of different low-priced fare products as perceived *a priori* by the consumer. Rather, it is the set of conditions and service amenities associated with a fare product that determines the market demand segment to which it will appeal.

Advance purchase and minimum stay requirements are designed to keep Type 1 and 2 consumers from making use of the lowest fares. Since the first super-savers were introduced with a 7-day advance purchase requirement, advance purchase periods have been lengthened gradually, culminating with the introduction by American Airlines in 1985 of the "Ultimate Super-Saver", which requires a 30-day advance purchase. Over the same period, the minimum stay conditions have been reduced to the point that most excursion-type fares now require only a Saturday night away from the originating point before return travel may commence. This requirement replaced previous 7-day minimum stay conditions because airlines found that the reduction led to little or no increase in diversion of high-fare travelers. At the same time, it increased accessibility of the low-fare product to price-sensitive (Type 3) consumers able to stay away over a weekend but not for a full week [18].

Any fare requiring a minimum stay by definition requires a round-trip ticket purchase. Some moderately reduced fare products currently have *only* a round-trip purchase requirement, perhaps in conjunction with an advance purchase condition. The objective of a round-trip purchase rule is to keep highly service-sensitive consumers (Type 1) who cannot commit themselves *a priori* to a particular return flight or who might have more complex itineraries from using this fare product. The round-trip purchase requirement, in conjunction with one or more other restrictions, is an attribute of one or more of the low-fare products offered in over 95 percent of domestic markets [19].

The three traditional fences — advance purchase, round-trip travel and minimum stay — are effective against diversion only if they are enforced by the airline. While it might be possible for consumers to side-step an advance purchase requirement by dealing with a travel agent who is willing to backdate ticket issuance transactions, the minimum stay rule is also susceptible to abuse by travelers. Many airlines allow a passenger to stand by for the return portion of a round-trip reduced fare, and some passengers will routinely stand by for the inevitable empty seat on return flights, never having had the intention of meeting the fare's original minimum stay conditions. Although it is up to the airline ticket and gate agents to prevent such abuse, few agents will deny a passenger with a ticket as long as empty seats remain on a departing flight, especially if the passenger was ticketed originally for a competing carrier's flight.

The abuse of both the minimum stay and advance purchase requirements, primarily by Type 1 and 2 consumers, prompted airlines to introduce monetary penalties for changing or cancelling flight bookings for the lowest fares. The objective of these penalties is to prevent those passengers who are unable to commit to exact departure and return flights far in advance from purchasing several advance purchase low-fare tickets, using the one that proves to be most convenient, and then obtaining a full refund for the unused tickets. The imposition of a monetary penalty also makes it more difficult for passengers to confirm reservations for an alternate return flight without being asked for an additional fare payment by the airline, although it is still possible to stand by without penalty in many cases.

Non-refundability of tickets for the lowest-priced fare products was implemented industry-wide in early 1987 with the introduction of "MaxSaver" fare products by the Texas Air Corporation. The introduction of total non-refundability represented both an additional restriction designed to prevent diversion of passengers from higher-priced fare products and an opportunity for airlines to reduce the revenue impact of passenger no-shows and cancellations. With a 100 percent cancellation penalty imposed on their lowest fares, most major carriers extended cancellation penalties of 50, 25, and 10 percent to their less restricted and higher-priced fare products.

Advance purchase, minimum stay and cancellation/change penalties enable airlines to prevent the diversion of Type 1 and 2 consumers to the lowest fares. Several other types of fare product restrictions have also been used by airlines to target individual market segments more precisely, particularly in distinguishing Type 1 from Type 2 consumers. Non-stop/connecting flight, day-of-week and time-of-day limitations on certain reduced fare products can be found in most domestic markets.

A price-sensitive consumer planning a trip with little time sensitivity will likely accept all three of these constraints on his/her travel itinerary in order to obtain the lowest possible fare. The degree to which a time-sensitive *and* price-sensitive (Type 2) consumer is able to conform will depend on how price-sensitive he/she is, and the degree to which the restrictions impinge on the desired travel times. Survey findings indicate that, apart from advance purchase conditions, restrictions limiting the availability of low-fare products to particular travel days and/or times are the most effective market demand segmentation technique [20]. In addition to performing a demand segmentation function, time of travel restrictions also help the airline to fulfill another goal of the original differential pricing strategy — spreading peak period demand to fill otherwise empty seats on less popular flights. The availability of the lowest 30-day advance purchase fares for travel on Tuesdays and Wednesdays, for example, helped to reduce day-of-week variation in total demand tremendously [21].

The limitation of certain reduced fares to one-stop or connecting service only is an example of how airlines practice value-based pricing and product differentiation. Booking, handling and carrying a passenger on more that one flight leg or over a circuitous routing will cost the airline more in variable costs than putting the same passenger on a non-stop flight. As long as seats are available for both itineraries, there is little cost-based rationale for charging $520 for an unrestricted non-stop fare from Boston to Los Angeles and $250 for an unrestricted (but capacity controlled) one-way fare that is available only on one-stop or connecting flights in the same market [22]. These two fare products are targeted at Type 1 and 2 consumers, respectively. The latter is more likely to accept a longer travel time to save over 50 percent of the full coach fare, and to change flight times and/or carriers if the capacity-controlled low-fare seats are sold out on the first flight(s) requested.

The fare conditions described above are all "fences" designed to differentiate fare products and prevent diversion of passengers whose service sensitivity causes them to view fare products with such fences negatively. Airlines have also used positive forms of product differentiation to segment market demand. First class and business class services priced at a premium above the full coach fare are targeted at consumers with little or no price sensitivity. While first class service has existed for decades, the concept of an intermediate class between first and coach in terms of in-flight service amenities has emerged since deregulation and is yet another attempt by airlines to segment total market demand more precisely.

Business class options are offered in almost all international markets by most airlines, where long-haul flights make more legroom and a higher level of inflight service very attractive to those already paying a full coach fare. Trans World Airlines offers its Ambassador Class product domestically as well, on its wide-bodied aircraft. Air Canada has recently introduced a similar Executive Class on long-haul domestic and Canada-U.S. routes. Both carriers price their business class products just slightly above the full coach fare, such that the premium is generally less than 10 percent. In many respects, these business class products have replaced the standard full-fare coach products designed originally for business travelers.

Several carriers have also experimented with providing full-fare coach passengers with added service amenities, ranging from separate check-in counters and VIP lounge privileges at airports to complimentary ground transportation services. United Airlines at one point guaranteed that a full fare passenger would not be denied boarding due to overbooking and, more recently, stopped giving advance seat assignments to passengers with the lowest discount fares. While U.S. carriers have met with resistance from passengers and travel agents in their attempts to enhance the full coach fare product, Canadian carriers have been doing just that for years. Air Canada's Connaisseur Class and CP Air's Empress Class are services to which only full coach fare passengers are entitled. Both carriers seat full-fare passengers in a separate area of the economy cabin and provide them with amenities like free drinks and movies, first choice of meals, and advance seat assignment.

The use of an increasing variety of low-fare restrictions and the inclusion of service amenities with full fare tickets represent attempts by the airlines to differentiate the range of fare products they offer, particularly those that will share seating in the coach compartment of the aircraft. Significant product differentiation is a major determinant of effective demand segmentation, which in turn allows the airline to practice differential pricing to maximize revenue. If the fare products offered by an airline are differentiated to the point that consumers understand why a reduced fare product is priced far lower that the full fare product, the airline's fare structure is entirely rational.

Unfortunately, this type of ideal fare structure can be found in very few markets. If one is found, it is unlikely to remain stable for any length of time. With the growth of hub-oriented route systems, most domestic markets are now served by several airlines, and the choice of carriers in a market often includes at least one low-fare, single-product competitor. Even if the price leader does not offer a level of service in a particular market, experience has shown that most of the competing carriers will match the low price with a similar product. Furthermore, sporadic "fare wars" and regional price battles serve to undermine the possibility of offering a standardized differentiated range of fare products in all markets. It is under such circumstances that airlines and consumers alike marvel at the complexity and incomprehensibility of airline industry pricing practices.

It is possible, nonetheless, to develop a generalized fare product typology that is representative of the range of fare products found in most domestic markets in the spring of 1987. Figure 1.4 shows the typical fare product categories offered by airlines, the primary market demand segments targeted by each product, as well as the service amenities, restrictions and relative price levels of each. Not all types of fare products offered by all carriers in all markets are included in Figure 1.4, as the restrictions on individual fare products will depend on the competitive conditions in each O-D market. Similarly, carriers might not offer all the product categories listed in markets where there is little effective competition.

Figure 1.4: Typology of Airline Fare Products, Spring 1987

FARE PRODUCT	TARGETED DEMAND	RESTRICTIONS ON PURCHASE AND/OR USE	SERVICE AMENITIES	APPROX. PRICE (% of COACH)
FIRST CLASS	Type 1	None	Separate cabin; wider seats; deluxe meals; priority check-in.	150
BUSINESS CLASS	Type 1	None	Separate cabin; less dense seating; free cocktails and movies.	110
FULL COACH FARE (Y)	Type 1	None	Coach cabin service; advance seat selection.	100
SUPER COACH FARES	Type 2	0-3 day advance purchase; 1-stop or connecting flights; limited seat availability.	Coach cabin service.	70-85
EXCURSION FARES	Type 2	7-14 day advance purchase; Sat. night minimum stay; round-trip purchase required.	Coach cabin service.	50-70
SUPER SAVER FARES	Type 3	21-30 day advance; Sat. night minimum stay; cancel/change penalties.	Coach cabin service.	30-50
"MAX" SAVER FARES	Type 3	2-day advance purchase; Sat. night minimum stay; non-refundable; very limited seats.	Coach cabin service.	20-30

First class, business class and full coach fare products are targeted at the Type 1 consumer, and are differentiated only by their service amenities and relative price levels. The most important attribute of all three of these fare products is their lack of restrictions. A passenger may purchase a ticket at any time prior to departure and make virtually unlimited changes to both the outbound and return flight itineraries, all without paying a penalty. At any time, unsold seats will always be available to passengers purchasing these unrestricted products.

Super coach fares and excursion fares are moderately priced products targeted at the Type 2 consumer. Super coach fares appear primarily in markets where a low-fare competitor offers unrestricted fares below the established coach fare. Depending on the extent to which a carrier wants to respond to the low-fare competitor, it might attach a minimal (3-day) advance purchase requirement or a one-stop/connecting flight limitation to its super coach fares to prevent Type 1 consumers from diverting to this lower fare. Most of the established carriers place capacity controls on their super coach fares to keep some seats available for full coach fare passengers right up to departure time.

Excursion fares as defined in Figure 1.4 are essentially the traditional round-trip reduced discount fare products introduced in the mid-1970's. Originally targeted at leisure travelers, these fares are being used more often by business travelers willing to extend a business trip over a weekend or to combine business and pleasure travel to obtain a lower fare.

The super-saver fare product category is targeted exclusively at the price-sensitive Type 3 demand segment. Under stable competitive conditions, these fare products will carry advance purchase, minimum stay and cancellation/change penalty conditions that only Type 3 consumers can meet. Such stability is short-lived in highly competitive markets, however, as carriers will reduce the severity of these fences to stimulate traffic or in response to similar actions by a competitor. The net result of such reductions in fence levels is inevitably a higher diversion rate of Type 2 and even Type 1 consumers to the poorly-restricted lowest fares.

The lowest-priced fare product category includes the "MaxSaver" fares introduced recently and available at least through May 1987 on all major U.S. carriers. These fare products are intended for Type 3 consumers, as they require ticket purchase within 24 hours after a reservation is confirmed and are totally non-refundable once purchased.

The advance purchase requirement, however, is only 2 days prior to departure, meaning that price-sensitive Type 2 consumers will try to obtain these fare products just days before a trip. The Saturday night minimum stay requirement will not be a deterrent to diversion in the many markets where the round-trip "MaxSaver" is priced below the lowest available one-way fare. The advance purchase requirement for these fare products is to be increased to 7 days for travel during the peak summer period.

Each of the product categories presented in Figure 1.4 can in fact contain several specific *fare basis codes* , or published fares, with their own set of rules, effective dates, and price levels, all in the same O-D market. For example, a "QE30X23S" fare basis code might be defined by an airline to describe a specific super-saver ("Q") excursion ("E") fare product that requires a 30-day advance purchase ("30") and applies for travel on Thursdays through Mondays ("X23"), but only on one-stop or connecting flights ("S"). A "QWE21" fare basis might be available in the same market at the same time, describing a different fare product that requires a 21-day advance purchase, weekend travel, and is available on all flights (subject to capacity controls). The coding of fare bases is entirely at the discretion of creative minds in airline pricing departments, although imitation has led to some degree of uniformity of fare basis codes among airlines and across markets.

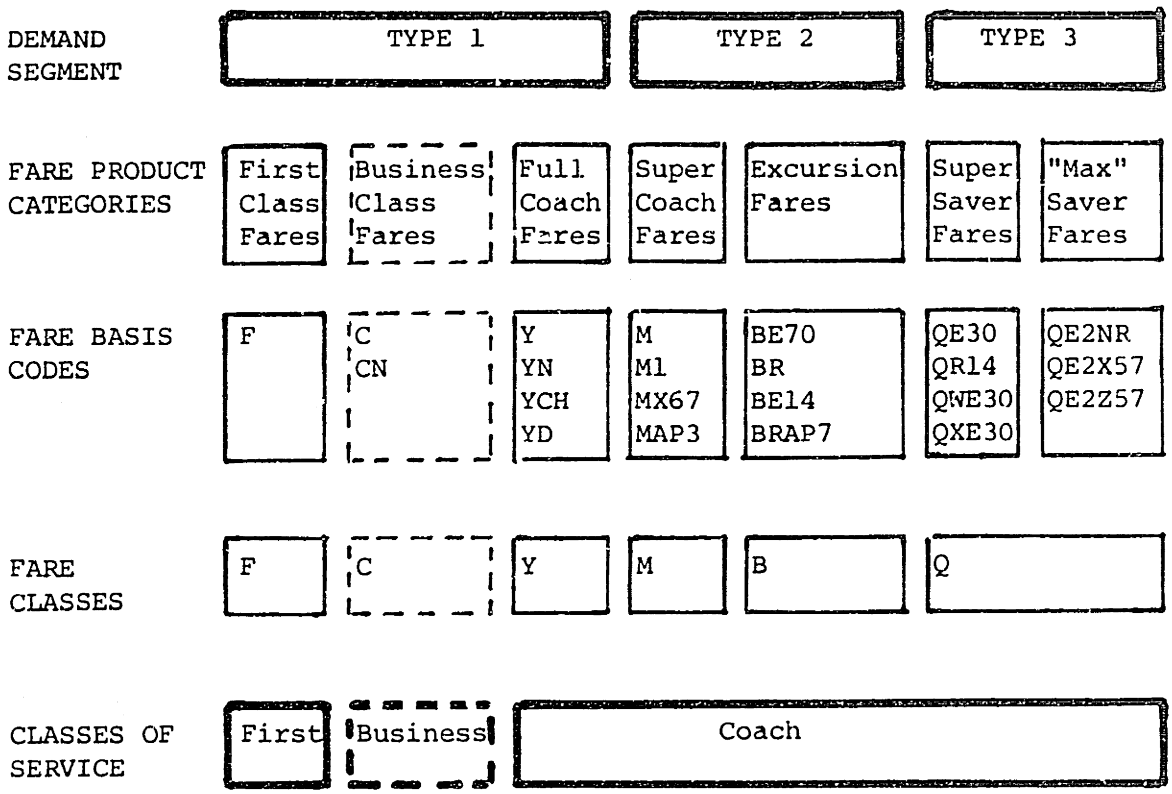
Fare bases are grouped by the airline into *fare classes* for the purpose of accepting and controlling bookings in its reservations system. Ideally, the reservations system fare classes would correspond exactly to the fare product categories described above, such that distinct seat inventories could be made available for each category. Most airline reservations systems, however, are limited to five primary fare classes for any one flight. The implications of this limitation for seat inventory control are discussed in Chapter Three. At this point, it suffices to note that airlines offering more fare product categories than they have reservations system fare classes must combine more than one product category into one fare class.

All the fare basis codes that comprise a particular fare class begin with the same letter, by convention. As a historical rule, separate physical compartments on an aircraft have been assigned their own fare class. Thus, a carrier offering first and business classes in addition to coach uses up three of five fare classes on Type 1 fare products. The letter codes used for each fare class can differ by carrier, although "F" for first, "C" for business and "Y" for full coach fare are industry standards. The lowest-priced fare products are commonly in Q-class, while all intermediately priced fare products might be listed as M, B, K, V, S, L, or H-class fares. Because not all fare products and fare classes are associated with distinct physical compartments on board the aircraft, passengers booked in two or more fare classes can in fact receive the same *class of service* in flight. On most carriers, all passengers paying full coach fare and less travel in the coach cabin of the aircraft and receive identical on-board service.

A summary of the relationship between market demand segments and airline fare structures as described in this section is provided in Figure 1.5. Note that each airline has flexibility with respect to which fare basis codes it groups into one fare class. Seat inventory control is necessarily related directly to the reservations system fare classes, as defined by each airline, since only the seats made available to a fare class "inventory" can be controlled. The airline must decide how many of the total coach compartment seats on a future flight it should make available to each of the fare classes sharing that compartment.

For the consumer wishing to make a reservation for a future flight, the availability of a specific fare product on the preferred flight departure depends on the availability of a seat in the associated reservations fare class inventory. If the fare class has been closed down to further reservations, the consumer is then faced with the option of accepting a different fare product in another fare class on the same flight or making inquiries into the availability of the same or similar fare product on other flights and/or carriers. This decision process is discussed in detail in the following chapter.

Figure 1.5: Fare Structure Relationships



Chapter 2

Consumer Decisions and Air Travel Demand

The availability of differentiated fare products, each associated with specific amenities, travel conditions and price levels, has changed the way in which consumers make air travel choices. Because many fare products are associated with restrictions that relate directly to the timing and nature of the trip being considered, consumer decisions with respect to air travel have become more closely tied to the total cost to be incurred on a trip. The consumer no longer buys a ticket for air transportation between A and B independent of other trip-related decisions. Selecting a particular fare product over another can have a significant impact on the total cost and timing of the trip.

This chapter explores the consumer decision process that has evolved with the institution of multiple-tier pricing strategies by airlines. With reference to classical microeconomic utility theory, it presents a conceptual model of how individuals choose from among the available fare products, alternative flights and competing carriers in an air transportation market. This choice process for individual consumers is then used to develop a description of the aggregate demand for a fare product on a particular flight, as viewed by the airline through its reservations system. The patterns of reservations demand described by the aggregate model are those that will influence the airline's ability to forecast demand and allocate seats to different fare products in a revenue-maximizing manner.

2.1 Individual Consumer Choice

A plausible model of individual choice for air travel decisions is one that is based on a valid theory of consumer decision-making and one which can be adapted to the context of transportation decisions in general and air travel options in particular. This section begins with a brief overview of the decision theories that seem to be best suited to the air travel choice problem, and then extends these theories into a qualitative choice model that deals explicitly with the selection of fare products and flights.

Classical micro-economic consumer theory makes use of the notion of individual utilities to explain the choice behavior of consumers faced with a range of goods and a limited budget to spend on these goods [23]. Consumer choice models based on the criterion of utility maximization assume that each individual, when confronted with a choice between "bundles" of goods, is able to compare alternatives and rank them in terms of his or her personal preferences. By using this "ordinal utility function", a rational consumer will then choose the bundle with the maximum perceived utility.

An important component of classical consumer theory is the assumption that each individual undergoes a *rational* decision process in making each and every choice of goods before making a purchase. This rational process requires the consumer to formulate explicitly his/her preferences for all possible combinations of goods that can be purchased, to identify all the alternatives that are available, to characterize each alternative in terms of its attributes and the weights assigned to each attribute, and to select the alternative yielding the greatest utility [24]. The rigid definitions of this rational decision process were relaxed by Simon [25], who introduced the notion of bounded rationality, recognizing the constraints on rational decision-making imposed by limitations on both information availability and human capabilities to process large amounts of information. A further relaxation of the underlying behavior in individual choice involved the concept of random utility [26], which takes into account the observation that individuals tend to choose the alternative that *appears* to maximize utility at the time the decision is made.

Regardless of how rigidly we interpret the notion, utility maximization seems to be a reasonable basis for the development of an individual choice framework for air travel decisions. Several other components of the classical micro-economic model of consumer choice can also be incorporated into this framework, although modifications to account for the unique nature of transportation demand are required. Most important of these is the derived nature of transportation demand [27]. That is, few individuals travel for the sake of travel itself. They travel to be at a particular place at a particular time. The benefits associated with travel are thus almost entirely attributable to the activities to be undertaken at the destination. Travel itself imposes costs or disutilities which must be incurred to realize the benefits of the trip.

A related characteristic of transportation is that it is purchased not for its quantity, but rather for its attributes [28]. This is also true of many commodities, but is particularly applicable to transportation. As mentioned in the previous chapter, consumers do not purchase a quantity of seat departures or available seat-miles as if they were a commodity. A trip taken by air consists, at a minimum, of a departure time from a specified origin and an arrival time at the desired destination. With respect to individual choice, the attributes of a particular travel option are what generate disutility. The disutility of a travel option will depend on the nature of the planned trip, the characteristics of the trip-maker, and on the attributes of the other available travel choices for that trip.

The "price" of a travel option is therefore not simply the monetary cost of the ticket. Various travel options can involve additional disutilities and monetary costs associated with the timing of the travel components of the trip, as well as the value of the time actually spent traveling. Given the derived nature of transportation demand, consumers will want to minimize travel time, cost, discomfort and inconvenience for a trip with a given level of perceived benefits. The concept of utility thus represents a generalized function that takes into account the pleasant and unpleasant components of making a trip and can form the basis of consumer choice [29].

The benefits of being at a destination are generally independent of the attributes of the available travel options. Travel choice depends on the minimization of the disutilities of traveling to the desired destination. The value of the utility function for a particular travel alternative is a measure of the degree to which that alternative is desirable to the consumer, relative to other alternatives. The utility function reflects trade-offs that the consumer is implicitly or explicitly using to compare the attributes of the various travel alternatives.

A descriptive model of individual choice for the air travel decision process must take into account not only the availability of heterogenous alternatives, but the fact that individual decision-makers will have different choice sets and can assign different utility values to the same attributes of the same alternative. The choice set for each individual will include the alternatives that are feasible to him/her, given the constraints on trip timing and costs. This choice set might only include a few alternatives if the consumer is unable or unwilling to obtain information about more options, or if the first option considered is found to be acceptable.

The air travel choice process for an individual is triggered by a perceived benefit of being at a particular destination to participate in activities there. The individual evaluates the travel options available for a given trip, and the costs or disutilities of each. If the benefits of being at the destination exceed the disutility of one or more travel alternatives, the alternative that minimizes disutility will be chosen. If the travel disutility of all alternatives exceeds the perceived benefits at the destination, the individual will choose not to make the trip.

We begin the development of a choice framework by assuming that the consumer wishes to select one of three fare products, F_1, F_2, F_3, offered on a single flight which happens to be the consumer's preferred flight. That is, we assume that the expected benefits of making the trip have been determined, the desired departure time and airline have been established, and the consumer must only choose among three fare products, all of which are available. With reference to the above discussion, we can postulate that the consumer will choose the fare product that minimizes his or her disutility of traveling to the destination.

Each fare product is defined by its *attributes*, which contribute to the total disutility of that fare product. Negative attributes create added inconvenience for the traveler, resulting in real or perceived costs that are higher than the dollar price paid for the fare product. The relevant attributes of each fare product may be grouped as follows:

1. *Price*, (P_i), the dollar price charged for fare product F_i;
2. *Service Amenities*, (S_i), the positive attributes related to the quality of service received in conjunction with the purchase of F_i. On-board service amenities such a meal quality, complimentary cocktails and headsets, and preferred seat assignments can differ among fare products, as can airport amenities like separate check-in facilities and VIP lounge privileges. Currently, the fare products offered by many

domestic U.S. airlines in a shared economy cabin involve few distinctions in service amenities.

3. *Restrictions,* (X_i), the conditions associated with F_i, including required advance booking and ticket purchase, cancellation and refundability limitations, effective and discontinued dates, minimum and maximum stay requirements, and applicability to particular times, days of the week and/or flight itineraries. Fare products priced lower than the standard full fare generally have one or more of these restrictions, the severity of which increases as the price decreases.

The total disutility of each fare product under consideration can be represented by the notion of its "generalized cost", denoted $Z(F_i)$ for fare product F_i. The generalized cost of each fare product to the consumer will be a function of its attributes, for example:

$$Z(F_1) = Z(P_1, S_1, X_1) \tag{2.1}$$

Each passenger will place different values or utilities on the relative importance of each of these attributes in deriving the generalized cost of each fare product option, depending on his or her set of values, as determined by both the price-sensitivity of the individual and the time-sensitivity of the trip being considered. The relative values of $Z(F_1)$, $Z(F_2)$, and $Z(F_3)$ may be different, requiring each individual to evaluate and rank the alternatives with respect to generalized cost, in order to identify the fare product that minimizes perceived disutility.

This simple example is unrealistic in that it assumes that all fare products for the passenger's first choice of departing and/or return flights are available. The choice model must therefore be expanded to incorporate the added generalized cost and reduced utility associated with the consumer's not being able to obtain the preferred fare product on his/her first choice of flights. There will be separate costs and disutilities associated with displacement from the preferred outbound and return flights and times. Both the outbound and return flights are included because most air travel decisions involve round-trip travel. The benefits of being at a particular destination must be weighed against the disutility of both outbound and return travel.

We assume that each consumer has determined the time period over which being at the desired destination and/or away from the origin will generate a benefit. On the basis of this time period, a "time window" of departure and return times can be deduced. Within each time window, there is some "ideal" set of departure times for the outbound and return flights, independent of any knowledge of actual flight schedules or seat availability.

The consumer begins the process of obtaining information (from a travel agent or from the airlines directly) about the actual alternatives that might be available for a particular trip. The specific way in which this choice set is developed will depend on the consumer's approach to selecting particular flights. For example, an infrequent air traveler might contact a travel agent to determine which flights and fare products are available. A more experienced air traveler might establish a choice set of preferred flight alternatives from published airline schedules before determining the availability of fare products on each.

Depending on the amount of information the consumer is willing to gather, this choice set might include as few as one or as many as tens of alternatives. If the first alternative determined to be available comes close enough to the "ideal" departure times for the individual and a fare product at an acceptable price level is available, other alternatives might not even be considered. If, on the other hand, the consumer is willing to gather additional information in an effort to find alternatives with a a lower generalized cost, the choice set can keep growing in size. Presumably, the consumer is ranking the alternatives as they enter the choice set, or at least keeping track of the alternative with the lowest disutility.

Given that particular fare product/flight combinations might be "sold out" or simply infeasible, the disutility associated with an alternative must include the costs related to the entire "travel alternative". Each alternative, T, represents a set of trip components, including an outbound flight itinerary, D, a return flight itinerary, R, and a fare product, F, the attributes of which were defined earlier. The D and R components of a travel alternative represent the attributes of a particular flight itinerary for outbound and return travel, respectively. These attributes include the actual flight departure time, t_d, relative to the "ideal" departure time, t_d^*, as determined by the consumer, and the enroute time of a particular flight itinerary, t_{er}, which is a function of the number of stops and connections, and the routing taken. The generalized cost of a particular outbound flight itinerary, D, can therefore be defined as:

$$Z(D) = Z(|t_d - t_d^*|, t_{er}) \tag{2.2}$$

The expression for $Z(R)$ will be identical.

Each travel alternative consists of three components, $T = \{F, D, R\}$, each of which contribute to the generalized cost of the alternative. Each travel alternative must consist of a *feasible* combination of fare product, outbound and return flight itineraries. The restrictions associated with a particular fare product, X_i, might constrain the combinations of D and R that may be included in the choice set. For example, an outbound flight next week cannot be combined with a fare product that requires a 30-day advance purchase into a feasible air travel alternative. The restrictions on a fare product could also exclude specific D or R components, for example, by requiring the fare product to be used only on connecting itineraries.

Given a range of outbound and return flight itineraries offered by competing airlines in a market, there is a large number of air travel alternatives, $T = \{F, D, R\}$, that can be constructed from the components F, D, and R. In terms of individual choice, however, all but a relatively small subset of these combinations will be eliminated by the timing constraints established by the consumer for a particular trip. These constraints are determined by the "time window" established for activities at the destination, over which the traveler will be able to realize a benefit, as well as by any time constraints on being away from the origin.

For each individual, then, the generalized cost of each travel alternative is a function not only of $Z(F_i)$, but of the disutility of the D and R components of the alternative:

$$Z(T) = Z\{F, D, R\} \tag{2.3}$$

The generalized cost of the F component includes the dollar price, P_i, of the fare product, and the disutility of meeting all of the restrictions, X_i, of that fare product, perhaps adjusted by the utility of any differences in service amenities, S_i, associated with the fare product. The generalized costs associated with both the D and R components include the dollar costs and disutilities of displacement from the "ideal" flight departure times, $|t_d - t_d^*|$, and of enroute time, t_{er}. These costs can include monetary expenditures in addition to the value of travel time, for example, increased hotel or meal expenditures.

The critical assumption of the choice model developed here is that individual passengers implicitly (or perhaps explicitly) ***rank*** the set of feasible air travel packages in their choice set with respect to the relative disutility of each. As mentioned earlier, an individual might not consider a full range of alternatives and is likely to seriously consider only the current option with the lowest disutility. The choice process can therefore be iterative, in that the consumer can be induced to expand his/her choice set as information about the (un)availability of preferred travel alternatives is gathered. The *availability* of a particular D or R component depends on the airline having seats to sell over the flight itinerary in question, in the fare class inventory from which fare product F must be sold. The ultimate choice is made by the consumer so as to minimize the total disutility of the travel required for a particular trip, *subject to* the availability of each successively less attractive travel alternative.

As an example, consider the hypothetical choice set $\{T_1, T_2, ..., T_n\}$ as a list of feasible travel alternatives formulated by an individual, ranked in increasing order of disutility. Similarly, $D_1, ..., D_n$ and $R_1, ..., R_n$ are the ranked outbound and return flight intineraries, while F_1 and F_2 are the relevant fare products offered on these flights. If the consumer prefers F_1 to F_2, then the rule of dominance dictates that the first choice of options for this consumer would be $T_1 = \{F_1, D_1, R_1\}$.

The first six travel alternatives in this hypothetical consumer's ordered choice set might look like this:

$$\begin{aligned} T_1 &= \{F_1, D_1, R_1\} \\ T_2 &= \{F_1, D_1, R_2\} \\ T_3 &= \{F_1, D_1, R_3\} \\ T_4 &= \{F_1, D_2, R_1\} \\ T_5 &= \{F_2, D_1, R_1\} \\ T_6 &= \{F_2, D_1, R_3\} \end{aligned} \tag{2.4}$$

We can infer from this ordered choice set that, for this hypothetical passenger, the marginal disutility of accepting the second and third-ranked return flight itineraries is smaller than the marginal disutility of accepting the second choice of outbound flight itineraries. Only after determining that D_2 is not available at F_1 will this consumer accept what is most likely a higher-priced F_2, but only on the first choice outbound and return flight itineraries. T_6 illustrates how a shift in the fare product component of the air travel alternative might rule out one or more flight itineraries (in this case R_2). The restrictions associated with F_2 might eliminate R_2 as a component of a feasible travel alternative.

Note that the relative disutilities of F, D, and R are determined in combination by the individual, meaning there need not be a systematic ranking of the separate trip components. The individual chooses the highest-ranked travel alternative that he is aware of at the time of booking or ticket purchase, subject to the *feasibility* and *availability* of the various combinations.

Given that one or more alternatives might not be available because all F seats have been sold on the D or R components of the trip alternatives in question, the choice process is simply one of systematically checking each ranked travel alternative in ascending order of disutility until a feasible *and* available combination is found. The extent to which an individual is willing to accept a lower-ranked travel alternative will be a function of the generalized cost of that alternative relative to the perceived benefits of being at the destination point. If we define the total of these perceived benefits to be W, then the individual will make the trip as long as the generalized cost of the highest-ranked feasible and available travel alternative, T^*, is less than W. That is, the trip will be made only if:

$$Z(T^*) \leq W \tag{2.5}$$

If more than one feasible and available travel alternative has a generalized cost less than W, then the individual will choose T^* so as to maximize $W - Z(T^*)$.

The value of W for a particular trip thus establishes a "cut-off point" in the set of ranked alternatives, below which an alternative will not be accepted by the consumer even if it is available. At this point the disutility of the travel alternative, due to displacement from preferred flight times and/or much higher monetary costs, is so great that the trip itself becomes of no value to the individual. The location of this cut-off point on the ranked list of alternatives will be determined by the characteristics of the trip and of the consumer, as defined by the market demand segmentation criteria described in Chapter One. That is, each $Z(T^*)$ is a function of the time-sensitivity of the trip in question and of the price-sensitivity of the individual with respect to that trip.

The criteria that define the market demand segments also affect the willingness of the individual to accept lower-ranked trip alternatives and the disutilities of shifts in the

F relative to the D and R components. For example, a consumer wishing to take a trip for business purposes to attend a meeting at a specified time and who cannot depart before a certain time nor return after a certain time will place a greater disutility on accepting lower-ranked flight itineraries than lower-ranked fare products. In the demand segmentation model, this trip would be classed as time-sensitive, and the consumer would be relatively insensitive to price.

Although we assumed initially that the consumer is indifferent between competing airlines, different flight itinerary components could in fact represent travel on different carriers. If the relative utilities of different flight itineraries are judged to be different by the consumer on the basis of attributes other than those associated with D or R, as defined above, the consumer exhibits a preference for one carrier over another. Such a preference can be incorporated into the fare product (F) component of the trip package, and will be reflected in its generalized cost through the level of service attributes (S_i) associated with a fare product offered by a particular carrier.

To distinguish fare products offered by different carriers, we can define F_{ik} to be fare product i offered by carrier k. The generalized cost of F_{ik} is then given by:

$$Z(F_{ik}) = Z(P_{ik}, X_{ik}, S_{ik}) \tag{2.6}$$

This formulation allows for the possibility that similar fare products offered by competing carriers in the same market will have different price levels and/or different restrictions, in addition to any perceived differences in service amenities. A preference for a particular carrier on the basis of "brand loyalty" or frequent flyer program considerations can also be reflected in different utilities assigned to S_{ik} for the same fare product offered by different carriers.

If the values of $Z(F_{ik})$ are determined to be different for the same fare product offered by different carriers, the consumer would then select the fare product/carrier combination that minimizes $Z(F_{ik})$, with all else being equal. It is more likely that distinguishing between identical fare products offered by competing carriers will involve different D and R components in the travel alternatives being considered. The consumer must then take into account the relative disutilities of the different flight itineraries on different airlines. It is possible that increased disutility of the D and R components involving the preferred airline will be outweighed by the lower value of $Z(F_{ik})$ relative to that of other airlines. That is, the utility of traveling on the preferred carrier can exceed the disutility of less convenient flight itineraries.

The individual choice process for air travel described in this section relies heavily on the notions of relative utilities and rational consumer behavior. It is important to emphasize, however, that very few potential air travelers will undergo an exhaustive enumeration and evaluation of alternatives before making a completely informed utility-maximizing choice. As mentioned, the number of alternatives considered and the way in which information is gathered can lead to the consumer making a less than optimal choice. Nonetheless, the choice process described remains valid insofar as the consumer will select the trip package alternative providing the lowest *perceived* disutility from among those *considered,* given *feasibility* and *availability* of that alternative.

It should also be emphasized that the framework presented here is a qualitative description of individual consumer choice in the context of current airline marketing and pricing practices. Development of quantitative models of individual choice for air travel would require empirical research that is well beyond the scope of this dissertation. This qualitative choice model is important, however, to an understanding of the reservations process as viewed from the airline's perspective. To understand aggregate reservations patterns, we must understand the individual choice process that generates these patterns.

2.2 Airline Reservations Framework

In accepting a passenger's reservation, the airline decrements the inventory of seats available in one of the fare classes established in its reservations system. Seat inventory control techniques are used to place limits on the maximum number of bookings that may be accepted in each reservations fare class. The aggregate outcomes of the choice process of individual consumers are thus viewed by airlines in terms of reservations totals and booking patterns by future flight leg and fare class. The objective of this section is to extend the model of individual choice, given a reservations system that presents available fare product and flight itinerary alternatives to the consumer.

If all fare products were always available, aggregate demand for each fare class on each future flight leg to be operated by an airline would be the sum of all requests from potential passengers for whom the relevant travel alternative $\{F, D, R\}$ ranks first in perceived utility. In reality, there is a limited number of seats available for each alternative in the choice set. Requests for different alternatives arrive over time before flight departure. If either the D or R components of an individual's preferred travel alternative are unavailable in conjunction with the desired F at the time the request is made, the potential passenger will be forced to consider and perhaps accept successively less desirable alternatives from his/her ordered choice set. For the airline, each shift by a potential passenger from one travel alternative to the next has important implications with respect to total bookings and expected revenues for future flights.

The airline receives a request from a potential passenger for a particular outbound flight (D), return flight (R), and fare product (F). If the passenger's preferred travel alternative is not available in its entirety, the passenger consider the next-best alternative. This next-best choice will include one or more components different from those in the initial alternative requested. Depending on which of these components change in the shift to a less desirable alternative, the impacts for the airline will differ.

If the next-best choice includes a different F component but identical D and R components, the potential passenger is willing to accept a fare product, with its price, restrictions, and service amenities, that has a higher generalized cost in favor of keeping the outbound and return flights of the initially requested alternative. For most passengers, a shift to the next-best alternative will mean a shift to a higher-priced fare product, given no change in the flights involved. Extremely price-sensitive passengers would of course be more willing to shift flights than accept a higher fare.

For the airline's purposes of seat inventory control, it is the upward shift in the fare product component associated with the next-best alternative in the consumer's choice set that is of most interest. Given the objective of maximizing total flight revenues, the airline must consider the possibility that a passenger denied a reservation in a low fare class will be willing to step up to the next highest available class for the same D and R flight itineraries. Stepping up to the next highest fare class represents what we shall call a ***vertical shift*** or "upgrade" in the reservations request, from the airline's perspective. Such a shift is desirable for the airline, since no potential traffic is lost and total revenues actually increase.

Unfortunately, not all potential passengers will be willing to make this vertical shift when denied a reservations request, depending on the composition of the next-best travel alternative in their ordered choice sets. Those not willing to accept a vertical shift might be willing to accept different departure and/or return flight components in their trip packages but with the same fare product component as the one requested initially. This will be the case for extremely price-sensitive passengers (Type 3) in particular. For the airline denying a reservations request, then, the probability that the denied passenger will make a vertical shift in his/her choice process is directly affected by the passenger's price-sensitivity and changes in the utility of the fare product components between the first-ranked and next-best alternatives.

If a passenger's request for a particular travel alternative is denied, the next best choice might involve the same fare product component, but a different departure and/or return flight itinerary. Acceptance of the next-best alternative in such a case can involve a shift of flights and/or carriers. If the passenger selects the same fare product for another flight on the same airline, the airline retains this passenger's business and revenue. Furthermore, the airline may sell seats that would otherwise have gone unsold on lower demand flights. This type of movement in the passenger's choice process can be termed a ***horizontal shift*** in preference, or "displacement" of the passenger from the originally requested flights.

The passenger denied a request and unwilling to accept a higher-priced fare product might instead select a departing and/or return flight involving a competing airline in the same market as his/her next-best option. This shift results in a loss of business and revenue to the denying airline, meaning it represents a ***booking loss,*** or "defection" to another carrier. A booking loss would also result in cases where the denied passenger has reached the cutoff point of disutility in his ordered choice set and decides not to make the trip. In either case, the denying airline loses a customer.

A passenger denied a request for a preferred travel alternative, regardless of how far down the ordered choice set it may be, will therefore always take one of three actions, resulting in three distinct impacts on the airline reservations process. A denied request will lead to one of the following consumer reactions:

1. a vertical shift, v, ("upgrade") from one fare class to another, on the same outbound and return flights as those requested initially;

2. a horizontal shift, h, ("displacement") to different outbound and/or return flights, but on the same airline and in the same fare class as requested initially;

3. a booking loss, l, either to another competing airline ("defection") or to a decision not to travel.

Of these reservations "responses" by consumers, vertical and horizontal shifts are clearly desirable from the perspective of the airline denying the original reservations request. A vertical shift will result in increased revenues, while a horizontal shift enables the airline to retain the potential passenger's revenue without accepting another reservation on a heavily-booked or high-demand future flight in the requested low-fare class. The combination of vertical and horizontal shifts by denied passengers has been termed "recapture" by American Airlines [30]. The probability that a denied passenger will shift either vertically or horizontally on the same airline is thus the "recapture rate".

Mathematically, the denial of a reservations request by the airline will elicit one of the three consumer responses, v, h, or l, with various probabilities, such that:

$$P(v) + P(h) + P(l) = 1 \tag{2.7}$$

The recapture rate, RR, will then be:

$$RR = P(v) + P(h) = 1 - P(l) \tag{2.8}$$

For the airline deciding whether to accept or deny a particular request for a fare class and flight, the magnitudes of these probabilities are extremely important, especially $P(l)$. Recapture eliminates the losses associated with refused requests. American Airlines has developed a model of reservations recapture and has made attempts to measure it. Their theoretical model assumed that "recapture can be estimated by redistributing demand from a closed or cancelled flight based on passenger preference models."[31]

Validation of such a recapture model requires observation of actual recapture rates, and American has found it difficult to measure recapture behavior accurately. A preliminary survey involving intervention in actual telephone reservations transactions was undertaken, but only 30 sample points were gathered. The results of this limited survey showed that:

1. passengers were "extremely flexible" in terms of flight times;

2. fares were more important in determining recapture than schedules;

3. very few passengers were aware of or asked for specific flights.

American's preliminary estimate of overall recapture rate was about 30 percent of denied requests.

The problem with these observations concerning recapture rate, apart from the small sample size and the drawbacks associated with inferring passenger behavior from a telephone transaction, is they do not recognize that recapture rates will differ depending on the market demand segment characteristics of the passenger involved. The willingness of a potential passenger to shift vertically or horizontally from the travel alternative requested initially will depend on the passenger's price sensitivity, the time sensitivity associated with the benefits of the trip, and on the specific point in the passenger's ordered choice set at which the reservations request is denied.

We can extend the individual choice model presented in Section 2.1 to incorporate the choice shifts that occur when a preferred travel alternative is not available. The objective is to "explain" the relative probabilities of each type of choice shift described earlier, with reference to the travel disutilities of the attributes of different travel alternatives, as evaluated by the consumer. The choice shift that an individual makes will be determined by the specific alternative(s) being considered in place of the unavailable alternative and the disutility associated with the change in one or more components of the travel alternative. The expected choice shift is therefore a function of the time constraints on the trip involved and of the price-sensitivity of the consumer for this trip, which in turn define the demand segment to which the consumer belongs over the course of the choice process for the trip at hand.

For any one denied reservations request, the choice shift that will be made by the consumer depends on the composition of the "next-best" travel alternative. This next-best alternative may be the one that offers the smallest increase in travel disutility over the unavailable alternative, as derived from a ranked set of alternatives by a completely informed consumer. In practical terms, the "next-best" alternative is often an alternative selected by either the airline reservations agent or a travel agent as being potentially acceptable to the consumer. In many cases, the consumer's ordered choice set will consist of as few as one alternative presented to him/her as being the "next best" option. If the consumer believes an alternative with a lower disutility than that of the alternative being presented might be available, the suggested "next best" option will be refused. The consumer will then call another airline, or deal with another travel agency.

The airline that is unable to accommodate a passenger with his/her preferred or requested travel alternative can offer the passenger a higher-priced fare product for the same flight itineraries, D and R, as those requested, or different flight itineraries for the same fare product, F, requested. The probability that the consumer will accept either of these two next-best options is the recapture rate, RR. The recapture rate reflects the expected choice shift behavior of similar passengers in similar situations. The choice shift behavior of each individual is determined by that individual's evaluation of the relative disutilities of the next-best alternatives presented to him/her, as discussed below.

We examine first the simple case in which the consumer deals directly with only one airline, and assume that no competing airlines are involved in the choice process. When this consumer is denied a reservations request for a preferred travel alternative, the airline will offer one or more "next-best" alternatives, different from the preferred alternative in the fare product component and/or one or both flight itinerary components. Given a next-best travel alternative which differs from the initially requested alternative only in terms of the fare product component, the consumer will accept this alternative only if the disutility of the change in fare product components is less than the increased disutility associated with all other known next-best alternatives. If no other alternatives are being considered by the consumer, or if no other alternatives are available, the next-best alternative will be accepted as long as the increased disutility of the new fare product component does not increase the total disutility of travel to the point that it exceeds the perceived benefits of the trip.

With reference to the notation introduced previously, a shift in the fare product component of a travel alternative from F_1 to F_2 on the same airline for identical flight itineraries will involve an increase in travel disutility for the consumer equal to:

$$Z(F_2) - Z(F_1) \tag{2.9}$$

This increase in disutility will be a function of the price difference between the two fare products, any changes in the restrictions involved, as well as any perceived or actual changes in service amenities:

$$Z(F_2) - Z(F_1) = Z(P_2, X_2, S_2) - Z(P_1, X_1, S_1) \tag{2.10}$$

Given only the option of accepting the less desirable fare product component or not making the trip, the consumer will make a vertical choice shift ("upgrade") if:

$$Z(F_2) - Z(F_1) < W - Z(T_1) \tag{2.11}$$

where T_1 is the initially requested (first choice) travel alternative and W represents the total perceived benefits of making the trip.

We now consider the case in which a consumer is presented one "next-best" alternative with different D and R components but an identical F component as that in the unavailable alternative. The increased disutility of the change in flight itineraries will be a function of the increased displacement from the consumer's "ideal" flight departure times and any changes in en route times, including perhaps the added inconvenience of making a connection. For a horizontal shift from D_1 to D_2 and from R_1 to R_2, the increased disutility of the next-best travel alternative is given by:

$$Z(D_2, R_2) - Z(D_1, R_1) \tag{2.12}$$

where the disutilities of each of the four travel components in the above expression are a function of $|t_d - t_d^*|$ and t_{er}.

As was the case with an isolated vertical shift, the consumer presented with *only* the option of making the specified horizontal shift or not traveling will accept the changes in the D and R components of the travel alternatives as long as:

$$Z(D_2, R_2) - Z(D_1, R_1) < W - Z(T_1) \tag{2.13}$$

Note that a horizontal shift can involve a change in either D or R, or both. Even if the same R component is included in both alternatives, for example, the disutility of the R component can change when it is paired with a different D component. The disutility of a travel itinerary is thus expressed as a non-separable function of both flight itinerary components.

Even when a consumer calls an airline directly, the "next-best" alternatives presented to him/her are unlikely to be this limited. The airline is more likely to present two or more options from which the consumer may select one. The set of options presented will generally include a vertical shift for the same flight itineraries as those requested, along with one or more horizontal shift options to other flight itineraries offered by the same airline, for which the preferred fare product is in fact available. Presented the option of making a vertical choice shift or a horizontal choice shift on the same airline, the consumer will select the alternative that provides the smallest increase in travel disutility, subject to an upper limit on this increase, as defined by $W - Z(T_1)$.

In the absence of any information about travel alternatives involving other airlines, the consumer will accept a vertical choice shift from fare product F_1 to F_2 rather than a horizontal shift from (D_1, R_1) to (D_2, R_2) if:

$$Z(F_2, D_1, R_1) - Z(F_1, D_1, R_1) < Z(F_1, D_2, R_2) - Z(F_1, D_1, R_1) \qquad (2.14)$$

A horizontal shift will be preferred if the relative disutilities in the above inequation are reversed. Equality of the two changes in travel disutility would indicate indifference between the two options.

The relative disutilities associated with a vertical as opposed to a horizontal choice shift for any one consumer and trip can be related to the time-sensitivity of the trip and the price-sensitivity of the consumer for that trip. The likelihood that vertical or horizontal shifts will be accepted will depend on the specific attributes of the changes in travel alternative components that are being considered. That is, the disutility of a vertical shift depends on both the price-sensitivity of the consumer *and* the increase in price involved (tempered by reductions in restrictions or increased amenities) with a *specific* pair of alternatives. Similarly, the disutility of a horizontal shift depends on both the time-sensitivity of the trip *and* the change in time displacement associated with a *specific* pair of travel alternatives. It is therefore difficult to generalize with respect to the values of $P(v)$ and $P(h)$ for any "typical" denied request.

It is possible, however, to generalize with respect to the relative values of $P(v)$ and $P(h)$ on the basis of the market demand segment definitions introduced in Chapter One. An extremely price-sensitive consumer on a time-insensitive trip will place a higher disutility on a vertical shift than on the time displacement associated with a horizontal shift (within limits), meaning that he/she is more likely to accept a horizontal than a vertical choice shift. When considering the "next-best" options presented by a single airline, then, the Type 3 demand segment will exhibit:

$$P_3(v) < P_3(h) \tag{2.15}$$

where the subscript "3" refers to the demand segment classification.

At the other extreme, a Type 1 passenger who is price-insensitive and planning a very time-sensitive trip will place a greater disutility on time displacement from the preferred flight itineraries than on increased price, such that:

$$P_1(v) > P_1(h) \tag{2.16}$$

This relationship depends on the magnitude of time displacement associated with the horizontal shift. For a single airline offering a horizontal shift to another of its own flights, we assume that this displacement will be in the order of several hours in most markets, and that the above relationship will generally hold.

Passengers belonging to the Type 2 demand segment are assumed to be relatively price sensitive although the trips they are planning are associated with time constraints. Given this sensitivity to both price and time displacement, it is difficult even to speculate about the relative levels of $P_2(v)$ and $P_2(h)$. And, given that the Type 2 demand segment definition can apply to a variety of consumers and trips, any estimates of these probabilities will be subject to substantial uncertainty.

The market demand segmentation criteria of price and time sensitivity can also be used to postulate relationships between the probabilities of vertical and horizontal choice shifts when consumers belonging to different segments are presented with the same "next-best" travel alternatives. Given the price and time sensitivity characteristics associated with Type 1, 2, and 3 consumers, respectively, the following relationships would be expected:

$$\begin{aligned} P_1(v) > \ & P_2(v) > \ P_3(v) \\ P_1(h) < \ & P_2(h) < \ P_3(h) \end{aligned} \tag{2.17}$$

These relationships should hold when considered separately, with other factors held constant. That is, when presented with the same set of "next-best" alternatives to one that is not available, a typical Type 1 passenger will be more likely to make a vertical choice shift than a Type 2 and, in turn, a Type 3 consumer. Similarly, Type 3 consumers will be most likely to make horizontal choice shifts, followed by Type 2 and Type 1 consumers, respectively.

Although horizontal and vertical shifts enable the airline to retain, or even increase, the revenue it receives from a denied passenger, airlines in competitive markets must be more concerned with $P(l)$, the probability that a denied request will cause the consumer to take a seat on a competing carrier, or not travel at all. The complexity of the choice shift process increases dramatically as we introduce the possibility of *booking loss,* which includes "defections" to competing carriers.

First, we must distinguish between consumers who choose to determine the availability of their initial or preferred alternative by dealing directly with an airline and those who deal with a travel agency. When dealing with a travel agency, the consumer still evaluates the next-best alternatives offered on the basis of minimizing increases in travel disutility, as described earlier. In this case, however, the next-best alternatives presented to the consumer are more likely to involve different airlines or even combinations of airlines. Furthermore, these next-best alternatives will generally involve lesser increases in disutility, both in terms of the fare product and flight itinerary components, than those presented by a single airline.

A next-best alternative can simply be a similar or identical fare product offered by a competing carrier on a very similar flight itinerary. The differences in travel disutility between alternatives can thus be very small, given a much larger set of alternatives from which the "next-best" one is selected. The greater number of next-best alternatives and the relatively small differences in travel disutility among them make any generalizations about $P(v)$ and $P(h)$ far more difficult to make in the travel agency situation.

The same conditions lead to the conclusion that $P(l)$ is very difficult to measure. The probability of booking loss is affected by whether the request is being denied by the airline or by a travel agent. All else being equal, there must be some reduction in $P(l)$ in the former case, since the airline reservations agent can encourage the consumer to accept a vertical or horizontal shift. This is one reason for the preference of most airlines that passengers book directly with them (the savings in travel agent commissions being another).

Like the probabilities of horizontal and vertical choice shifts, the probability of booking loss is determined by the specific attributes of the changed components of each pair of travel alternatives being considered. The travel alternatives involved are in turn affected by the source of information on availability (airline or travel agent), as described above. The probability of booking loss will also be a function of how many of the consumer's successive requests have been denied already, particularly when dealing directly with the carrier. Generally, each less desirable travel alternative offered by the same airline will represent a greater increase in travel disutility than would be the case when alternatives involving several carriers are involved. All else being equal, the consumer denied several requests by the same airline is likely to try another airline.

The mathematical formulations of the utilities relevant to consumer choice when a preferred travel alternative is unavailable can be extended and generalized to account for most, if not all, possible "next-best" alternative scenarios. The utility valuations and decision rules will be similar to those presented earlier for the simplified one-airline case, since the consumer will be making the same decisions, but with more information about travel options.

For each travel alternative T_A that is presented to the consumer as a "replacement" for the *current* preferred alternative T^* that has been found to be unavailable, the consumer's decision whether to accept T_A will be based on the increased travel disutility, $Z(T_A) - Z(T^*)$, relative to $W - Z(T^*)$. The consumer will choose the T_A that minimizes $Z(T_A)$, subject to the constraint that $Z(T_A^*) \leq W$.

For each travel alternative T_A of which the consumer is aware, the increase in travel disutility over that of the initially request but unavailable alternative is given by:

$$Z(T_A) - Z(T^*) = Z(F_A, D_A, R_A) - Z(F^*, D^*, R^*) \tag{2.18}$$

As described earlier, the disutility of different fare products will be related primarily to the price-sensitivity of the consumer for a particular trip. The disutility associated with any changes to the fare product restrictions and/or service amenities will also have an effect, as will consumer preference for a particular carrier. The disutilities associated with different D and R components of travel alternatives will be primarily time-related.

The market demand segment to which a consumer belongs for a particular trip can therefore be related to the consumer's expected choice shift behavior in the case of unavailability of one or more requested alternatives. Market segment criteria alone, however, may not be sufficient in predicting choice shift behavior, because the "choice set" with which the consumer is presented or of which he/she is aware will differ from one case to the next and have an impact on the ultimate choice shift made.

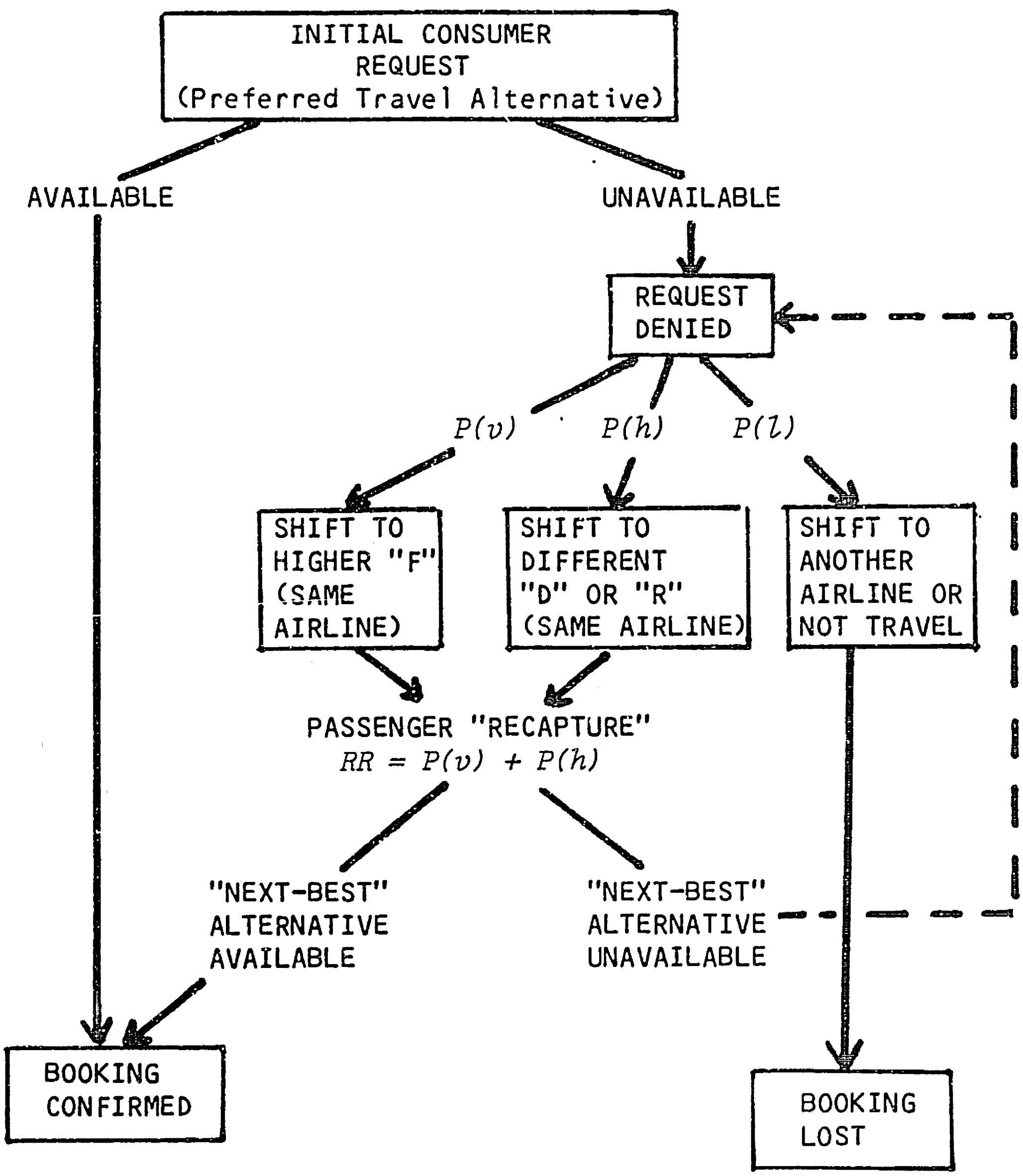
Passenger choice shift behavior can differ not only from one individual to another, but for the same individual planning different trips, and even for the same individual and trip at different points in the process of determining the availability of various travel alternatives. The probabilities of recapture and booking loss represent an average rate or expectation that must be estimated from an analysis of actual choice shifts. The difficulty lies not only in the survey methods required to sample intended behavior, but also in the need to identify similar consumers who are planning similar trips. Further theoretical consideration of the consumer choice process is required, along with much more extensive empirical studies of choice shifts.

The choice shift process described here can be summarized in a reservations decision tree framework, as shown in Figure 2.1. A request for a particular travel alternative can either be accepted or refused by the airline, based on the availability of the requested fare product class and flight itinerary. If accepted, a booking is confirmed for the consumer. If denied, there is a probability that the consumer will be *recaptured* by accepting either a vertical or horizontal shift in travel alternatives, $RR = P(v) + P(h)$. There is also a probability, $P(l) = 1 - RR$, that the denied passenger will represent a *booking loss* to the airline.

To this point, our airline reservations framework has focussed on *denied* requests stemming from unavailable travel alternatives. The assumption made implicitly is that each accepted request, or booking, will result in a ticket sale and, ultimately, a revenue passenger carried on a future flight. In reality, an accepted reservations request does represent a booking that decreases the seat inventory allocated to a fare class and flight. Each booking, however, will translate into a filled seat that generates revenue for the airline only if the passenger actually purchases a ticket and shows up for the booked flight.

Our model of the airline reservations process should therefore incorporate the probabilities associated with each booking becoming a revenue passenger on a future flight. The probabilities of relevance are:

Figure 2.1: Consumer Choice Shift Process



1. the probability that a booking will be cancelled prior to flight departure;

2. the probability that a booked passenger will fail to appear for a scheduled flight departure (i.e., be a "no-show").

Both of these probabilities will be affected by whether the passenger has purchased a ticket subsequent to having a booking confirmed for a future flight. A passenger that accepts a "next-best" alternative suggested by one airline may still call another airline to make another booking, and ultimately buy a ticket and travel on the competing carrier. As was the case with choice shift probabilities, these probabilities can also be affected by the demand segment of the passengers and the restrictions associated with the fare products booked.

The probabilities of cancellation prior to flight departure and of passengers not showing up at departure time are in essence both "cancellation" probabilities. Airlines distinguish the two rates because a cancelled booking prior to departure frees up a reservations "space" in the seat inventory that has the potential of being resold to another passenger. It is only possible to sell a seat suddenly left empty by a no-show at departure time to "stand-by" passengers, if they are present.

The pre-departure cancellation rate has been treated in past research as having the Markovian property [32,33]. That is, the probability of cancellation for any reservation on a particular day prior to departure is assumed to be independent of when that reservation was made initially. This assumption of "memoryless" cancellation rates was validated by Rothstein in 1968 [34], but fare product attributes and, in turn, reservations behavior have changed considerably since that time.

There exists a probability that a booking will be cancelled on any particular day prior to departure, and the "memoryless" assumption seems reasonable for this *cancellation rate*. That is, on day 17 prior to departure, there is no reason to expect that a particular booking made 35 days out is more likely to be cancelled than one made 25 days out. There also exists a ***cumulative cancellation probability*** for any one booking that, under this assumption, increases with the number of days before departure that the booking was made. Thus, a booking made on day 35 before departure has a higher overall probability of being cancelled than one made on day 20, all else being equal.

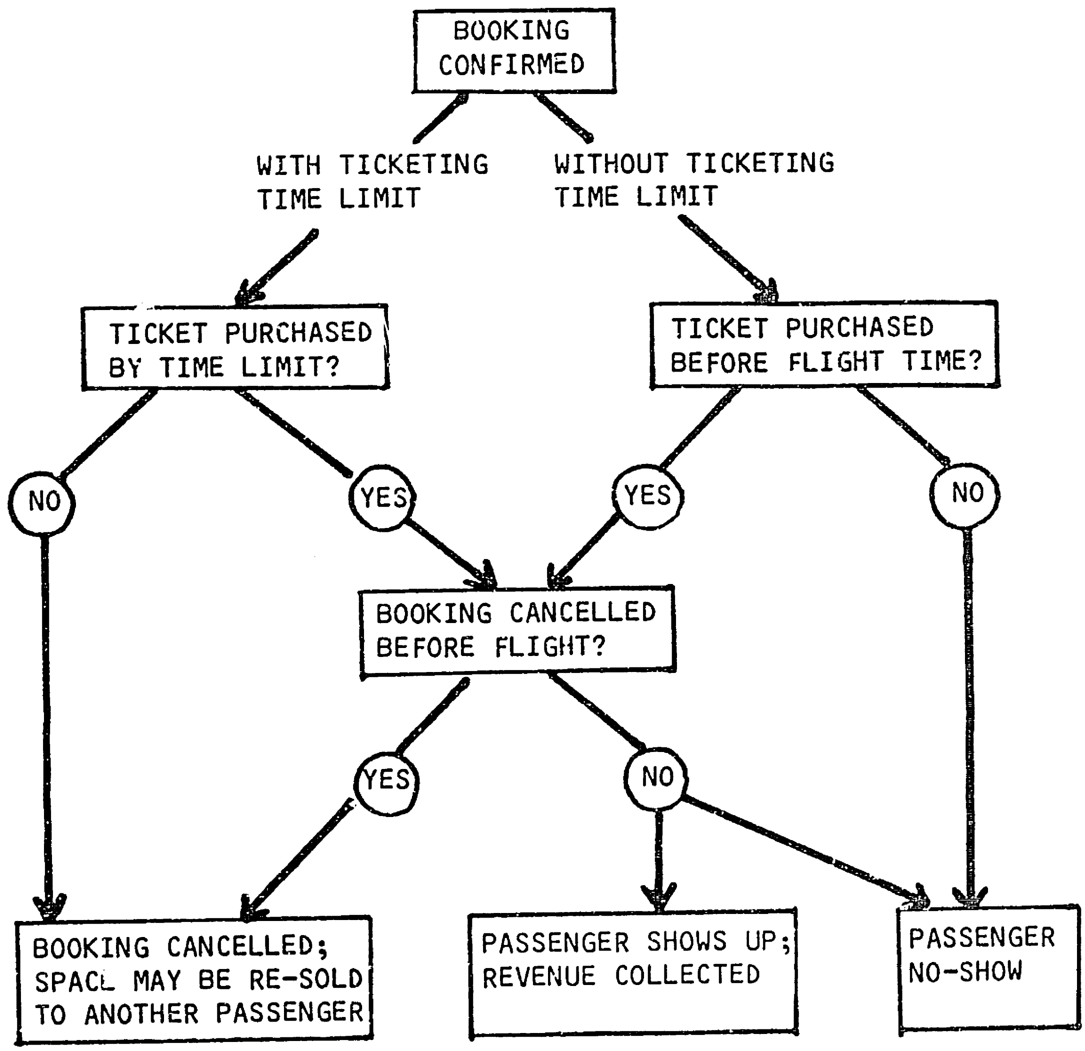
The length of time before departure that a booking is made, however, is no longer the only determinant of that booking's cumulative cancellation probability. The imposition of advance purchase requirements and cancellation penalties after ticket purchase on certain low-priced fare products has introduced additional elements that can affect cancellation probabilities. Furthermore, many airlines now stipulate time limits before which either tickets must be purchased or the booking must be reconfirmed, even for "unrestricted" fare products. If the ticketing time limit, be it arbitrary or part of the fare product's restrictions, is not met, the airline's reservations computer will cancel the booking automatically.

We can extend the decision tree of Figure 2.1 past the point at which a booking is confirmed to examine cancellation and no-show probabilities under different scenarios. As shown in Figure 2.2, the first major determinant of the cumulative cancellation probability for a booking is the presence or absence of a ticketing time limit. In either case, ticket purchase leads to the same set of subsequent branches in the decision tree, and the same set of probabilities. A failure to purchase a ticket for the booking, however, leads to distinct results in the two cases. Given a ticketing time limit, the consumer who does not purchase a ticket will have the booking cancelled, whereas the consumer who fails to purchase a ticket in the absence of any time limit will ultimately be counted as a "no-show" should he/she decide not to make use of the reserved space.

It is difficult even to speculate about the relative cumulative cancellation probabilities of bookings with and without ticketing time limits imposed, without additional information about the fare products being booked. As discussed earlier, the choice shift process and the consumer's knowledge of other travel alternatives can result in a booking that will not lead to a ticket purchase. Airlines seem to have recognized this, as they now impose ticketing time limits with greater regularity. Automatic cancellation of bookings that were never intended to be used also enables the airline to derive better estimates of "true" no-shows.

Without a ticket having been purchased, no additional information is available to refute the assumption that cumulative cancellation probabilities are simply a function of length of time before departure that the booking is made. When a ticket is purchased, however, we can illustrate how cancellation probabilities can differ depending on the attributes of the fare product purchased. On one hand, the length of time in advance

Figure 2.2: Cancellations and No-shows



of flight departure that a ticket is purchased should still have a positive relationship with the cumulative cancellation probability, because travel plans made far in advance are more likely to change. On the other hand, the imposition of restrictions on the fare product of partial or total non-refundability should reduce the cancellation probability once the ticket has been purchased. It would seem reasonable, then, to expect that a booking for a fare product with a cancellation penalty is less likely to be cancelled than one for a fully refundable fare product. This is especially true for the price-sensitive passengers most likely to purchase heavily restricted fare products.

The inclusion of cancellation penalties and, more recently, non-refundability rules in the restrictions on low-priced fare products suggests that different cumulative cancellation probabilities could well be exhibited by bookings for different fare products. The degree to which these probabilities will in fact differ depends on the advance purchase and cancellation restrictions imposed on different fare products. It will also depend on the consumer and trip characteristics of those making use of each type of fare product. With the recent introduction of non-refundable low-priced fare products, the potential for observing significantly different cancellation probabilities across fare products has increased further.

The same reasoning applies to no-show rates across fare classes at departure time. No-show probabilities are bound to be higher for reservations made for fare products for which tickets are less likely to have been purchased and those with no cancellation penalties. A reservation for a very low-priced fare product, on the other hand, is more likely to have been made and tickets purchased well in advance. The price-sensitivity of the "typical" low-fare passenger, combined with a cancellation penalty on the low fare, should result in a lower no-show rate for the more restricted fare products.

These arguments suggest that, in seat inventory control, the probabilities of reservations cancellation and passenger no-show should be distinguished on the basis of the fare product booked, for total expected flight revenues to be maximized. The airline's ability to make such a distinction will depend on its success in differentiating fare products by imposing advance purchase requirements and cancellation restrictions, as well as on its success in separating pre-departure cancellations from "true" no-shows.

The airline reservations framework described in this section is based on the assumption that passengers requesting and booking seats for different fare products will behave differently both with respect to their responses to a denied request, and in terms of their cancellation and no-show actions. The relationships between market demand segments, fare products, and the reservations process itself should thus be considered in the development of a comprehensive approach to seat inventory control.

The framework developed in this chapter describes passenger choice behavior in the context of current U.S. domestic air travel markets and its impact on the airline reservations process. The probabilistic components of this reservations process highlight the uncertainty faced by airlines in accepting or denying reservations requests. Although difficult to measure empirically, these probabilities must be recognized by the airline wishing to determine the limits on the number of seats to make available to different fare classes so as to maximize expected revenues. The application of these concepts to actual seat inventory control methods will be presented in subsequent chapters of this dissertation.

Chapter 3

The Seat Inventory Control Problem

As discussed in the preceding two chapters, differentiated fare products are targeted at distinct market segments of the total demand for air travel. Individuals in each of these segments will emerge from the air travel choice process with preferences for different types of fare products. To make the most effective use of the differential pricing strategies that they have adopted, most airlines practice seat inventory control to limit the number of seats that may be sold as each of the fare products offered. In airline reservations systems, limits are placed on the number of seats available in each fare class or booking class, which can contain several fare products. Controlling the mix of fare products to be sold for a particular flight has come to be regarded by some airline managers as by far the most important aspect of marketing, more important than the actual prices charged for each fare product [35].

This chapter provides an overview of the seat inventory control problem from the airline's perspective. The problem is described in the context of current airline operations, given existing route systems, fare structures and reservations systems. The types of data available to most airlines for dealing with this problem are outlined and alternative approaches to limiting the seats available to various fare classes are discussed. The second section of this chapter then summarizes the state of current practice in seat inventory control in the airline industry, and identifies the constraints and limitations that airlines are working to overcome.

3.1 Problem Definition and Context

The differential pricing strategies adopted by U.S. airlines in recent years in an attempt to increase total revenues also represent a response by major carriers to intensified price competition from low-cost new entrants or older airlines revived by pricing aggressiveness. By offering a limited number of seats at the lowest fares advertised by price leaders, established airlines can at least appear to be competitive in price and might even be able to fill otherwise empty seats with stimulated demand. The availability of different fare products generating different revenue levels has led many airlines to become interested in "yield management". *Yield management,* taken literally, involves the application of strategies that affect the airline's overall yield, usually with the objective of increasing it.

While yield is an important airline revenue measure, it should not be used as the only indicator of an airline's ability to generate revenues and, in turn, its profitability. High yields and low load factors could be far less profitable for an airline than lower yields and higher load factors. Yields and traffic levels tend to be inversely related, such that the attainment of both high yields and high load factors can be an elusive goal for airlines, particularly in competitive markets.

The *yield* or average revenue per passenger-mile collected by an airline that offers a multiple-tier fare structure will depend on both the price levels associated with, and the number of seats made available to, each fare class. Yields can be increased either by increasing price levels or by reducing the proportion of seats sold in the lowest fare product categories. Neither action, however, will guarantee an increase in total operating revenues for the airline. An increase in yield caused by reducing the availability of low-fare seats must be weighed against a potential loss of traffic.

For the profit-maximizing airline, an objective of maximizing total revenues is more appropriate than one of maximizing overall passenger yield. Given a commitment to operate a scheduled flight with a fixed operating cost, and acknowledging the very low marginal costs of carrying additional passengers on that flight, an airline that maximizes total flight revenues will in fact maximize operating profits. "Revenue management" is thus a more appropriate term than yield management, because maximizing yield does not necessarily maximize profit.

Revenue management involves both pricing and seat inventory control. Under ideal conditions, airlines would have sufficient market information to maximize total revenues by finding an optimal combination of prices and seats associated with each of the fare products to be offered on a future flight. In practice, airlines must take into account the pricing actions of their direct competitors, particularly those actions that affect the price-sensitive market segments. In the past, airlines have matched competitors' fares in a very predictable manner, especially the lowest fares. More recently, some airlines have recognized that some differences in service levels (e.g., non-stop versus connecting flights) offered in a market can allow them to maintain a price differential, even for the lowest-priced fare products. Nonetheless, matching of price levels for all fare products is virtually universal among direct competitors.

The competitive nature of airline pricing means that, as a strategy in revenue management, changes to price levels and/or fare product attributes can seldom be implemented successfully by any one airline. American Airlines attempted to introduce a more "rational" fare structure to its domestic markets several times in the recent past. Each introduction was followed by reactions and revisions on the part of American's competitors, with the result that the ultimate changes were not nearly as "rational" as those originally proposed by American. Not all airlines can have even this muted impact on fare product offerings, because most airlines do not have the national market presence of American. The introduction of revised fare product attributes, even by a large airline, can be subverted by the independent pricing actions of a smaller competitor in one or more markets, causing a wider reaction amongst its competitors in other markets, which can in turn spread throughout national markets.

While pricing is clearly an important component of revenue management, no one airline can predict the impact of new prices on its own revenues without taking the reactions of its competitors into account. Seat inventory control, on the other hand, is a tactical component of revenue management that is entirely under the control of each individual airline and is hidden from consumers and competitors alike. The seat inventory control actions of competitors, to the extent that they can be determined, must still be taken into account, as making too few seats available to low-priced fare products might result in a loss of traffic to a competitor with more low-fare seats available. (It is possible for airlines to determine when a competitor's fare class is "sold out" for a particular flight from availability messages sent between reservations systems). Nevertheless, seat inventory control has the potential of increasing the total revenues expected from flights on a departure-by-departure basis, something that would be far more difficult through pricing actions.

The remainder of this dissertation examines the seat inventory control component of the airline revenue management problem. For the reasons cited above, the price levels associated with each of the fare products, and thus the fare classes in the reservations system, are assumed to be given and constant over the period during which most of the reservations are accepted for a future flight departure. The seat inventory control problem, then, is to determine the optimal (revenue-maximizing) limits on the sale of available seats in the various fare classes being offered on a future flight departure.

The seat inventory control problem for a particular flight is defined to a great extent by the equipment utilization decisions made in the airline scheduling process. From the outset, the decision to operate a particular type of aircraft on a flight routing can have implications for seat inventory control, given that anticipated demand and aircraft size are unlikely to match exactly. Routing constraints and passenger flow imbalances can mean that the most appropriate size of aircraft cannot always be physically assigned to a flight departure. A significant excess or shortage of seats relative to total demand for a particular flight departure will reduce or increase the need, respectively, for strict control of low-fare sales. Seat inventory control in essence provides an opportunity to adjust for imperfections in the longer-range schedule design process. Ironically, seat inventory control can at the same time hide true demand data from those responsible for long-range schedule planning, as increased availability of low-fare seats can increase total traffic and overall load factors.

Seat inventory control can be approached with respect to individual flight legs, the airline's entire route network, or some sub-unit of the schedule such as a particular connecting complex at a given time at an airline's hub airport. The flight leg approach is by far the simplest and is currently used by most airlines, as discussed in the following section. Demand on the flight leg for each fare class is estimated and seats are allocated among the classes for each future flight leg separately. The demand for a fare class is considered to be the total number of passengers that will request a seat on a particular flight leg and fare class as part of their origin-destination itinerary.

The leg-based approach to seat inventory control is directed toward maximizing flight leg revenues, not necessarily total system revenues for the airline. For a two-leg flight, (e.g., Boston - New York - Miami) a leg-based approach to seat allocation might limit the number of seats sold on the first leg to the lowest fare class, say class Q. This inventory of Q-class seats would be shared, for example, by the Boston-New York local Q-class passengers paying $29 and the Boston-Miami through Q-class passengers paying $99. If demand for the local fare is high, the airline could be losing the revenue from potential through passengers to Miami. To maximize system revenues, it is necessary to distinguish between these two requests for Q-class seats on the first leg.

An origin-destination or market-based approach to seat allocation is required to overcome such revenue-loss problems. Seats would be allocated to the fare class/passenger itinerary combinations available on each flight leg, with seats being protected for those combinations generating the greatest revenues. Complex network formulations of all fare class/passenger itinerary combinations and their expected revenues would be required to find the system-wide optimal seat allocations. In more practical terms, airline computer reservations systems would have to be re-programmed to manage seat inventories on the basis of passenger O-D itineraries rather than by simple flight leg and fare class availability. Nevertheless, several major carriers have plans to pursue this approach to inventory control, with American Airlines being among the leaders in this development.

The complexity of the problem, even at the single-leg level of analysis, has increased tremendously with the development of hub-and-spoke route networks by most large airlines. A large air carrier can operate over 1000 flights per day, serve several thousand O-D markets and offer five fare classes in each market. American Airlines, for example, now serves over 2700 origin-destination markets, compared with about 300 before deregulation [36]. For any one particular flight departure to its Dallas/Ft. Worth hub, passengers typically can be booked into one of at least five fare classes to one of more than fifty destinations. There can thus be over 250 possible fare class/destination combinations for each seat on such a flight leg, each of which will generate different revenue levels for the airline. With reservations for future flights being accepted up to 11 months in advance, the size of the seat inventory control problem can become unmanageable.

Clearly, no airline is in a position to make separate seat inventory control decisions about each of the tens of thousands of fare product/itinerary possibilities it offers each day. Much of the battle in the development of an effective seat inventory management process involves balancing the aggregation of O-D markets and/or fare classes offered on a flight, necessary to keep the size of the problem manageable, against the disaggregation necessary to enable the airline to control the availability of seats in different fare classes in specific markets.

At the level of the individual flight leg, seat inventory decisions must be made within the constraints imposed by the airline's network, schedule and reservations system capabilities. The aircraft assigned to a particular flight departure is known and, in turn, the number of seats available in each compartment (first, business, coach) can be regarded as fixed. Seat inventory decisions must also take into account the overbooking levels to be used for each flight departure, as determined from an analysis of the costs associated with oversales and subsequent denied boardings of confirmed passengers, as well as the costs in lost revenue associated with "no-shows" and unused seats. Because the problem involves the management of available reservations "spaces" as opposed to physical seats on the aircraft, the interaction between the fare class mix of passengers booked and the number that ultimately show up for a flight can have significant revenue implications.

Managing the inventory of available reservations spaces on a future flight leg is therefore a process which occurs in the context of a pre-determined departure time and aircraft type, and which is generally subordinate to other capacity decisions involving the distribution of on-board space among physical compartments and, in current practice at least, the targeted limits for overbooking the flight. Furthermore, in most instances, the fare products (and thus the fare classes) as well as their respective prices can be assumed to be given and constant throughout the booking period for the flight.

Finding the optimal limits on the number of bookings that may be accepted in a particular fare class on a future flight leg requires estimates to be made of both the expected demand for each fare class and the average revenue associated with each class. Whether these estimates are based entirely on historical patterns or derived from a forecasting model, data from past flights are required. For a forecast of leg demand, information on booking levels prior to departure and actual boardings by fare class, by flight leg, and by day of week must be extracted from the reservations system and stored for seat inventory control decision support purposes. Most airlines have implemented database management systems that perform this function. Estimating the revenue associated with bookings in each fare class can be more difficult, since revenue data are usually collected and summarized independently of the reservations system. Thus, the sophistication of an airline's seat inventory control system depends not only on the particular approach used to determine optimal fare class limits, but on the quality of the data retrieval, forecasting and estimation methods that provide the input data required for such calculations.

The actual process of seat inventory control for a future flight departure can be as simple as a one-time setting of booking limits on discount fare classes at the start of the reservations process for that flight, and taking no further action as reservations are accepted. A more sophisticated approach would take into account the information provided by actual reservations as they are accepted, through monitoring of booking data, and then adjust fare class booking limits as flight departure approaches. A comprehensive effort to manage seat inventories and improve yield would be a dynamic one in which traffic and booking histories are used to set initial fare class booking limits, actual bookings in each fare class are monitored relative to these limits, and frequent adjustments are made on the basis of an analysis of past data, current bookings, and forecasts of future bookings for the flight.

Setting, monitoring and adjusting fare class booking limits is a process which requires both technical capabilities and human expertise. Management information systems are required to retrieve, summarize and analyze historical reservations and traffic data. The judgement of seat inventory control analysts assigned by the airline is also necessary to determine the extent to which historical data can be applied validly to current conditions in each market. Clearly, rapid changes to competitive conditions in the industry or in particular markets, or unusual operational events in a market, reduce the value of historical data. Still, there is a wealth of information available to carriers in their reservations systems which can be used to assist in seat inventory control by reducing the manual effort required by analysts while improving their decision making process.

The process employed by an airline to manage its seat inventories will be affected by the range of tools and resources that it has available for this purpose, which will in turn be determined by the importance that corporate decision-makers attribute to improving revenue management. From the outset, the seat inventory control system employed will also depend on the characteristics of the airline's network and on its fare structure. An airline offering a single fare level for all seats on flights serving point-to-point markets on a non-stop basis clearly need not be concerned about sophisticated seat inventory control techniques. At the other extreme, a carrier with multiple fare classes on flights into and out of large connecting hub complexes can benefit immensely from improved seat inventory control.

Most established (i.e., pre-deregulation) airlines are closer to the latter extreme, and as such are extremely interested in revenue management. Among such airlines, there is a range of effort devoted to, and a range of sophistication in, seat inventory control. As described in the following section, current practice in this area is evolving quickly, yet a strong emphasis on human expertise in making seat inventory control decisions remains.

3.2 A Survey of Seat Inventory Control Practices

Representatives of nine large North American airlines were interviewed between August 1985 and March 1986 to determine the present status and future of seat inventory control at each carrier. The airline representatives were understandably reluctant to provide specifics as to the booking limits and other criteria used in managing their seat inventories. It was possible, however, to develop an understanding of the process adopted by each carrier, including the organizational structures involved, reservations system and data retrieval capabilities, as well as the methods used to monitor bookings and control the sale of seats in low fare classes.

3.2.1 Organizational Issues

Seat inventory control and revenue management are closely related to a range of other functions found in current airline corporate structures, such as pricing, marketing, reservations, overbooking and payload control. None of the airlines surveyed have been able to combine all the functions critical to revenue management into a single unit, although several major carriers have moved in this direction. Pricing and overbooking control were the most frequently named functions to have been incorporated into the revenue management unit, which in turn was most commonly found in the airline's marketing or market planning department. Coordination with the remaining related functions that for various reasons could not be included in the same department, poses a problem for most of the airlines surveyed.

There seems to be a consensus among those involved in the process that the preferred place for a seat inventory/revenue management group is in the airline's marketing department. The close relationship of seat inventory control to the reservations, pricing and fare product development functions make the marketing department a logical place for the revenue management function. At many of the airlines surveyed, however, those responsible for setting and monitoring flight overbooking levels remain in separate departments responsible for reservations and sales in some cases and payload control/operations in others. There can thus be substantial resistance and turf-based conflicts with respect to the transfer of the overbooking function to an ever-expanding marketing department.

The personnel responsible for monitoring and adjusting booking limits throughout the reservations process — seat control analysts — make up the largest group in most revenue management units. There were substantial differences among the carriers surveyed in the use of seat control analysts, both with respect to the number of analysts employed and the degree to which the analysts are responsible for specific markets.

Although the number of analysts employed does not necessarily reflect the sophistication of an airline's seat inventory control process, it does reflect the resources devoted to the problem. The number of agents "required" to manage a carrier's seat inventory would seem to be related to the number of flight legs operated each day and to the proportion of a carrier's flight legs that operate in highly competitive markets or during peak periods of demand.

Table 3.1 illustrates the range of human resources devoted by the nine airlines surveyed to the flight monitoring and booking limit adjustment process. When the carriers were compared in terms of daily flights per seat control analyst, the largest carriers proved to be in the middle of this range, while the high and low extremes were provided by the smaller carriers. Carriers F and H realize that the resources they currently devote to seat inventory control are inadequate, and have plans to expand their efforts in the near future. Carrier G had recently undergone such an expansion, along with a major reorganization of its revenue management group, in order to monitor its flight leg bookings more closely.

For those carriers with relatively few analysts working on seat inventory control, booking limit monitoring and adjustment is almost entirely an *ad-hoc* process, perhaps

Table 3.1: Comparison of Seat Control Analyst Staffing Levels

North American Airlines, March 1986

AIRLINE	NUMBER OF AGENTS(1)	NUMBER OF AIRCRAFT(2)	DAILY FLIGHT DEPARTURES(2)	DAILY FLIGHTS PER ANALYST
A	25	200 +	1,000 +	60
B	35	200 +	1,000 +	30
C	35	200 +	1,000 +	40
D	15	150-200	1,000 +	90
E	20	150-200	500-1000	30
F	10	100-150	500-1000	110
G	25	100-150	0-500	20
H	5	0-100	0-500	150
I	20	0-100	0-500	15

(1) Approximated to nearest 5.
(2) Ranges are given to protect airline anonymity.

targeted at selected markets and flight legs. The carriers with proportionately more analysts generally have a more systematic process in which teams of analysts are responsible for groups of markets and/or flight legs. The carrier with the highest number of analysts relative to its daily departures has taken the notion of specialization to the extreme, making each seat control analyst responsible for all flight legs that serve a particular market or set of routes. Each agent must consider historical data, competitors' actions and current trends to both set initial fare class booking limits and make adjustments as reservations are accepted. These agents are then held accountable for the traffic mixes and revenue levels achieved on their own routes.

Overall, the airlines surveyed face very similar problems in organizing their revenue management groups. While some carriers have progressed more rapidly than others in unifying seat inventory control and related activities, coordination of functions, resistance to organizational change and even intra-departmental organization are still issues that all face as their revenue management functions evolve.

3.2.2 Reservations and Decision Support Systems

The effectiveness of an airline's seat inventory control process depends on the capabilities of its reservations system and on its ability to extract useful historical reservations and traffic data to assist in making booking limit decisions. As was the case with organizational structures, many carriers are currently taking steps to make their reservations systems more responsive to seat inventory control needs and to develop more sophisticated data retrieval and analysis methods.

Most of the airline reservations systems in place today are deficient in several areas of relevance to seat inventory control. For example, the number of fare class "buckets" into which bookings can be logged is a major limitation for most of the carriers surveyed. At least one bucket is required for each of the physical compartments on an aircraft. Any remaining buckets are used by most carriers as sub-classes of a shared coach compartment. Currently, most systems limit fare class bookings by flight leg and some have the capability of limiting sales to local passengers in favor of through and connecting passengers that generate more total revenue. The large number of fare product/O-D market combinations possible on a single flight leg makes it desirable for the airline to be able to take reservations in a large number of fare classes and to limit sales in specific markets when necessary.

Airline attempts to make use of more fare class buckets to control their seat inventories by passenger O-D itinerary have been impeded by the existing standardized interline distribution system. The major reservations systems used by travel agents currently display seat availability and accept bookings in a maximum of five classes. Given the need to exchange availability and booking information with other reservations systems, airlines are constrained in improving their own systems by the need to maintain this standardization.

Several carriers are nonetheless changing their own reservations systems to accept bookings in a much larger number of reservations buckets. In one example of what will be the "new generation" of reservations systems, eight primary fare classes will have up to five subordinate buckets for use in controlling low-fare bookings in particular origin-destination markets. The expanded reservations systems will allow an airline to limit, for example, the number of extremely low-priced seats in selected markets without closing down the entire fare class on a flight leg to additional bookings. Existing reservations systems do not permit the airline to differentiate between passengers requesting similar fare products for very different itineraries on the same flight leg, as these passengers are booked in the same fare class.

The development of such origin-destination based reservations systems is a top priority in the area of revenue management for several of the airlines surveyed. American Airlines has acknowledged that it in fact has such a system already in place, and that the focus of its seat inventory control process has changed to one of "selling the system" [37]. The goal is to manage fare class inventories with respect to the revenues generated by the passengers on local, through, and connecting itineraries, all on the same flight leg. The decision models required to achieve this goal are far more complex than those required for simple leg-based seat inventory control, and are still in the developmental stages. The mathematical techniques needed to make effective use of these expanded reservations systems will be discussed in detail in Part Two of this dissertation.

Although a few airlines are well-advanced in the redesign of their reservations systems for more effective seat inventory control, many others are still several years away from implementing such changes. And, even if most airlines succeed in expanding their reservations systems, it is likely that flight availability will continue to be displayed, and bookings made, in one of the five primary fare classes. The additional sub-classes are intended to be "hidden" from travel agents and airline reservations staff, and are to be used only by airline seat control analysts. The industry standard of five booking classes for any one flight leg is therefore likely to remain intact for some time, until a "new generation" of reservations systems has been implemented widely.

The implication of the five fare class limit imposed by the interline standards is that all airlines hoping to upgrade their own systems are constrained by the need to remain compatible. Carriers that operate flights with both First Class and Business Class compartments are particularly constrained by this compatibility problem. With distinct reservations classes required for each of the physical compartments on the aircraft (including the coach cabin), only two fare classes remain for control of reduced fare sales. In markets where there exist, for example, 30-day advance purchase fares, 14-day advance purchase fares, and unrestricted "super-coach" discounted one-way fares, the availability of only two discount fare categories can force the airline to aggregate rather dissimilar fare products and to give up some of its control over reduced fare sales.

In addition to the number of fare classes, primary and subordinate, in an airline's reservations system, the structure of these fare classes is perhaps equally important to the needs of effective seat inventory control. The structure of fare classes in the system involves the relationship between the classes, specifically their interdependence in terms of applying booking limits. For seat inventory control purposes, it is desirable to have a *nested* hierarchy of fare classes, nested in ascending order of revenue value to the airline. For example, several airlines offer four fare classes in their coach cabins — Y, M, B, and Q — with Y being the full coach fare class and Q being the lowest-revenue promotional fare class. With a nested structure, it is possible to prevent bookouts in higher fare classes while seats are available in lower classes. Setting a limit on one fare class will implicitly limit the cumulative total of reservations accepted in that class and all lower classes. It is therefore possible for reservations in a higher-revenue fare class to cause lower-revenue classes to be closed down, even if the lower fare class limits have not been reached. In a nested hierarchy of fare classes, requests for higher-revenue seats automatically pre-empt those for the lowest-revenue seats, in contrast to a non-nested system in which a distinct inventory of seats is allocated to each fare class.

Over half of the carriers surveyed currently have a totally or partially nested fare class hierarchy in their reservations systems. Other carriers have their coach cabin fare classes independently nested in a cumulative Y fare class, with little or no interdependence among the reduced fare classes within the coach cabin. The airlines planning to upgrade and expand their reservations systems agree that a nested fare class hierarchy would be preferable from the seat inventory control perspective. In the expanded systems, each sub-class could be nested within the respective primary class, which would in turn be nested within the cumulative coach cabin reservations bucket.

The capabilities of existing reservations systems place a technical constraint on the improvement of seat inventory control techniques. Ironically, individual airlines' efforts to improve their own reservations systems must take into account the direction adopted by other airlines, particularly those with the dominant systems.

An airline's reservations system also plays an important role in providing the data required as decision support for seat inventory control. Decisions must be made with respect to initial fare class booking limits, which may require revision on the basis of actual reservations received and forecast demand. Both the initial limits and demand forecasts must be derived at least in part from historical booking patterns and traffic data for the same or similar flights. The capability to retrieve and summarize relevant historical data is thus crucial to seat inventory control.

The decision support function of the airline reservations system is distinct from its capabilities in terms of the number of fare classes and their relationships. Consequently, the data management and analysis functions required for decision support in seat inventory control can be developed and even used independently of the reservations system itself, although linkages for recovering raw data from the system must be maintained. Of the carriers surveyed, those with computer systems large enough to handle the extra load are developing their decision support tools for use in conjunction with their reservations systems. On the other hand, several carriers expressed concern that the decision support functions would tax their already overburdened reservations systems and significantly increase delays for both reservations agents and customers. These carriers are contemplating the purchase of separate software *and* hardware for seat inventory control decision support purposes.

At least two "yield management packages" are being marketed to airlines by software companies. Both Control Data Corporation (CDC) and Sperry Rand offer systems that extract data from airline reservations systems and "offer ways to manage the multiplicity of fares" [38]. The CDC's "MARKSMAN" airline yield management system "boils down the vast mass of data to manageable proportions and presents it to the agent in a form which allows decisions to be made." [39] While it also offers an ability to monitor actual bookings relative to historical patterns, the CDC system is essentially a statistical data management software package developed for seat inventory control applications. This and other "packages" on the market are not designed to determine optimal booking limits, as they do not have the ability to forecast demand or revenues, nor do they make use of revenue or traffic optimization routines.

The carriers surveyed have all considered investing in such packages, and at least one has already followed through. This carrier finds its package to be a valuable tool, and has based an expansion of its revenue management efforts around the capabilities and requirements of the package it purchased. At least two of the other carriers have developed and implemented their own versions of these packages for use on their existing computer systems. The remaining carriers are either considering the packages available, or are just beginning to develop customized database management systems.

Regardless of which package is purchased or to what extent software is developed in-house, efficient and usable database management systems are essential to seat inventory control. These systems must provide data to the seat control analysts in a form that will enable them to determine the best response to changes in booking patterns as departure time approaches. Because the development of such systems is well within existing computer technologies, and given that packages can simply be purchased, airline managements are eager to invest in this aspect of revenue management. In fact, it is safe to say that the areas currently receiving the greatest amount of resources and effort from airlines interested in improving their seat inventory control process are reservations and database management systems.

3.2.3 Setting and Monitoring Booking Limits

The organizational structure of an airline's revenue management unit, together with the information tools available to it, provide a foundation for the tasks of setting, monitoring and adjusting fare class booking limits so as to maximize flight revenues. It is in this component of the seat inventory control process that differences among airlines in terms of sophistication are most apparent. These differences stem in part from the organizational and information issues discussed above, but also reflect varying amounts of emphasis placed by each airline on setting initial fare class limits, monitoring actual reservations relative to these limits, and then making necessary adjustments.

At the simplest level, setting the initial booking limits for reduced fare classes can be done on an "across-the-board" or default value basis. The use of default values for all flights operated in certain types of markets or with particular aircraft types requires little in the way of resources, but does not take into account important differences in passenger mixes and booking patterns between markets or even between flights in the same market. Greater precision can be achieved by setting lower fare class limits by market, day-of-week, even time-of-day of future flight departures, but this requires substantial analytical effort.

If the airline's seat inventory control process consists solely of setting fare class limits at the start of the booking process, then a more detailed approach is certainly preferable. All the carriers surveyed have progressed beyond this simplest level of seat inventory control, although several have done so only in the recent past. Most airlines have recognized that greater benefits from their seat inventory control efforts can be realized by developing a reservations monitoring and booking limit adjustment process than by intensifying their efforts to improve the accuracy of the initial limits.

That is not to say that the initial limits placed on different fare classes are not important. Any differentiation among markets or flights can help to reduce the amount of intervention required later in the process and is better than no differentiation at all. One of the larger carriers surveyed uses some default values for groups of flights and then identifies specific markets, flights, and/or periods of operation that might require more careful attention. Another carrier that assigns responsibility for different routes to individual agents leaves the setting of initial limits to them. The default limits automatically entered into the reservations system for this same carrier are the actual passengers boarded by fare class for the same flight, one year earlier. Most of the remaining carriers use aggregate approaches in which limits are initially set according to aircraft type and/or broad market groupings.

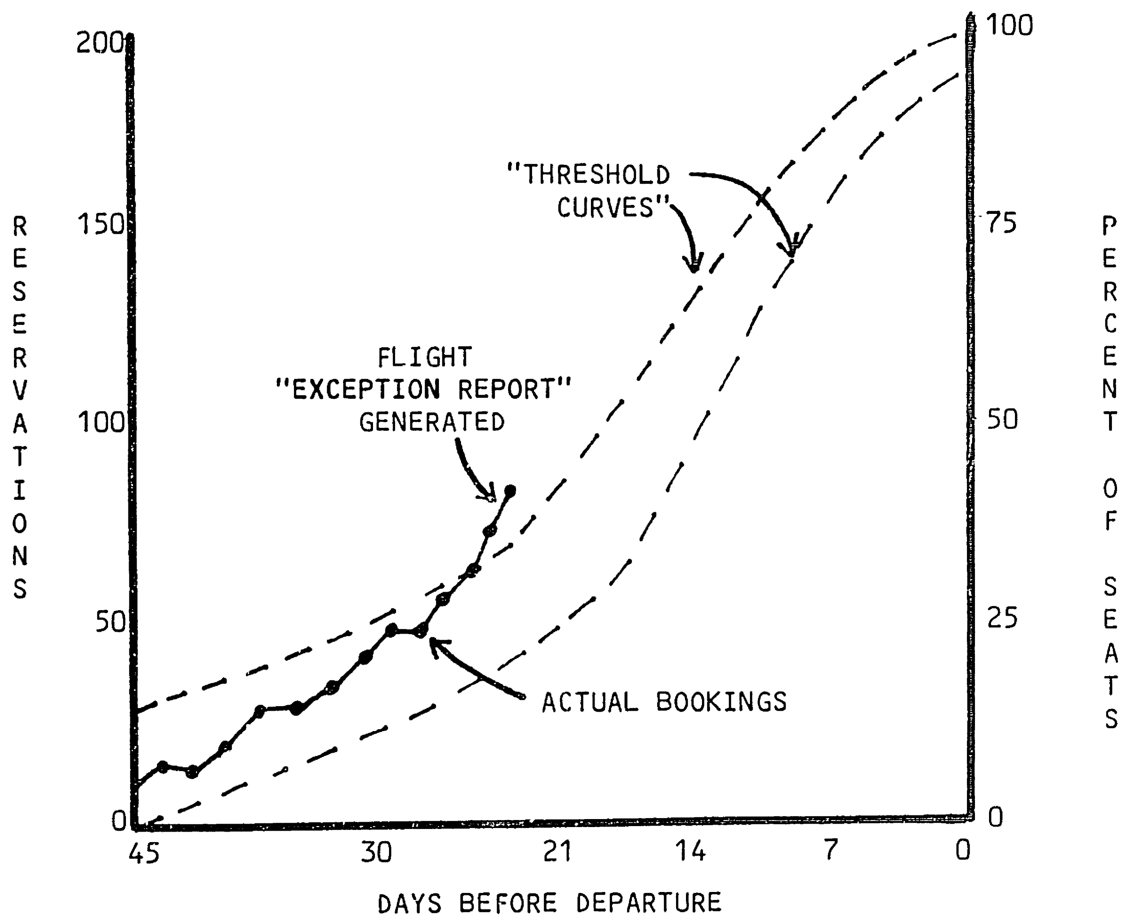
A relatively simple approach to setting initial booking limits can be counter-balanced by a more sophisticated system of monitoring cumulative bookings by fare class, relative to these limits. The monitoring function can easily be automated through the airline's reservations system. Even the simplest of systems can be programmed to generate reports listing the future flights for which the number of accepted reservations approaches the fare class booking limits. All of the carriers surveyed have routines in place that can perform this function, although not all of them have a seat inventory control process in place that makes full use of the monitoring reports.

More sophisticated computer routines can be designed to flag flights for which actual reservations levels meet or approach one of several different types of booking limit values. For example, a flight might be flagged by the monitoring system if the number of bookings accepted in a lower fare class approaches the class limit, if total low-fare bookings exceed some percentage of total seats, and/or if the total number of unbooked seats drops below some pre-set value. The most advanced of the automated monitoring systems found among the airlines surveyed makes use of all three monitoring methods. That is, flights are flagged when any one reduced fare class reaches it own booking limit, when the nested low fare classes reach a cumulative limit, or when the total number of unbooked seats remaining in the coach cabin falls below some targeted level.

The airline reservations systems surveyed all monitor actual bookings relative to pre-set and static fare class limits, with varying degrees of sophistication. An improvement to the monitoring process is provided by the CDC yield management package, among others. On the basis of historical booking trends for a flight or group of flights, the CDC package generates "booking threshold curves" which show the expected range for cumulative bookings at any point before departure (see Figure 3.1). The flights for which actual bookings stray outside the range of these curves are then flagged by the system and listed in periodic reports. The dynamic monitoring capability provided by the CDC package can be used in conjunction with the static fare class booking limits described above.

The least advanced aspect of revenue management and seat inventory control at all of the airlines surveyed is that of booking limit adjustment to maximize flight revenues. This task involves modifying, when necessary, the initial booking limits on each fare class to take into account both actual bookings received and additional bookings expected over

Figure 3.1: Example of "Threshold Curve" Monitoring



the time remaining to flight departure. This is the most important component of seat inventory control, yet it remains dependent on *ad-hoc* human judgement rather than systematic analysis. Resource limitations and a lack of practical models for making optimal decisions about changes to fare class booking limits as the flight departure approaches have kept the adjustment task at a low level of sophistication.

When an airline's reservations monitoring system flags a flight for which actual bookings approach any one of the limits or the threshold established for that flight, a decision must be made either to increase the availability of seats in the relevant fare class or to allow the system to close it down to additional reservations. This decision is currently being made by seat control analysts on the basis of experience and judgement at every airline surveyed, although the development of decision support tools is designed to reduce the amount of guesswork involved. Differences among airlines in adjusting booking limits stem to a greater extent from the number and responsibilities of the seat control analysts than from differences in the actual techniques used to make adjustment decisions.

At least two carriers are hoping to make the adjustment process more systematic by developing an automated routine for maximizing expected flight revenues at various points in time before departure. One such algorithm would cause lower-yield fare classes to be closed down automatically when the expected revenue from bookings in higher-yield classes exceeds the revenue from selling additional seats in the lower-yield class. Some computational problems must be overcome, however, before such an algorithm can be implemented and be of practical use.

Even if an optimization algorithm were to be used by an airline, the need for human judgement in seat inventory control cannot be eliminated entirely. Any optimization models would be probabilistic in nature and would make use of forecasts based on historical data. Changes in the competitive environment of airline markets and the occurrence of unexpected events that affect recent flight booking patterns are but two examples of variables which cannot be accounted for in such algorithms. The objective in developing optimization models for seat inventory control is to allow analysts to focus their efforts on these exceptional variables by making routine tasks more systematic.

Seat inventory control and revenue management in the airline industry is at an intermediate level of sophistication, although there exist substantial differences between individual carriers. The increased importance to airline profitability of effective revenue management has prompted many carriers to invest in improvements to the decision support tools required, and some are exploring the possibilities of making the process more systematic with the help of mathematical optimization and forecasting techniques. The objective of most carriers at this point in time is to increase the amount and usefulness of the information available to those responsible for making seat inventory control decisions.

The above discussion of seat inventory control practice suggests that a range of approaches are being used by airlines to address essentially the same problem. The seat inventory control problem, at the level of the single flight leg, is to determine the booking limits on each fare class that will maximize total revenues for a future scheduled flight departure. The static problem is to establish these fare class limits once at the start of the booking process, taking into account the uncertainty associated with expected bookings by fare class, to the extent possible. The dynamic problem is to revise these initial limits on the basis of the additional information provided by actual bookings as departure day approaches, once again incorporating the probabilistic nature of future demand.

In Part Two, mathematical approaches for solving the seat inventory control problem are considered. Throughout the review of past work and the development of new approaches, the distinction between the static and dynamic problems will be emphasized, as will the importance of incorporating probabilistic demand into any solution approach. The discussion of mathematical approaches will be based on the concepts developed thus far. The demand segmentation and pricing practices described in Chapter One define the current environment in which seat inventory control is practised. The consumer choice and airline reservations frameworks presented in Chapter Two will be used in the development of solution approaches that take into account realistic demand behavior. Above all, the practical constraints on seat inventory control, as described in Chapter Three, will guide the development of a seat inventory control framework that can be adopted readily by most airlines.

PART TWO
Mathematical Models for Seat Inventory Management

The relatively low level of sophistication in seat inventory management as practised by airlines is due in part to the recent realization of its importance to airline revenues and profitability. It is also attributable to a lack of practical models for making optimal seat inventory control decisions. Although there has been a substantial amount of theoretical and empirical research devoted to airline seat inventory control, the results have in most cases been large-scale optimization models that solve simplified representations of the actual problem faced by airlines. Overall, the development of practical models for determining the number of seats to make available in each fare class on a future flight simply did not keep pace with the rapid changes in airline marketing and pricing practices that have transpired since deregulation.

It is the objective of Part Two to develop a practical mathematical framework for airline seat inventory control that can be applied to the fare structures, reservations systems, and route networks of the current airline marketplace. Chapter Four begins this process with an overview of the mathematical approaches that have been considered in previous work. The distinction between seat inventory control for separate and nested fare class inventories is explained first. Past work on methods for allocating available seats among distinct fare class inventories is then reviewed, with a focus on mathematical programming and network formulations of the problem. The evolution of the "marginal seat" decision rule approach to limiting fare class seat availability in nested reservations systems is then summarized.

This "marginal seat" approach is the basis for the development of a more current mathematical framework in Chapter Five. A decision model for determining revenue-maximizing booking limits for multiple fare classes in a nested reservations system is presented for the single flight leg inventory control problem. Applications of this model to both static and dynamic seat inventory management decisions are described. The model is then extended to incorporate airline overbooking of seats due to passenger no-shows. A method for including the possibility of passenger choice shifts in the booking process into the model is also described. Finally, examples of model results are provided for a hypothetical flight leg.

Chapter Six examines the EMSR model in the context of actual applications to airline reservations systems. The demand inputs and the assumptions required when the model is applied to the data provided by existing leg-based airline reservations systems are discussed first. The impacts on the EMSR output results of the most important demand assumptions are described, and the sensitivity of these outputs to each of the input variables is considered. Finally, the potential application of the EMSR approach to seat inventory control systems designed to limit seat availability by passenger itinerary, currently under development, is introduced.

Chapter 4

An Overview of Previous Research

As a prelude to the extension of existing models and the development of more practical tools for airline seat inventory control in subsequent chapters, this chapter presents an overview of the mathematical approaches developed previously. The differences between the separate and nested fare class seat inventory control problems are examined first, and methods for solving simple representations of the former are introduced. Operations research models for determining optimal seat allotments among distinct fare classes in more complicated representations are then reviewed, including recent work on network-based solutions for revenue maximization. The final section of this chapter examines the evolution of the marginal seat revenue approach to determining optimal booking limits on fare class inventories, including past applications to dynamic booking limit adjustment.

4.1 Distinct Versus Nested Fare Class Inventories

The relationship between fare class inventories or reservations "buckets" in an airline's reservations system affects the way in which the seat inventory control problem is represented mathematically and, in turn, the solution methods that are most appropriate. The simplest reservations system structure involves distinct and separate inventories for each fare class. The booking limits on each "bucket" must sum to the total capacity of the shared cabin in such systems. When overbooking is involved, the booking limits on each bucket must sum to the overall limit on reservations for the shared cabin.

The problem of allocating a limited number of seats or reservations spaces among distinct fare class buckets is similar to that of allocating on-board space among two or more physical compartments on an aircraft. In both cases, the objective is to maximize total revenues subject to a capacity constraint. The distribution of space among physical compartments on an aircraft is a medium- to long-term decision in airline operations, given that seat pitches and widths are difficult to change for each flight departure. This problem has received attention primarily from aircraft manufacturers [40].

The use by some European airlines of easily moveable dividers or "bulkheads" that enable the airline to determine the physical boundary between the business and economy class cabins just prior to flight departures presents a revenue maximization problem that is even more similar to that of allocating seats among fare classes. Considerations such as overbooking of each physical class and the possibility of passenger upgrades or downgrades in the event of oversales are important in this problem, and have been addressed most recently by Madsen [41].

Although similar to both of these problems, seat allocation among distinct fare class buckets that will share a common compartment involves some subtle differences. Because no distinction is made between passengers booked in different fare classes in terms of on-board service, there is no possibility of upgrades or downgrades from one fare class to another in the case of oversales. The problem of oversales involves the *total* capacity of the shared cabin. Another difference is that a single seat may be allocated to a fare class inventory, whereas at least a row of seats must be allocated to a physical compartment, even with moveable bulkheads. Finally, an airline may offer seats in as many fare classes as can be handled by its reservations system, while there is a practical limit on the number of physical compartments and levels of on-board service that may be provided on a single aircraft.

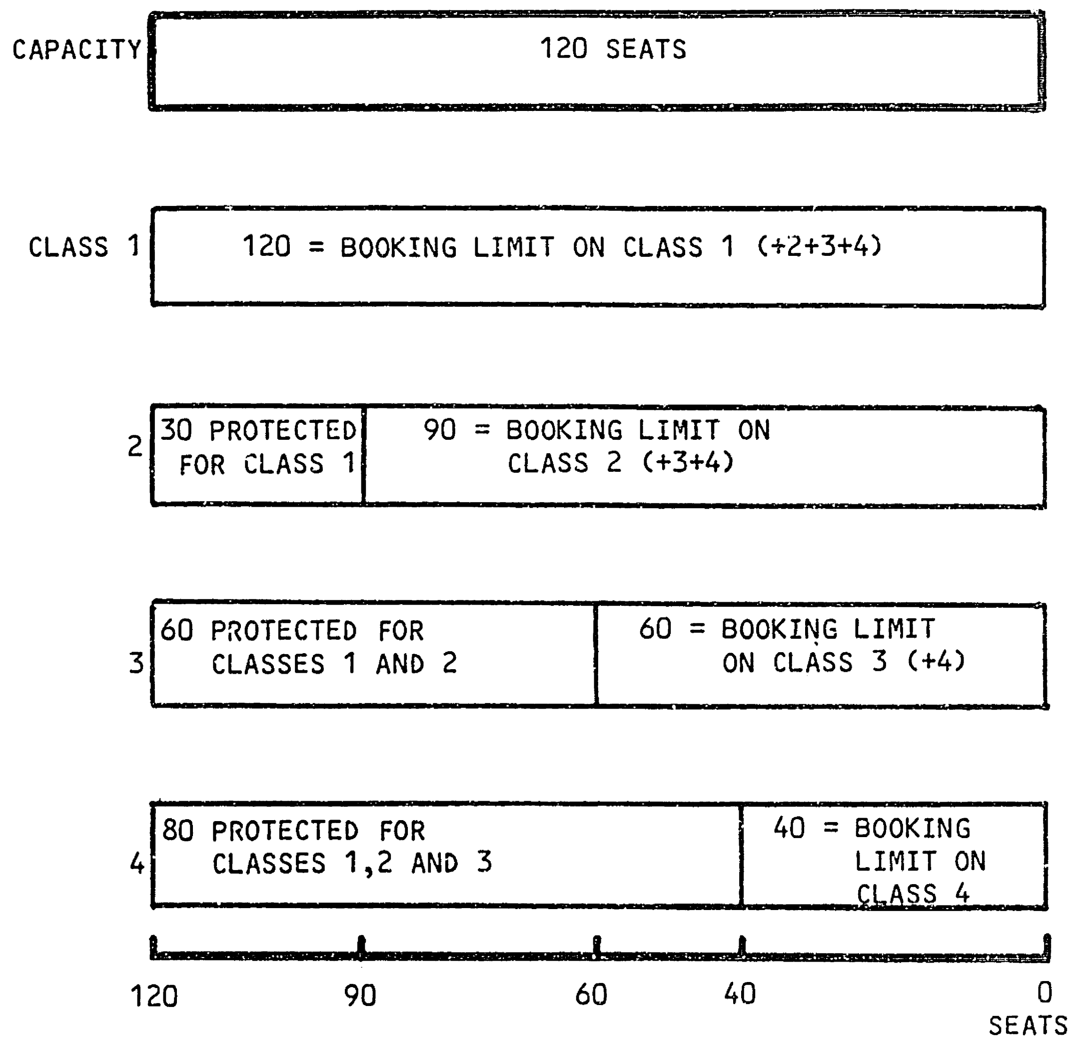
Seat allocation among distinct fare classes that share a common inventory of identical seats is nonetheless essentially a problem of distributing available seats among "invisible" compartments that are defined only in the reservations system. In fact, the seat allocation problem has been referred to as the "electronic bulkhead problem". Once a seat is assigned to a fare class inventory, it may be booked only in that fare class, or remain unsold.

In contrast, a "nested" reservations system is one in which the fare class inventories are structured such that a high-fare request will not be refused as long as any seats remain available in lower fare classes. A nested reservations system is thus binding in its limits on lower fare classes, but its limits are "transparent" from above (for higher fare classes). The nesting applies to each successively lower fare class, such that the limit on the lowest class, say Class 4, applies to bookings in Class 4 only, whereas the limit on Class 3 applies to the *total* of Class 3 and 4 bookings. Sales to Class 3 can in essence take over the inventory set aside for Class 4. The booking limits that apply in a completely nested reservations system are illustrated for a four-class example in Figure 4.1.

The distinction between separate and nested fare class inventories is important both to the way in which the seat inventory control problem is represented and to the mathematical methods that are used to determine optimal seat allocations or fare class booking limits. Equally important is the distinction between the static problem, in which fare class booking limits are applied at the start of the booking process for a future flight, and the dynamic problem, in which booking limits may be revised as actual bookings are accepted. In the latter case, the length of the interval between revisions determines the degree to which the differences between distinct and nested fare class inventories affect optimal booking limits. Request-by-request revision of limits on distinct inventories eliminates the expected revenue advantages associated with nested inventories. The longer the interval between revisions, however, the greater the importance of finding the optimal booking limits, particularly in the case of nested fare classes.

The following section reviews the solution approaches that have been presented in past work for seat allocation among distinct fare class inventories. The discussion centers on the static problem, although applications of the methods described to the dynamic problem are introduced. The subsequent section then presents an overview of the "marginal seat" model for limiting seats sold in low fare classes. This approach has been developed in the context of the dynamic seat inventory control problem, but is equally applicable to the static problem for nested fare class inventories.

Figure 4.1: Booking Limits in a Nested Reservations System



4.2 Seat Allocation Among Distinct Fare Classes

The tools of differential calculus, Lagrangian multipliers, mathematical programming and network optimization have all been applied to the allocation of a given number of seats among two or more distinct fare class inventories, depending on the complexity of the particular problem representation involved. To correspond with the distinct fare class buckets, the assumption that the demand levels for each fare class are also separate and not correlated has been made in conjunction with most applications of these tools. Furthermore, the degree to which stochastic demand is incorporated into the solution methods has differed in past work. This section provides an overview of methods for allocating seats among distinct fare class inventories and their application to dynamic booking limit revision processes.

Apart from the characteristics of demand for different fare classes, assumptions with respect to passenger no-shows or cancellations as well as the expected reactions of passengers denied a reservations request are required when mathematical models for finding optimal seat allotments are considered. In general, past work has been based on the following simplifying assumptions:

1. Demand for each fare class is separate and not correlated;
2. A booking accepted by the airline will translate into a revenue passenger carried, meaning that there are no cancellations or no-shows;
3. A rejected request represents a loss of revenue to the airline.

In the allocation of seats among distinct fare class buckets, no assumption as to the booking rates in the different fare classes is required.

It is possible to find the optimal seat allotments for distinct fare class inventories through a relatively straightforward application of differential calculus when only two fare classes on a single flight leg are involved. The seats allocated to each fare class (S_1 and S_2) must share a total capacity of C seats. To find the value of $S_1 = C - S_2$ that will maximize total expected revenues, $\overline{R}$, for the flight, we differentiate $\overline{R}$ with respect to S_1 and set to zero:

$$\overline{R} = \overline{R}_1(S_1) + \overline{R}_2(C - S_1) \tag{4.1}$$

$$\partial\overline{R}/\partial S_1 = 0 \tag{4.2}$$

That is, seats are allocated between the two fare classes such that the marginal expected *total* revenue with respect to additional seats in one class or the other is equal to zero. At optimality, total expected flight revenues cannot be increased by taking a seat from class 1 and allocating it instead to class 2. The expected marginal revenue of the last seat allocated to each class will be equal across the two classes:

$$\partial\overline{R}/\partial S_1 = \partial\overline{R}/\partial S_2 \tag{4.3}$$

This analysis can be extended to three or more distinct fare class inventories, where S_i^* for each fare class can be found to maximize total revenues. With more than two S_i variables, the problem becomes more difficult to solve analytically, but can be formulated as a constrained revenue maximization problem and solved with the Lagrangian multiplier method. The optimality conditions for the multiple fare class problem would then be:

$$\frac{\partial\overline{R}}{\partial S_i} = \frac{\partial\overline{R}}{\partial S_j} = \lambda, \qquad \text{for all fare classes } i \neq j \tag{4.4}$$

where λ, the Lagrangian multiplier, equals the expected marginal revenue for the last seat allocated to each fare class. As in the two-class case, seats are allocated among fare classes such that the total expected marginal revenue is equal across all relevant fare classes. Under the Lagrangian multiplier formulation, a single, non-linear equation must be solved for λ, and it can be shown that this solution is a unique optimum [42].

Further expansion of the problem to multiple fare classes and multiple-leg flights or even connecting flight operations requires the application of mathematical programming and related techniques to find the optimal fare class seat allotments numerically. All of these methods, as reviewed below, make use of the principle of equating marginal revenues of the last seats allocated to each inventory in determining the revenue-maximizing seat allotments.

The potential application of mathematical programming techniques to the seat allocation problem was considered by Mathaisel and de Lamotte in 1982 [43]. A mathematical program of the airline's system-wide seat inventory control problem would maximize total operating income by allocating capacity to each O-D market and fare class on each feasible passenger routing throughout the system. Mathaisel and de Lamotte realized that such a program would include prices, capacity and passenger demand as variables in the operating income objective function, making the revenue term non-linear.

Although this non-linearity problem could conceivably be overcome by splitting the fare and traffic components of the objective function, and then using an iterative solution technique in a dual objective goal programming approach, Mathaisel subsequently scaled down the scope of the original formulation considerably [44]. When supply variables, including operating expenses and price are fixed, the mathematical program to maximize total carrier revenues contains only one major supply-side decision variable -- the number of seats to be allocated to class i in O-D market m using routing r on each flight leg.

Even this reduced formulation of the system-wide revenue maximization problem cannot easily be solved. A traffic assignment algorithm would be required to allocate demand in a particular O-D market along different feasible routings (through- op flights as opposed to one or more connecting flights). With respect to solution methods, Mathaisel suggested that several alternatives might be pursued, including network flow and mathematical programming techniques.

The network flow approach was pursued by Glover *et. al.* [45], who postulated that there is "some number of passengers at each fare class on each flight segment that will optimize revenue" for the airline. They developed a network-based seat allocation model for Frontier Airlines, which was designed to find the mix of passenger itineraries (PIs) flowing over the airline's network in various fare classes that would maximize total daily revenues. The model required inputs of the demand and revenues expected for each PI from j to k in fare class i, and produced outputs of the optimal number of seats to be allocated to that PI/fare class combination.

In this network formulation, a set of forward arcs represented the flight legs to be operated between points in the airline's system (nodes on the network). Flow on these arcs was associated with passenger loads on each leg, and was constrained by the leg capacities. A set of backward arcs represented all possible PI/fare class combinations. Flow on the backward arcs consisted of the number of seats to be allocated by PI/fare class, constrained only by the total demand for each PI/fare class combination. The network was solved as a maximum flow (revenue) problem with side constraints to maintain feasibility (i.e., eliminate cycles) when connecting flights were included.

Glover's model for Frontier accommodated up to 600 daily flights and 30,000 PIs in five fare classes. The formulation required 200,000 variables and 3,000 constraints, and took several hours to solve, making interactive use impossible. A more serious shortcoming of the model is its reliance on deterministic demand estimates for each PI/fare class combination. No consideration of probabilistic demand or spill was included.

The importance of accounting for probabilistic demand can be illustrated with a very simple example in which 100 seats must be allocated among three fare class buckets on a single flight leg. In Table 4.1, a comparison of the optimal seat allotments, expected loads and revenues is presented for inputs of deterministic demand and stochastic demand with coefficients of variation equal to 0.3 and 0.5. The total demand exceeds the total capacity of the aircraft, meaning some passengers will be refused. The optimal seat allotments do in fact vary, depending on the assumed variability in demand More importantly, the deterministic solution overestimates total expected revenues relative to the variable demand solutions, and this difference increases as the variability of demand increases.

Analysts at both Boeing Aircraft [46] and McDonnell-Douglas [47] have addressed the problem of incorporating probabilistic demand into mathematical programming formulations of the seat allocation problem. Because the expected revenue function for each class is non-linear given a constant fare and stochastic demand, a simple linear program is an inaccurate representation of the problem. McDonnell-Douglas analysts proposed a formulation of the single-leg seat allocation problem that makes use of binary decision variables in a linear integer programming framework. Each variable, X_{ik}, represents the combination of fare class i and seat k on the flight leg. A 150-seat aircraft, four fare class problem would thus require 600 such decision variables. Associated with each X_{ik} is the expected marginal revenue from selling the kth seat in class i, denoted $m_i(k)$, derived by multiplying the average fare level in class i by the probability of selling k or more seats in that class. Thus, the effects of probabilistic demand on total revenues are incorporated in this formulation.

With an available total capacity of n seats, the formulation of this linear program is as follows:

$$MAX \quad \overline{R}(n) = \sum_i \sum_k X_{ik} \cdot m_i(k) \tag{4.5}$$

Table 4.1: Optimal Seat Allocation — Deterministic vs. Stochastic Demand

3 Distinct Fare Class Inventories; Capacity 100 seats.

	CLASS 1	CLASS 2	CLASS 3	TOTAL
FARE ($)	100	75	50	—
MEAN DEMAND	25	40	50	115
CASE 1: DETERMINISTIC DEMAND (k = 0)				
Optimal Allocation	25	40	35	100
Expected Loads	25	40	35	100
Expected Revenues	$2500	$3000	$1750	$7250
CASE 2: STOCHASTIC DEMAND (k = 0.3)				
Optimal Allocation	27	38	35	100
Expected Loads	22.9	34.1	33.8	90.8
Expected Revenues	$2290	$2561	$1688	$6539
CASE 3: STOCHASTIC DEMAND (k = 0.5)				
Optimal Allocation	29	40	31	100
Expected Loads	21.8	32.0	27.8	81.6
Expected Revenues	$2176	$2402	$1389	$5966

subject to:

$$\begin{aligned} \sum_i \sum_k X_{ik} &\leq n \\ 0 \leq X_{ik} &\leq 1 \end{aligned}$$

The solution to this linear program will be integer, with the X_{ik} variables corresponding to the n largest values of $m_i(k)$ equal to 1, and all other $X_{ik} = 0$.

The problem of a non-linear objective function was also addressed by D'Sylva of Boeing Aircraft [48], who used a piece-wise linear approximation of the expected revenue curve in a linear programming formulation. He found that five to ten binary decision variables could replace the 200 used previously for a 200-seat aircraft in approximating the expected revenue function of each fare class. D'Sylva used this approach to extend Glover's algorithm to include stochastic demand. An arc was added to the network representation for each of the straight-line approximations used in the total expected revenue objective function. A comparison of solutions to the probabilistic and deterministic formulations showed that the latter overestimated expected revenues by about 12 percent. Furthermore, the best variable demand solution produced a five percent higher expected revenue than the deterministic solution.

At McDonnell-Douglas, Wollmer [49] also applied the network approach to the multiple fare class/multiple flight leg problem, based on the mathematical programming formulation in equation (4.5). Using the single-leg formulation as a starting point, he developed a seat allocation algorithm to find the revenue-maximizing booking limits more efficiently than conventional LP solution methods. Since the expected revenue term, $m_i(k)$, is a decreasing function of k, seats may be allocated optimally with an incremental or marginal approach. A relatively simple routine to find the n largest $m_i(k)$ values across all fare classes will determine the revenue-maximizing combination of fare class booking limits.

This seat allocation algorithm was extended to a two-leg flight in which the expected revenue for local as opposed to through passengers must be compared across relevant fare classes. Wollmer demonstrated that the same incremental approach can be used to find the optimal solution for the two-leg case, and he suggested that multiple-leg flights could also be handled in this way. That is, the algorithm would identify the maximum expected marginal revenue for each possible combination of fare classes and passenger itineraries over the legs of the flight, and then allocate a seat to the combination with the largest value.

When the linear program was enlarged to the two-leg problem, each X_{ik} represented a seat k which could be allocated to a specific fare class/passenger itinerary (PI), i. A two-leg flight offers three possible passenger itineraries (A-B, B-C, A-C). With four fare classes, there would thus be 12 feasible fare class/PI combinations. We define A(1), A(2), and A(3) to be the set of fare classes available to passenger itineraries A-B, B-C, and A-C, respectively. Then, with n_1 and n_2 unbooked seats available on flight legs 1 and 2, the linear program becomes:

$$MAX \quad \overline{R}(n_1, n_2) = \sum_{A(1)} \sum_{A(2)} \sum_{A(3)} m_i(k) \cdot X_{ik} \tag{4.6}$$

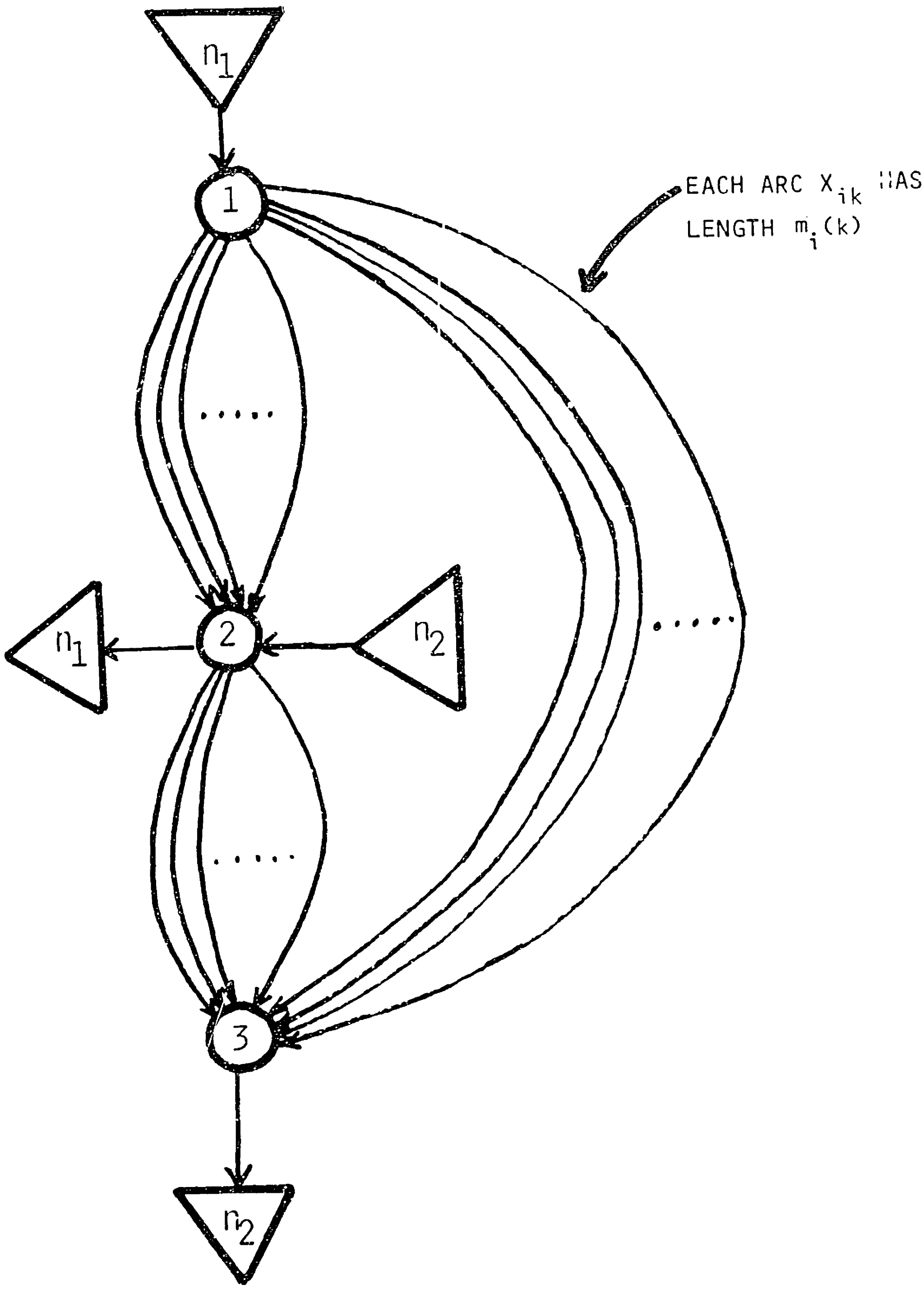
subject to:

$$\begin{aligned} \sum_{A(1)} \sum_{A(3)} \quad X_{ik} &\leq n_1 \\ \sum_{A(2)} \sum_{A(3)} \quad X_{ik} &\leq n_2 \\ 0 \leq \quad X_{ik} &\leq 1 \end{aligned}$$

Wollmer noted that this program satisfies the conditions of a transshipment or network flow problem, as represented in Figure 4.2. The network shown has three nodes, corresponding to each point served by the two-leg flight. Each X_{ik} in $A(1)$ is represented by an arc from node 1 to node 2, with a length equal to $m_i(k)$. Each arc has an upper bound on flow of 1 and a lower bound of 0. Total network revenues are maximized by finding the longest (highest expected revenue) paths from 1 to 3 and sending one unit of flow along each path, subject to the leg capacity constraints.

The two-leg case presented above would require 1800 arcs (3 PIs, 4 classes, 150 seats) in a complete network formulation for a 150-seat aircraft. As the formulation is expanded to multiple flight legs and/or passenger itineraries involving connecting flights, the number of X_{ik} variables (arcs) required increases rapidly. By using the network characteristics and the decreasing $m_i(k)$ property of the seat allocation problem, however, Wollmer was able to reduce dramatically the size of the formulation required for solution purposes. He suggested that only the shortest (lowest revenue) arc with an

Figure 4.2: Network Representation of Seat Allocation Problem



existing flow of one and the longest (highest revenue) unused arc for each fare class/PI combination need to be considered at any one time.

A solution algorithm for this network problem therefore need only solve a series of longest path problems for a relatively small network. At each iteration, the expected marginal revenue of each fare class/PI combination in the current reduced network must be determined, the longest path identified, and a seat allocated to that path. The marginal expected revenues (lengths) associated with the arcs on this longest path would then be updated and the procedure repeated. Wollmer and Stroup at McDonnell-Douglas have continued work on the seat allocation problem, extending the mathematical programming formulations to an airline connecting hub operation [50]. The network characteristics of these formulations and the expected marginal revenue approach to optimal seat allocation are being used to develop solution algorithms more efficient than the integer programming methods described above.

Work is also under way on dynamic applications of these optimization models to the reservations process, to revise booking limits as reservations are accepted and flight departure day approaches. Making seat inventory decisions dynamically with the help of mathematical programming techniques requires an assessment to be made of the value of accepting a current reservation request relative to the decrease in expected total revenue associated with removing one seat from the available inventory on the flight leg(s) requested. The expected total revenue for the network with n seats available on the flight leg requested was defined to be $\overline{R}(n)$. Dynamic applications of the network formulation thus involve a comparison of $MAX\ \overline{R}(n)$ with $MAX\ \overline{R}(n-1)$, given n unbooked seats on a flight leg for which a request is received. If a request for fare class and O-D itinerary i is received, it should be accepted as long as:

$$f_i > MAX\ \overline{R}(n) - MAX\ \overline{R}(n-1) \tag{4.7}$$

where f_i is the revenue associated with the current request. In other words, the marginal expected network revenue from retaining the *nth* seat on the flight leg requested for possible sale in higher fare classes and/or for other itineraries is compared to the revenue associated with the current request.

This comparison of the expected revenue differential for each incremental unbooked seat with a "certain" revenue from the current reservations request is extendable in conceptual terms to the most complex of network formulations. The drawbacks of such an approach, however, make its practical use by airlines unlikely in the near future. The computing capabilities required to run network optimization models in order to make expected revenue assessments for each reservations request received by an airline would be enormous. Even with the reduced network formulations proposed by Wollmer, finding $MAX\ \overline{R}(n)$ and $MAX\ \overline{R}(n-1)$ for each fare class/PI request would be a major burden on even the most advanced airline computer systems.

A more fundamental drawback is the amount of input data required to determine expected revenues from various points in time prior to departure. Unless it is assumed that the rate of future reservations requests is independent of the number of bookings accepted up to day t before departure, this dynamic approach would require a massive amount of reliable historical data to predict future reservations by fare class from any point in time, ***conditional on*** actual bookings. Wollmer has in fact made the independence assumption in his work, an assumption that we shall examine in a subsequent chapter.

Given that most airlines in practice only make revisions to fare class booking limits periodically, the solution to the static seat allocation problem must apply over the entire interval between revisions. For applications in which this interval is substantial, the mathematical programming and network formulations described in this section are in fact solving a simplified seat inventory control problem. All of the formulations described find the revenue-maximizing allocation of seats to distinct and separate reservations "buckets", either by fare class or by fare class/PI combination. As described in the first section of this chapter, most airlines have the lowest fare class buckets nested within the higher fare classes, such that a high-fare request will not be denied as long as lower fare classes are still open. Mathematical programming approaches will therefore not necessarily give the true revenue-maximizing seat allotments for a ***nested*** reservations system.

4.3 Evolution of the "Marginal Seat" Model

The optimization models reviewed in Section 4.2 all equate the marginal revenues of incremental seats allocated to distinct fare class inventories so as to maximize total expected flight revenues. This principle of equating marginal seat revenues has been applied by numerous researchers to the problem of dynamic booking limit revision and incorporated into iterative solution approaches. The assumptions of separate demands for each fare class, no cancellations or no-shows, and no possibility of recapturing a refused reservations request have been made in virtually all of the work reviewed in this section. The assumption of probabilistic demand, however, has been incorporated effectively in most cases. Dealing with the dynamic problem requires an assumption concerning the order in which and the rates at which requests for different fare classes are received. The assumption that all low-fare requests are received first has been the most common, although alternatives have been examined, as will be discussed.

The request-by-request dynamic decision approaches reviewed in this section can also be applied to the static nested fare class booking limit problem. The nested fare class problem is essentially the same as the distinct fare class problem in cases where seat allotments for the latter can be revised on a "real-time" (i.e., request by request) basis. This constant revision allows unexpectedly high demand for the highest fare class to be accommodated, as does a nested reservations system. For this reason, many of the applications of the "marginal seat" approach to revenue maximization reviewed in this section are valid for either representation of the problem.

The marginal seat model was applied in a dynamic reservations context to the two-class, single flight leg seat inventory control problem by Littlewood in 1972 [51]. He suggested that revenues could be maximized by "closing down" the low-fare class to additional bookings when the certain revenue from selling another low-fare seat is exceeded by the *expected* revenue from saving that seat for a potential high-fare passenger. That is, low-fare passengers paying f_2 should be accepted as long as:

$$f_2 \geq \overline{P}_1(S_1) \cdot f_1 \tag{4.8}$$

where $\overline{P}_1(S_1)$ is the probability of selling all S_1 remaining seats to high-fare passengers, and f_1 is the higher fare level. The smallest integer value of S_1 that satisfies the above condition is the number of seats that should be protected for class 1 in a nested fare class system in order to maximize expected revenues.

In 1973, Trans World Airlines analysts expanded on Littlewood's simple model [52]. Their model formulation, which is in fact equivalent to Littlewood's, showed the optimal allotment of high-fare seats in the dynamic case to be $(C - S_2^*)$, where:

$$f_2 = f_1 \cdot \int_{C-S_2^*}^{\infty} p_1(r_1)dr_1, \qquad 0 \leq S_2^* \leq C \tag{4.9}$$

This formulation includes the probability density of reservations requests for class 1, $p_1(r_1)$, which when integrated from $C - S_2 = S_1$ to infinity produces $\overline{P}_1(S_1)$, as defined in Littlewood's model.

The term "differential revenue" was used by Richter of Lufthansa in 1982 in reference to the loss in total expected revenue when low-fare passengers ultimately deny space to higher-fare passengers [53]. The differential revenue from allocating an additional seat to a low-fare passenger was defined as the difference between the additional revenue realized from the low-fare sale and the revenue lost from prospective high-fare passengers, as follows:

$$DR = f_2 \cdot \overline{P}_2(S_2) - f_1 \cdot \overline{P}_1(C - S_2 + 1) \tag{4.10}$$

The expected marginal seat revenues for the two fare classes are equal when DR is set to zero. The optimal value of S_2 must therefore satisfy:

$$\frac{f_1}{f_2} = \frac{\overline{P}_2(S_2)}{\overline{P}_1(C - S_2 + 1)} \tag{4.11}$$

This "differential revenue" model in fact generates the optimal seat allotments for two *distinct* fare classes, in which a seat allocated to the lower fare class can actually deny a seat to a high-fare passenger. Richter demonstrated, however, that the optimal limit on the lower fare class in a dynamic application is a function of the relative fare levels, total capacity, and the demand for the high-fare class only. It is *not* a function of low-fare demand, although the distribution of low-fare demand and the booking process assumed will influence the expected total revenue for the flight. For dynamic applications, the optimal value of S_2 must satisfy:

$$f_2/f_1 = \overline{P}_1(C - S_2 + 1) \tag{4.12}$$

This optimality condition is thus equivalent to Littlewood's "simple" model.

The conclusion that the density of low-fare demand will not affect the optimal number of seats to be allocated to the higher fare class is important in the *nested* fare class problem. The implication is that, at any point in time, the airline should *protect* S_1^*

seats for potential high-fare demand, to the point at which the expected revenue from an additional protected high-fare seat equals the actual fare level in the lower fare class. At this optimal point, the airline is indifferent between the "certain" revenue of accepting another low-fare booking and the "expected" revenue from protecting that seat for a high-fare passenger who is yet to appear. This concept will be central to the development of a mathematical framework for more complex representations of the nested fare class problem in Chapter Five.

Richter also applied his conclusions to a conditional booking or "priority waitlist" model in which the carrier offers an additional price reduction to low-fare passengers willing to be placed on a priority waitlist once the low-fare seat allotment has been filled. Such passengers would be subject to "bumping" by subsequent high-fare demand, and would then be entitled to a pre-specified level of compensation. The differential revenue formulation was used by Richter not only to find the optimal booking limits for confirmed and wait-listed passengers, but also to determine the optimal discount incentive and compensation levels for the wait-listed passengers.

A sensitivity analysis of the simple model under the assumption that low-fare passengers book first was performed by Mayer of El Al in 1976 [54]. He showed that the greater the difference between f_1 and f_2, the more sensitive the total expected flight revenue will be to a non-optimal allocation of seats. The decrease in expected revenue will be smaller when too many seats are allocated to low fare passengers than when too few seats are offered.

Mayer also changed the "early-bird" assumption of the simple model, assuming instead that low-fare passengers would book first in each of many *periods* before departure. He further assumed that demand in period t is independent of demand in previous periods $t+1$, $t+2$, ...etc., and defined $V_n(t)$ to be the maximum revenue from period t until departure, given n seats remaining at the beginning of the period. Mayer then applied a dynamic programming framework to find $V_n(t)$ and the corresponding optimal allotment of low-fare seats for every relevant value of n and t. Dynamic programming allowed for the possibility that not all low-fare seats would be booked first, or in any particular period t. The optimal booking policy derived from this formulation would allow as many high-fare seats as requested to be booked in each period, as long as any seats remain available.

A sensitivity analysis of the expected revenue generated by the multi-period model as opposed to the simple model showed that use of the multi-period model reduced the revenue impacts of erroneous seat allocations. On the other hand, Mayer found that the choice of initial seat allotments did not benefit significantly from the dynamic programming framework. Consequently, he suggested that the simple model be used for finding the optimal initial seat allotments, and that the multi-period model be used to modify the initial allotments in light of actual bookings made.

The question of how the assumption that low-fare passengers book first affects the optimal seat allotments and expected revenue levels was also addressed by Titzer and Griesshaber in 1983 [55]. The "early-bird" low-fare booking process was compared with a simulated parallel process of booking in which the rate of low-fare requests decreases and the rate of high-fare requests increases as departure time approaches. The simulation varied the means and variances of the low-fare and high-fare demand distributions for both booking behavior assumptions so that results from several scenarios could be compared.

The results of the sensitivity analysis showed the maximum expected revenue level to be reached at approximately the same optimal low-fare seat limit, regardless of which booking behavior assumption is made. The parallel booking process simulation produced higher total expected flight revenues at the optimal seat allotment, simply because high-fare passengers were able to book early and fill a greater proportion of seats on heavily-booked flights, an illustration of the benefits associated with nested fare classes. The total revenue curves as a function of low-fare seat limits were found to be relatively flat around the optimal points, such that a 20 percent deviation from the optimal limit was required to cause a noticeable revenue decrease. The revenue curves for the parallel booking process were even flatter than those for the "low-fare first" scenario, confirming Mayer's findings that total expected revenues are less sensitive to erroneous seat allotments when a mixed booking process is involved.

The expected marginal revenue concept was applied by Buhr of Lufthansa in 1982 to the problem of allocating seats on a two-leg flight (A to B to C) to either local or through passengers, although only one fare class was considered [56]. The buckets for each origin-destination market were treated as distinct inventories. Buhr defined the expected "residual" revenue, E, from allocating an additional seat to a passenger flying from A to C as:

$$E_{AC}(S_{AC}) = \overline{P}_{AC}(S_{AC}) \cdot f_{AC} \tag{4.13}$$

where f_{AC} is the average fare for A to C passengers and $\overline{P}_{AC}(S_{AC})$ is the probability of selling more than S_{AC} seats to passengers from A to C.

Demand for each origin-destination market — A-C, A-B, and B-C on a two-leg flight — was assumed to be independent, and conditions were imposed such that:

$$S_{AB} = S_{BC} = C - S_{AC} \tag{4.14}$$

Total flight revenues would be maximized when:

$$E_{AC}(S_{AC}) = E_{AB}(S_{AB}) + E_{BC}(S_{BC}) \tag{4.15}$$

subject to the capacity constraint. An iterative solution method was used to find the optimal values of S_{AC} and S_{AB}.

Buhr suggested that the allotment of seats among fare classes be performed as a second step in the process, given the optimal allotments for local and through passengers on each leg. He acknowledged that varying the allotment of low-fare seats for different passenger itineraries could change the expected revenue levels on each leg of the flight, but did not address this problem.

Buhr's formulation was extended to a more complicated problem by Ken Wang of Cathay Pacific in 1983 [57]. He set out to develop a model for determining the optimal seat allotments for multiple leg flights and multiple fare classes, with each origin-destination and fare class combination representing a distinct seat inventory. He suggested that optimal limits for each fare class in each origin-destination city-pair served by the flight could be found by allocating seats incrementally to the "string" of combinations with the highest total expected revenue. For a flight that serves j O-D pairs and offers k fare classes in each O-D city-pair market, each feasible string of combinations has an expected marginal revenue, EMR, of:

$$EMR = \sum f_{jk} \cdot P[r_{jk} > b_{jk}] \tag{4.16}$$

for the (j, k) pairs in the string, where f_{jk} is the fare yield for fare class k in O-D market j, and $P[r_{jk} > b_{jk}]$ is the probability that another request for (j, k) will be received given b_{jk} bookings accepted.

The terms of the right hand side are summed over all feasible combinations of sequential (j,k) pairs for the flight. For example, on a flight operating A to B to C, a feasible sequence would be to allocate the marginal seat to a low-fare class from A to B and to save that same seat for a high-fare passenger from B to C. In total, there would be 12 feasible combinations for a one-stop flight with three fare classes. Wang's approach requires the feasible combinations to be ranked in terms of expected marginal revenue as each seat is allocated incrementally. This iterative solution process clearly has limitations in terms of its application to larger-scale formulations of the problem.

The applications of the "marginal seat" principle described above have succeeded in incorporating probabilistic demand explicitly into the seat inventory revenue maximization problem. The simple decision rule presented by Littlewood and subsequent researchers determined optimal fare class limits for dynamic revision of booking limits, for two fare classes on a single flight leg. As mentioned, the "marginal seat" principle can be applied to the static problem when *nested* fare class inventories are involved. Applications of the same principle to multiple fare classes and flight legs, however, have not addressed the problem of nested fare classes, as evidenced by the works of Buhr and Wang.

Problem formulations with multiple fare classes, flight legs and passenger itineraries have seen the application of mathematical programming and network solution methods for revenue maximization. The biggest shortcomings of these approaches relate to their inability to deal in a practical way with setting static limits for nested fare classes and to incorporate probabilistic demand. Both of these shortcomings can be overcome with sheer computer power. Optimization models that generate optimal seat allotments for distinct inventories can be applied dynamically by re-running the optimization at frequent intervals, thereby taking into account the properties of a nested reservations system. The same models can incorporate probabilistic demand for each O-D and fare class combination if enough variables and/or arcs are added to the formulation. It would be impractical, however, for most airlines to commit this magnitude of computer resources to the problem.

None of the mathematical models reviewed here address the practical considerations introduced in Part One that can play an important role in determining the optimal booking limits for different fare classes. The relationship between overbooking and seat inventory control has been overlooked. Furthermore, refused requests are not necessarily lost to the refusing airline. The refused passenger may be accommodated in the requested fare class on another flight of the same airline, or might in fact agree to accept a reservation in a higher fare class on the originally requested flight. There is also the possibility of significant correlation of demand levels among fare classes. All of these considerations can have a considerable impact on the optimal fare class booking limits, and complicate the derivation of these limits. Nonetheless, a realistic solution model must take these factors into account.

The size of the seat inventory control problem and the volatility of the airline competitive environment dictate that any optimization model be both efficient and adaptable to changing conditions. Decision rules that can be used dynamically to limit bookings on flight legs or in specific markets might be a more practical approach to improving seat inventory control than large-scale optimization models. The emphasis in any future model development must be on matching the solution to the practical constraints of the problem, including reservations system capabilities, data availability and the nature of airline competitive practices. Above all, efforts must be made to incorporate more realistic demand assumptions and to find optimal booking limits for reservations systems with interdependent or nested fare classes.

Chapter 5

Expected Marginal Seat Revenue Model

The mathematical approaches reviewed in the previous chapter suggest directions for the development of a revenue maximization framework that addresses the seat inventory control problem being faced by airlines in the current market environment. Finding the optimal booking limits on multiple fare classes, even for a single flight leg, will in practice require an approach different from those of past works. It is the objective of this chapter to incorporate probabilistic demand into a decision model for seat inventory control that can be applied to multiple fare classes on a single flight leg, in a nested reservations system.

The probabilistic nature of the seat inventory control problem is examined first, and the use of probabilistic concepts is illustrated in a very simple model for seat allocation between two distinct fare classes. The "marginal seat" principle discussed in the previous chapter is then extended to account for multiple fare classes in a completely nested reservations system, and an "expected marginal seat revenue" model for making revenue-maximizing seat inventory decisions on the basis of historical demand is presented. This framework is then expanded to take account of actual bookings accepted for a future flight leg, making it a dynamic seat inventory control model.

Section 5.4 begins the process of incorporating more realistic demand behavior into the EMSR decision framework. The uncertainty associated with passenger no-shows is introduced and the model is extended to account for flight overbooking. The probability of a refused passenger accepting a vertical shift to a higher fare class is incorporated to illustrate the impact of passenger choice behavior on revenue maximization methods. Finally, examples of EMSR model results for set of hypothetical demand and revenue inputs are presented.

5.1 A Probabilistic Approach

The airline seat inventory management problem is probabilistic because there exists uncertainty about the ultimate number of requests that an airline will receive for seats on a future flight and, more specifically, for the different fare classes offered on that flight. In this section, the probabilistic concepts relevant to seat inventory control are defined, and incorporated into a static model for determining optimal seat allotments for distinct fare class inventories.

The total demand for a particular flight, on average, will fluctuate systematically in cycles described by day of week and season of the year. There will also be stochastic variation in demand around the expected values, among similar flights sampled consistently over a homogeneous period of time. This stochastic demand for a future flight departure can be represented by a probability density function. Past analyses have generally assumed a Gaussian (normal) distribution of total demand for a flight, with means and variances that depend on the market being studied and on the nature of its traffic [58].

For the purposes of this discussion, we will continue to assume Gaussian demand densities. Furthermore, we assume initially that the demand densities for different fare classes are not correlated significantly and the number of requests received for the various fare classes during different periods before flight departure are not correlated. These assumptions will be discussed further and tested empirically in Chapter Six.

The notion of probabilistic demand is central to the airline seat inventory control problem, as the expected number of requests for each fare class must be estimated from historical distributions of demand. We define $p_i(r_i)$ to be the probability density function for the total number of requests for reservations, r_i, received by the airline for seats in fare class i by the close of the booking process for a scheduled flight leg departure. This probability density function is derived from historical data for the same or similar flights, and is assumed to remain valid for future operations of the flight, at least in the short run and in the absence of changes to exogenous factors.

Also important is the capacity constraint on the total number of seats available on a flight leg. The number of seats allocated to a particular fare class, S_i, might not exceed the number of requests for that fare class, resulting in rejected demand, or "spill". We can thus define a cumulative probability that all requests for a fare class will be accepted, as a continuous function of S_i:

$$P_i(S_i) = P[r_i \leq S_i] = \int_0^{S_i} p_i(r_i)dr_i \tag{5.1}$$

Conversely:

$$P[r_i > S_i] = \int_{S_i}^{\infty} p_i(r_i)dr_i = 1 - P_i(S_i) = \overline{P}_i(S_i) \tag{5.2}$$

The probability of receiving more than S_i requests for fare class i, or the probability of spill occurring, is therefore $\overline{P}_i(S_i)$. An illustration of the relationship between $p_i(r_i)$ and $\overline{P}_i(S_i)$ is provided in Figure 5.1.

We define the number of requests accepted for fare class i to be b_i (bookings), and the number of refused requests to be l_i (spill). Given the number of seats allocated to each fare class, S_i, then:

$$\begin{aligned} &\text{if } r_i \leq S_i, \quad b_i = r_i \quad \text{and} \quad l_i = 0 \\ &\text{if } r_i > S_i, \quad b_i = S_i \quad \text{and} \quad l_i = r_i - S_i \end{aligned} \tag{5.3}$$

The average or expected number of bookings in class i, given a seat allocation of S_i, is therefore:

$$\bar{b}_i(S_i) = \int_0^{S_i} r_i \cdot p_i(r_i)dr_i + S_i \cdot \overline{P}_i(S_i) \tag{5.4}$$

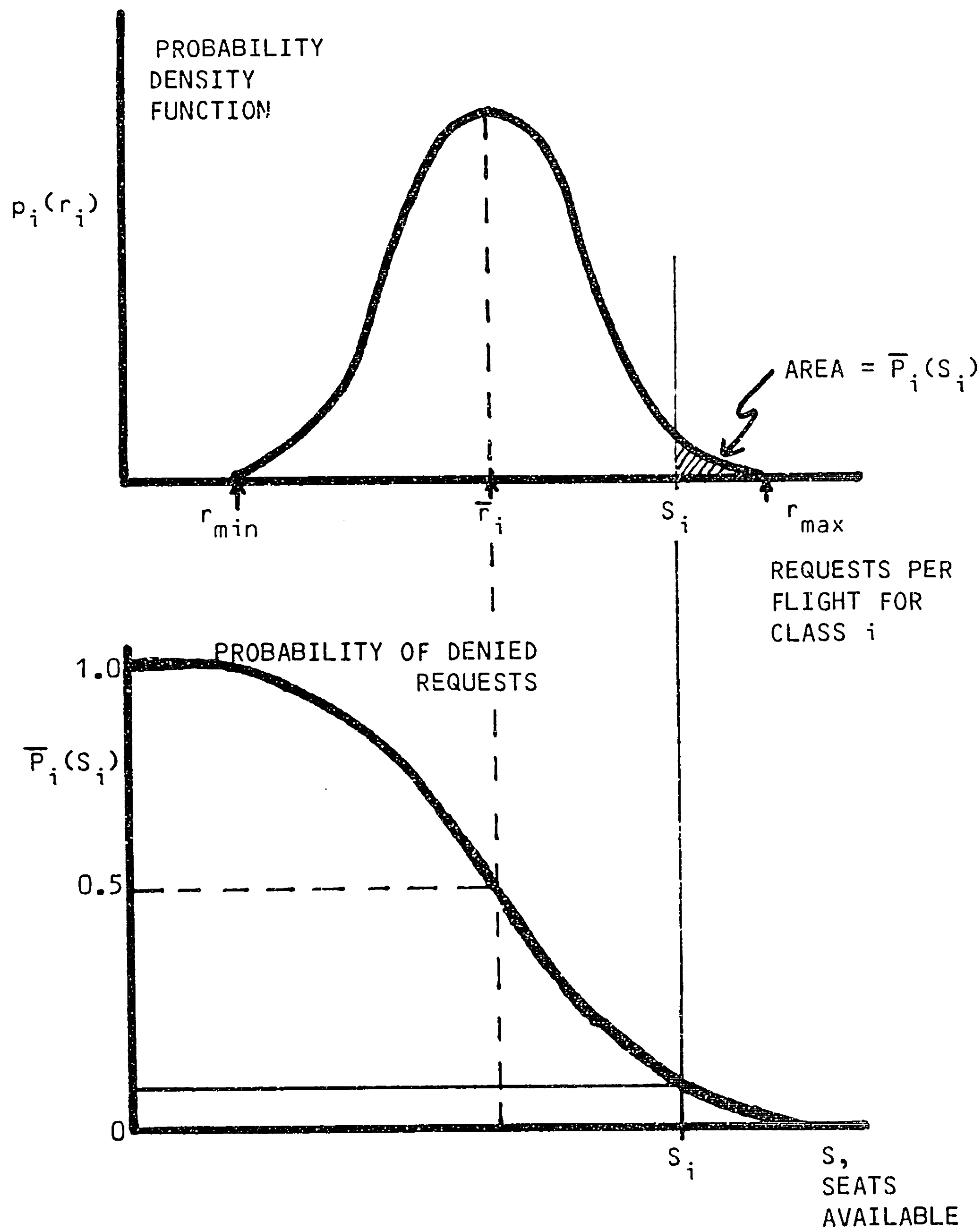
The average or expected number of passengers refused (bookings lost) is:

$$\begin{aligned} \bar{l}_i(S_i) &= \int_{S_i}^{\infty} (r_i - S_i) \cdot p_i(r_i)dr_i \\ &= \int_{S_i}^{\infty} r_i \cdot p_i(r_i)dr_i - S_i \cdot \overline{P}_i(S_i) \end{aligned} \tag{5.5}$$

From the above expressions, we see that:

$$\bar{b}_i + \bar{l}_i = \bar{r}_i \tag{5.6}$$

Figure 5.1: Typical Probability Distributions of Requests



The probability of any given request being refused in class i (the refusal rate) is simply:

$$PRR_i = \bar{l}_i / \bar{r}_i \tag{5.7}$$

It should be emphasized that PRR_i, the expected rate of request refusals, *does not* equal $\overline{P}_i(S_i)$, the probability that at least one request will be refused and the fare class will sell out.

The above discussion has referred to continuous densities of demand. In practice, the number of requests for a fare class and the number of seats offered are integer values. In applying these probabilistic concepts to models for determining fare class booking limits, the solutions must be constrained to be discrete values, even though the demand distributions may be represented by continuous density functions.

The implications of these probabilistic concepts for mathematical modelling of the problem become apparent in the simplest of revenue maximization models. We can expand on the static model for seat allocation introduced at the start of section 4.2 to include probabilistic demand explicitly. As before, we want to determine the optimal allocation of seats between distinct (i.e., non-nested) fare classes, subject to the total capacity constraint. The total capacity of the cabin to be shared among i fare classes is C, such that:

$$C = \sum_i S_i \tag{5.8}$$

Let f_i be the average fare received by the airline when a reservation request for fare class i is accepted. We want to find the integer values of S_i that will maximize total expected revenue, $\overline{R}$, for a flight:

$$\begin{aligned} \overline{R}_i(S_i) &= f_i \cdot \bar{b}_i(S_i), \qquad \text{for all i} \\ \overline{R} &= \sum_i \overline{R}_i \end{aligned} \tag{5.9}$$

subject to the capacity constraint.

In a two-class example, we have fare classes 1 and 2 with relative fares f_1 and f_2. To find the value of $S_1 = C - S_2$ that will maximize total expected revenues, $\overline{R}$, for the flight, we differentiate $\overline{R}$ with respect to S_1 and set to zero:

$$\begin{aligned} \overline{R} = \overline{R}_1(S_1) + \overline{R}_2(C - S_1) = &\quad f_1 \cdot \bar{b}_1(S_1) + f_2 \cdot \bar{b}_2(C - S_1) \\ \partial \overline{R} / \partial S_1 = &\quad 0 \end{aligned} \tag{5.10}$$

The expected marginal seat revenue for each class, $EMSR_i = \partial \overline{R}/\partial S_i$, will also be equal across classes at optimality, but will not necessarily be zero, due to the imposed capacity constraint. The capacity constraint could also result in a possible "corner" solution, in which all available seats are allocated to fare class 1, meaning $\partial \overline{R}/\partial S_1$ will not equal zero.

The expected marginal seat revenue of the $S_i th$ seat in fare class i, $EMSR_i(S_i)$, is simply the average fare level in that class multiplied by the probability of selling S_i or more seats:

$$EMSR_i(S_i) = f_i \cdot \overline{P}_i(S_i) \tag{5.11}$$

The optimal values of S_1 and S_2 in the case of two distinct fare class inventories must satisfy:

$$EMSR_1(S_1^*) = EMSR_2(S_2^*) \tag{5.12}$$

These optimal values of S_1 and S_2 will depend on the parameters of the probability densities of expected demand for each fare class, the relative fares or revenue levels, and the total capacity available. The relationships between $\overline{R}$, S_i and C are illustrated for the simple two-class model in Figure 5.2. $\overline{R} = \overline{R}_1 + \overline{R}_2$ is maximized when S_1^* out of a total of C seats are allocated to the higher-priced fare class 1.

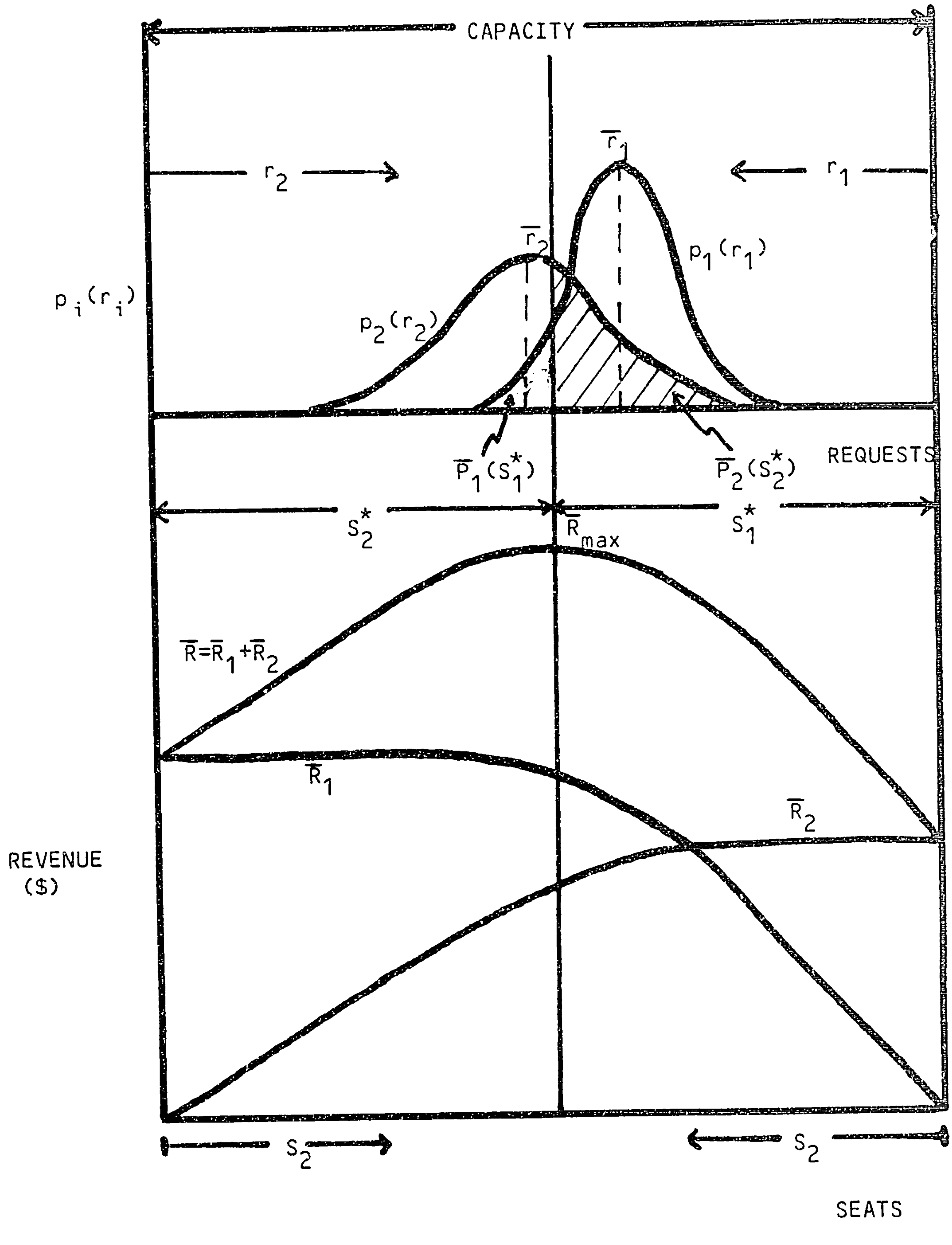
When this analysis is extended to three or more distinct fare classes, S_i^* for each fare class can be found to maximize expected total revenues given the densities of demand and relative fare levels, subject to a capacity constraint. With more than two S_i variables, the problem becomes more difficult to solve analytically, but can be formulated as a constrained revenue maximization problem and solved with the Lagrangian multiplier method. The optimality conditions for the multiple fare class problem would then be:

$$EMSR_i(S_i^*) = f_i \cdot \overline{P}_i(S_i^*) = \lambda, \quad \text{for all i} \neq \text{j} \tag{5.13}$$

where λ, the Lagrangian multiplier, equals the expected marginal seat revenue for each fare class. As before, seats are allocated among fare classes such that the expected marginal revenue is equal across all fare classes.

The values of S_i^* derived from this model represent the optimal allotments of the available seats to different fare classes based on expected demand for each fare class. The demand for each fare class is assumed to be independent of that for other classes,

Figure 5.2: Optimal Seat Allocation, Two Demand Classes



and the optimal seat allocation is made only once, at the beginning of the booking period for a flight. In reality, demand for different types of fares might not be independent, as high demand for one fare class could be associated with high demand for another. More importantly, this static seat inventory management model does not account for the dynamic nature of the reservations process in which actual bookings accepted for a particular flight might provide valuable additional information, as the time to departure decreases, about the ultimate number of requests that can be expected for that flight.

This simple model allocates a given number of shared seats among distinct or "stand-alone" reservations fare classes so as to maximize total expected flight revenues. Optimal seat allotments were derived under the assumption that any one seat could be made available to one fare class or another, but not both. In a nested reservations system, the possibility that more than the "allotted" number of seats may be sold to higher fare classes suggests that these formulations do not accurately represent the problem faced by many airlines. Any model that is used to find a fixed "allotment" of seats among fare classes in a nested reservations system overlooks the lost revenue potential when unexpectedly high demand causes high-fare requests to be refused when seats are available in lower classes.

5.2 Expected Revenues in Nested Fare Classes

One objective of seat inventory control, as introduced in Chapter One, is to limit the number of seats sold at less than the full coach fare. The classical surplus seat management concept was based on the protection of seats for high-fare passengers and made only those seats that would ultimately remain empty available to discount passengers. While many of the assumptions of the surplus seat system have become obsolete, this basic notion of protecting seats for higher-fare passengers remains valid.

It is especially valid in light of the characteristics of the reservations process for a future flight. The nature of discretionary price-sensitive travel (Type 3 consumers) and the application of advance purchase restrictions to the lowest discount fare products mean that the lowest fare classes will be requested and will book up first, well before the majority of requests for the highest fare class are received. The seat inventory control problem is therefore to determine how many seats ***not to sell*** in the lowest fare classes and to retain for ***possible*** sale in higher fare classes closer to departure day.

In a nested reservations system, seat inventory control must therefore be directed toward finding the ***protection levels*** for higher fare classes which can be converted into ***booking limits*** on lower fare classes. Each protection level is the ***minimum*** number of seats that should be retained for a particular fare class (and available to all higher fare classes). Each booking limit is the ***maximum*** number of seats that may be sold to a fare class (including all lower fare classes with their own, smaller, booking limits). The booking limit on the highest fare class is thus the total capacity of the shared cabin. The protection level for the highest fare class is the difference between its booking limit and the booking limit of the next lowest class.

The seat allocation models of the previous chapter assumed independent fare class demand densities to correspond with the distinct fare class inventory assumption. For the time being, the assumption of no relationship between demand levels for different fare classes will be retained. In other words, we will assume that, along the lines of the market demand segmentation model developed earlier, demand for a particular fare class can be identified and distinguished from other classes. Furthermore, we assume that there is no possibility of vertical or horizontal choice shifts on the part of a consumer denied a flight/fare class request, meaning a refused request is a booking loss to the airline. On the other hand, an accepted booking represents certain revenue for the airline, as we assume that no booking cancellations or passenger no-shows occur.

These assumptions are retained to allow the development of a seat inventory decision model to focus on the concepts of probabilistic demand and expected revenues in deriving booking limits. Once a simplified framework has been presented, the possibility of consumer choice shifts and booking cancellations or no-shows will be incorporated.

With the help of the probabilistic concepts introduced in Section 5.1, the expected marginal revenue approach to seat allocation can be extended to multiple fare classes in a nested reservations system. Given the historical density of requests for a fare class i and, in turn, the expected bookings as a function of S_i, the expected revenue from S_i seats available in class i is:

$$\overline{R}_i = f_i \cdot \bar{b}_i(S_i) \tag{5.14}$$

where f_i is the net revenue or average fare collected from passengers booked in class i.

We defined $EMSR_i$ to be the expected marginal seat revenue for class i when the number of seats available to that class is increased by one. The relationship between expected marginal seat revenues and expected total revenues for a fare class is illustrated in Figure 5.3. Note that $EMSR_i(S_i)$ depends directly on $\overline{P}_i(S_i)$, the probability that the number of requests will exceed the number of seats available.

To begin with a simple example, we consider a single-leg flight for which bookings will be accepted in two nested fare classes, 1 and 2, having average fare levels f_1 and f_2, respectively. In order to maximize total expected flight revenues, the reservations process should give priority to class 1 passengers. Class 1 will have the total available capacity of the shared cabin, C, as its booking limit, BL_1. The seats protected from class 2 and available exclusively to class 1 will be denoted S_2^1.

The optimal protection level $S_2^{\cdot}$ for class 1 is the largest integer value of S_2^1 that satisfies the following condition:

$$EMSR_1(S_2^1) \geq f_2 \tag{5.15}$$

The expected marginal seat revenue of protected seats in class 1 is equated to the average fare level of class 2 to find the optimal protection level for class 1. Graphically, the optimal value of S_2^1 is the point at which the $EMSR_1(S_1)$ curve intersects f_2, as shown in Figure 5.4(a).

The revenue-maximizing protection level for Class 1 is determined in the EMSR model by finding the value of S_2^1 that satisfies:

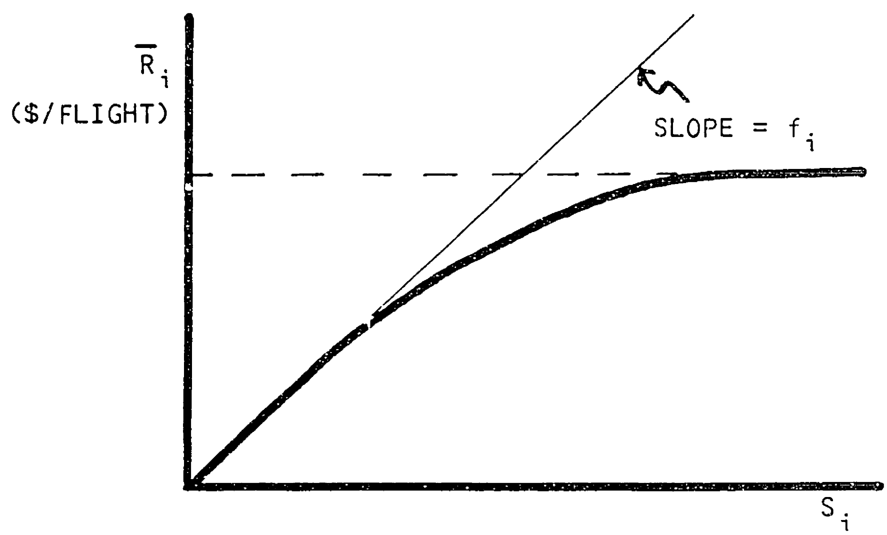
$$EMSR_1(S_2^1) = f_1 \cdot \overline{P}_1(S_2^1) = f_2 \tag{5.16}$$

This optimal protection level is ***not*** a function of the lower fare class demand density in the static case where nested buckets are involved. It *is* a function of the ratio of f_2 to f_1 and of the parameters of the high-fare demand density assumed from historical data. As noted in the previous chapter, Richter reached this same conclusion with respect to a dynamic reservations control process.

This solution will maximize *expected* revenues in cases where the booking limit is set only once at the start of the reservations process (static seat inventory control). A

Figure 5.3: Total and Marginal Seat Revenue Curves

(a) EXPECTED TOTAL REVENUE BY CLASS



(b) EXPECTED MARGINAL SEAT REVENUE BY CLASS

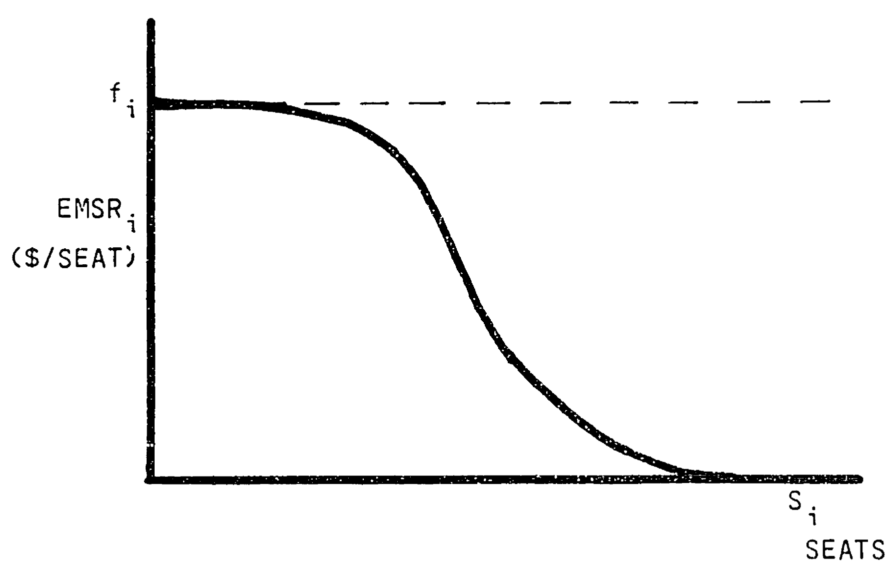
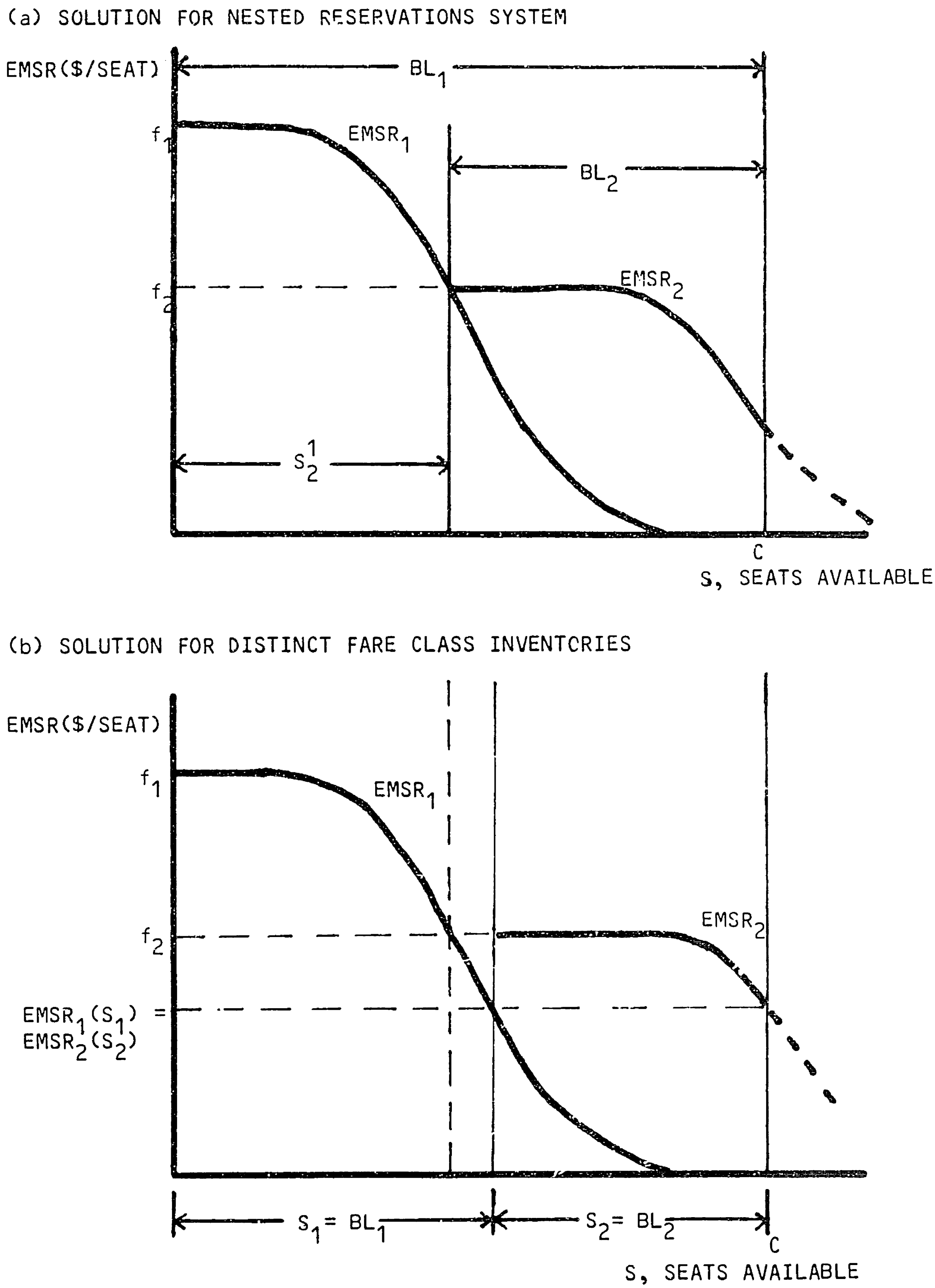


Figure 5.4: Maximizing Expected Revenues, Two Class Example



class 2 request will be rejected only when BL_2 is reached, at which point the expected revenue for all remaining seats will be greater than the class 2 average fare. If class 2 requests never reach BL_2, the unsold seats will be available for unexpectedly high class 1 demand. In any event, the *expected* revenue per seat from class 1 requests in excess of S_2^1 is below f_2, in the absence of additional information on actual bookings for the future flight being managed. Total expected revenues therefore cannot be increased by making BL_2 smaller. If BL_2 is increased, the airline will lose the difference $EMSR_1(S_1) - f_2$ in expected revenue for each seat less than S_2^1 protected for class 1.

The difference between this nested fare class solution and the "stand-alone" fare class solution is apparent in Figure 5.4(b). In the latter, seats are allocated between classes such that:

$$EMSR_1(S_1) = EMSR_2(S_2) \tag{5.17}$$

at optimality, subject to the capacity constraint. The result is that the final seat allocated to class 1 could well have an $EMSR$ lower than f_2, in which case unexpectedly high class 2 demand would be refused in favor of keeping that seat open for possible class 1 demand even though it has a lower expected revenue in the class 1 inventory.

Solution methods that treat class 1 and class 2 seats as distinct inventories can therefore produce less than optimal results when applied in a nested reservations system context. The differences in expected revenues between the two solutions, applied to nested fare classes, can be illustrated with a simple example. Table 5.1 summarizes the optimal limits and expected revenues for three solutions, for the case of a two-class shared cabin with a capacity of 10 seats. Identical independent Gaussian demand densities are assumed for the two classes. Fares are \$100 and \$60 for class 1 and class 2, respectively.

The "stand-alone" optimal solution allocates 6 seats to class 1 and 4 seats to class 2. When applied to distinct fare class inventories, the total expected revenue of this solution is \$684. When the "stand-alone" solution is applied to nested fare classes, determining the total expected revenue requires an assumption to be made with respect to the order in which bookings are made. In this example, it is assumed that all class 2 seats are requested before any class 1 requests are made. This assumption defines the lower bound on the estimate of total expected revenues in a nested system, since any assumption involving a parallel booking process would give class 1 requests greater access to class 2 seats.

Table 5.1: Optimal Limits and Expected Revenues

2 Fare Class Inventories, Single Flight Leg; Capacity 10 Seats.

	CLASS 1	CLASS 2	TOTAL
INPUT DATA:			
Fare ($)	100	60	—
Mean Demand	5	5	—
Std. Deviation	2	2	—
CASE 1: "STAND-ALONE" Solution, Distinct Fare Classes			
Optimal Allocation	6	4	10
Expected Revenue	$465	$219	$684
CASE 2: "STAND-ALONE" Solution, Applied to Nested Fare Classes			
Class 1 Protection	6	—	—
Class 2 Limit	—	4	—
Expected Revenue	$471	$219	$690
CASE 3: "EMSR MODEL" Solution, Nested Fare Classes			
Class 1 Protection	5	—	—
Class 2 Limit	—	5	—
Expected Revenue	$477	$255	$702

Under the assumption that class 2 passengers book first, we must consider the joint probabilities of not selling class 2 seats to class 2 passengers and selling these seats instead to excess class 1 demand. Thus, application of the stand-alone solution to nested fare classes will result in a total expected revenue higher than $684. The increment is equal to the expected marginal seat revenues of the 7th, 8th, 9th and 10th seats available to class 1, weighted by the probability that a maximum of 3, 2, 1, and 0 seats, respectively, will be sold to class 2 passengers. Total expected revenues therefore increase to $690, or by 0.9 percent, when the stand-alone solution is applied to nested fare classes.

The total expected revenue will generally be higher for nested fare classes than for distinct fare class inventories, given the same booking limits on the lower fare class, as long as there is some probability that not all class 2 seats will be booked by class 2 passengers and that class 1 demand will exceed the class 1 allotment. As mentioned earlier, total expected revenues will be higher still when classes 1 and 2 are assumed to book concurrently.

The third and final case summarized in Table 5.1 involves the EMSR model solution for the same simple example. The EMSR solution is different from the stand-alone solution, protecting 5 seats for class 1 and limiting class 2 sales to 5 seats. The total expected revenue for the EMSR solution in a nested system is $702, an increase of 1.7 percent over the previous case.

This simple example demonstrates how the EMSR solution for nested fare classes can in fact generate higher total expected revenues than the application of the "stand-alone" optimal allotments to a nested fare class system. The extent of this revenue difference will depend on the demand densities and booking process assumed. The logic of accepting all lower fare requests as long as the lower fare exceeds the expected marginal revenue of the higher fare class seat being "displaced" suggests that the EMSR solution will generate at least as much in total expected revenues as the stand-alone solution applied to nested fare classes, for a range of demand and booking assumptions.

Extension of the EMSR model to more than two fare classes on a single flight leg simply requires that more comparisons of expected marginal revenues be made among the relevant classes. Looking first at a 3-class example, we assume that $f_1 > f_2 > f_3$ and that these fare classes are hierarchically nested, as before. In this problem, we must find not only the number of seats to be protected for class 1 relative to class 2, we must also establish protection levels for classes 1 and 2 relative to class 3.

No change is required in the determination of S_2^1. We need only protect seats for the exclusive use of class 1 up to the point at which $EMSR_1(S_2^1) = f_2$, as before. Only S_2^1 seats will have an *expected* revenue greater than f_2, meaning all remaining seats may be made available to class 2. The optimal booking limit on class 2 (and lower classes) thus remains:

$$BL_2 = C - S_2^1 \tag{5.18}$$

With 3 fare classes, however, additional revenue comparisons involving class 3 must be made to determine BL_3. Using the EMSR approach, we determine two protection levels:

1. seats protected for class 1 from class 3, S_3^1;

2. seats protected for class 2 from class 3, S_3^2.

The optimal values of S_3^1 and S_3^2 must satisfy:

$$\begin{aligned} EMSR_1(S_3^1) &= f_3 \\ EMSR_2(S_3^2) &= f_3 \end{aligned} \tag{5.19}$$

respectively. The total protection level for classes 1 and 2 relative to class 3 can be taken as the sum of the two, and:

$$BL_3 = C - S_3^1 - S_3^2 \tag{5.20}$$

That is, all seats with an expected marginal revenue greater than f_3 should be held back from sale to class 3. Otherwise, any request for a class 3 seat may be accepted.

The values S_3^1 and S_3^2 are used in the EMSR model to determine an optimal BL_3, but will not actually appear in the reservations system. In essence, the incremental number of seats protected for class 2 above that required for class 1 alone is equal to $BL_2 - BL_3$. We shall refer to this number as the *nested protection level* for class 2, and denote it NP_2. The nested protection level for class j is thus given by:

$$NP_j = BL_j - BL_{j+1} \tag{5.21}$$

The nested protection levels can also be expressed in terms of the EMSR model seat protection outputs, S_j^i:

$$NP_j = \sum_{i \leq j} S_{j+1}^i - \sum_{i<j} S_j^i \tag{5.22}$$

These expressions for NP_j apply to all fare classes higher than the lowest fare class. In the EMSR framework, the lowest fare class does not have a protection level *per se*, but rather a booking limit equal to the number of seats that remain after all upper classes have been protected. With a capacity of C seats and k fare classes, then, the values of NP_j must satisfy:

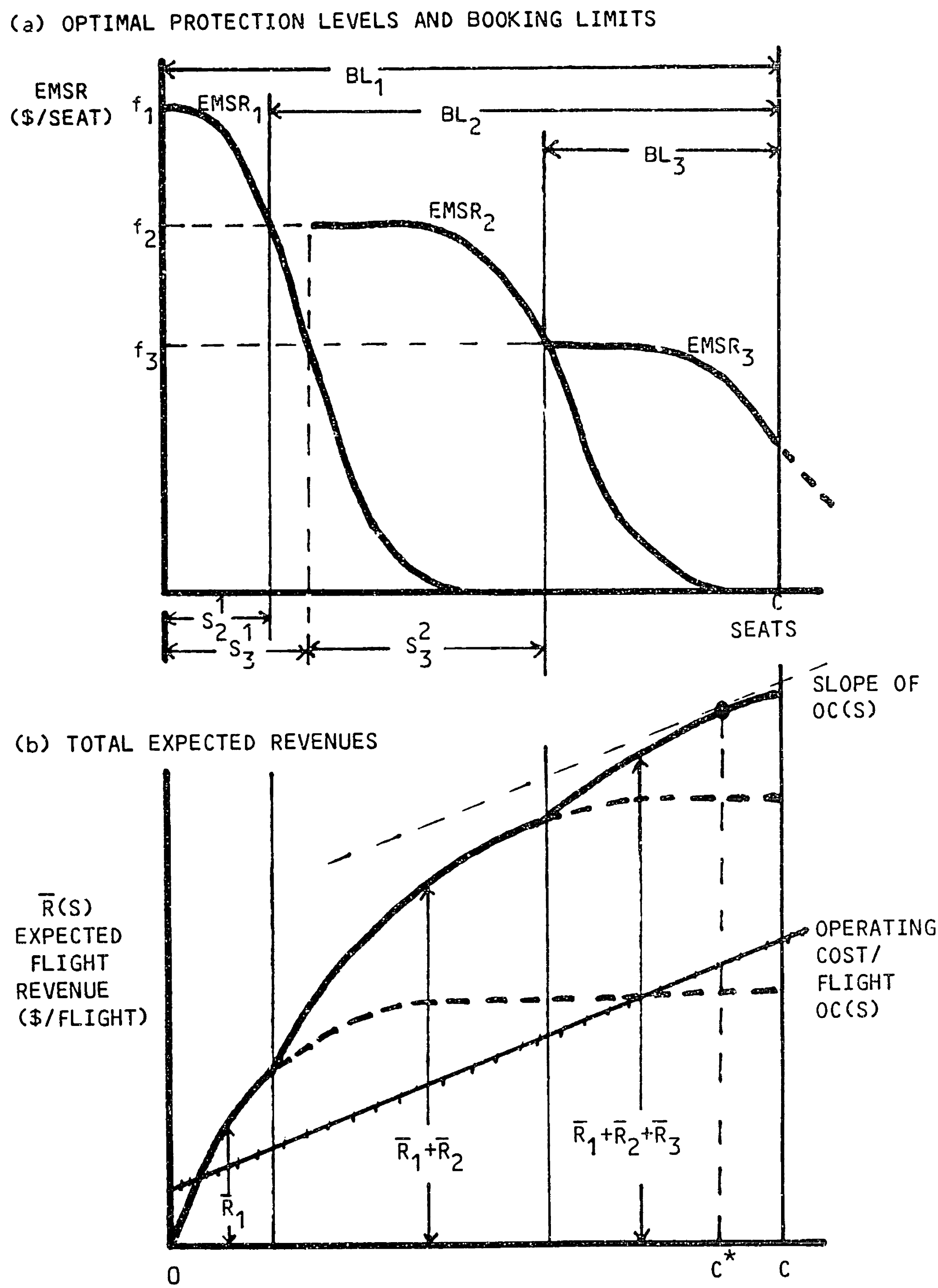
$$C = \sum_{j<k} NP_j + BL_k \tag{5.23}$$

The EMSR protection levels S_j^i, nested protection levels NP_j, and booking limits BL_j, are shown graphically for a three-class example in Figure 5.5(a). Figure 5.5(b) shows the relationship between the booking limits as derived from the EMSR model and the "stacked" configuration of total expected revenues by fare class in a nested reservations system. Total expected revenues are shown as a function of total seats available on the aircraft. As available seats are increased beyond S_2^1, seats become available to both classes 1 and 2, and total expected revenues reflect a sharp increase over those of Class 1 alone. The slope of the total expected revenue curve increases suddenly at these points due to the increased probability of selling the marginal seat made available to the lower fare class. The result is an "envelope" of stacked expected revenue curves.

This analysis of total expected revenue as a function of available seats can also be applied to aircraft sizing considerations. As shown in Figure 5.5(b), if operating costs per flight as a function of seating capacity on an "elastic" airplane can be estimated, an optimal seating capacity can be found. With three fare classes offered, operating income will be maximized when marginal flight revenues equal marginal flight costs, or where the total expected revenue curve has the same slope as the operating cost curve (i.e., at the point C^*). This slope is the marginal seat operating cost.

Given that many U.S. domestic airlines make use of at least four fare classes to manage their coach cabin seat inventories, we shall present the EMSR model solution to the nested 4-class problem. Six comparisons between upper class expected revenue

Figure 5.5: EMSR Solution for Nested 3-Class Example



curves and lower class fare levels are required. The optimal S values satisfy the following conditions:

$$\begin{aligned} EMSR_1(S_2^1) &= f_2 \\ EMSR_1(S_3^1) &= f_3 \\ EMSR_1(S_4^1) &= f_4 \\ EMSR_2(S_3^2) &= f_3 \\ EMSR_2(S_4^2) &= f_4 \\ EMSR_3(S_4^3) &= f_4 \end{aligned} \tag{5.24}$$

The resultant optimal booking limits on each fare class are then:

$$\begin{aligned} BL_1 &= C \\ BL_2 &= C - S_2^1 \\ BL_3 &= C - S_3^1 - S_3^2 \\ BL_4 &= C - S_4^1 - S_4^2 - S_4^3 \end{aligned} \tag{5.25}$$

In the general case of k fare classes offered on a flight leg, the optimal values of S_j^i must satisfy:

$$EMSR_i(S_j^i) = f_j, \quad i < j, \quad j = 1, \ldots, k \tag{5.26}$$

The total number of comparisons required for k nested fare classes is given by:

$$\frac{k(k-1)}{2} \tag{5.27}$$

These optimal protection levels in turn determine the optimal booking limits on each fare class j:

$$BL_j = C - \sum_{i<j} S_j^i \tag{5.28}$$

It is possible that one or more values of BL_j derived from these equations might be negative, in which case class j should not be offered at all if expected revenues are to be maximized. In such a case:

$$BL_j = MAX[0, \quad C - \sum_{i<j} S_j^i] \tag{5.29}$$

The EMSR model was presented in the context of a reservations system with complete hierarchical nesting of fare classes that share a common seat inventory. The decision rules can be applied to any number of nested fare classes with no change to the model's basic structure. It is important to recognize that not all reservations systems are nested in this way, or might not be nested at all. Nesting is preferable in seat inventory management, because there is no difference in the physical seats or on-board service being sold to different fare classes. Airlines without nested fare class systems are denying themselves the flexibility of accommodating unexpectedly high demand levels in high-fare classes and, in turn, are losing potential revenues.

In "stand-alone" fare class seat allocation, the mathematical programming methods described in Chapter Four will in fact produce optimal seat allotments for the static seat inventory problem. For nested fare classes, however, determining the number of seats to be protected for higher fare classes and the booking limits on lower fare classes requires the decision approach of the EMSR model. In reservations systems with partially nested fare class structures, the concepts of expected marginal seat revenues, protection levels and optimal booking limits remain valid, and the EMSR model can be adapted easily. In all cases, the EMSR model is suited to a dynamic seat inventory control process, as described below.

5.3 Dynamic Applications of the EMSR Model

The EMSR model outlined in the preceding section provides a decision framework for finding the revenue-maximizing booking limits on nested fare classes for a single flight leg. In its simplest application, this decision framework provides a solution to the *static* seat inventory control problem, in which booking limits must be set for a future flight at a particular point in time prior to departure, usually at the start of the reservations process. The static application requires estimates of the demand densities (from a sample of historical flights) and of the relative revenue levels associated with each of the fare classes to be offered on the future flight.

The same EMSR decision framework can be applied to a seat inventory control context in which booking limits may be revised on a regular basis as the flight departure day nears. In such a situation, additional information is available in the form of actual bookings already accepted for the future flight. Because an actual booking in any fare class will almost certainly (barring cancellations and no-shows) translate into a revenue passenger occupying a seat, incorporating actual bookings into the EMSR decision framework can reduce the uncertainty associated with the estimates of expected demand used as input.

In developing a dynamic seat inventory control decision model that makes use of the EMSR calculations, we will continue to make the simplistic assumptions with respect to demand densities, refused requests, cancellations and no-shows, as outlined in the previous section. Furthermore, we add the assumption that there is no relationship between the booking rates of different fare classes or among time periods before departure. As before, the reservations buckets are nested hierarchically in descending order of fare level.

The availability of actual booking information by fare class is of little practical value to the EMSR model unless an of estimate future requests expected over the remaining time before departure can be derived from historical data. In the static problem, the EMSR model required an estimate of *total* expected requests by class. In the dynamic case, estimates of future requests at various times before departure are required to calculate optimal protection levels for the unbooked seats still available for the flight.

Dynamic application of the EMSR framework in essence involves repetitive use of the static model described in the previous section, but with revised input data. The objective is to determine the optimal fare class limits for the time period remaining to departure, irrespective of the (non-) optimality of the booking decisions already made. Each EMSR calculation in the dynamic case is thus based on a static assessment of expected fare class revenues from that point in time, based on the most recent available demand information for the flight leg.

In the previous section, the EMSR decision model derived the optimal booking limits, BL_i, for each fare class on the basis of the input values of f_i and $\overline{P}_i(S_j^i)$. For the dynamic problem, the estimates of bookings to come must be generated from densities of demand by fa ss from a historical sample of requests made between day t before departure and the day of departure. We define r_i^t to be the number of requests made for class i between days t and 0 before departure, meaning:

$$r_i^t \leq r_i \tag{5.30}$$

by definition. The probability density of requests from day t onward is then $p_i(r_i^t)$, and the probability of receiving S or more requests for class i in the time remaining to departure is $\overline{P}_i^t(S)$.

On any day t prior to flight departure, the inputs required by the EMSR model are the average fare or revenue levels for each fare class, f_i, which may or may not remain constant over the booking period, and the estimates of $\overline{P}_i^t(S)$ for all relevant values of S, derived from the $p_i(r_i^t)$ densities. The optimal seat protection level for class 1 relative to class 2 for the period remaining before departure is $S_2^1(t)$, such that:

$$EMSR_1^t[S_2^1(t)] = f_1 \cdot \overline{P}_1^t(S_2^1) = f_2 \tag{5.31}$$

This optimal protection level for day t can be used to find the revised optimal booking limit on class 2, as follows:

$$BL_2(t) = C - b_1^t - S_2^1(t) \tag{5.32}$$

where b_1^t is the number of bookings already accepted in class 1 up to day t before departure. The maximum number of seats still available is $C - b_1^t$, and $S_2^1(t)$ of these seats are protected for class 1. In essence, actual bookings are "protected" along with additional seats required to accommodate expected bookings to come in a revenue-maximizing manner.

For more than two fare classes, comparisons must be made between the $EMSR(S)$ values of all upper classes relative to the average fare levels of lower classes, as before. These comparisons involve the demand densities of future requests from the current day t. Each comparison of a higher fare class i with a lower fare class j generates an optimal value of $S_j^i(t)$ that satisfies:

$$EMSR_i[S_j^i(t)] = f_i \cdot \overline{P}_i^t(S_j^i) = f_j \tag{5.33}$$

The revised booking limits for day t take actual bookings into account:

$$BL_j(t) = C - \sum_{i<j} S_j^i(t) - \sum_{i<j} b_i^t \tag{5.34}$$

The nested protection levels, $NP_j(t)$, for successively lower fare classes are derived as in the initial case, but with actual bookings included:

$$NP_j(t) = BL_j(t) - BL_{j+1}(t) \tag{5.35}$$
$$= \sum_{i \le j} S^i_{j+1}(t) - \sum_{i<j} S^i_j(t) + b^t_j$$

As before, $BL_j(t)$ is constrained to be greater than or equal to zero. It is also constrained in this case to be no lower than the actual number of bookings already accepted in class j and all lower classes up to day t. Because requests for lower fare classes are generally received earlier in the booking process than full-fare requests, it is possible that the revised $BL_j(t)$ derived from the EMSR calculations will be lower than the number of bookings already on hand in classes j and lower. With no possibility of cancellations or no-shows assumed, the revised booking limit on class j then becomes:

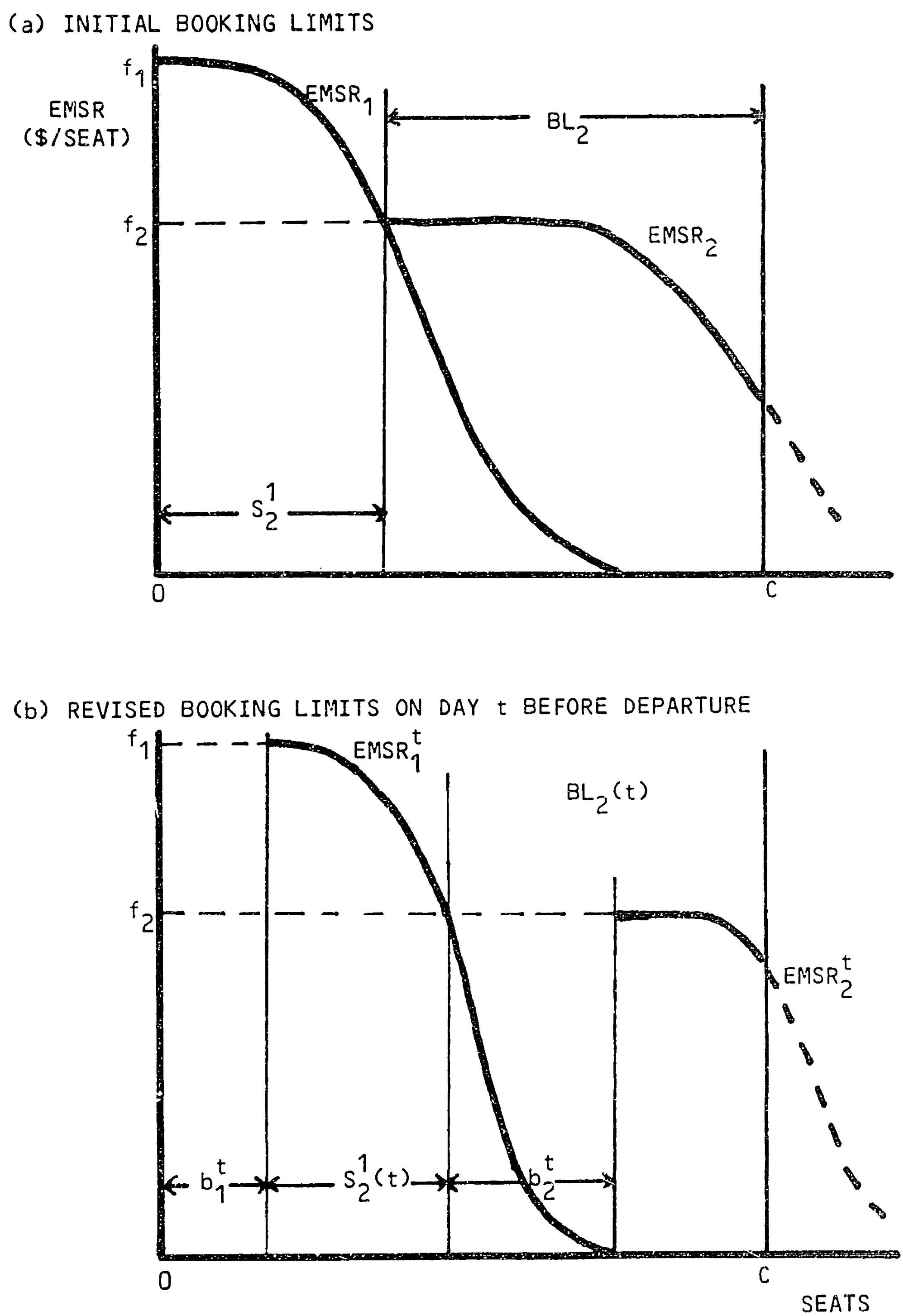
$$BL_j(t) = MAX[C - \sum_{i<j} S^i_j(t) - \sum_{i<j} b^t_i, \quad \sum_{k \ge j} b^t_k, \quad 0] \tag{5.36}$$

Figure 5.6 illustrates how actual bookings in a class receiving protection can cause the booking limit on a lower class to be revised downward in a dynamic seat inventory control process. The initial booking limit on Class 2, BL_2, shown in Figure 5.6(a), is reduced to $BL_2(t)$ in Figure 5.6(b), due to the actual bookings on hand in class 1 at day t and the need to protect additional seats for the expected future requests for class 1 before departure.

5.4 Passenger No-shows and Flight Overbooking

The EMSR decision framework has to this point overlooked the possibility that a reservation might not translate into a passenger being carried on the flight in question. With each accepted booking assumed to represent certain revenue, the model has focused on the problem of managing the physical seat inventory on the aircraft. As described in Chapter Two, there is some probability that a booked passenger might not be carried on the flight for which the reservation was made. Whether the original booking is cancelled prior to departure or the passenger simply fails to appear at departure time, the outcome from the airline's perspective is the loss of a revenue passenger for that particular flight.

Figure 5.6: Dynamic Application of EMSR Model, Two Nested Classes



This loss is the opportunity cost of having removed a seat from the available inventory and then not receiving any revenue for it.

The airline industry practice of controlled overbooking of flights above the physical capacity of the aircraft has evolved with the objective of minimizing such costs. Overbooking analysis is performed to determine the extent to which a future flight should be overbooked so as to minimize the sum of the lost revenues associated with empty seats and the costs of denying boarding to passengers with confirmed reservations.

Early work on this problem treated the overbooking level as a fixed number that applied throughout the booking process for a future flight [59]. More sophisticated models allow for bookings to be cancelled throughout the reservations process, such that the optimal overbooking level for a future flight is a function of the time left to departure. These overbooking models require estimates of the demand densities and cancellation probabilities for future flights. Another approach was proposed by Rothstein, who developed a model for maximizing passenger revenues subject to a limit on the proportion of denied boardings [60].

For the purposes of seat inventory control, it is important to recognize the potential interaction between the fare class mix of passengers booked for a flight and overbooking, one which could have significant revenue implications for the airline. In this section, the EMSR decision framework is extended to account for flight overbooking. It is not the intent of this work to develop an overbooking model. Rather, the objective is to illustrate how overbooking proportions by flight and even by fare class can be incorporated into the EMSR framework for deriving revenue-maximizing fare class booking limits.

Most airline overbooking models generate estimates of the overbooking percentage that should be applied to the total number of seats, by physical compartment on the aircraft. This overbooking percentage can be a function of both expected cancellation rates before departure and no-show probabilities at departure time. Given that cumulative cancellation probabilities generally decrease as time to departure decreases, it is common for the overbooking percentages to be reduced as departure day nears. For the purposes of this discussion, we define the "overbooking factor" to be OV, where $OV \geq 1.0$, and where OV applies at the time of the EMSR calculations.

With the introduction of the possibility that a booking made in fare class i will not generate revenue f_i with certainty, the EMSR decision framework must be adjusted.

The demand inputs required are still estimates of the densities of *requests* for each fare class. The difference here is that each accepted request (booking) in a fare class cannot be treated as if the revenue associated with that class will always be realized. The problems of estimating request densities from actual booking data will be discussed in Chapter Six. At this point, it suffices to note that the expected revenue associated with accepting a request will be lower than the actual fare level in that class, because of the possibility of cancellation or no-show.

The overbooking factors for a fare class determine the extent to which the expected revenue from a booking is reduced due to this uncertainty. Given overbooking factors OV_1 and OV_2, the optimal protection level for class 1 from class 2, S_2^1, must satisfy:

$$\overline{P}_1(S_2^1) \cdot f_1 \cdot \frac{1}{OV_1} = f_2 \cdot \frac{1}{OV_2} \tag{5.37}$$

The revenue levels of the two classes are in essence "deflated" by the assumed overbooking factors. We are thus reducing the expected marginal revenue of each incremental request for class 1 for which the possibility of protection is being evaluated.

The value of S_2^1 that satisfies (5.37) is the protection level for class 1 from class 2, expressed in terms of the number of *reservations spaces* (as opposed to physical seats) that should be protected for exclusive use of class 1 passengers. Because the demand inputs to the EMSR framework for all fare classes will be in terms of expected requests, the calculated protection levels for all upper fare classes will be in terms of reservations spaces, with the relevant overbooking factors already incorporated. The generalized decision rule for the EMSR framework with overbooking factors is to protect S_j^i seats for class i from class j such that:

$$EMSR_i(S_j^i) \cdot \frac{1}{OV_i} = \overline{P}_i(S_j^i) \cdot f_i \cdot \frac{1}{OV_i} = f_j \cdot \frac{1}{OV_j} \tag{5.38}$$

The derivation of optimal overbooking limits, BL_j^*, from this revised EMSR decision rule is complicated by the fact that the optimal protection levels are expressed in terms of reservations spaces (overbooking factors included), while the capacity of the aircraft is in terms of physical seats. The simplest case occurs when all the OV_i values are equal across fare classes, so that the OV factors drop out of the above formulation, reducing it to the original formulation in (5.26). The overbooking limit on the total capacity of the shared aircraft cabin, C^*, is simply:

$$C^* = BL_1^* = OV \cdot C \tag{5.39}$$

Booking limits for each subordinate fare class in the coach cabin may be derived after an overbooking target is established for the total capacity of the cabin. The protection levels generated by the EMSR model are for *reservations spaces* rather than physical seats, and the overbooking limits on each fare class j are given by:

$$BL_j^* = C^* - \sum_{i<j} S_j^i \tag{5.40}$$

The net result is that each fare class may be overbooked by the same percentage and C^* will be the same regardless of the fare class mix actually booked for any particular flight.

The fact that fare classes are designed to appeal to different air travel demand segments suggests that passengers booked in each fare class might exhibit different no-show behavior. As discussed in Section 2.2, it is plausible that passengers in lower fare classes will be more likely to show up for booked flights than those in higher fare classes for several reasons, the most important being the cancellation penalties associated with the lowest excursion fares. Few airlines currently have the detailed no-show data required to confirm this hypothesis empirically.

The application of different overbooking levels, OV_j, to each fare class j would be relatively straightforward in a non-nested fare class structure, as each BL_j^* would simply be the number of reservations spaces allocated to class j. In a nested structure, it is possible for all seats to be booked in the highest fare class, for example, in which case the total overbooking limit derived from different OV_j values might not be correct. That is, a total cabin overbooking limit based on an expected fare class mix of bookings might not be high enough in the extreme case where all seats are taken by the highest fare class with the highest no-show rate.

In a nested fare class structure with protection levels S_j^i derived from the EMSR model, it is not possible to determine analytically the fare class mix that can be expected, in the absence of assumptions concerning the relative booking rates of each of the fare classes. Application of different overbooking factors to each fare class at the start of the booking process, before any bookings are accepted, requires some estimate of an ultimate fare class mix of bookings. If we assume that, for a high demand flight, the lowest fare classes will book up first, then the simplest estimate of total bookings by fare class, given the EMSR protection levels, would be the nested protection level derived for each successively lower fare class, NP_j.

These estimates of the maximum number of bookings that can be expected in each of the upper fare classes are already expressed in terms of reservations spaces. For the lowest fare class, class 4 in this case, we must determine the number of physical seats that remain before inflating this number to incorporate overbooking. The incremental protection levels for each of the upper fare classes must be deflated by their respective OV_i values and subtracted from the physical capacity of the cabin. The difference is then inflated by OV_4 to find the overbooking limit on class 4:

$$BL_4^* = OV_4[C - \frac{1}{OV_1}NP_1 - \frac{1}{OV_2}NP_2 - \frac{1}{OV_3}NP_3] \tag{5.41}$$

Note that, with equal OV_i values for all fare classes, this equation reduces to (5.40), for $j = 4$.

Having found BL_4^*, we can derive the overbooking limits on each of the upper fare classes by adding the incremental protection levels, as calculated from the EMSR decision model:

$$\begin{aligned} BL_3^* &= BL_4^* + NP_3 \\ BL_2^* &= BL_3^* + NP_2 \\ BL_1^* &= BL_2^* + NP_1 \end{aligned} \tag{5.42}$$

An equivalent expression for the overbooking limit that applies to the entire coach cabin is:

$$C^* = S_4^1 + S_4^2 + S_4^3 + BL_4^* \tag{5.43}$$

since the sum of the top 3 nested protection levels is simply the total protection for the upper classes from class 4. The aggregate overbooking factor on the shared cabin, OV, is the ratio of total reservations spaces to total seats, C^*/C.

The total authorized overbooking limit, in the absence of actual bookings for the future flight, is thus determined by the total protection level calculated for the upper fare classes. The number of physical seats that remain for the lowest class must be determined, inflated by the relevant overbooking factor, and added to the summed upper class protection levels. It is possible that the sum of all the upper class protection levels, when deflated to physical seats, will still exceed the physical capacity of the shared cabin. In this case, BL_4^* must be constrained to be greater than or equal to zero. Furthermore, in the subsequent derivation of BL_j^* for upper classes, the same constraint applies.

These overbooking limits take into account the possibility of different no-show rates by fare class, and are based on an estimate of the ultimate fare class mix of bookings. As bookings are accepted in different fare classes, these overbooking limits can be refined to reflect the actual fare class mix for a particular flight. The dynamic adjustment capabilities of the EMSR model, as described in the previous section, enable us to revise both the fare class protection levels and overbooking limits as departure day nears.

Recall that the dynamic application of the EMSR model revises fare class booking limits for the physical seats on the aircraft periodically to account for actual bookings accepted and to protect additional seats as necessary. Without no-shows or overbooking, the booking limit on class j at day t before departure was given by:

$$BL_j(t) = C - \sum_{i<j} S_j^i(t) - \sum_{i<j} b_i^t \tag{5.44}$$

With the introduction of overbooking factors by fare class, both the protection levels and actual bookings are expressed in terms of reservations spaces, not physical seats. The ***total*** incremental seat protection level for each successively lower fare class at day t can be defined as the sum of the nested protection level calculated from the EMSR model for the period remaining to departure, $NP_j(t)$, and the number of bookings already on hand in that fare class, b_j^t.

Starting with class 1, the EMSR formulation generates a protection level for the period from day t to departure, $S_2^1(t)$, based on estimates of requests still to come in class 1. Actual bookings are added to this protection level, and the total protection for class 1 on day t is then the difference between the overbooking limits on class 1 and class 2:

$$BL_1^*(t) - BL_2^*(t) = S_2^1(t) + b_1^t \tag{5.45}$$

For class 2, the nested protection for future requests above that required for class 1 is given by $NP_2(t)$. This incremental protection level and actual bookings in class 2 are added to give the difference between the overbooking limits on class 2 and class 3:

$$BL_2^*(t) - BL_3^*(t) = NP_2(t) + b_2^t \tag{5.46}$$

Similarly:

$$BL_3^*(t) - BL_4^*(t) = NP_3(t) + b_3^t \tag{5.47}$$

To this point, the differences between adjacent overbooking limits have been defined in terms of nested protection levels for future requests plus actual bookings already on hand for the fare class in question. The overbooking limit on the lowest fare class, class 4 in this example, depends on the total of the protection levels and actual bookings in the upper fare classes. Thus, we inflate whatever seats remain unprotected and unsold by OV_4. Along with upper class nested protection levels, actual bookings in the top 3 classes must be deflated by their respective OV_j values to determine how many *physical* seats remain unprotected and unsold. The optimal overbooking limit on class 4 on day t is then given by:

$$BL_4^*(t) = OV_4\{C - \sum_{i<4} \frac{b_i^t}{OV_i} - \frac{1}{OV_1}[NP_1(t)] - \frac{1}{OV_2}[NP_2(t)] - \frac{1}{OV_3}[NP_3(t)]\} \tag{5.48}$$

Note that class 4 does not have any "protection level" for its own use, and that $BL_4^*(t)$ is a function of protection levels and bookings in upper classes, relative to C. Actual bookings in class 4 are also not included in the derivation of $BL_4^*(t)$. It is possible that b_4^t could prove to be ***higher*** than $BL_4^*(t)$ at some point in time prior to departure in cases where $BL_4^*(t)$ has been revised downward in the dynamic adjustment process. Most airline reservations systems will allow such an inconsistency to be retained so that cancelled class 4 bookings do not "open up" spaces for additional class 4 bookings.

The coach cabin physical capacity, C, does not enter into the overbooking limit computation except in conjunction with class 4. This is a result of the EMSR approach of allowing any unprotected seats to be made available to the lowest fare class. The complete expression for the total cabin overbooking limit as derived from the above equations is as follows:

$$C^*(t) = BL_1^*(t) = \sum_{i<4} S_4^i(t) + \sum_{i<4} b_i^t + OV_4[C - \sum_{i<4} \frac{b_i^t + NP_i(t)}{OV_i}] \tag{5.49}$$

This formulation for $C^*(t)$ can readily be modified to accept different OV_i values as a function of time before departure. Several overbooking models developed in the past make use of cancellation probabilities to adjust overbooking factors downwards as departure day approaches. If such an overbooking approach is to be used, the relevant $OV_i(t)$ factors may be included in the above formulation.

5.5 Incorporating Passenger Choice Shifts

Having extended the EMSR formulation to incude the possibility that an accepted booking might not translate into a revenue passenger carried, we now turn to the issue of passenger choice shifts during the booking process. As described in Section 2.2, a refused reservations request for a particular future flight and fare class does not always result in a booking loss to the airline. Depending on the individual consumer's choice process, the unavailability of a desired flight and fare class can lead to:

1. a vertical shift to a higher fare class, same flight;
2. a horizontal shift to a different flight, same fare class and airline;
3. a booking loss to the refusing airline.

For the purposes of managing the seat inventory for a single flight leg, the probability of interest is that of a vertical shift in fare classes on the same flight. The probability that a passenger refused a request for fare class i will accept a booking in the next highest fare class $(i-1)$ was defined in Section 2.2 to be $P_i(v)$. Although the probability of horizontal shift, $P_i(h)$, and the overall recapture rate is also important to the airline wishing to maximize system revenues, our focus on flight leg inventories requires us to assume that a horizontal choice shift is equivalent to a booking loss. In this section, the EMSR formulation is extended to include the possibility of vertical shift on the part of the refused passenger. Only shifts from the requested fare class to the next highest fare class will be considered. We assume that, given a competitive market and a fare structure that segments market demand effectively, a vertical choice shift of two or more fare classes will be unlikely.

The EMSR formulation finds the optimal number of seats to be protected for the exclusive use of the highest fare class, and in turn for each successively lower nested fare class. There is some probability that a refused class 2 request will in fact become a class 1 booking, due to a vertical choice shift on the part of the denied consumer. Defined as $P_2(v)$, this probability is assumed to apply equally to all class 2 requests, regardless of time before departure.

When a class 2 request is received by the airline and a booking accepted, a revenue of f_2 is realized, in the absence of overbooking and no-shows. If the request is refused, the expected revenue associated with the denied passenger accepting a vertical shift in fare classes is:

$$P_2(v) \cdot f_1 \tag{5.50}$$

We want to find the incremental protection level required for class 1 so as to take into account this potential class 1 revenue when class 2 is closed.

From the basic EMSR formulation, a protection level for class 1 from class 2 of S_2^1 will still be required. Additional seats protected for class 1, V_2^1, can be taken either by a refused class 2 passenger or a class 1 passenger. The expected revenue from a class 1 passenger in the V_2^1th additional protected seat is:

$$EMSR_1(S_2^1 + V_2^1) = f_1 \cdot \overline{P}_1(S_2^1 + V_2^1) \tag{5.51}$$

If the upgrade probability, $P_2(v)$, is greater than zero, the incremental expected revenue associated with potential vertical shifts from class 2 may be realized if the seat is not purchased by a class 1 passenger.

The combined expected marginal seat revenue for the V_2^1th seat protected for class 1 is thus equal to f_1 multiplied by the probability that a class 1 request will be received for that seat *or* a vertical choice shift is accepted, given that BL_2 is reached. The optimal value of V_2^1 must therefore satisfy:

$$f_1\{\overline{P}_1(S_2^1 + V_2^1) + P_2(v) - P_2(v)\overline{P}_1(S_2^1 + V_2^1)\} = f_2 \tag{5.52}$$

where S_2^1 is the protection level for class 1 in the absence of vertical choice shifts, as before. Factoring and re-arranging the probability terms in the above equation produces the following equivalent condition:

$$EMSR_1(S_2^1 + V_2^1) \cdot [1 - P_2(v)] + P_2(v) \cdot f_1 = f_2 \tag{5.53}$$

The combined expected revenue from each additional seat protected for class 1 will be greater than or equal to f_2, given that BL_2 is reached. If BL_2 is not reached, this additional protection will have no impact.

The total protection level for class 1 from class 2 is then $(S_2^1 + V_2^1)$ and the booking limit on class 2 (without overbooking) is:

$$BL_2 = C - S_2^1 - V_2^1 \tag{5.54}$$

Overbooking factors may be incorporated as before, in which case the total EMSR protection level $(S_2^1 + V_2^1)$ would be treated the same as S_2^1 in the calculation of overbooking limits.

The determination of BL_3 to include the probability of vertical choice shifts from class 3 to 2 is very similar to the approach described above. The basic EMSR model sums S_3^1 and S_3^2 to find the total protection for classes 1 and 2 from class 3. These values remain intact, but we want to find the incremental protection required to account for vertical shifts from class 2 to 3. Because we assume that only shifts of one fare class are possible, the optimal value of V_3^2 will depend on $P_3(v)$, and must satisfy:

$$EMSR_2(S_3^2 + V_3^2)[1 - P_3(v)] + P_3(v) \cdot f_2 = f_3 \tag{5.55}$$

The corresponding booking limit on class 3, without overbooking, is:

$$BL_3 = C - S_3^1 - S_3^2 - V_3^2 \tag{5.56}$$

In general terms, the incremental protection required for any class i to account for the probability of vertical choice shift from class $j = i + 1$, is V_j^i such that:

$$EMSR_i(S_j^i + V_j^i)[1 - P_j(v)] + P_j(v) \cdot f_i = f_j \tag{5.57}$$

The general formula for the resultant booking limit on class j is:

$$BL_j = C - \sum_{i<j} S_j^i - V_j^i \tag{5.58}$$

As before, the nested protection levels are defined by the differences between successively lower booking limits. Overbooking factors may be applied to these protection levels to find the optimal overbooking limits, as described in the previous section, with $(S_j^i + V_j^i)$ replacing the protection levels for the relevant classes.

The impact of including one or more $P_i(v)$ values in the EMSR formulation will be an *increase* in the protection levels for each of the higher fare classes. Each lower fare class will see its booking limit decrease by the incremental protection level required to account for the possibility of vertical choice shifts to the next-highest fare class. The magnitude of this decrease will depend on the relative magnitudes of the $P_i(v)$ values estimated or assumed, highlighting the importance of this probabilistic element to the EMSR framework.

5.6 Examples of Model Results

The previous sections of this chapter have presented the single-leg Expected Marginal Seat Revenue model in terms of the mathematical formulations required to find optimal fare class seat protection levels and booking limits. This final section illustrates the application of these mathematical formulations to a future flight leg departure. Based on historical demand and revenue data for a hypothetical flight leg, the outputs of the EMSR model are presented for the static case (initial booking limits), in a dynamic limit adjustment example, with overbooking factors included in the calculations, and, finally, with upgrade probabilities taken into account.

The example used throughout this description is that of a medium-haul flight leg to be operated with a B-737 aircraft. Four fare class inventories (Y, M, B, Q) are to share the 107 seats in the coach cabin, and the reservations system is assumed to be completely nested from the lowest to the highest fare class. Table 5.2 shows the demand and revenue estimates used as inputs to the EMSR framework to establish "initial" fare class booking limits. These initial limits are set before any bookings are accepted for the future flight, or at least well before the number of actual bookings in any fare class would threaten the most conservative authorized booking limits.

The prorated flight leg revenue estimates by fare class shown in Table 5.2 reflect the hierarchy and range of revenue levels characteristic of an actual competitive market. These revenue inputs will remain constant throughout this example, unlike the demand inputs. The demand estimates shown are for *total* expected requests by fare class, as would be derived from a historical sample of past operations of the same flight leg under similar conditions. In the dynamic application of the EMSR model, these demand estimates will change as the time to departure decreases. The means and standard deviations in each case define an assumed Gaussian demand density for each fare class.

Table 5.2 also shows the EMSR protection levels, nested protection levels and initial booking limits for the flight. These outputs are derived from preliminary estimates of total expected requests by fare class, and do not take into account actual bookings received. The limits would apply to this flight until bookings to date and estimates of future requests begin to differ significantly from the original estimates of total requests.

Table 5.2: EMSR Example — Initial Booking Limits

Capacity = 107 seats

INPUTS:

	Y	M	B	Q
Total Expected Requests	20.3	33.4	19.3	29.7
Standard Deviation	8.6	15.1	9.2	13.1
Average Revenue ($)	105	83	57	39

CALCULATIONS:

$S_M^Y = 14$; $BL_M = C - S_M^Y = 107 - 14 = 93$

$S_B^Y = 20$, $S_B^M = 27$; $BL_B = C - S_M^Y - S_B^M = 60$

$S_Q^Y = 24$, $S_Q^M = 35$, $S_Q^B = 15$; $BL_Q = C - S_Q^Y - S_Q^M - S_Q^B = 33$

EMSR MODEL RESULTS:

	Y	M	B	Q
Nested Protection	14	33	27	—
Nested Booking Limits	107	93	60	33

That is, the dynamic revision process should begin when, according to historical information, more bookings should have been received in one or more of the upper fare classes. It should also be initiated when actual bookings in one or more of the upper fare classes exceed or fall below the number expected up to that point in time. Because the EMSR model does not protect seats for the lowest fare class, expected booking rates for that class need not be taken into account, and actual bookings will be limited to the initial authorized level (i.e., 33) until the first revision run is made.

In this example, day 35 before departure was judged to be the first time that booking limits should be revised on the basis of a change in the estimates of future requests. The estimates of future requests used as inputs reflect the assumption of no correlation between actual bookings and future requests. The possibility of deriving conditional estimates of future requests based on actual bookings is explored in Chapter Six.

Table 5.3 shows the mean number of requests for each fare class expected from day 35 to departure, which, by definition, will be lower than or equal to the initial total expected requests. New protection levels are derived from the EMSR framework to accommodate this "remaining" demand at day 35. Actual bookings on hand at day 35 are then added to these protection levels to find the "total protection" required for each upper fare class. The actual bookings shown in Table 5.3 for day 35 reflect a higher than expected demand for M class. The revised booking limits are thus lower than the initial limits for both B and Q classes, due to the M-class demand.

Table 5.4 shows the progression of expected requests still to come and actual bookings on hand, along with revised protection levels and booking limits for a dynamic application of the EMSR model on days 28, 21, 14 and 7 before departure. Higher than expected M-class bookings in each period lead to successive increases in the total M protection derived by the EMSR framework and, in turn, successively lower booking limits on B class. The total protection required for B class remains relatively stable. The booking limits on Q class decrease due to the increased protection for M and the stable protection for B.

The downward revision of the Q booking limit demonstrates how relatively infrequent revision runs can allow too many seats to be sold in a low fare class. The Q booking limit is revised from its previous level of 27 down to 24 on day 21, by which time 27 Q-class bookings have already been accepted. The result is essentially "negative availability" in

Table 5.3: EMSR Example — Day 35 Booking Limit Revision

	Y	M	B	Q
INPUTS:				
Expected Future Requests Day 35 to Departure	19.0	27.5	13.7	8.2
Standard Deviation	8.1	14.8	7.1	7.5
EMSR MODEL RESULTS:				
Nested Protection	13	27	22	—
Actual Bookings	1	9	5	18
Total Protection	14	36	27	—
Nested Booking Limits (Revised)	107	93	57	30
Seats Available (Nested)	74	61	34	12

Table 5.4: EMSR Example — Dynamic Revisions Day 28 to Day 7

	Y	M	B	Q
DAY 28:				
Expected Future Requests	16.2	23.8	12.6	4.1
Standard Deviation	7.9	13.3	5.5	6.6
Nested Protection	10	24	20	—
Actual Bookings	3	15	8	25
Total Protection	13	39	28	–
Revised Booking Limits	107	94	55	27
DAY 21:				
Expected Future Requests	12.9	22.0	11.0	3.3
Standard Deviation	6.9	11.9	6.2	7.1
Nested Protection	8	22	18	—
Actual Bookings	6	19	10	27
Total Protection	14	41	28	—
Revised Booking Limits	107	93	52	24
DAY 14:				
Expected Future Requests	9.7	19.4	7.9	0.8
Standard Deviation	5.8	11.6	5.9	5.4
Nested Protection	6	18	15	—
Actual Bookings	6	24	13	27
Total Protection	12	42	28	—
Revised Booking Limits	107	95	53	25
DAY 7:				
Expected Future Requests	6.0	17.3	5.8	0.3
Standard Deviation	3.7	10.9	4.8	4.3
Nested Protection	4	14	13	—
Actual Bookings	8	30	15	27
Total Protection	12	44	28	—
Revised Booking Limits	107	95	51	23

Q class, a condition which worsens closer to departure. In an actual booking process in which cancellations occur, bookings on hand in Q class might decrease to the authorized limit. No additional bookings could be accepted in Q class, unless Q cancellations reduce actual bookings to less than the relevant Q booking limits, at which point Q class would "open up" again.

The occurrence of "negative availability" in this example stems from the assumed interval between revisions of 7 days. The booking process is in essence left unattended between revisions. More frequent revisions would enable the EMSR model to be more responsive to changes in expected booking trends and actual bookings on hand. The problem with more frequent revisions is one of deriving more frequent accurate estimates of future requests from historical data. Frequent revisions would also not eliminate the possibility of too many low-fare seats being sold early in the process, with the result that unexpectedly high demand for higher fare classes might not be accommodated. Because the EMSR protection levels are based on maximizing *expected* revenues, it is inevitable that the booking limits for an individual flight might not result in optimal revenues.

Dynamic application of the EMSR framework allows the airline to manage the booking process, taking into account bookings on hand and expected requests still to come. This management process becomes more responsive and precise with more frequent revisions driven by demand estimates from more detailed historical data. Demand estimates based on forecasts rather than simple historical averages would provide further improvement, as will be discussed in Chapter Six.

To illustrate how overbooking factors may be incorporated into the booking limit calculations, we employ the same flight leg example as above. The application of different fare class overbooking factors is demonstrated for the calculation of initial limits (without actual bookings) in Table 5.5(a), and for the day 21 revision in Table 5.5(b), as one example of how actual bookings affect the overbooking limits. Table 5.5(a) shows the assumed overbooking factors by fare class and the calculation of optimal initial *overbooking* limits for the example flight leg. In the absence of actual bookings, the nested protection levels, NP_j, are added to the overbooking limit on the lowest fare class (Q), which is derived as explained in Section 5.4. Note that the nested protection levels are slightly lower than those in Table 5.2, because each of the upper fare class demand estimates have been subjected to adjustment by an overbooking factor that is higher than that

Table 5.5: EMSR Example — Overbooking Limits

(a) INITIAL OVERBOOKING LIMITS

	Y	M	B	Q
Overbooking Factors, OV_j	1.30	1.25	1.20	1.10
Nested Protection, NP_j	13	31	25	—
Overbooking Limits, BL_j^*	126	113	82	57

Weighted Cabin Overbooking Factor = 126/107 = 1.178

(b) OVERBOOKING LIMITS AT DAY 21

	Y	M	B	Q
Overbooking Factors, OV_j	1.30	1.25	1.20	1.10
Nested Protection, $NP_j(21)$	7	21	14	–
Actual Bookings, b_i^{21}	6	19	10	27
Overbooking Limits, $BL_j^*(21)$	127	114	74	50

Weighted Cabin Overbooking Factor = 127/107 = 1.187

assumed for Q-class requests. Fewer "reservations spaces" are thus protected for these upper classes. Furthermore, the transition from physical seats to reservations spaces, while keeping the same inputs of expected requests as used initially, frees up both seats and reservations spaces for Q class, relative to Table 5.2. The "weighted" overbooking factor on the entire shared coach cabin at this point is 1.178.

Table 5.5(b) shows the calculation of revised overbooking limits at day 21 prior to departure, taking into account the seat protection levels from the EMSR model and actual bookings. Actual bookings are added to the nested protection levels, and overbooking limits are calculated starting with BL_Q^*. The weighted overbooking factor on the coach cabin increases to 1.187, primarily because the M-class bookings are higher than expected.

The final step in this example of EMSR results involves the inclusion of upgrade probabilities to account for vertical choice shifts in the consumer decision process. The upgrade probabilities are included in the calculation of EMSR seat protection levels for the initial case in which no overbooking factors are assumed (Table 5.2). Table 5.6 shows how these initial limits would change if we could estimate the probability of vertical choice shift from each of the three lowest fare classes. The upgrade probability is assumed to be lowest for shifts from Q to B, higher to shifts from B to M, and highest for shifts from M to Y.

Relatively low upgrade probabilities are assumed for Case 1 in Table 5.6. Nonetheless, the nested booking limit on M class is reduced by 2 seats when compared with the initial limits from Table 5.2, and by a total of 4 seats for B class. The booking limit on Q drops by only 2 seats, because Q class requests are assumed to have a lower probability of upgrading. Case 2 shows the results when the upgrade probabilities are assumed to be higher. The nested booking limits decrease further for all fare classes lower than Y, with the protection for B class remaining constant due to the low probability of Q upgrades relative to B and M upgrades.

This chapter has presented the Expected Marginal Seat Revenue framework for determining optimal protection levels and booking limits in a nested multiple fare class reservations system for a single future flight leg. The basic model was extended to

Table 5.6: EMSR Example — Initial Limits With Upgrade Probabilities

	Y	M	B	Q
INITIAL LIMITS (from Table 5.2)				
Nested Protection Levels	14	33	27	—
Nested Booking Limits	107	93	60	33
CASE 1:				
Upgrade Probabilities, $P_j(v)$	—	0.30	0.20	0.10
EMSR Model Results:				
Nested Protection Levels	16	35	25	—
Nested Booking Limits	107	91	56	31
CASE 2:				
Upgrade Probabilities, $P_j(v)$	—	0.40	0.30	0.20
EMSR Model Results:				
Nested Protection Levels	17	37	25	—
Nested Booking Limits	107	90	53	28

dynamic applications, to include overbooking factors, and to take into account the probability of vertical choice shifts on the part of the consumer denied a request for a low fare class. The model and its extensions were applied to a hypothetical future flight leg and a set of historical demand and revenue data. The EMSR framework is based on certain assumptions with respect to the demand for different fare classes and over time. These assumptions and the sensitivity of the model to them are examined in the following chapter.

Chapter 6

EMSR Model Assumptions and Sensitivity

The preceding chapter presented the Expected Marginal Seat Revenue decision framework, its revenue-maximizing rationale, and the mathematical formulations that determine optimal seat protection levels and fare class booking limits. Its emphasis was on how the EMSR framework can be used to manage the booking process for a future flight leg dynamically, and examples of results based on hypothetical data demonstrated the model's application. The intent was to describe the EMSR model and its outputs in a relatively simple (i.e., single flight leg) context.

In this chapter, the EMSR model is examined in greater detail, specifically the effects of input data and assumptions on the validity and performance of the model. The demand assumptions required in practical applications of the EMSR model are considered first. The sensitivity of model outputs to variations in the input variables (both demand and revenues) is then examined. The chapter concludes with a discussion of the potential application of the EMSR approach to the "virtual nesting" seat inventory control systems being developed by several major airlines.

6.1 Demand Inputs and Assumptions

Throughout the presentation of the basic EMSR decision framework and its extensions to include overbooking and consumer choice shift probabilities, no explicit mention was made of the assumptions concerning the demand for different fare classes. The model was presented under the assumption that there exists some "correct" estimate of the demand density for each fare class for each point in time at which the EMSR model is to be run. This density represents the random variation around the expected number of requests which cannot be explained by systematic patterns in demand such as seasonal, day of week, or time of day.

In practice, it is the process of estimating the demand densities for total expected requests or requests to come for each fare class that triggers the need to make assumptions about the characteristics of these densities. Analysis of empirical data from historical information will inevitably comprise a major part of this estimation process, and will contribute to the airline's estimate of the density of expected requests by fare class for input into the EMSR framework. The estimation process, including the determination of the most relevant historical sample from which estimates should be derived, the separation of systematic variation from stochastic variation, and perhaps the adjustment of preliminary estimates through the use of forecasting methods, can in fact become a far more complicated task than actually deriving optimal limits from the EMSR model.

A fundamental assumption of the EMSR framework is that demand for each group of fare products (represented in the reservations system by a fare class) is distinct and separable, as explained by the market demand segmentation model presented in Chapter One. Given a variety of fare products with different restrictions, price levels and service amenities, consumers will choose one on the basis of their own price-sensitivity and the time-sensitivity of the trip being considered. Adequate fare product differentiation is critical to distiguishing market segments and to keeping the demands for each fare product type separate.

In practice, this ideal market scenario might not be replicable or might change with changing airline marketing and pricing policies. Making use of the EMSR framework thus requires the airline to make assumptions about the characteristics of the estimated demand densities to be used as inputs to the model, and to consider the possibility that the fundamental assumption of distinct and separable demands for different fare classes might not hold.

The specific estimation issues that must be addressed on the demand side of the EMSR equation include:

1. The shape of the probability density assumed for each estimate of expected requests;

2. The extent to which any correlation between demand levels in different fare classes exists and should be accounted for in the EMSR model;

3. The extent to which correlation between booking activity in the same class over time is important to the dynamic application of EMSR.

This section examines each of the above issues briefly, considers their impact on the EMSR model and presents some empirical evidence in each case.

6.1.1 Demand Density Shapes

No explicit mention of the assumed shape of the demand density was made in Chapter Five simply because virtually *any* probability density function will work in the EMSR formulations. As long as $EMSR_i(S)$ is decreasing in S, the comparison of the expected marginal revenue from protecting a seat for possible sale in a higher fare class with the certain revenue from the relevant lower fare class will result in an optimal protection level. Recall that the value of $EMSR_i(S)$ is determined by:

$$EMSR_i(S) = \overline{P}_i(S) \cdot f_i \tag{6.1}$$

where

$$\overline{P}_i(S) = P[r_i \geq S] \tag{6.2}$$

For any probability density function, the probability of receiving *exactly* S requests for fare class i is by definition greater than or equal to zero, meaning that the cumulative probability of receiving more than S requests will always be a decreasing function of S.

For the purposes of deriving optimal protection levels with the EMSR framework, an assumption concerning the shapes of the demand densities is nonetheless required. As mentioned previously, Boeing studies have in the past assumed Gaussian (normal) densities of total demand for a particular flight operated on a given day of the week over a homogeneous period of time [61]. The EMSR diagrams in the previous chapter were in fact drawn to approximate Gaussian demand densities.

For the EMSR model, however, we need estimates of the densities of expected requests by fare class, as opposed to the total requests for a flight. Previous empirical research by this author addressed the question of whether demand densities by fare class differ significantly from total flight demand densities, at least to the point of making the Gaussian assumption invalid [62]. Day of departure reservations totals by fare class were analyzed from a sample of Trans World Airlines (TWA) transcontinental non-stop flights in 1983. The objective of the study was to accept or reject the assumption of Gaussian demand densities for each fare class by using a sample of total reservations by fare class as a proxy for total requests. All flights for which fare class booking limits were reached, thereby constraining the estimate of total requests, were edited from the original sample.

In addition to editing out "constrained" flights, the aggregation of flights into homogeneous sample subsets proved to be critical. Analysis of the distribution of request totals by fare class aggregated over flights, markets, or seasons of the year brought the Gaussian assumption into question, as positive skewness was caused by extremely high demand flights (like those that operated during holiday periods) and simply by aggregation over dissimilar periods of time. As mentioned at the start of this discussion, the estimated density of demand for a particular future flight and fare class should represent the stochastic or unpredictable variation in expected requests, and should not reflect systematic variation over time or across markets.

Disaggregation of the dataset to a more homogeneous time period (e.g., low season) by flight number and day of week allowed a more detailed statistical evaluation of the distribution of data points within each subset. Extreme values were removed from each subset when they could reasonably be associated with travel peaks surrounding major holiday periods (e.g., Thanksgiving). Many of the smaller sample subsets conformed reasonably well to the Gaussian model, or at least did not deviate enough from the model to permit statistical rejection of the normality assumption. Combining similar days of the week to increase sample sizes allowed goodness of fit tests to be performed. In general, in the absence of capacity constraints (either fare class booking limits or physical capacity), the data fit the Gaussian model well.

An important finding of this analysis was that fare class demand density shapes as determined by the standard deviation relative to the mean (i.e., the coefficient of variation or "k-factor") differed noticeably among days of the week, even for the same fare class. Furthermore, removing outliers from day of week subsets resulted in a significant reduction in k-factors, in many cases to levels lower than the 0.33 approximation that has traditionally been assumed for total flight demand. Table 6.1 shows the differences in means and coefficients of variation by day of week for the "original" and "edited" subsets of full fare coach class (Y) reservations from three flights. The general pattern demonstrated is one of smaller k-factors on high demand days of the week and *vice versa* . For the purposes of seat inventory control and the EMSR framework, then, it is important that airlines recognize the systematic patterns in demand by fare class across days of the week for a particular flight leg, not only in terms of the level of demand (mean), but also in terms of its variation (standard deviation) relative to that mean.

The conclusion of this analysis of demand densities by fare class was that the empirical evidence did not allow statistical rejection of the assumption of Gaussian densities. The analysis highlighted the importance of deriving an estimate of the density that represents the random variation in demand for a particular fare class on a given future flight by disaggregating historical data to remove as much of the systematic variation as possible.

6.1.2 Correlation of Fare Class Demands

The fundamental assumption of distinct and separable demand for each fare class on a future flight leg is reflected in the EMSR framework by an assumed independence (i.e., no correlation) between the total expected demand levels in each of the fare classes. The possibility that individual denied requests for a low fare class will ultimately be "upgraded" to the next higher fare class can be incorporated into the EMSR model, as described in Section 5.5. Nonetheless, the EMSR approach is based on the notion that there exists an initial or "natural" demand for each fare class.

This subsection considers the impact of correlated fare class demands on the EMSR model outputs, presents some empirical evidence related to this issue, and discusses approaches to overcoming the problems presented by such correlation, should it exist. We begin this discussion by noting that under the assumption of no correlation between the demand densities for, say, classes 1 and 2, the optimal protection level for both classes from class 3 could be derived from the EMSR equivalently by assuming a *joint* demand density for the sum of class 1 and class 2 demands. As an example, assume that the class 1 and 2 individual demands are identically distributed, as described in Table

Table 6.1: Editing of Day of Week Data Subsets

Coach (Y) Class Passengers.

		MON	TUE	WED	THU	FRI	SAT	SUN
(a) Flight 37, PHL/LAX, Sept.-Dec. 1983								
Original	$\overline{X}$	28	28	35	30	45	8	36
	k	0.24	0.35	0.33	0.38	0.35	0.36	0.32
Edited	$\overline{X}$	28	30	39	31	47	8	37
	k	0.24	0.29	0.19	0.22	0.29	0.30	0.24
(b) Flight 38, LAX/PHL, Sept.-Dec. 1983								
Original	$\overline{X}$	43	28	34	37	48	23	39
	k	0.33	0.36	0.41	0.44	0.33	0.52	0.42
Edited	$\overline{X}$	45	29	33	37	53	23	37
	k	0.22	0.31	0.27	0.22	0.22	0.38	0.38
(c) Flight 61, BOS/SFO, Sept.-Dec. 1983								
Original	$\overline{X}$	50	50	52	53	72	36	63
	k	0.40	0.28	0.28	0.29	0.27	0.52	0.24
Edited	$\overline{X}$	49	51	56	56	76	34	65
	k	0.29	0.27	0.19	0.20	0.20	0.44	0.28

$\overline{X}$ = mean of data subset
k = coefficient of variation of data subset

6.2(a), and that the relative revenue levels of the three classes are \$100, \$80, and \$60, respectively.

If we run the EMSR model for the individual densities for classes 1 and 2, we find the total protection from class 3 to be:

$$S_3^1 + S_3^2 = 18 + 17 = 35 \tag{6.3}$$

To determine the effect of correlation between these two demand densities on the EMSR protection levels, we can input a joint density of demand to find the total protection from class 3. If no correlation is present, then the parameters of the joint density, assuming a normal distribution, are simply the sums of the individual density means and variances:

$$\bar{r}_{1+2} = \bar{r}_1 + \bar{r}_2 \tag{6.4}$$

$$\hat{\sigma}_{1+2}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 \tag{6.5}$$

where $\bar{r}_i$ is the estimated mean expected demand for class i and $\hat{\sigma}_i^2$ is the estimated variance of demand for class i.

The problem with using a joint demand density for two fare classes in the EMSR model involves determining the appropriate revenue level, f_{1+2}, to be used in calculating the expected marginal revenue of protecting an additional seat from class 3. In the example shown in Table 6.2, a simple average of class 1 and 2 revenues is used as an estimate of the revenue associated with the joint demand density. No weighting was performed because the individual densities are identical. The problem of joint revenue levels becomes more complicated when this is not the case. The EMSR total protection for the joint demand density in this case of no correlation turns is 35 seats, the same as the summed protection in the case of individual demand densities.

If the correlation between the demand levels for classes 1 and 2 is in fact significantly different from zero, the mean of the joint density will still be the sum of the individual means, but the joint variance will be given by:

$$\hat{\sigma}_{1+2}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + 2Cov(r_1, r_2) \tag{6.6}$$

where $Cov(r_1, r_2)$ is the covariance between r_1 and r_2, which determines the value of the correlation coefficient, $\hat{\rho}$:

$$\hat{\rho} = \frac{Cov(r_1, r_2)}{\hat{\sigma}_1 \cdot \hat{\sigma}_2} \tag{6.7}$$

Table 6.2: Impact of Correlation of Fare Class Demands

(a) Separate Demand Densities for Classes 1 and 2.

	CLASS 1	CLASS 2	CLASS 3
INPUT DATA:			
Fare ($)	100	80	60
Mean Demand	20	20	—
Std. Deviation	8	8	—

EMSR Total Protection from Class 3:

$$S_3^1 + S_3^2 = 18 + 17 = 35$$

(b) Joint Demand Density for Class $(1+2)$.

Correlation of Demand Class 1,2 ($\hat{\rho}$)	Std. Deviation of Joint Demand Density ($\sigma_{\hat{1+2}}$)	EMSR Protection from Class 3 (S_3^{1+2})
+ 1.0	16	32
+ 0.75	15	33
+ 0.5	13.9	33
+ 0.25	12.65	34
0	11.3	35
- 0.25	9.8	35
- 0.5	8	36
- 0.75	5.66	37
- 1.0	0	40

In Table 6.2, a range of values for the estimated correlation coefficient and the corresponding EMSR joint protection levels are presented. The lower and upper limits on the joint protection level are defined by perfect positive and negative correlation, respectively. The effect on optimal protection levels of significant correlation between the demand densities is determined by the standard deviation of the joint demand density for classes 1 and 2. With positive correlation, the standard deviation of the joint density is higher than it would be with no correlation, and the optimal protection levels are lower due to this increased uncertainty. With strong negative correlation, the standard deviation of the joint density drops, causing the optimal protection levels to increase. In the example shown, perfect negative correlation results in a zero standard deviation for the joint density, and the optimal EMSR joint protection level is simply the sum of the mean demands.

Given that significant correlation between fare class demands will affect the optimal seat protection levels, two questions should be addressed:

1. To what extent is such correlation actually found in airline demand data?
2. How can the EMSR demand inputs be adjusted to account for such correlation?

A partial answer to the first question is provided by the results of an empirical analysis of Western Airlines reservations data, performed as part of this research.

A sample of total reservations by fare class from flights in three O-D markets over a six-month period was used in the analysis. A low season period was selected to minimize the number of flights that would have to be edited from the sample because of fare classes having reached their booking limits. This editing was necessary to isolate the data points that were not constrained by the imposition of booking limits. The objective of the analysis was to determine whether any significant correlation between fare class demand levels could be found for the types of homogeneous sample subsets that would be used to estimate demand densities for the EMSR model in practical applications.

Correlation matrices were generated for each O-D market and a wide variety of subsets of the total data sample in each market. As was the case with the test of the normality assumption for estimated demand densities, this correlation analysis could only be applied validly to sample subsets in which little or no systematic variation remained. As an example of why systematic variation would bias the estimated correlation coefficients, consider the case of grouping two dissimilar days of week into one sample subset—a very high demand day and a very low demand day, in all fare classes. A correlation matrix for this sample subset would inevitably show very high positive correlation coefficients between fare classes because low class 2 demand would always be paired with low class 3 demand on one day of the week, while high class 2 demand would be paired with high class 3 demand on the other day of the week.

For each correlation coefficient estimated, a test was applied to determine whether ρ was significantly different from zero at the 90 percent level of confidence, given the sample size of each particular subset. The analysis was complicated by the presence of only one or two significant correlation coefficients in the same matrices, out of the six pairs of classes being considered. In this example, data from four fare classes were analyzed — Y, M, B, Q in descending revenue order. The focus of the analysis was on subsets consisting of reservations data from the same flight, same day of week, over a relatively homogeneous period of time.

Table 6.3 summarizes the results of this correlation analysis for several of the subsets tested. The general findings, as reflected by the examples shown, were a preponderance of zero or insignificant correlation coefficients between fare class demand levels. A few significant positive and negative results were obtained, but with no apparent pattern. Aggregation of data subsets to include different days of the week and/or different flights in the same market resulted in significantly positive correlation between fare class demands in virtually all cases tested. Disaggregation of data to eliminate systematic variation, on the other hand, led to almost no evidence of significant correlation, positive or negative. It should be pointed out that at least some of this absence of significant correlation is attributable to the small sample sizes of the disaggregated subsets.

The implication for the EMSR model is that it may be used as described in Chapter Five on individual fare class demand densities, unless empirical analysis reveals significant and consistent correlation patterns between fare class demand levels. On the other hand, the presence of some significant correlation levels even in the disaggregated subsets shown in Table 6.3 and the presence of others not statistically significant but substantially different from zero in the disaggregate matrices suggests that it might be necessary

Table 6.3: Emp[illegible]l Analysis of Fare Class Demand Correlation

(+ or − indicates significant correlation at 90 % confidence)

SAMPLE	n*	(Y,M)	(Y,B)	(Y,Q)	(M,B)	(M,Q)	(B,Q)
(a) Boise - Salt Lake City							
All flights, 6 months	368	+	+	+	+	0	+
Flight 048	85	0	+	0	−	−	0
Flight 354	112	+	0	0	0	−	0
October	74	+	+	+	+	+	+
February	55	+	0	+	+	+	+
October, Flt. 048	20	0	0	0	−	0	0
October, Flt. 354	18	0	0	0	+	0	0
Flt. 048, Mondays	8	0	0	0	0	−	0
Flt. 048, Wednesdays	12	0	0	0	−	0	0
Flt. 048, Fridays	8	0	0	−	−	0	0
(b) Los Angeles - Oklahoma City							
All flights, 6 months	292	+	+	−	0	−	+
Flight 593	149	0	0	0	0	−	+
Flight 595	143	+	+	0	0	−	+
October	62	0	0	0	+	+	0
February	49	0	+	0	0	0	0
October, Flt. 593	31	0	0	0	+	0	0
October, Flt. 595	31	0	0	0	0	0	0
October, Tuesdays	10	0	0	−	+	0	0
October, Thursdays	10	0	0	0	0	0	0
October, Sundays	8	0	0	0	+	+	+

* n = sample size, with "constrained" flights deleted

to take whatever small amount of correlation that exists into account. Forecasting methods that make use of dependent variables related both to time (systematic variation) and actual demand in other fare classes to predict demand for a given fare class might be used to incorporate the correlation of fare class demands into the EMSR inputs. We shall return to the issue of reservations forecasting at the end of this section.

In summary, significant correlation between demand levels for different fare classes can affect the optimal protection levels for the classes involved, and the EMSR model outputs based on distinct and uncorrelated densities might not be correct in such a case. The fundamental assumption of demand segmentation between fare product types implies that no correlation between fare classes should be evident, in theory. Empirical analysis performed as part of this research found that correlation of demand between fare classes was not significant for properly defined homogeneous data samples.

6.1.3 Booking Activity Correlation Over Time

The final demand assumption considered in this section is that of independence (no correlation) between the estimates of future requests and actual bookings already in hand in the dynamic application of the EMSR model. The dynamic application described and demonstrated in Chapter Five assumes no relationship between the expected number of future requests and the bookings already confirmed for the same future flight. Because the revised protection levels at day t before departure are derived from the estimated densities of future requests by fare class, the relationship (or lack thereof) between these densities and actual bookings on hand is important to proper application of the model.

If the estimated densities of future requests can in fact be assumed to be independent of actual bookings on hand in the same fare class, then simple historical means and standard deviations of booking activity from the same day t to departure for a sample of similar flights will suffice as inputs to the EMSR model. Otherwise, forecasting methods that take into account actual bookings (in the same or other fare classes) as well as other time-related variables may be needed to improve the estimates of future requests. As was the case with the two previous demand assumptions discussed, this subsection considers the possible effect of violating the independence assumption on the EMSR framework, presents some empirical analysis results, and discusses possible solutions to the problem.

In the initial (static) application of the EMSR framework, the density of total expected requests by fare class is estimated by the mean, $\bar{r}_i$, and standard deviation, $\hat{\sigma}_i$, derived from a historical sample. In the dynamic application of the model, these inputs are replaced by an estimate of future requests and actual bookings on hand. This revised estimate includes b_i^t, the number of bookings on hand in class i at day t, and $\bar{r}_i^t$, the mean number of requests still to come from day t to departure, with a standard deviation of $\hat{\sigma}_i^t$.

At any point t, the number of bookings on hand is a given quantity for the particular future flight under consideration. If there is significant correlation between b_i^t and r_i^t, we can use this information to improve our estimates of $\bar{r}_i^t$ and $\hat{\sigma}_i^t$ over those derived from historical information alone. For normally distributed bookings and future requests, the expected value and variance of the latter *given* a value of b_i^t is determined by:

$$E[r_i^t|b_i^t] = \bar{r}_i^t + \hat{\rho}\frac{\hat{\sigma}_r}{\hat{\sigma}_b}(b_i^t - \bar{b}_i^t) \tag{6.8}$$

$$Var[r_i^t|b_i^t] = (1 - \hat{\rho}^2)\hat{\sigma}_r^2 \tag{6.9}$$

Significant correlation between actual bookings and future requests can thus substantially affect the conditional estimates (mean and standard deviation) of requests still to come and, in turn, the revised EMSR protection levels. The impact on $\bar{r}_i^t$ of a given value of b_i^t is determined by the correlation coefficient $\hat{\rho}$, as well as by the value of b_i^t relative to its own mean, as shown in equation (6.8). Just as important, the information about correlation could be used to improve the estimate of requests to come by reducing the variance of the estimate. As shown in equation (6.9), any significant correlation will act to reduce the variance of r_i^t in a conditional probability density.

An empirical analysis of the correlation between bookings made up to day t before departure and bookings made after day t was performed for a historical sample of Western Airlines flights. Correlation matrices were generated for two reduced fare classes (M and B), for t equal to 35, 28, 21, 14, and 7 days before departure. Flights in two markets were sampled over the 3-month period 1/1/86 to 3/31/86. Bookings in M-class and B-class were analysed separately, and flights for which the class being analysed (or any higher class) had sold out (booked out) during the booking process were edited out of the sample to isolate "unconstrained" booking behavior. The results of this analysis could well have been different had these flights not been excluded. Any conclusions, however, would have been far more speculative, given the difficulty of estimating "true" demand from constrained flights.

Correlation coefficients were tested for significance at the 90 percent level of confidence. Correlation matrices were produced by fare class for the total sample in each market, and then for a wide range of sample subsets grouped by month, day of week and flight number. As was the case with the analysis of correlation of total demand between fare classes, most of the sample subsets aggregated over different flights and days of the week produced correlation coefficients greater than zero.

When the sample subsets were disaggregated by day of week, the overall finding was one of zero or positive correlation of booking levels between time periods prior to departure. Significant negative correlation was evident in some cases, particularly for the more restricted B fare class. Mean bookings in B class actually decreased in many instances from day t to 0, for $t \leq 14$, due to advance booking and ticket purchase limitations of 14 or more days on B class fare products during the sample period. Such negative correlation was evident consistently on days of the week that experienced relatively high demand for these reduced fares (e.g., Saturdays).

Although the occurrence of significant correlation between booking activity by fare class in different time periods before departure was by no means overwhelming or consistent, the empirical evidence suggests that information about such correlation could provide useful additional information to the estimates of future requests for particular fare classes or in particular markets. This information can be incorporated into forecasting models designed to predict future requests on the basis of the number of bookings already on hand, as well as other time-related independent variables. A Master's thesis completed by Joao Sa in February 1987 [63] deals extensively with the topic of forecasting future requests for the purposes of improving the demand inputs required by dynamic applications of seat inventory control models like EMSR.

Sa developed a generalized forecasting model structure that could be applied to different flights, markets, and seasons, at different times in the booking process. Based on multiple regression analysis, the generalized forecasting model predicts future requests in a particular fare class from day t to departure. The independent variables that showed the greatest explanatory power were time-related (i.e., day of week, week of year traffic index), indicating the importance of systematic variation in predicting total and partial demand for a flight. Independent variables related to bookings on hand in the same or different fare classes, on the other hand, were not as consistent in their explanatory power, although they proved to be very significant in selected markets. In all the markets that Sa examined, the generalized forecasting model provided less variable estimates of future requests than simple historical averages.

It is therefore possible to make use of information on the correlation between booking activity in different time periods prior to flight departure, as well as the historical correlation of demand levels in different fare classes, to develop a model to forecast expected future requests more accurately. The success of such a forecasting effort will depend on the strength of these relationships, which empirical analysis has shown can vary substantially from market to market and over time.

As noted in the introduction to this section, it is the application of the EMSR model to a particular data context that requires assumptions to be made with respect to the demand for different fare classes. Empirical analysis conducted as part of this research suggests that, in many contexts, the simplifying assumptions of separable, independent and normally distributed requests are in fact valid. Clearly, there are also contexts in which significant correlation, particularly between booking activity over time prior to departure, indicates that adjustments to the demand inputs could produce more accurate EMSR outputs. As is the case with all optimization models, the EMSR model is only as good as the inputs used.

6.2 Model Sensitivity to Input Variables

The assumptions made with respect to the demand for different fare classes can have an impact on the estimates of request densities used as input to the EMSR model, as discussed above. Different estimates of demand can in turn change the outputs of the EMSR model, namely, the optimal protection levels and booking limits for each fare class. This section continues the discussion of the performance of the EMSR framework by examining the sensitivity of the results to the input variables, those related to both demand and revenue.

This discussion centers on the sensitivity of the optimal seat protection level for class 1 from class 2, S_2^1. In the EMSR decision rule, the value of S_2^1 that maximizes expected revenues in class 1 will be determined by the ratio f_2/f_1 (where $f_2 < f_1$), and by the parameters of the demand (request) density for class 1, estimated by $\bar{r}_i$ and $\hat{\sigma}_i$. The optimal value of S_2^1 must satisfy:

$$\overline{P}_i(S_2^1) = f_2/f_1 \tag{6.10}$$

The relationship between $\overline{P}_i(S_2^1)$ and the estimated parameters of the demand density is defined by the cumulative normal probability distribution. For the purposes of this discussion, we will refer to the standardized normal distribution, represented by Z [64].

That is, if the fare ratio in equation (6.10) is 0.75, then we must find the value of Z which has a probability of 0.75 of being exceeded, based on the estimated parameters of the class 1 request density. From a table of standardized cumulative normal probabilities, we determine $Z = -0.675$, and we solve the following for S:

$$Z = \frac{S - \bar{r}}{\hat{\sigma}} \tag{6.11}$$

or:

$$S = \bar{r} + Z\hat{\sigma} \tag{6.12}$$

We can vary each component of the above equation separately to determine how the seat protection level S will be affected. The impact of variations in $\bar{r}$ are the most straightforward to interpret. For a fixed Z and $\hat{\sigma}$, each unit change in $\bar{r}$ will result in a one unit (i.e., one seat) change in the optimal number of seats protected. All else being equal, then, an increase in the mean demand for a fare class will result in an equivalent increase in EMSR protection level.

In most airline applications, however, the absolute value of $\hat{\sigma}$ will increase with $\bar{r}$, while the coefficient of variation remains more stable. If we assume a constant coefficient of variation, $k = \hat{\sigma}/\bar{r}$, then the expression for S becomes:

$$\begin{aligned} S &= \bar{r} + Zk\bar{r} \qquad (6.13) \\ S &= \bar{r}(1 + Zk) \end{aligned}$$

The optimal value of S thus remains a linear function of $\bar{r}$ for a constant Z and k.

A more interesting sensitivity analysis emerges when we consider variations in $\hat{\sigma}$, holding the other input variables constant. The optimal value of S is also a linear function of $\hat{\sigma}$, although the relationship between Z and $\hat{\sigma}$ becomes important in this case. As mentioned, the value of Z is determined by the ratio f_2/f_1 and its translation to a standardized normal cumulative probability distribution. The value of Z can be either positive or negative, depending on the fare ratio involved:

$$\begin{aligned} \text{if} \quad f_2/f_1 &> .50, \quad Z < 0 \\ \text{if} \quad f_2/f_1 &< .50, \quad Z > 0 \end{aligned} \tag{6.14}$$

For $f_2/f_1 = .50$, Z will equal zero and the optimal S will simply equal $\bar{r}$. Varying $\hat{\sigma}$ will have no effect.

In terms of the effect on S of varying $\hat{\sigma}$, we must distinguish between cases where Z is negative and positive. For negative values of Z (where the lower fare is greater than one-half the higher fare) and a fixed $\bar{r}$, a decrease in $\hat{\sigma}$ will result in a higher value of S while a higher value of $\hat{\sigma}$ will act to reduce the optimal protection level. For positive values of Z, which occur when the lower fare is less than one-half the higher fare, the converse is true. These relationships are illustrated in Figure 6.1.

The importance of reducing the estimated standard deviation of expected demand as much as possible by editing out systematic variation and/or making use of forecasting models is demonstrated in Figure 6.2. With $\bar{r}$ fixed at 20, optimal S values are plotted in Figure 6.2(a) against a range of $\hat{\sigma}$ estimates, for specific fare ratios. Clearly, the range of optimal S values decreases dramatically for coefficients of variation below 0.5.

The same example can be used to illustrate the impact of different fare ratios (and in turn Z values) on the optimal S. As introduced above, the value of Z decreases with f_2/f_1, meaning that the closer f_2 is to f_1 proportionately, the smaller the optimal protection level for class 1. Figure 6.2(b) shows how, for a given k-factor or value of $\hat{\sigma}$, S is a non-linear function of the fare ratio. It also confirms that for a range of fare ratios between 0.90 and 0.10, the range of optimal protection levels is much smaller for lower estimates of the standard deviation of demand.

To summarize, both the demand and revenue inputs to the EMSR framework can affect the calculated optimal seat protection levels significantly. The estimates of mean demand by fare class have the greatest direct impact on the model results. The estimated variance of expected fare class demand in turn determines the range of output results for

Figure 6.1: EMSR Protection Levels for Different Estimates of Demand Variation

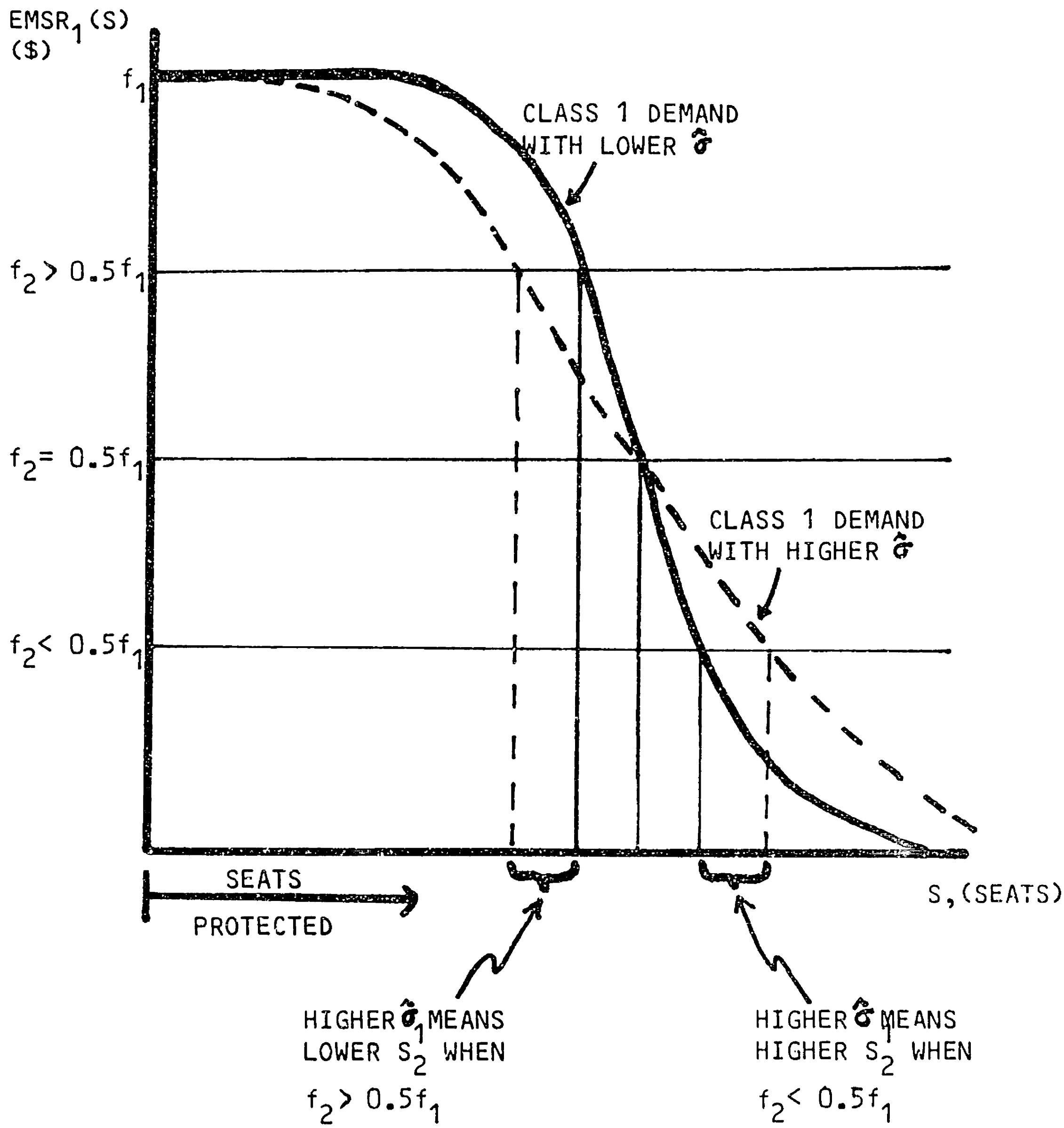
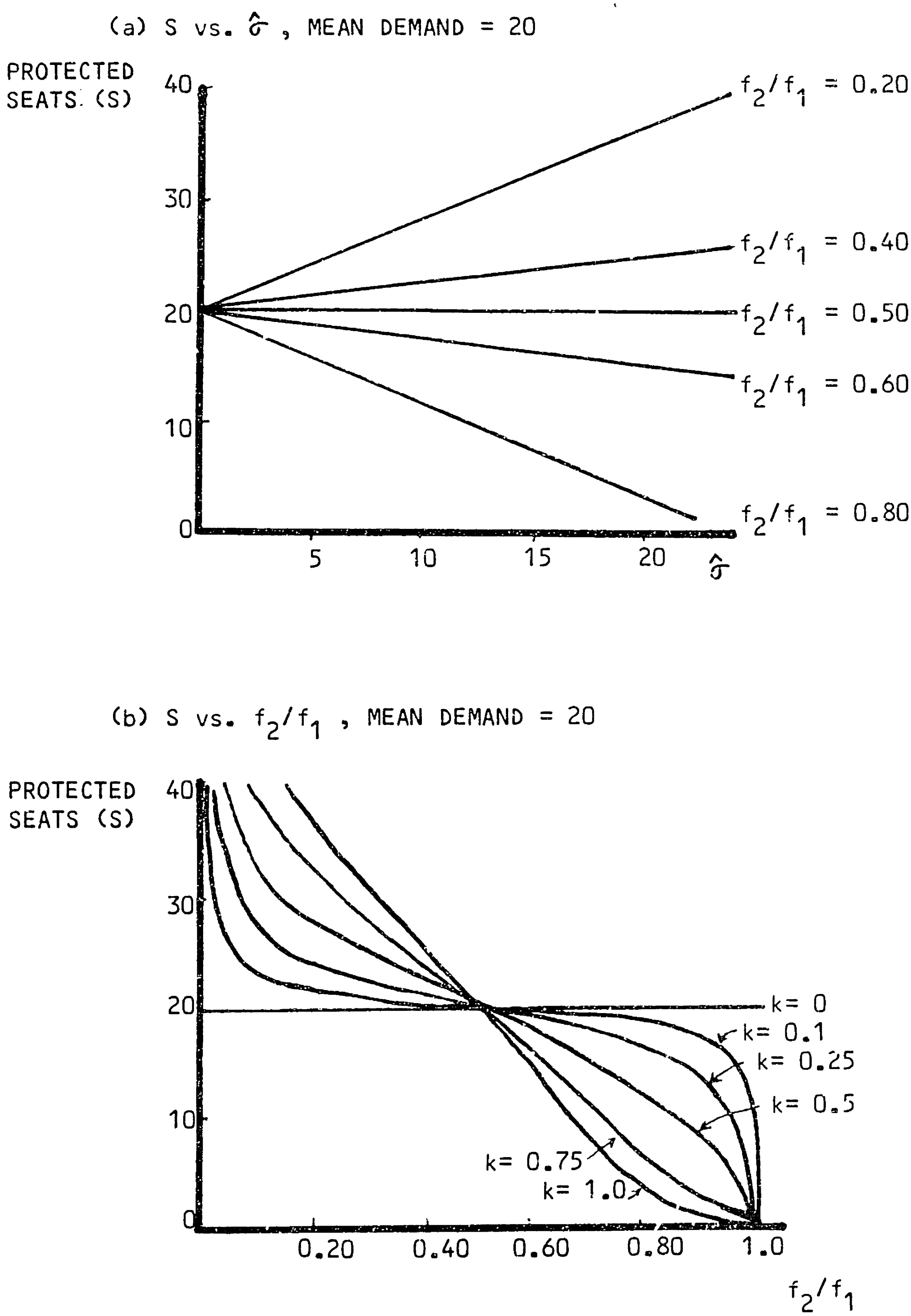


Figure 6.2: EMSR Protection Levels Plotted Against Input Variables



given fare ratios, reflecting the impact of increased uncertainty about future fare class demand. The sensitivity analysis of the EMSR model results presented in this section highlights the importance of accurate estimates of fare class demand levels and their variation.

6.3 Applications to O-D Seat Inventory Control

The flight leg-based approach to seat inventory control was selected for the development of the EMSR framework both because of its relative simplicity and because airline reservations systems maintain seat inventories and determine seat availability by fare class for each flight leg. As mentioned in Chapter Three, several major airlines are in the process of re-programming their reservations systems to manage seat inventories on the basis of the different passenger itineraries that might use a particular flight leg. The objective of such system development is to allow the airlines to take into account the impact of requests for different fare class/passenger itinerary combinations on total system-wide revenues rather than simply the prorated revenues associated with an individual flight leg.

This section considers the application of the EMSR approach to the more sophisticated inventory management structures currently being programmed into airline reservations systems. The systems being developed make use of a concept known as "virtual nesting" in determining the availability of seats in different fare classes for a specific origin-destination itinerary.

The fare class/flight leg approach to seat inventory control does not distinguish between passengers in the same fare class on different itineraries with different ticket values. The example of a $29 local Q-class fare product and a $99 Q-class fare product for a longer itinerary that must share the same initial flight leg was introduced earlier. In such situations, it is possible for all Q fare class seats to be booked by the local demand, denying space to the through demand that might have generated a much higher total revenue for the airline. This can occur under a leg-based inventory control approach when the demand for Q-class seats for local travel on the continuing leg of the flight is low. Seats on the second leg can remain unsold at the same time that all the Q-class seats on the first leg have been booked by local passengers. Total revenues for both legs could have been increased by protecting some Q-class seats on the first leg for through passengers.

Maximizing total expected revenues over an airline's complete system of routes and flights operated over a particular period of time requires a solution approach that takes into account the impact on the rest of the network of accepting a particular request for a fare class/passenger itinerary combination. Making an optimal decision at any point in time involves a comparison of the maximum expected revenues with and without the current request, as described in Section 4.2. Extremely large network formulations of the problem would be required, and the computing capabilities needed to perform repetitive optimization runs would far exceed those of most, if not all, airlines.

The alternative being pursued by several major carriers involves maintaining a primary inventory of five to eight fare classes, while establishing a relationship between each fare class/itinerary combination and one of a much larger number of "hidden" inventory classes internal to the reservations system. The approach has been called "virtual nesting" because a complete hierarchy of nested inventory classes exists solely within the computer system. Users such as reservations agents and travel agents will continue to see fare class availability by flight leg, as they do now.

The objective of reconfiguring the inventory classes in a reservations system is to give the airline greater control over the seats made available to passengers that generate different revenue levels for the airline. Virtual nesting by O-D ticket revenues enables the airline to prevent local passengers from taking all the seats in a fare class in favor of through passengers generating a higher total revenue. The potential benefits of taking through passengers over those with a higher local yield have been recognized by airlines for years, as several existing reservations systems have "local sales inhibitors" that allow limits to be placed on local passengers on a flight leg. This ability to limit local sales has become especially important with the development of large connecting hubs by the major airlines.

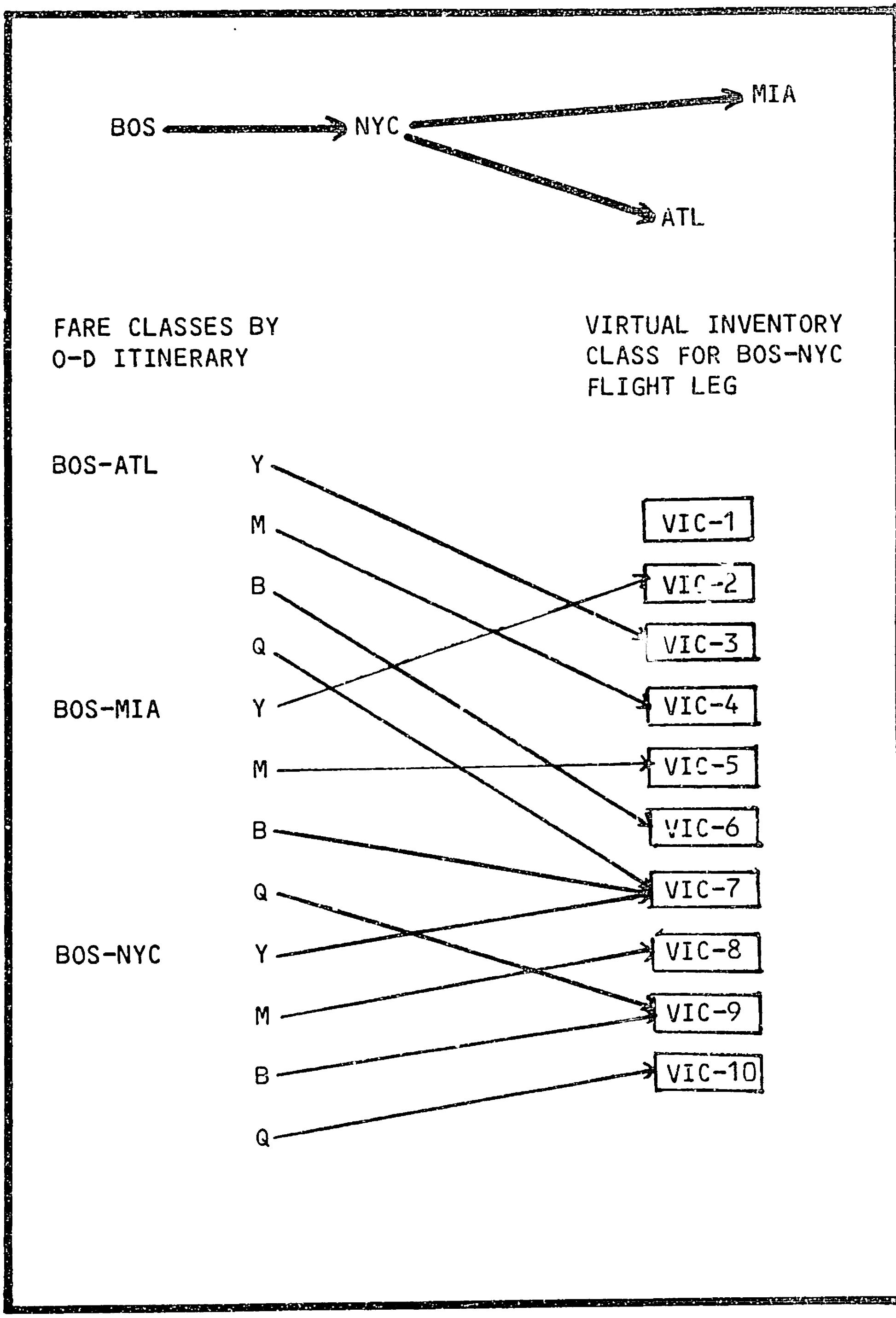
Under the virtual nesting concept, the availability of a particular fare class on a leg will be determined by the availability of the corresponding "virtual inventory class", or VIC, associated with a specific flight leg and O-D itinerary. Thus, Q-class seats might be available on a flight leg from Boston to New York when an ultimate destination of Atlanta is involved, whereas they might not be available in conjunction with connecting travel to, say, Miami. As shown in Figure 6.3, this determination of seat availability depends on the assignment of actual fare classes by O-D pair to VICs for each flight leg. In the example shown, the value of a Q-class booking in the BOS-ATL market has been judged by the airline to be higher than that of either a Q-class or B-class booking in the BOS-MIA market, and is assigned to VIC-6. Therefore, when the booking limit on VIC-8 for the leg BOS-NYC is reached, no more Q-class seats for travel from Boston to Miami would be available. Q-class seats for travel from Boston to Atlanta would remain available until the booking limit on VIC-6 is reached.

The basic EMSR framework can be applied with little modification to the management of flight leg seat inventories defined by a virtual nesting system. For each flight leg, the VICs must be nested hierarchically such that the top inventory class cannot be shut down to further bookings without closing down all lower-ranked classes, and the VICs must be ranked according to some measure of the value of each to the airline in terms of revenue. The EMSR approach of protecting seats for the top classes and establishing nested booking limits on the lower classes can then be applied directly to the virtual nesting system. A booking limit would be determined for, say, VIC-4 in Figure 6.3, and any request for a fare class/itinerary combination corresponding to VIC-4 will be accepted as long as the total bookings in VIC-4 and all lower VICs do not exceed BL_4 as derived from the EMSR model.

The EMSR approach may be applied to almost any aggregation of fare products into virtual inventory classes, as long as a density of expected demand can be estimated for each inventory class that must share a common inventory of seats and each category of demand can be associated with a particular revenue value to the airline. In virtual nesting based on O-D destination ticket revenues, for example, the VICs are defined for each flight leg on the basis of actual fare levels charged for each fare class/itinerary combination. Thus, a Boston to Miami M-class fare level of $250 is assigned a value of $250 for both the BOS-NYC and NYC-MIA flight legs to be traversed by the passenger, in the example introduced in Figure 6.3. No proration of revenues by flight leg is performed. Each possible O-D itinerary and fare class combination that uses the initial flight leg from Boston to New York City will be assigned to virtual inventory class.

The EMSR framework can then be applied directly to this hierarchy of virtual inventory classes, given estimates of expected demand for, and the average revenue level

Figure 6.3: The Virtual Nesting Concept



associated with, each VIC. Optimal protection levels and booking limits can be derived for each VIC, for each flight leg. Dynamic booking limit adjustment could also be performed by using the EMSR framework as before. Actual bookings in each VIC and incremental seat protection levels required for expected future requests in the same VIC are summed to determine the nested protection levels and booking limits. The only difference is that the demand and revenue inputs are estimates for each VIC offered on a flight leg, not simply each fare class.

The aggregation of different fare classes and therefore fare products into virtual inventory classes can complicate the inclusion of overbooking factors and upgrade probabilities into the EMSR framework. In the case of overbooking factors, aggregation of O-D itineraries into a single VIC can combine demand from different fare classes that might exhibit different no-show and cancellation behavior. The overbooking factors used as input to the EMSR framework will have to be adjusted to reflect the mix of bookings by fare class expected and accepted in each VIC. Further theoretical development of the model is required to address the question of optimal overbooking limits on virtual inventory classes.

Incorporating fare class upgrade probabilities into the EMSR framework as applied to a virtually nested inventory structure can be even more problematic. Fare class upgrades occur between adjacent fare classes on a given O-D itinerary request but adjacent fare classes for a specific O-D itinerary are not likely to be assigned to adjacent VICs. The approach for incorporating passenger choice shift probabilities into the EMSR framework described in Section 5.5 therefore cannot be applied directly to a virtually nested system. It might be possible, however, to adjust the availability levels determined from the EMSR formulations for a particular itinerary to account for upgrade potential.

Application of the EMSR framework to virtual inventory classes defined by O-D ticket values represents a heuristic approach to system revenue maximization. Its biggest shortcoming is that it can lead to obviously non-optimal solutions. The EMSR model will always favor longer-haul itineraries associated with higher total ticket values, even though the per-mile revenue may be much lower than that of two local itineraries that could have occupied the same seat. Implementation of such a virtually nested system, therefore, requires monitoring and human intervention to override the calculated booking limits in cases where local demand is high on both legs of a requested itinerary and the sum of the local fares is higher than the through fare being requested. Further theoretical development of the EMSR approach should be directed at making this intervention process more systematic.

In summary, the EMSR framework can be applied to the virtually nested inventory systems currently under development as a heuristic solution approach for system revenue maximization. Modifications and further development to account for overbooking factors, fare class upgrade probabilities and cases with high levels of local demand could improve this heuristic substantially. Even without major modifications, the EMSR approach can give an airline increased control of its seat inventories and the ability to approximate the impacts of different fare class/itinerary requests on system revenues. Until airlines are able to make practical use of the large network optimization routines described in Chapter Four, the use and further development of this heuristic approach offers the potential for increasing total revenues through seat inventory control, one which can be realized in a relatively short time.

The Expected Marginal Seat Revenue decision model developed and analysed in the preceding chapters offers airlines a systematic approach to determining fare class booking limits under existing reservations system capabilities. The basic principles of the model may be applied with varying levels of sophistication and automation to a comprehensive seat inventory control system. The way in which the EMSR framework is applied in practice will be determined to a large extent by the availability of data required as inputs, as will be discussion in the following chapter. The underlying decision approach of the EMSR framework can nonetheless be applied in conjunction with a range of estimation methods and demand assumptions. With the development of more sophisticated systems for managing seat inventories, the basic principles of the EMSR approach can be extended to take into account system revenue impacts.

PART THREE
System Development and EMSR Model Applications

The final part of this dissertation returns to the practical side of airline seat inventory management. Having developed a quantitative decision model for seat inventory control in Part Two, we now examine the application of the EMSR framework in an actual airline environment. Several of the practical considerations in seat inventory management were introduced in Chapter Three, namely, the need for human intervention in the process, the limitations imposed by existing reservations system capabilities, and the importance of data availability. It is the objective of Part Three to discuss the implementation of the EMSR model as part of an automated seat inventory management system, in light of these and other practical considerations.

Chapter Seven describes the implementation and testing of an automated booking limit system that incorporates the EMSR leg-based model, based on actual experiences with Western Airlines. The ways in which the development of the automated system was affected by company policies and operating procedures, as well as data availability, will be discussed. An evaluation of the system implemented at Western will then be described, and the revenue impact results presented.

Chapter Eight concludes this dissertation by presenting an overview of the findings and contributions of this work, namely the lessons learned from developing a quantitative decision model for seat inventory control and implementing it as part of an automated system. Finally, the seat inventory management issues that require further theoretical and empirical research are outlined.

Chapter 7

EMSR Model Implementation and Testing

The development of the EMSR decision framework, and its dynamic adjustment, overbooking, and choice shift components in particular, was guided to a large extent by the constraints and capabilities of the existing fare class/flight leg approach to seat inventory control used by virtually all airlines. Many of these constraints were discussed in the survey of current practice presented in Chapter Three. The impact of these constraints on the development of a more systematic method for setting fare class limits and its incorporation into an automated booking limit control system, however, did not become apparent until the EMSR model was actually applied to an airline operating environment.

Under a research agreement with Western Airlines, such an automated booking limit system (ABLS) was developed and implemented during 1986. The decision model programmed into the system to determine optimal fare class booking limits for future flight leg departures was based on the EMSR approach developed in Chapter Five. By the end of 1986, the basic EMSR revenue maximization model and a dynamic booking limit adjustment routine had been programmed into Western's ABLS. For a number of reasons, not the least of which was the fact that Western Airlines would no longer be operating on its own after March 31, 1987, further system development to include overbooking factors and upgrade probabilities was not possible. Nonetheless, the revenue impact of a completely automated approach to setting fare class booking limits relative to the manual and *ad-hoc* methods used previously was tested for a sample of actual flights during the first three months of 1987.

This chapter first examines seat inventory control system development and implementation issues, based on experiences at Western and discussions with representatives of several other major airlines. The impacts of airline marketing strategies and operating procedures on system development are outlined. A separate sub-section is then devoted to the reservations and revenue data requirements of the EMSR framework, with an emphasis on the ways in which the operating policies and procedures discussed previously can affect data availability. The second section of this chapter describes the Automated Booking Limit System developed at Western during 1986, which incorporates the EMSR approach to seat protection. Finally, a revenue impact test of the EMSR dynamic framework conducted on a sample of Western flight legs is described, including the test methodology and the results.

7.1 System Development and Implementation Issues

As discussed in Section 3.2, the seat inventory control process adopted by an airline is closely related to several other functions in its corporate structure. This fact alone suggests that the development and implementation of a system designed to make seat inventory control more systematic and/or more automated will be constrained to some degree by the airline's organizational structure and operating procedures. In addition, given the relationship of seat inventory control to pricing, together with the effect of both strategies on yield and, more importantly, total revenues, changes to seat inventory control practices can have implications for the airline's marketing policies and overall corporate philosophy.

Airline policies and operating procedures can in turn affect the quantity and the characteristics of the demand and revenue data available for input to decision models like the EMSR framework. We have seen that the validity and accuracy of the data used as input is extremely important to the EMSR model outputs. While it might be possible to improve data availability simply by improving computerized data collection and retrieval capabilities, changes to the policies and procedures that generate the input data might also be required.

7.1.1 Airline Policies and Procedures

The development and implementation of an improved method for controlling seat inventories is intended to contribute to one or more objectives being pursued by the airline's management. The broadest objective established for most companies is that of long-run profit maximization. In the context of airline revenue management, profit maximization requires the maximization of total revenues given what is essentially a fixed level of operating costs (in the short run, at least).

The improvement of seat inventory control methods is an important *strategy* in effective revenue management. Together with the effectiveness of the pricing strategy adopted by the airline, the effectiveness of its seat inventory control process determines the extent to which the objective of revenue maximization is attained. A well-defined objective for the airline's seat inventory control strategy is thus an essential basis from which improvements to the methods employed can be made. Yet, it is still the case within the managements of several major airlines that this type of single clear objective for seat inventory management does not exist.

The absence of a well-defined objective is attributable at least in part to the traditional airline industry emphasis on *yield* and, specifically, on *yield management*. As discussed in Chapter Three, however, high yields do not necessarily lead to high total revenues. Yield reflects the mix of passengers being carried by an airline, but does not reflect the proportion of empty seats for which no revenue is received.

In recent years, the industry and several carriers in particular have realized that revenue maximization, not yield maximization, is the most appropriate overall objective for the relatively short run strategic focus of seat inventory control. In fact, the director responsible for pricing and inventory management at one of these carriers has forbidden any references by his staff to "managing yields". The development by several airlines of inventory systems designed to manage seats by total O-D revenues provides further evidence of the acceptance of the revenue maximization objective.

Even with an overall objective of revenue maximization, improving an airline's seat inventory control process will be subject to the sometimes conflicting objectives of the critical functions closely related to seat inventory control. These critical functions can include fare product pricing and definition, marketing and promotional program planning, reservations procedures, travel agent sales, and airport check-in procedures. The potential impact of each of these functions on the revenue maximization objective and, in turn, system development, are discussed below in the context of incorporating the EMSR framework into a seat inventory control system.

The objective of applying the EMSR model to a leg-based fare class inventory system is to maximize total expected revenues for each future flight departure. The model calculates the number of seats that should be protected for each upper fare class and not made available to fare classes with lower expected per-passenger revenues. The concept is straightforward, yet each of the critical functions listed above can act to reduce the effectiveness of the EMSR model in maximizing expected flight revenues.

Pricing and fare product "development" (i.e., the definition of the restrictions and amenities associated with each fare product) is the function that is most closely related to seat inventory control, and the most likely to be found in the same organizational unit of the airline. Nonetheless, it is possible for the objectives of seat inventory control to be undermined by uncoordinated pricing actions and definitions of fare product attributes. For example, inconsistency in the categorization of fare products into the fare class inventories can reduce the effectiveness of seat inventory control.

For seat inventory control to have an impact on total revenues, the seats made available to a particular fare class inventory should be associated with a similar average revenue level, relative to the other fare classes. In the case of a system based on prorated leg revenues by fare class, categorizing different fare products into fare classes should be done so that all fare products with similar prorated average revenues are grouped into the same fare class. Grouping fare products on the basis of similar prorated revenue levels can result in similar fare products for different O-D markets being grouped into different fare classes, causing apparent inconsistencies in the classifications.

Using the criterion of relative revenue averages alone can also result in inconsistencies with respect to fare product restrictions such as advance purchase requirements. It might seem counter-intuitive for the pricing staff to categorize a fare product with a longer advance purchase requirement into a higher fare class than a lower-priced fare product with a shorter or no advance purchase requirement. Along the same lines, those responsible for pricing and fare product categorization might feel compelled to replicate the fare class categories of competitors to reduce confusion on the part of travel agents. For a seat inventory control system to have a positive impact on total revenues, however, the pricing function of the airline must adhere to the criterion of relative revenue levels in determining fare class categories.

Fare product categorization is particularly important at the top end of the fare class hierarchy, that is, the "Y" fare class intended to accommodate all "full coach fare" bookings. At most airlines, the Y fare class is also used to book various other fare products at revenue levels far below the full coach fare. The result can be extreme "pollution" of the Y fare class demand data gathered from historical flights and a prorated Y-class average revenue level that falls below one or more of the supposedly lower fare classes.

Currently, the Y fare class inventory is used by most airlines to book not only the full coach fare product, but some or all of the following reduced or zero-fare products as well:

- children's or senior citizen's fares;
- "Visit USA" fares for foreign travelers;
- airline and travel industry reduced fares (up to 90 percent off the full coach fare);
- frequent flyer award ticket travelers, paying reduced or zero fares;
- "positive space" airline employees paying zero fare.

The rationale behind booking these fare products into the Y fare class involves issues of seat availability. Being at the top of the nested inventory hierarchy, Y seats will be available as long as any seats remain unsold. Access to the Y class inventory for a fare product thus ensures maximum seat availability. Tradition, as well as competitive considerations, have acted to keep the fare products listed above in the Y fare class at most airlines. Children's and senior citizen's fares are unrestricted equivalents to the full coach fare, and pricing departments will argue that they should have access to any available seats as a result. "Visit USA" fares are booked into the Y fare class because many foreign carriers have only the capability to request coach cabin space for their clients.

The three remaining fare product groups listed have unrestricted access to the Y class inventory at many airlines because of corporate policies and competitive posturing. Airline and travel industry reduced fares "have always been" booked in Y class, and a departure from this practice by one airline could alienate, for example, travel agents. Similarly, discounted and free frequent flyer program ticket awards are booked in Y class because those responsible for these programs in the airline insist that:

1. If such awards were restricted to a less available fare class, it would reduce the value of these awards relative to those offered by competing carriers' frequent flyer programs.

2. Frequent flyer award tickets represent a substantial amount of revenue already collected by the airline from the individual involved, and should therefore not be regarded as contributing no revenue in the seat inventory control process.

Both arguments are valid, given the competitive nature of airline marketing, and they illustrate clearly how the objectives of seat inventory control can conflict with those of other airline functions. The objective of seat inventory control is to save seats for the highest revenue passengers, while the frequent flyer program awards these valuable seats for no tangible revenue.

The impacts on seat inventory control, and the EMSR framework in particular, are twofold. First, the demand data for the Y fare class can be inflated by reduced-fare or even zero-fare passengers, causing the model to protect more seats for Y class than would be optimal. If the dilution of the Y-class revenues is so severe that its average prorated revenue falls below that of other classes, on the other hand, the EMSR model will not protect any seats for the Y fare class, regardless of the demand levels. It can be difficult for those responsible for seat inventory control to explain to upper management why no seats are being protected for the "highest" fare class. Overall, the use of the Y fare class to book other reduced-fare products simply makes it impossible for the EMSR model to determine a revenue-maximizing protection level without making arbitrary assumptions about the fare products being booked in Y class and their revenues.

Seat inventory control system development must either be accompanied by significant changes to corporate marketing policies or take place around existing policies. Most realistic is a compromise between these two extremes. The technical components of the system can be designed to take the potential impacts of existing policies into account. In the above cases of data pollution, for example, changes to the reservations and data management systems to edit reduced-fare bookings and revenues from the Y-class data set might be possible. The seat inventory control system could then determine protection levels on the basis of "true" Y-class demand and revenue data. Unfortunately, unless the availability of the Y-class inventory is also limited to "true" Y-class demand, reduced fare bookings will still infiltrate the top fare class, taking away seat availability from the full Y-fare passengers.

Reservations and overbooking procedures represent two additional airline functions that can adversely affect the development of a revenue-maximizing seat inventory control system. As described earlier in this thesis, the overbooking levels applied to a particular flight departure are closely related to the booking limits placed on each of the fare class inventories and, ultimately, to the actual fare class mix of passengers carried on that flight. Chapter Five described how overbooking factors can be incorporated directly into the EMSR model to derive optimal *overbooking* limits by fare class. At most airlines, overbooking limits are currently set by a separate group responsible solely for overbooking. The development of a seat inventory control system that makes use of overbooking factors by fare class will therefore affect the function of this group in the airline.

Implementation of an EMSR-based system will actually increase the need for analysis of no-show and cancellation rates, particularly by fare class. Development of such a system should thus result in an expansion of the overbooking function, but it might also require a substantial change in the responsibilities of long-time employees that have performed this function primarily on the basis of their own judgement. The overbooking function must be coordinated and even united with the seat inventory control function. Given the close relationship between the two functions and their potential revenue impacts, they simply cannot be left separate if an effective system is to result.

Reservations procedures, while not nearly as important to seat inventory control as the overbooking function, can still undermine the effectiveness of seat inventory control. For example, it is common for airline reservations "supervisors" to have the discretionary power to accept bookings above the limit specified in the reservations system for a fare class. While this type of discretionary authority is necessary to deal with exceptional circumstances, it is possible that this authority is being used arbitrarily and too often, undermining the objectives of seat inventory control. As part of the system development process, airline procedures in the reservations area might have to be re-assessed.

Such a re-assessment should be accompanied by a review of travel agency reservations and ticket sales procedures. Under existing procedures, it is possible for travel agents to book passengers in a higher available fare class when a low fare class is closed, and then to issue tickets to the passengers for a low-priced fare product. Unless the airline monitors travel agency sales closely and then follows through by warning or even billing offending travel agencies for the difference in ticket values, such practices also undermine the effectiveness of a seat inventory control system. Even though only a small minority of travel agencies are likely to be involved in intentional misticketing on a regular basis, major airlines are careful not to offend these agencies and risk losing their business by chastising or penalizing them.

Part of the monitoring required to catch misticketing, as well as most of the enforcement required, must take place at the airport during the passenger check-in process. The airline's check-in procedures not only reflect its policy with respect to fare product rules and travel agent misticketing, they can be an important determinant of the actual boarding and no-show data to be made available to the seat inventory control system. First, passenger check-in at the airport represents what might be the only opportunity for the airline to reconcile the booked fare class, the fare product's restrictions, and the price paid. If a passenger is booked in Y class but presents a Q-class ticket, the details should at a minimum be noted and reported by the check-in agent. At a maximum, the passenger should be denied boarding and/or asked to pay the difference between the fare class booked and the ticketed price.

Most airlines' airport check-in agents find it difficult to deny boarding to a "confirmed" passenger, regardless of the ticket he/she is holding, as long as it is valid and empty seats are available on the departing flight. Most airlines are reluctant to impose policies requiring such passengers to be turned away, again because of a fear of losing business and goodwill. The same arguments apply to cases in which check-in agents are presented either with tickets issued for travel on another carrier, or with tickets issued for travel on a restricted fare product. Short term gains may be realized when check-in agents allow virtually any ticketed passenger to "stand-by" if empty seats are available, particularly if the passenger holds another carrier's ticket. Each stand-by passenger accepted in contravention of the restrictions of the fare product purchased (e.g., Saturday night stayover required), however, undermines the longer term revenue-maximizing objectives of differential pricing and seat inventory control. Several carriers have in the recent past strengthened their policies with respect to this issue, but experiences to the contrary suggest that such policies must be understood by check-in agents and enforced more consistently.

Check-in procedures, together with the capabilities of the airline's reservations system to perform the check-in process, also determine the characteristics of the post-departure data that will be available for input to the seat inventory control system. The data requirements of such a system will be discussed in greater detail in the following subsection. At this point, we should note the impacts on data resulting from passenger check-in under the policies and procedures described above. For example, a case of misticketing can cause the post-departure data to show one fewer passenger boarded in the booked class and one additional passenger in the ticketed class, depending on the computer check-in procedures used. The use of first class upgrade coupons by passengers booked in lower fare classes can cause further discrepancies. Without reservations system capabilities to note such discrepancies separately, the post-departure data will generate biased estimates of no-shows, for example. For some airlines, the capability to distinguish between passengers boarded in different fare classes (as opposed to by physical compartment on the aircraft) is itself relatively recent, and one which is critical to seat inventory control system development.

The preceding discussion provides an overview of the critical functions with the greatest direct impact on seat inventory control system development. There could well be numerous others, depending on the policies and procedures of the specific airline involved. The intent of this discussion has been to illustrate the extent to which the objectives of functions throughout the airline can affect those of seat inventory control. Under ideal conditions, the policies and procedures of the airline should all correspond exactly to the revenue maximization objective of seat inventory control. In practical terms, however, those responsible for system development and implementation must at least be aware of the functions that might undermine the objectives of this development. While efforts to coordinate policies should be undertaken, seat inventory control system development can also require a review of these functions, given their impacts on system performance and data availability. As outlined in this section, seat inventory control system development must be coordinated with the airline's pricing, fare product development, marketing, promotional programs, overbooking, reservations, sales and passenger check-in functions.

7.1.2 Data Availability and Estimation Methods

The EMSR decision framework developed in Chapter Five requires two types of input data for a particular future flight leg departure — estimates of expected demand and its variation by fare class, and estimates of the average prorated leg revenue associated with a passenger booking, again by fare class. The sensitivity analysis presented in Section 6.2 demonstrated the importance of accurate estimates of both demand and revenue for input to the EMSR decision model. This sub-section considers the data requirements of the EMSR model and the extent to which computer capabilities and operating procedures can affect the data that is actually available. Data availability will in turn determine the methods that may be used to transform the data into the estimates required as inputs.

The demand inputs required by the basic EMSR model are estimates of the total number of requests expected for a future flight leg departure, by fare class. Because the model takes into account stochastic variation in demand, an estimate of the variance in total requests around the expected value is also required. For dynamic applications of the EMSR framework, estimates of partial demand by fare class in the form of requests still to come from day t to departure are required, as well as estimates of the variation of this partial demand. Furthermore, the number of actual bookings already accepted for the particular future flight leg being considered are also necessary demand inputs.

Implementation of the EMSR model as part of a seat inventory control system necessitates an evaluation of what data are or can be made available, and the estimation process to be used. Under ideal circumstances, airline reservations systems would be able to provide the demand data required by the EMSR model in terms of the number of requests actually received for a particular flight leg during the reservations process, the number accepted as bookings, and the number ultimately ticketed and carried as revenue passengers, including any upgrades and discrepancies that arise during check-in. The ideal reservations system might even provide information about denied requests and whether a vertical or horizontal choice shift on the part of the denied consumer led to his/her recapture by the airline.

Most airline reservations systems do not approach this level of data sophistication, and few are likely to in the near future. Currently, most reservations systems log total bookings by fare class for a future flight leg. Database management systems have been developed by many airlines to generate extracts of the current booking levels and booking limits by fare class on a daily basis. These extracts become part of an evolving historical database of total bookings by day before departure for all flight legs for which reservations are being accepted. Once the flight has departed, a complete booking history for that flight leg can be retrieved from the database. Figure 7.1 shows an example of such a historical build-up of reservations from one airline's database.

From this type of database, total bookings by fare class for a departed flight can be extracted, as can partial bookings for any selected range of days prior to departure (e.g., day 21 to departure). The difference in total bookings between any two observations in the build-up history represents the *net change* in booking levels in that fare class. Thus, in Figure 7.1, we can determine that a net gain of 10 bookings in the Q class was realized between days 35 and 14 prior to the departure of the flight. We know nothing, however, about the actual number of bookings accepted and the number cancelled over the same period. No existing airline reservations system currently has the capacity to log both bookings and cancellations as separate data items.

For the purposes of the EMSR model, then, existing reservations systems, even with the "sophisticated" database management software developed or purchased by many airlines, are able only to provide a historical record of booking levels by fare class by day prior to departure, as well as total bookings by fare class, net of intervening cancellations for any period during the reservations process. These reservations systems are also able to provide counts of the actual bookings on hand for future flight departures.

The EMSR model, for initial and dynamic applications, requires estimates of the mean and standard deviation of *requests* by fare class. These estimates may be derived from a sample of past operations of the same flight, or similar flights on the same leg. If this historical sample includes only departed flights for which no fare class booking limits were reached during the reservations process (implying that no requests were refused), net booking levels may be used directly in the estimation of requests by fare class. A

Figure 7.1: Historical Booking Build-Up for a Departed Flight

HISTORICAL OR FUTURE LEG BOOKING
CITY PAIR: SEA / SLC

DAYS PRIOR	DAY	SEATS SOLD VSD	MSD	BSD	QSD	ZSD	TOT
	**ACTUAL	6	14	17	56	0	93
00	SU	4	19	29	60	0	112
01	SA	4	18	29	60	0	111
02	FR	4	20	33	61	0	118
03	TH	2	21	30	61	0	114
04	WE	2	22	28	58	0	110
05	TU	1	21	27	50	0	99
06	MO	1	20	27	49	0	97
07	SU	1	20	26	47	0	94
08	SA	1	20	26	49	0	96
09	FR	1	17	30	44	0	92
10	TH	1	16	30	43	0	90
11	WE	1	12	30	44	0	87
12	TU	1	12	46	42	0	101
13	MO	1	11	46	40	0	98
14	SU	1	11	46	40	0	98
15	SA	1	11	46	39	0	97
16	FR	1	11	46	39	0	97
17	TH	1	7	43	38	0	89
18	WE	1	7	41	36	0	85
19	TU	0	7	25	36	0	68
20	MO	0	7	23	36	0	66
21	SU	0	7	23	36	0	66
22	SA	0	7	23	36	0	66
23	FR	0	7	23	36	0	66
24	TH	0	6	23	34	0	63
25	WE	0	6	23	34	0	63
26	TU	0	6	23	32	0	61
27	MO	0	5	20	32	0	57
28	SU	0	5	20	32	0	57
29	SA	0	5	20	32	0	57
30	FR	0	5	20	32	0	57
31	TH	0	4	20	37	0	61
32	WE	0	2	18	37	0	57
33	TU	0	2	19	35	0	56
34	MO	0	2	15	31	0	48
35	SU	0	2	14	30	0	46
36	SA	0	2	14	30	0	46
37	FR	0	2	14	29	0	45
38	TH	0	2	14	27	0	43
39	WE	0	2	13	27	0	42
40	TU	0	2	15	27	0	44
41	MO	0	2	13	24	0	39
42	SU	0	2	13	24	0	39
43	SA	0	2	12	24	0	38
44	FR	0	2	13	27	0	42
45	TH	0	1	4	16	0	21

sample of observations from the recent past might be selected, with simple means and standard deviations of the fare class booking totals in the sample used as reasonable proxies for total requests for each fare class. The same estimation procedure can be used for partial booking levels, as required for input to the EMSR dynamic adjustment routine.

It is not likely that all the flights in the historical sample will have booked up without reaching one or more of the fare class booking limits. In such cases, one or more requests for a particular fare class and flight are likely to have been refused by the airline, and the disposition of these refused requests cannot be determined from the available data. The net booking levels in the reservations system database thus represent a *constrained* estimate of total requests for a particular flight leg and fare class. It is possible to use statistical methods to derive unconstrained estimates of requests by fare class, given the booking levels for each observation in the sample and knowledge of whether these booking levels were in fact constrained by a fare class booking limit. The statistical methods used can be as simple as graphical extrapolations on "normal probability paper", or can involve more complicated maximum likelihood estimation techniques [65].

The premise of either method is to estimate what the mean and standard deviation of requests *would have been* in the absence of booking limit constraints, based on the assumption of normally distributed request totals. As a general rule, these procedures can only be applied validly when fewer than one half of the observations in the sample were subject to a constraint. For a flight leg and fare class that reached its booking limit consistently during the period sampled, arbitrary approximations of unconstrained request totals by fare class may be required. Such a situation may signal a need to increase seat availability, which in turn would allow better data to be collected.

For both the initial and dynamic applications of the basic EMSR framework, then, the necessary demand estimates can be derived from the existing data available to most airlines from their reservations systems. The EMSR model extensions to incorporate overbooking factors and fare class upgrade probabilities, however, require demand data inputs that are not readily available to most airlines. In the development of a seat inventory management system that includes these extensions, the potential to improve data availability should be reviewed, as should the procedures that generate the data.

With respect to the overbooking factors required by one extension to the basic EMSR decision model, a simple approach of comparing actual boardings by fare class with bookings just prior to departure to estimate no-show rates might be considered. Issues that can arise in making data available for even this simple approach to no-show estimation, however, include:

1. How close to departure is the current "snapshot" of final bookings by fare class being taken, and when should it be taken?
2. To what extent does the existing check-in process affect the relationship between "final bookings" and actual boardings by fare class?

More sophisticated approaches to estimating cancellation and no-show rates, and then modeling the optimal overbooking factors on each class throughout the reservations process require far more detailed data than that provided currently by the majority of airline reservations systems. As mentioned earlier, total bookings and total cancellations on each day before departure rather than net bookings by fare class would be required by models that determine optimal overbooking limits by fare class as a function of time left to departure. Efforts to improve data availability for such models should not only focus on the development of computer capabilities to log, manipulate and report greater quantities of data, they should also examine the meaning and validity of these data, and the impacts of the procedures that generate them.

The extension to the EMSR model that incorporates passenger choice shift probabilities poses even greater data requirements than that of overbooking factors. Major re-design of existing reservations systems would be required to provide accurate flight-by-flight and class-by-class estimates of upgrade behavior and total passenger recapture. Without the ability to collect this type of data automatically through its reservations system, the airline must derive estimates based on surveys of passenger behavior or comparisons of historical flight build-ups for which booking limits were reached.

The demand data available from most existing airline reservations can be extensive when extracted to a database on a daily basis. For the purposes of seat inventory control and the EMSR framework in particular, the type of booking data available currently can be adequate if valid estimation methods are used. Improvements are possible, of course, to the amount and detail of data collected by the reservations system and then extracted to a historical database. In all cases, the airline procedures that generate the data should be a primary consideration in making such improvements.

In contrast to the volume of historical booking data that airlines can extract from their reservations systems, the availability of detailed revenue data is extremely limited at many airlines. Few carriers maintain revenue databases that can provide revenue information by flight leg and fare class. In fact, several major airlines have no idea of the revenue they collect for specific flights, and must resort to periodic samples of collected ticket coupons to obtain detailed revenue information. Developed before deregulation, the revenue accounting systems at most airlines were designed to report total system revenues, perhaps broken down by source of sale (e.g., travel agencies, airline ticket offices). Little effort was made to gather and store revenue data by flight and fare class, let alone by passenger itinerary.

With the evolution of seat inventory control practices, the need for better revenue data has become critical. In the EMSR framework, the revenue associated with bookings in different fare classes is just as important an input to the decision model as the estimates of fare class demand. For the model to be used effectively, average prorated leg revenues by fare class and city-pair are required at a minimum. Historic revenue data by day of week and flight number might provide more accurate estimates of the average revenues that can be expected for each booking accepted for a future flight departure. As the development of inventory systems based on O-D ticket values progresses, much more detailed revenue data by passenger itinerary will be required.

When the basic EMSR framework is applied to the existing leg-based management of seat inventories by fare class, some estimate of the relative revenue value of passengers booked in each fare class for the future flight under consideration is required. A simple estimate may be provided by an assessment of the relative price levels being charged for fare products in each of the fare classes that share the coach cabin inventory. The shortcoming of this approach is that is does not reflect the actual mix of fare products and O-D itineraries that are booked into each fare class, because relative price levels among fare products can differ substantially from one O-D itinerary to another on the same flight leg. Furthermore, the passengers actually being booked into each fare class might be using fare products that are far different from the non-discounted fares published for each fare class. The dilution of Y-class revenue (and demand) data by numerous reduced-fare products provides an example of this problem. Fare product *usage* in each fare class is an important consideration.

Another approach to estimating revenues by flight leg and fare class involves the airline actually sampling ticket coupons collected at departure time and/or tickets reported sold to build a revenue database. Total O-D ticket values can then be prorated by flight leg and categorized by fare class, according to the fare basis code on each ticket coupon. At least one major airline has developed this type of revenue database, which can provide the EMSR framework with direct estimates of average fare class revenues prorated by flight leg. Adjustments to these estimates might be required to account for expected or observed changes to relative revenue levels of different fare classes that are not reflected in the most recent historical information in the database. Similarly, substantial changes in the O-D itinerary mix of passengers traveling on a flight leg might require adjustment of the historical revenue averages. In either case, however, the aggregation of many itineraries and fare products into a single flight leg and fare class revenue average means that minor changes to fare product prices and their usage in a small number of markets will not have a significant impact on the aggregate estimates.

This aggregation, which is characteristic of the flight leg/fare class approach to both revenue estimation and seat inventory management, can result in reduced control over its seat inventories for the airline. Aggregation of revenue data by flight leg and fare class will result in revenue averages that are not as different in relative terms as the prices being charged for different fare products in individual O-D markets. The O-D nesting of virtual inventory classes, described in Section 6.3, is an attempt to increase the differentiation in revenue estimates and, in turn, to give the airline greater ability to discriminate between requests of different revenue value to the airline.

The amount of revenue data required for applications of the EMSR approach to O-D virtual nesting far exceeds that being collected by most airlines. American Airlines, a leader in the development of virtual seat inventory control, has also developed a revenue database that is updated to include tickets collected from departed flights within days of departure. This revenue database can be used to generate passenger counts and average ticket revenues by passenger itinerary, fare class, and fare product. Control of virtual inventories nested by O-D ticket revenues does not require proration of on-line revenues by flight leg.

The characteristics of the demand and revenue estimates required as input by the EMSR or other decision models for seat inventory control are thus determined by the data being collected and stored by the airline as well as its format, and by the procedures that created the data originally. The development of a seat inventory control system can in the short run be constrained by the availability of data. The greater the extent to which the available data do not correspond to the input needs of the EMSR model, the greater the need for estimation procedures and assumptions to generate the required inputs. For the airline wishing to realize rapid improvements to its seat inventory control process by implementing a decision approach like that of the EMSR model, however, working within the constraints of the available data and using estimation methods might be the only alternative.

In the longer run, seat inventory control system development should be related to, and even direct, the development of data collection and retrieval systems. Significant changes might be pursued in terms of the type of demand and revenue data being collected, as well as policies and procedures that can affect the purity of the data. At the same time, the capabilities of the reservations and database management systems should also be reviewed.

The discussion in this section has focused on the system development and implementation issues which are likely to arise when an airline decides to apply what appears to be a relatively straightforward mathematical approach to its seat inventory control process. While the EMSR model is not overly complex in theoretical terms, incorporating it into a seat inventory control system within an existing airline operating environment can be more complicated. The system development process provides an opportunity to evaluate the impacts of existing policies and procedures not only on the planned system, but on the overall objectives of seat inventory control. Consideration of the input data required by the system presents an opportunity to assess the usefulness of the data being collected by the airline and to develop a more comprehensive database management system. While it might be necessary to implement a seat inventory control system around the constraints posed by existing conditions, the potential for reducing or eliminating these constraints should not be overlooked.

7.2 Automated Booking Limit System (ABLS)

Many of the issues discussed in the previous section had to be addressed during 1986 by Western Airlines, as it prepared to implement an Automated Booking Limit System (ABLS) for seat inventory management. These procedural and data availability constraints were in fact reviewed by Western management, and action was initiated with respect to several of the problems identified over the course of the year-long development effort. Still, there remained numerous policy issues and data availability limitations that could not be resolved, meaning that the system developed at Western was, to a certain extent, built around existing procedures and capabilities.

The ABLS developed at Western incorporated the EMSR model in its static and dynamic forms, as described in Sections 5.2 and 5.3 of this dissertation. The objective of developing ABLS was to make the process of setting and adjusting fare class booking limits for future flight departures more systematic and to automate it as much as possible. It was hoped that the implementation of an automated system would reduce the manual effort required on the part of a relatively small staff of seat control analysts, allowing them to increase their analysis efforts for the small proportion of flights requiring closer attention. Inclusion of the EMSR decision model in the system was intended to provide the analysts with specific recommendations of what the fare class booking limits *should be,* based on a systematic evaluation of the input data.

Given a leg-based inventory class structure in the Western reservations system and availability of both demand and revenue data by fare class and flight leg for historical flights, the initial focus of ABLS development was on the flight leg/fare class problem. That is, the EMSR model used estimates of the density of expected requests by fare class for the flight leg under consideration, and average fare class revenues prorated for all flight legs serving the same city-pair. Future development plans included a shift to an origin-destination seat inventory system.

ABLS was developed to consist of two parallel components or routines: "batch" and "on-line". The batch routine was designed to set and periodically revise fare class booking limits for all future flight leg departures, based on booking data from a sample of recent departures of the same flight leg on the same day of week. Estimates of the mean and variance of the total and partial requests by fare class were derived from the booking data, and adjusted when necessary to account for "constrained" demand due to booking limits having been reached on the sample flights. Revenue averages extracted from Western's prorated revenue database for the most recent available sample period were input directly into the EMSR formulations.

Fare class booking limits were calculated for a future flight leg 90 days before departure, using the initial EMSR model. These fare class limits were revised weekly thereafter, up to and including 6 days before departure. The batch routine operated on a day-of-week rotation, handling all flight legs scheduled to depart on all future Tuesdays up to 90 days out, for example, on Wednesday mornings. Each weekly revision run between day 90 and day 41 made use of the initial (static) EMSR model, incorporating the additional data provided by the most recent (i.e., the previous day's) departure of the same flight leg into the demand density estimates, as well as any changes to the relative revenue averages.

Starting on day 34 prior to departure and weekly thereafter, a dynamic adjustment routine based on the dynamic application of the EMSR model took over the revision process. These batch revision runs not only incorporated the most recent input data available, they re-calculated the optimal seat protection levels required for expected requests still to come by fare class, as estimated from historical build-up patterns for the same flight leg. Actual bookings were added to these protection levels, from which revised booking limits on each fare class were derived.

The batch routine was intended to be self-monitoring, in that reports would automatically be issued for inspection by the seat control analysts, detailing the future flight legs for which booking limits had been or were about to be reached, as well as flight legs for which the batch routine had revised the previous fare class limits by a substantial amount. The analysts could then examine the "flagged" flights and intervene in the process before the revised limits were actually loaded into the reservations system.

The "on-line" routine allowed the analysts to intervene and to run the EMSR decision model for an individual flight leg and day of week. This routine was designed to display the input data used by the EMSR model to the analysts. It allowed the analysts to make exclusions of specific past demand data or to change the input data assumptions outright. The EMSR model calculations could then be performed on-line for the revised inputs and, if the results were judged to be reasonable by the analysts, the resultant booking limits would be loaded into the reservations system instead of those recommended initially by the batch routine. The on-line routine also enabled the analysts to compare the batch routine recommendations for booking limits with current limits and actual bookings already on hand for each future departure of the flight leg on the same day of the week (up to 90 days out). The EMSR recommended protection levels and booking limits could simply be overridden manually before being loaded into the reservations system.

ABLS was thus designed to allow user (seat control analyst) intervention in the application of the EMSR decision model to future flights. This capability was especially important in light of the many "imperfections" in the system. As of the end of 1986, ABLS did not include a demand forecasting model or an ability to make seasonal adjustments to demand and revenue estimates. Furthermore, no upgrade probabilities were included in the derivation of optimal fare class booking limits. These limitations required that analysts have the capability to override the recommended limits on flight legs where upgrade potential was thought to be significant and in markets or during periods for which the recent historical data did not provide a valid estimate of the revenue or demand conditions expected for future departures of the same flight leg.

The EMSR dynamic adjustment routine was implemented into the Automated Booking Limit System in December 1986. Further system development was not possible at Western Airlines because of the announced merger with Delta Air Lines. Fortunately, this pause in system development activities allowed the revenue impacts of ABLS, and of the EMSR dynamic adjustment routine in particular, to be tested for a sample of actual Western flight legs, as described below.

7.3 EMSR Revenue Impact Test

This final section of Chapter Seven describes a performance evaluation of the Automated Booking Limit System (ABLS) developed at Western Airlines, conducted during the first three months of 1987. The testing methodology is described first and its limitations are discussed. The revenue impacts of ABLS seat inventory control relative to the established manual method are then presented, and their implications for further system development are assessed.

7.3.1 Test Methodology

A performance evaluation of ABLS at Western Airlines began on December 22, 1986. The system users (i.e., the seat control analysts) were informed that specific flight legs operating on specific days of the week would have their fare class booking limits set automatically by ABLS, which included a dynamic adjustment routine based on the EMSR decision model. The automated system (or "batch" routine) began setting limits for a single day of the week for each of 21 flight legs. Both the flight legs and days of week involved in the test were known to the analysts because they had to refrain from applying their manual seat inventory control methods to these flights. The group of 21 flight leg/day of week combinations for which ABLS would be allowed to set and revise fare class booking limits without analyst intervention comprised the BATCH test group.

To give the analysts an opportunity to become familiar with ABLS and to interact with it, they were told to use the system for each of the 21 flight legs in the test, but on a specified day of week for each leg, different from the BATCH day of week. The seat control analysts could run the on-line version of ABLS for these flight legs/days of week and then adjust the EMSR recommended booking limits on the basis of their own judgement before loading them into the reservations system. Ultimately, interactive use of ABLS was to be the norm, with the combination of automation and human intervention resulting in what should have been higher flight leg revenues than what the automated system alone could generate. For the purposes of this test, however, this ONLINE test group was not intended to provide a measure of the joint performance of the system and the analysts. The objective of this test was to compare the completely automated system of seat inventory control used for BATCH test flights with the existing manual methods used for CONTROL test flights.

For each of the 21 flight legs in the test sample, then, five days of the week remained, for which booking limits would be set manually by the analysts, based entirely on their own judgement, as they had been for years. One of these remaining days of the week was selected to represent a CONTROL flight for each of the BATCH flights in the test. The specific day of week to be used as a basis for comparison with the BATCH day of week was unknown to the analysts. In fact, for the first two months of the test, the existence of any CONTROL flights at all was not known to the analysts.

The city-pairs and flight legs to be used in the test were selected on the basis of expected demand levels high enough to cause fare class booking limits to be reached in one or both of the test groups (i.e., BATCH or CONTROL). If fare class booking limits were never reached for the flight leg departures in the test because of very low demand in all fare classes, it would not be possible to determine the impacts of the different seat inventory control methods. The primary consideration in selecting flight legs for the test was therefore an expectation of relatively high demand, preferably in more than one fare class.

Historical traffic data by flight leg from the fall of 1986 (the most recent data available at the time) and for January through March 1986 (corresponding to the three months of the planned test in 1987) were examined and the highest load factor flight legs identified. Actual bookings already on hand as of December 15 for each of these flight legs were then reviewed for the test period. All flight legs for which one or more fare classes had already booked out on specific days of the week for departures during the test period were eliminated from further consideration. This step effectively removed all winter vacation markets from the test sample, since many of the flight legs considered already had one or more low fare classes sold out for the February and March peak period departures.

For each of the city-pairs/flight legs remaining in the sample, a more detailed analysis of historical booking patterns, load factors, and fare class mix of passengers was performed, including a comparison of days of the week. For a flight leg to be selected for the test, it had to exhibit stable and relatively high historical load factors, as well as a similar fare class mix of passengers for three days of the week (i.e., BATCH, ON-LINE, CONTROL). The final test sample included 21 flight legs operating between 19 city-pairs, as listed in Table 7.1. These city-pairs include a variety of stage lengths and involve primarily mixed business/pleasure markets on flight legs feeding Western's Salt Lake City or Los Angeles connecting hubs.

The three days of the week associated with each flight leg in the sample were assigned to the three test groups randomly. When minor differences in historical load factors between the days were apparent, this assignment was performed on the basis of the load factor rankings of the three days. For example, the highest ranked day of week was assigned to the BATCH group for one flight leg, to the CONTROL group for the next flight leg, and so on. This was an attempt to reduce the potential of systematic bias in demand levels by day of week favoring one test group over another.

Table 7.1: ABLS Flight Legs in Test Sample

FLIGHT LEG	CITY PAIR	DISTANCE (mi.)
BOI/SLC	Boise - Salt Lake City	291
DEN/SLC (2)	Denver - SLC	381
DFW/SLC	Dallas/Ft. Worth - SLC	989
IAD/SLC	Washington(Dulles) - SLC	1839
LAS/LAX	Las Vegas - Los Angeles	218
LAX/ABQ	L.A. - Albuquerque	660
LAX/PHX	L.A. - Phoenix	355
LAX/SMF	L.A. - Sacramento	372
MSO/SLC (2)	Missoula - SLC	436
MSP/SLC	Minneapolis/St. Paul - SLC	991
ORD/SLC	Chicago(O'Hare) - SLC	1254
PDX/SLC	Portland - SLC	630
SEA/SLC	Seattle - SLC	693
SLC/IAD	SLC - Washington(Dulles)	1839
SLC/JFK	SLC - New York(Kennedy)	1979
SLC/PDX	SLC - Portland	630
SLC/PHX	SLC - Phoenix	507
SLC/SMF	SLC - Sacramento	533

(2) – Two flight numbers included in test from this city-pair.

The test sample thus consisted of 21 flight legs. For each flight leg, all departures on one day of the week would have fare class booking limits set automatically by the ABLS "batch" routine, while all departures on another day of the week would be managed manually by seat inventory control analysts, thereby providing the "control" group for comparison purposes. A third day of the week for each flight leg would be handled through an "online" process in which analysts could use ABLS to assist in setting fare class limits. The primary focus of this evaluation of ABLS performance, however, was on the traffic and revenue levels of the BATCH group relative to the CONTROL group.

Several problems with interpreting the results of this test were identified even before the test began. The format of the test was constrained by the design of the system being tested, while comparisons of traffic and revenue impacts would be limited by the uncertainty associated with the demand for future flights and by the nature of any post departure assessment of the impact of different fare class booking limits on flight performance.

From the outset, the day of week rotation programmed into ABLS dictated that a day of week approach to comparative testing be used. The BATCH flights for a particular flight leg would be those departing on Tuesdays, for example, throughout the duration of the test, while all CONTROL flights might be Thursday departures. Although the days of the week for each flight leg were selected originally on the basis of similar historical loads and booking patterns, unexpected consistent differences in demand for one day in the test could give a systematic advantage for one test group over the other on that flight leg. Valid comparisons of the two test groups for the leg under such conditions would be difficult to make. It would have been preferable to assign different days of the week to each test group for different weeks during the test period.

A second interpretation problem involved the previously mentioned fact that the effects of different fare class booking limits can only be evaluated for flights that actually reached one or more of these limits in at least one of the test groups. If no fare class limits are ever reached for a set of flight legs, any difference in loads or revenues will be attributable solely to variations in demand between the departed flights. Even though the flight legs in the test sample were selected on the basis of high load factors for historical periods similar to the test period, there was no guarantee that unexpectedly low demand would not be observed on some of the test flights due to changing market conditions.

Finally, the inability to subject the same flight departure to different seat inventory control methods under exactly the same demand conditions meant that much of the analysis of the test results would be speculative. Even with a complete history of the fare class limits applied and the booking levels by day before departure, we could only speculate about what *might have* occurred in the booking process for the same flight departure had different fare class limits been applied. A comparative analysis of similar flights thus would not provide conclusions with certainty, although it was expected that the evidence of revenue impacts of different limits would allow reasonable conclusions to be reached. This is the major problem faced by airline managers hoping to measure the effects of seat inventory control practices — there is no way to determine exactly what revenues or loads would have been realized in the absence of seat inventory control or under different methods.

In light of these interpretation constraints, the results of this test had to be scrutinized on a departure-by-departure basis, to assess how different fare class booking limits *likely* affected actual bookings in each fare class, as well as total flight revenues. It was decided from the outset that aggregate revenue and traffic totals by test group would not be used to compare the performance of ABLS relative to the manual method of seat inventory control. Furthermore, all flight departures for which no fare class booking limits were ever reached would be excluded from the comparisons. The objective in interpreting the test results was therefore to identify BATCH and CONTROL flight departures that would have had similar booking patterns in the absence of different fare class limits.

Only flights departing after January 18 were included in the evaluation to ensure that they had been exposed to differences in booking limits for at least a month prior to departure. For the purpose of making direct comparisons between flights with as similar a demand patterns as possible, BATCH-CONTROL flight pairs were identified from the same week of the test period. It was felt that comparing a BATCH flight from one week with a CONTROL flight from another might introduce time-dependent variations in both demand and test group performance.

7.3.2 Revenue Impact Results and Assessment

The results of the revenue impact test were evaluated in terms of the differences in fare class mixes of passengers carried, load factors, and total flight revenues between the BATCH and CONTROL test groups. Actual boarding data by fare class were provided by Western's reservations and traffic database for departed flights, as were detailed pre-departure booking and fare class limit histories. Revenue levels associated with each boarded passenger on a test flight were the prorated leg revenue averages by fare class used as inputs to the EMSR calculations, extracted from Western's revenue database. The assessment of revenue impacts was based on a comparison of *flight pairs* (i.e., BATCH and CONTROL) that departed during the same week on the same flight leg.

Each flight pair presented a potential basis for comparison of the revenue and traffic impacts of ABLS relative to the manual method of seat inventory control used at Western. Post-departure results and booking histories were collected for a total of 210 flight pairs (21 flight legs, 10 test weeks). The major task in analysing the results, as introduced above, involved identifying the flight pairs that demonstrated differences in loads and/or revenues due to the application of different fare class booking limits, with all else being equal.

This identification process involved making judgements as to the similarity of the booking patterns of the two flights in each flight pair, based on a number of criteria. Final reservations and boarding totals by fare class as well as complete booking histories for each flight pair were examined. From the outset, flight pairs for which no fare class booking limits were reached for either flight were eliminated from further consideration. In spite of efforts to select flight legs and days of the week for which high demand was expected during the test period, more than half of the flight pairs were eliminated for this reason.

Identification of flight pairs for which at least one fare class booking limit was reached during the reservations process did not ensure a valid comparison of similar flights. Each remaining flight pair had to be judged as providing a "valid" comparison of the effects of different fare class limits or not, based on the historical booking build-ups of the flights involved. Acceptance of a flight pair as a valid comparison was based on the criteria described below.

A difference in load factors between the two flights in a valid flight pair should be attributable to differences in one or more fare class booking limits. For example, the lower load factor flight might have had lower fare class limits applied to a lower fare class, constraining total bookings for that flight relative to the total bookings accepted for the other flight with a higher limit on the same fare class. Significant differences in load factors that could not be explained by differences in fare class booking limits indicated a higher level of total demand for one flight over the other, resulting in rejection of that flight pair as a valid comparison.

In cases where booking limits on the same fare classes were reached for both flights, bookings in the next higher class should reflect a presence or lack of "upgrade" behavior, rather than a difference in total demand for the two flights. For example, bookings on one flight in a pair might have reached the booking limit on the lowest fare class (Q) of, say, 40, while bookings in the next highest class for the same flight reached a total of 30. For this flight to provide a valid comparison to the other flight in the pair, which had a higher unreached limit of 60 on its Q-class bookings, total bookings in B and Q classes for the second flight must be greater than or equal to the 70 accepted for the first flight. If the total demand patterns are the same for the two flights, a constraint on Q demand on the first flight should result in fewer total B and Q bookings than that accepted for the unconstrained B and Q classes on the second flight.

Even if the two lowest classes were judged to provide a valid comparison of similar demand processes, differences in booking totals for higher fare classes would result in rejection of the flight pair. The general rule used in selecting valid flight pairs was that any difference in load factors and fare class mix of passengers between the flights must be explainable by differences in the fare class limits applied during the booking process.

Determining the similarity of the demand process for the two flights in each flight pair considered also required a comparison of the pre-departure booking build-up process of each flight. For the two flights to be judged similar, the rate of booking level growth in the various fare classes had to be similar at various points in time prior to departure, in the absence of booking limit constraints. This task was made all the more difficult by the dynamic revision of booking limits performed on the BATCH flights. While the CONTROL fare class limits exhibited relative stability over the course of the booking process for each flight, perhaps changing once or twice, BATCH fare class limits were revised weekly for all fare classes. It was thus possible for a fare class to be closed down at one point during the booking process, only to be re-opened later due to ABLS calculating a higher revised fare class booking limit.

The best way to illustrate the characteristics of a valid flight pair comparison is to present specific examples from the test. Figure 7.2 shows the historical booking build-up for a BATCH and CONTROL flight pair from the BOI/SLC flight leg. The asterisks indicate that the fare class had reached its booking limit and was therefore unavailable on that particular day prior to departure. For the BATCH flight, no limits were reached during the booking process, whereas the CONTROL flight reached its Q limit 10 days prior to departure. On day 10, total bookings for the CONTROL flight actually exceeded those for the BATCH flight in all fare classes. The BATCH flight accepted a total of 71 more bookings, 32 of which were in Q class. The CONTROL flight accepted only 35 more bookings from day 10, none of which were in Q class. The net result was 20 more bookings for BATCH than CONTROL, and a load factor of 83 percent relative to 75 percent, respectively. The revenue totals for the two flights showed a 5.3 percent advantage for the BATCH flight.

A second example of a flight pair included as a valid comparison is provided by Figure 7.3, which shows booking histories for another flight pair, on the SLC/BOI flight leg. On the BATCH flight, the Q limit was not reached until day 6, at a level of 61, with a total booking level of 96. Thirty more bookings were accepted in higher fare classes between day 6 and departure day. In contrast, the CONTROL flight reached its Q-class limit on day 37 before departure, at a level of 13. This flight pair was judged to provide a valid comparison because total bookings on day 37 for the two flights were very similar and because the total booking rate thereafter was also similar, taking the constrained Q-class on the CONTROL flight into account. That is, differences in the booking rates between the flights for the higher fare classes could reasonably be attributed to the closed Q class on the CONTROL flight.

The final result for this flight pair again showed a substantially different fare class mix of passengers booked and boarded. The BATCH flight sold out on the day of departure and went out full (some passengers might have been denied boarding), whereas the CONTROL flight realized an 87 percent load factor. In this case, the differences in fare class mix on the two flights resulted in a 5.1 percent higher total revenue for the CONTROL flight, even though the BATCH flight carried more passengers.

The identification of valid flight pairs for purposes of comparison depended on judgements of whether the demand process of each flight in the pair appear to be similar. It

Figure 7.2: Valid Flight Pair Comparison – Example 1

HISTORICAL OR FUTURE LEG BOOKING
CITY PAIR: BOI / SLC

DAYS PRIOR	DAY	BATCH FLIGHT SEATS SOLD YSD	MSD	BSD	QSD	TOT	CONTROL FLIGHT SEATS SOLD YSD	MSD	BSD	QSD	TOT
	**ACTUA_	3	26	15	69	113	2	30	14	56	102
00		2	37	17	84	140	3	41	17	59 *	120
01		2	29	15	85	131	3	42	16	61 *	122
02		2	28	15	82	127	3	41	17	63 *	124
03		2	30	17	84	133	4	34	17	65 *	120
04		1	27	12	85	125	5	31	18	65 *	119
05		0	16	12	82	110	5	20	16	60 *	101
06		0	15	12	76	103	4	17	14	61 *	96
07		0	9	10	67	86	3	12	11	59 *	85
08		0	7	10	56	73	3	12	11	58 *	84
09		0	7	10	57	74	2	12	11	59 *	84
10		0	7	10	52	69	1	14	11	59 *	85
11		0	5	10	46	61	1	14	11	53	79
12		0	5	11	42	58	1	14	10	52	77
13		0	5	6	39	50	1	12	9	43	65
14		0	5	6	36	47	1	11	9	38	59
15		0	4	5	35	44	1	11	9	37	58
16		0	4	5	35	44	1	11	9	36	57
17		0	4	5	35	44	1	11	9	36	57
18		0	3	4	35	42	1	9	10	33	53
19		0	3	4	33	40	1	9	8	26	44
20		0	0	0	0	0	1	7	8	25	41
21		0	3	4	33	40	1	7	6	24	38
22		0	3	6	33	42	1	7	6	24	38
23		0	3	6	33	42	1	7	6	23	37
24		0	3	6	33	42	2	7	6	20	35
25		0	2	6	32	40	2	7	6	21	36
26		0	2	4	32	38	2	6	5	15	28
27		0	2	4	31	37	2	5	3	13	23
28		0	2	3	29	34	2	5	3	8	18
29		0	2	3	29	34	2	5	3	8	18
30		0	2	3	30	35	2	5	3	8	18
31		0	2	3	30	35	1	5	3	9	18
32		0	2	3	30	35	1	5	2	9	17
33		0	2	2	27	31	1	3	2	8	14
34		0	0	2	25	27	1	2	2	9	14
35		0	0	2	25	27	1	2	1	5	9

Figure 7.3: Valid Flight Pair Comparison – Example 2

HISTORICAL OR FUTURE LEG BOOKING
CITY PAIR: SLC / BOI

DAYS PRIOR	BATCH FLIGHT SEATS SOLD YSD	MSD	BSD	QSD	TOT	CONTROL FLIGHT SEATS SOLD YSD	MSD	BSD	QSD	TOT
**ACTUAL	8	37	7	55	107	6	57	18	12	93
00	7*	50	9	60*	126*	5	60	20	12*	97
01	4	42	11	60*	117	6	61	20	13*	100
02	2	39	10	61*	112	2	56	24	13*	95
03	2	35	8	59*	104	1	55	24	13*	93
04	2	34	8	57*	101	1	52	23	12*	88
05	2	34	9	60*	105	1	52	25	12*	90
06	1	25	9	61*	96	2	52	25	13*	92
07	1	22	7	59	89	1	41	25	12*	79
08	1	21	6	56	84	0	33	23	13*	69
09	1	15	4	43	63	0	27	21	13*	61
10	1	12	4	41	58	0	22	17	13*	52
11	1	12	4	41	58	0	18	15	13*	46
12	1	12	4	40	57	0	18	15	11	44
13	1	10	4	37	52	0	18	15	11	44
14	1	8	4	41	54	0	18	14	11	43
15	1	8	3	38	50	0	12	14	12	38
16	1	7	3	26	37	0	10	14	14*	38
17	1	7	3	25	36	0	9	14	12	35
18	1	7	3	27	38	0	6	14	12	32
19	1	7	3	22	33	0	6	14	12	32
20	1	9	3	18	31	0	6	14	12	32
21	1	8	3	17	29	0	7	12	14*	33
22	1	7	3	15	26	0	7	11	13*	31
23	1	4	3	14	22	0	5	9	13*	27
24	1	4	4	14	23	0	5	9	13*	27
25	1	4	4	14	23	0	3	7	13*	23
26	1	4	4	13	22	0	3	7	13*	23
27	1	4	4	11	20	0	3	7	13*	23
28	1	4	4	11	20	0	3	8	9	20
29	1	4	4	11	20	0	3	8	13*	24
30	1	6	2	11	20	0	2	7	12*	21
31	1	5	1	12	19	0	4	5	13*	22
32	1	5	1	12	19	0	3	4	13*	20
33	1	5	1	12	19	0	3	4	13*	20
34	1	5	1	12	19	0	3	4	13*	20
35	1	5	1	12	19	0	3	2	14*	19
36	1	5	1	14	21	0	4	2	14*	20
37	1	5	1	10	17	0	3	2	13*	18
38	1	5	1	9	16	0	3	1	11	15
39	1	5	1	9	16	0	3	1	11	15

would have been extremely difficult to identify flight pairs with *exactly* the same booking patterns, which also met the various other selection criteria described above. The objective of flight pair selection was to identify BATCH and CONTROL flights operating on the same flight leg during the same week that exhibited reasonably similar demand processes and demonstrated the impacts, both positive and negative, of the Automated Booking Limit System (ABLS).

Even with a relatively liberal application of the criteria described above, the final set of valid flight pair comparisons proved to be relatively small. Out of the 210 possible flight pairs in the test, approximately two-thirds were eliminated because of unexpectedly low demand. Winter flight cancellations and a change in aircraft type on one of the flight legs in the test also contributed to the number of flight pairs eliminated from the outset. Systematic day of week demand differences or simply different booking build-up processes eliminated about one-half of the remaining flight pairs, leaving 36 valid flight pair comparisons (72 flight departures) for an assessment of ABLS impacts.

The flight pairs showing positive and negative revenue impacts of ABLS are listed in Tables 7.2 and 7.3, respectively. For each flight pair listed, the fare class mix of passengers as well as the flight load factors and the percentage difference in flight revenues (BATCH over CONTROL) are shown. The passenger mix is the number of passengers actually boarded by fare class, and the load factor is the percentage of seats on the aircraft filled for that specific flight departure. The total revenues by flight have been omitted for reasons of data confidentiality.

The asterisks in both figures indicate that the associated fare class limit was reached by bookings for that flight during its reservations process, and that the closing of the fare class had a significant impact on how the flight proceeded to book up relative to the other flight in the pair. It is therefore possible, especially for flights that booked out very close to departure, that a 100 percent load factor might be indicated without asterisks being shown for one or more of the fare classes. The asterisks imply a *significant* constraint on the booking process for the relevant fare classes.

Table 7.2 details the 25 flight pairs in which ABLS was judged to have a positive impact on total flight revenues under similar demand conditions. The average impact per flight pair in this group amounted to a 14.3 percent higher revenue for the BATCH flight over the CONTROL flight. When weighted by the revenue levels and loads associated

Table 7.2: ABLS Positive Revenue Impact

TEST WEEK	FLIGHT LEG	TEST GROUP	Y	M	B	Q	L.F.(%)	REVENUE IMPACT
1	LAX/ABQ	B	1	7	12	87 *	100	+ 13.1 %
		C	0	7	31	47 *	79	
2	SLC/BOI	B	5	58	22	19	97	+ 17.1 %
		C	9	41 *	27	16	87	
2	SLC/PHX	B	8	13	20	84	92	+ 3.3 %
		C	0	20	50	44 *	84	
3	LAX/PHX	B	10	3	32	75 *	88	+ 13.6 %
		C	5	3	36	62 *	78	
4	LAX/PHX	B	17	3	56 *	28 *	97	+ 10.2 %
		C	4	13	38 *	51 *	99	
5	MSP/SLC	B	9	7	9	97 *	90	+ 15.1 %
		C	11	14	11	61 *	71	
5	SLC/BOI	B	5	36	30	32	96	+ 20.7 %
		C	0	33	25 *	25 *	78	
6	BOI/SLC	B	2	44	27	60	98	+ 7.1 %
		C	17	41 *	17 *	44 *	88	
6	MSP/SLC	B	2	9	25	100 *	100	+ 29.9 %
		C	1	20	19	69 *	80	
7	LAX/SMF	B	6	9	14	70	93	+ 28.7 %
		C	6	15	10	43 *	69	

Table 7.2: ABLS Positive Revenue Impact (cont'd)

TEST WEEK	FLIGHT LEG	TEST GROUP	Y	M	B	Q	L.F.(%)	REVENUE IMPACT
7	BOI/SLC	B	3	26	15	69	83	+ 5.5 %
		C	2	30	14	56 *	75	
7	SLC/PHX	B	0	6	4	127 *	100	+ 7.1 %
		C	6	16	44	49 *	85	
7	LAS/LAX	B	4	1	8	62	63	+ 22.6 %
		C	5	4	9	41 *	49	
7	LAX/PHX	B	1	1	28	64 *	88	+ 35.7 %
		C	0	6	14	49 *	64	
8	LAX/PHX	B	18	10	41	38 *	100	+ 15.1 %
		C	5	10	30	62 *	100	
8	IAD/SLC	B	5	4	50	77 *	100	+ 14.7 %
		C	3	6	41 *	68 *	87	
8	SLC/PHX	B	6	4	13	118 *	100	+ 4.0 %
		C	13	12	54	44 *	90	
8	MSP/SLC	B	0	2	19	110 *	96	+ 29.8 %
		C	5	16	5	71 *	71	
9	IAD/SLC	B	7	8	61	61 *	100	+ 8.0 %
		C	6	9	48 *	66 *	95	
9	LAS/LAX	B	3	4	1	69	64	+ 23.6 %
		C	8	7	1	43 *	49	

Table 7.2: ABLS Positive Revenue Impact (cont'd)

TEST WEEK	FLIGHT LEG	TEST GROUP	Y	M	B	Q	L.F.(%)	REVENUE IMPACT
9	SLC/IAD	B	4	6	40	86 *	100	+ 3.4 %
		C	13	17	42 *	53 *	92	
9	LAX/ABQ	B	10	9	16	71 *	99	+ 12.1 %
		C	2	3	15	83	96	
9	PDX/SLC	B	10	21	42 *	34 *	100	+ 2.7 %
		C	4	24	30 *	39 *	100	
10	SLC/BOI	B	3	45	17	42 *	100	+ 6.2 %
		C	8	46	25	13 *	86	
10	SLC/IAD	B	8	9	35	85 *	100	+ 9.0 %
		C	4	17	46 *	50 *	86	

with each flight pair, the aggregate positive impact on total revenues amounts to a 12 percent advantage for the BATCH flights. The average flight load factors for all flights in this positive impact group were also 12.3 percentage points higher for the BATCH flights than for the CONTROL flights, due to the application of ABLS-generated fare class booking limits.

A scan of the fare class mixes and load factors in this group reveals that, in the majority of the flight pairs, the BATCH flights carried a far greater number of passengers in the lowest fare classes than did the CONTROL flights, with a higher total revenue. The flight pairs in this group illustrate cases in which the revenue benefit of selling more low-priced seats outweighed that of closing down the lowest fare classes in the hope that denied requests would result in fare class upgrades. Note, however, that in many of the flight pairs, lower Q-class limits on the CONTROL flight did in fact increase the number of B passengers carried.

The relatively high booking limits on the lower fare classes set by ABLS for the BATCH flights can be attributed to at least two factors. First, the prorated leg revenue averages used as inputs by ABLS did not differ radically among the fare classes for many of the flight legs in the test. This was the case for flight legs in highly competitive, short-haul markets served by Western, in which the price levels of the fare products being sold in adjacent fare classes can differ by as little as $10 (e.g., $49 for a Q-class fare versus $59 for a B-class fare). When such similar price levels are aggregated and prorated over flight legs, the fare class revenue ratios used by the EMSR formulation contribute to relatively low protection levels for the higher fare classes, leaving more seats available to the lowest fare class, Q.

The second factor is related to the first, and involves the lack of any measure of upgrade potential in the EMSR formulation used in this test. Whereas the seat control analysts at Western could could set fare class limits on CONTROL flights with some expectation of upgrade potential, the BATCH limits reflected an assumption of zero upgrade potential. As a result, there is evidence in many of the flight pairs listed in the positive impact group of upgrade behavior for the CONTROL flights with constrained fare class demands, as indicated by the asterisks. For the flight legs operating in markets in which the price levels of fare products in adjacent fare classes were similar, or in which Western faced little effective competition, the actual upgrade potential appeared to be significant. The flight pairs for which the inability of ABLS to take into account fare class upgrade potential led to a negative revenue impact are listed in Table 7.3, and will be discussed later.

The higher total revenues for the BATCH flights in Table 7.2 can be explained in the majority of the flight pairs listed by higher load factors stemming from more liberal booking limits on lower fare classes. This effect is readily apparent from the fare class mixes and asterisks shown for these flight pairs. Several of the flight pairs on the list, however, merit further explanation:

- *Week 2, SLC/BOI:* Although the number of passengers carried in B and Q classes were similar, the number in M class differed substantially due to a lower M-class limit for the CONTROL flight.

- *Week 2, SLC/PHX:* Substantial differences in Y-class boardings do not have as great an impact on total revenues in this test as would be expected. Recall that the Y-class pollution of demand and revenue data generally causes the Y-class prorated revenue average to be lower than that for M or even B classes.

- *Week 4, LAX/PHX:* In the only flight pair for which the BATCH flight showed a lower load factor than the CONTROL flight, higher total revenues were realized on the former because of *lower* Q limits and substantial upgrades from Q to B. This result was also observed in Week 8, for the same flight leg.

- *Week 5, SLC/BOI:* Both B and Q classes were constrained on the CONTROL flight, with little upgrade activity, relative to the BATCH flight.

- *Week 7, LAS/LAX:* For short-haul, highly competitive markets like Las Vegas - Los Angeles, very little upgrade activity was observed when Q class was constrained, especially for relatively low load factor flights like this one.

- *Week 7, LAX/PHX:* In this case, the higher number of B passengers for the BATCH flight is due to a relatively low Q-class limit having been applied by ABLS early in the booking process. The Q-class limit on the BATCH flight was then increased late in the process, when expected bookings in the higher classes did not materialize. The lower Q-class limit on the CONTROL flight applied throughout the booking process, keeping total bookings lower.

- *Week 9, LAX/ABQ:* This pair deserves special mention because it provides an example of how the seat control analysts began to adjust their own practices given exposure to the ABLS recommended limits. In this case, the CONTROL flight actually carried 83 Q-class passengers, a number much higher than the Q-class limits applied for the same CONTROL flight earlier in the test period.

- *Week 9, PDX/SLC:* The pair demonstrates how the limits set for fare classes higher than Q (in this case, B) can be just as important to total revenues as the Q-class limits.

To summarize the results of the positive impact flight pairs, the BATCH limits on lower fare classes were substantially higher than those for CONTROL flights, with a few exceptions. The CONTROL limits on the lowest fare classes were generally *too low* to maximize total flight revenues. On the other hand, there is some evidence that the BATCH limits on the lower fare classes might have been too high, giving passengers that would have been willing to purchase higher-price seats access to lower fare classes. This finding is supported by the results for the 11 flight pairs presented in Table 7.3, which demonstrate the negative impact on revenues of ABLS.

In Table 7.3, all of the flight pairs once again show a BATCH flight load factor greater than or equal to that of the CONTROL flight, but a total flight revenue that was lower for the BATCH flight in each case. The revenue advantage for the CONTROL flights stemmed from lower booking limits on Q class and upgrade movements to higher revenue fare classes. The asterisks in this figure demonstrate this phenomenon clearly. Flight pairs that merit further explanation include:

- *Week 1, LAX/SMF:* The upgrade movement here involved a vertical shift from Q to M class, likely because the B-class fare product offered in the local market served by this flight leg was not as attractive an upgrade alternative at the M-class product (i.e., same price, more heavily restricted). The same observation applies to the mix of bookings for Week 7, SLC/BOI.

- *Week 7, IAD/SLC:* The BATCH flight reached its limits in both B and Q classes, but these limits were higher than those for the CONTROL flight in the same classes.

Table 7.3: ABLS Negative Revenue Impact

TEST WEEK	FLIGHT LEG	TEST GROUP	Y	M	B	Q	L.F.(%)	REVENUE IMPACT
1	LAX/SMF	B	5	3	12	85	98	– 4.0 %
		C	8	28	12	54 *	95	
2	LAX/ABQ	B	1	1	12	90	97	– 0.5 %
		C	1	4	37	49 *	85	
3	SLC/PHX	B	2	6	13	114	99	– 4.0 %
		C	4	25	51	43 *	90	
4	SLC/PHX	B	3	8	18 *	107 *	100	– 10.7 %
		C	7	18	67 *	44 *	100	
5	LAX/ABQ	B	4	5	9 *	89 *	100	– 10.7 %
		C	4	14	39 *	43 *	93	
5	SLC/PHX	B	4	4	11	109 *	94	– 2.7 %
		C	2	12	59	46 *	88	
7	SLC/BOI	B	8	37	7	55 *	100	– 5.1 %
		C	6	57	18	12 *	87	
7	IAD/SLC	B	7	3	53 *	73 *	100	– 2.8 %
		C	13	23	38 *	61 *	99	
8	SLC/BOI	B	3	36	21 *	46 *	99	– 15.3 %
		C	22	52 *	22 *	10 *	99	
9	SLC/PHX	B	1	7	12	116 *	100	– 6.0 %
		C	3	19	63 *	46 *	96	
10	SLC/PHX	B	3	4	13	116	100	– 5.0 %
		C	8	23	60 *	35 *	93	

- *Week 8, SLC/BOI:* An example of upgrade movement from M to Y class is provided by the CONTROL flight in this pair.

For the 11 flight pairs in this group showing negative revenue impact, the average revenue difference per flight pair amounted to a 6 percent shortfall for the BATCH flight relative to the CONTROL flight. The weighted difference in aggregate revenues for this group amounted to a 5.9 percent shortfall for the BATCH flights. The difference in average flight load factors was 5.7 percentage points in favor of the BATCH flights, many of which in fact departed with no empty coach seats.

The overall conclusion that can be drawn from the flight pairs shown in Table 7.3 is that the BATCH flights definitely did have booking limits on the lower fare classes that were *too high* relative to the CONTROL flights. Upgrade activity contributed to higher total revenues for the CONTROL flights in this group, in spite of higher load factors for the BATCH flights. Most of the flight pairs listed, however, belonged to one of three flight legs in the test, suggesting that these flight legs experience a significant upgrade potential that might not be realized on other flight legs in different markets.

The aggregate revenue and load factor results for all 36 flight pairs considered to reflect a valid comparison of ABLS versus the manual method of seat inventory control are summarized in Table 7.4. An overall positive revenue impact of 6.2 percent was realized for the BATCH flights over the CONTROL flights in this comparison. The average flight load factor advantage for BATCH flights amounted to 10.3 percentage points. This discrepancy between the magnitude of the increase in load factor and total revenues is explained by the lower overall yield realized on the BATCH flights than on the CONTROL flights. The shortfall in overall yield was outweighed, however, by the increase in total revenues, providing evidence that yield maximization does not necessarily mean revenue maximization.

The potential positive impact on total flight revenues of a more systematic approach to seat inventory control than the application of booking limits based on analysts' judgements was clearly demonstrated by this initial test of a relatively basic version of Western's Automated Booking Limit System (ABLS). This conclusion is supported by the many other examples of individual flights for which dynamic revision of fare class limits by ABLS had a significant positive impact on total flight revenues, but for which the

Table 7.4: Summary of Test Results

	BATCH	CONTROL	ABLS IMPACT
ABLS Positive Revenue Impact (25 flight pairs)			
TRAFFIC			
Total Passengers	2,886	2,508	+ 15.1 %
RPMs	2,178,983	1,902,789	+ 14.5 %
LOAD FACTORS			
Average per Flight	93.8 %	81.5 %	+ 12.3 pts.
Weighted Average	97.0 %	84.7 %	+ 12.3 pts.
REVENUES			
Total Flight Revenues	$ 210,774	$ 188,217	+ 12.0 %
Aggregate Yield (cents)	9.67	9.89	- 0.22
ABLS Negative Revenue Impact (11 flight pairs)			
TRAFFIC			
Total Passengers	1,336	1,262	+ 5.9 %
RPMs	830,604	792,123	+ 4.9 %
LOAD FACTORS			
Average per Flight	98.9 %	93.2 %	+ 5.7 pts.
Weighted Average	99.1 %	94.5 %	+ 4.6 pts.
REVENUES			
Total Flight Revenues	$ 85,501	$ 90,866	- 5.9 %
Aggregate Yield (cents)	10.29	11.47	- 1.18
GRAND TOTALS (36 flight pairs)			
TRAFFIC			
Total Passengers	4,222	3,770	+ 12.0 %
RPMs	3,009,587	2,694,912	+ 11.7 %
LOAD FACTORS			
Average per Flight	95.4 %	85.1 %	+ 10.3 pts.
Weighted Average	97.5 %	87.3 %	+ 10.2 pts.
REVENUES			
Total Flight Revenues	$ 296,275	$ 279,082	+ 6.2 %
Aggregate Yield (cents)	9.84	10.36	- 0.52

corresponding CONTROL flight did not reach its limits or did not exhibit a similar demand pattern.

While the final results of this test show a positive revenue impact due to the application of the quantitative decision approach of the EMSR framework, there are several *caveats* with respect to the test results that should be emphasized. First, this test compared the impact on loads and revenues of an automated system for seat inventory control relative to the manual methods being used at Western Airlines. We can only assume that the judgement and expertise of the analysts involved in this test is representative of that of analysts at other airlines. It is possible that analysts with better training and skills, or even a larger number of analysts paying closer attention to individual flights could have reduced the the observed differences between the BATCH and CONTROL test flights.

On the other hand, given that many airlines currently rely on the judgement of relatively few seat control analysts, as discussed in Chapter Three, this test succeeded in demonstrating that there are advantages associated with a more systematic decision approach. It should also be noted that ABLS was designed to enable analysts to apply the EMSR-derived booking limits or to intervene in instances where factors not accounted for in the ABLS routine might affect the optimal booking limits for a future flight. This test compared a strictly automated approach with a strictly manual approach, meaning the joint potential of automation with human intervention was not assessed.

While the performance of flights for which analysts set fare class limits on the basis of the ABLS recommended limits (i.e., the ONLINE test group) was not evaluated in detail, the results for the interactive group seem to be mixed. The analysts at Western were not trained formally with respect to the use of ABLS in an interactive environment. The result was substantial variation between analysts, across flight legs, and over time in the extent to which the recommended limits were actually applied. As the test progressed, there was in fact evidence of a learning curve effect on the part of the analysts, which also affected the CONTROL flights in some cases. Given repeated exposure to the ABLS recommended limits for ONLINE flights, the analysts began applying remarkably similar limits to some CONTROL flights.

The second *caveat* involves the version of ABLS that was used for this test. The system tested included only the basic EMSR formulations for calculating initial fare class limits and revising them periodically in light of changes to the input data and actual bookings for a future flight. The EMSR model extensions to include upgrade probabilities and to take into account overbooking factors were not incorporated into ABLS. Furthermore, the data inputs were limited to those available from Western's reservations and revenue databases. The prorated revenue averages reflected substantial data pollution in many cases, while the demand inputs used were simple historical averages from recent operations of the same flight leg. No seasonal adjustment or growth trend forecasting was performed.

The period between ABLS revision runs for each future flight was chosen to be 7 days, which proved to be too long for many of the flight pairs examined in this test, particularly within the last two weeks of the booking process. More frequent revisions would have reduced the BATCH limits on low fare classes substantially in many of the flight pairs. In spite of these shortcomings of the system tested, analysis of the performance of ABLS with respect to specific flight pairs clearly illustrated the benefits of an automated seat inventory control system that can adjust booking limits dynamically.

The proportion of flights for which an impact was observed presents the third major *caveat* with respect to the test results. It was known from the outset that different methods of seat inventory control will only have an observable impact when demand for at least one of the fare classes available on a future flight is relatively high. The greater the demand for a flight, the greater the impacts of effective seat inventory control methods. The test results described above generally involved flights with load factors above 80 percent. In general, about one-quarter of an airline's flight legs depart with such high load factors, meaning that the positive revenue impacts observed in this test would be reduced when spread over all of the flight legs operated by an airline. However, high load factor flights generate a disproportionately high percentage of total airline revenues, and it is the high load factor flights that require the greatest amount of seat inventory control attention.

The shortcomings of the system tested and of the test itself highlight the difficulty of measuring the impacts of seat inventory control on airline traffic and revenues. The performance evaluation of ABLS completed at Western Airlines as part of this research nonetheless demonstrated that there exist potential benefits in automated seat inventory control over manual methods. These benefits could be even greater with further improvements to the estimation methods and decision models employed. The EMSR decision framework developed earlier in this dissertation provided the quantitative approach used by the automated system to derive the fare class booking limits applied to the BATCH flights. The specific formulations used in ABLS, however, did not incorporate fare class upgrade probabilities. The test results provided numerous examples of the importance of upgrade potential to revenue maximization. Furthermore, the demand and revenue inputs to the EMSR framework had shortcomings that could be overcome through the development of better estimation techniques.

The potential of automated seat inventory control could be enhanced significantly with the application of the EMSR decision approach in an interactive system for seat inventory management. Properly trained analysts could intervene to account for the shortcomings of even the basic version of ABLS that was tested. Operational and time constraints made the evaluation of such an interactive system impossible at Western. It is nonetheless clear that, even with continued research and development of quantitative methods for seat inventory control, the effectiveness of an automated system can only improve with the intervention of skilled analysts to handle unusual circumstances.

Chapter 8

Conclusions for Future System Development

The preceding chapters have dealt with the many interrelated aspects of airline seat inventory management, from the theoretical underpinnings of the problem through the development of a quantitative decision framework and its implementation. The focus of this work has been on the existing constraints of the seat inventory control problem faced by airlines and their impacts on practical solution approaches. References have also been made to the ways in which some of these existing constraints might be overcome and solution methods improved through further seat inventory control system development.

This chapter concludes this dissertation by relating the characteristics of the seat inventory control problem, current and future, to those of airline systems designed to manage seat inventories and maximize revenues. The findings and contributions of this research are summarized first. Lessons from the development and implementation of the EMSR decision framework into an automated booking limit system are also discussed, with an emphasis on the most immediate needs for system development. Finally, some of the most important unanswered theoretical questions surrounding the practice of differential pricing of airline fare products and limiting the availability of seats sold at different price levels are presented as directions for further research.

8.1 Research Findings and Contributions

Much of the discussion in this dissertation has emphasized the way in which the methods employed to manage seat inventories are related to the marketing policies, operating procedures, and technical capabilities of an airline. The argument has been made that improvements to the seat inventory control process must be made in the context of the practical constraints posed by these elements. At the same time, the development of better seat inventory control methods presents an opportunity to eliminate some of these constraints.

The realization of significant improvements to airline seat inventory control systems will in the long run involve a recursive development process. Existing constraints determine the techniques which may be implemented in the short run, while the implementation of new methods may suggest innovations that will eliminate one or more limitations of the existing approach. This section summarizes the findings and contributions of this research in the context of this process.

From the outset, the need for airlines to practice seat inventory control came about because of the differential pricing and market demand segmentation practices that have evolved in the air travel marketplace. The importance of effective demand segmentation through the definition of fare product attributes was described in Chapter One. The price sensitivity of the consumer and the time sensitivity of the trip being considered were the two criteria used to segment demand in the model described in this thesis. These criteria were incorporated into a qualitative model of the consumer choice process for air travel, and extended to reflect the reservations process as viewed by airlines. This descriptive framework, based on the concepts of travel disutility minimization, provides a foundation for the development and estimation of quantitative consumer choice models.

Further work in this area will require significant improvements to airline data availability. Data collection practices in the airline industry are surprisingly primitive, yet the potential for the retrieval of useful data from airline reservations systems is enormous. Millions of transactions in the form of flight availability inquiries, bookings, cancellations, and ticket purchases occur every day. Currently, even the most advanced airline reservations systems can only provide running totals of cumulative bookings by fare class for a future flight, net of all booking and cancellation activity that occurs between the times when data is extracted to historical databases. Activities at flight check-in can be as important to the accuracy of the data collected as the booking process itself, yet few airlines' systems have the capability to provide detailed information about no-shows, "go-shows", misticketed passengers, and class of service upgrades.

Even more surprising than the lack of detailed reservations and traffic data is the virtual non-existence of detailed revenue data at several major airlines. While a few carriers have developed revenue databases, most others must rely on sporadic samples of ticket coupons collected from passengers to estimate revenues and yields by flight leg, O-D market, and fare class. Having focused on demand data while taking revenue data for granted, airlines are now beginning to realize that making optimal seat inventory control decisions requires revenue data that corresponds in the level of detail to the demand data being collected.

The limitations on data availability, as well as many of the practical constraints for seat inventory control, can be traced to the characteristics of the dominant reservations systems and distribution channels that have evolved in the airline industry. The way in which reservations systems display seat availability and keep track of bookings (currently by flight leg and fare class) defines the seat inventory control "problem". The way in which the problem is defined in turn determines the most appropriate mathematical approaches and solution methods. For at least the next several years, much of the airline industry will still employ reservations systems that control seat inventories by flight leg and fare class, with the fare classes nested within a shared inventory of available seats.

The decision approach of the Expected Marginal Seat Revenue (EMSR) model developed in this dissertation can be applied with little modification to most existing reservations systems with multiple nested fare classes. The theoretical basis of the EMSR model is a focus on expected future requests by fare class and the derivation of optimal *protection levels* for upper fare classes, a requirement of nested fare class inventories not addressed by previous work. The introduction of fare class overbooking factors and upgrade probabilities into the derivation of these protection levels represent further theoretical contributions to seat inventory control research.

By the turn of the decade, several of the largest airlines will have completed the transition to seat inventory control on the basis of a "virtual nesting" approach that limits seat availability by flight leg and passenger itinerary/ticket revenue, not simply by fare class. The data requirements of this "next generation" of seat inventory control systems will be substantially different from those discussed above, and extensions to the mathematical methods already developed will be required. As discussed in Chapter Six, the basic decision approach of the EMSR framework can be extended to reservations systems that employ virtual nesting.

Practical contributions of this work include the application of the EMSR decision model to a dynamic revision process that incorporates actual bookings received for a future flight into the adjustment of fare class protection levels and booking limits. The revenue impacts of using the EMSR dynamic adjustment routine were demonstrated in the evaluation of automated seat inventory control at Western Airlines. Given that most airlines still rely almost exclusively on human judgement in making seat inventory control decisions, there is great potential for using mathematical models to assist in this decision process. Airlines must recognize, however, that even the most sophisticated mathematical approaches will require the intervention of skilled analysts.

The practical use by airlines of optimization or forecasting models will depend on the detail and accuracy of the data required, the availability of which will be limited, for the immediate future, at least. The solutions or estimates derived from such models will be probabilistic in nature and based on historical patterns of demand. Even with a "perfect" set of data inputs, there will always be variables that cannot be accounted for in mathematical models, including rapid changes to the competitive environment in one or more markets and the occurrence of unusual events that can affect flight bookings. The objective in the development and application of quantitative methods for seat inventory control is to incorporate them into systems that make routine calculations as systematic as possible so that analysts can focus their efforts on these exogenous variables.

The experience of implementing a quantitative decision process at Western highlighted the requirements of future seat inventory control system development. Many of the short-term requirements relate to improving the estimates of fare class demand and revenues that must be derived from existing airline data. The greatest needs involve the inclusion of fare class upgrade probabilities and differential overbooking factors into the decision model. Even if substantial improvements to data availability for this purpose are realized, empirical analysis of these data will be required to develop reliable estimation methods. Directions for further theoretical and empirical work will be discussed in greater detail in the final section of this chapter.

Further development of forecasting methods to predict future demand for flight departures and fare classes is also an important system development requirement related to the implementation of the EMSR approach. Given the inherent variability of air travel demand and the volatility of airline markets, the development of generalizable yet accurate forecasting models will be difficult. The application of the EMSR framework nonetheless requires that statistical and econometric methods be used to improve the estimates of, and reduce the uncertainty associated with, the demand inputs.

The development of seat inventory control methods and revenue management systems is certain to be an on-going process, the requirements of which will change as the definition of the seat inventory control problem evolves. In the short run, the development of improved data collection and retrieval systems is critical to realizing the potential benefits of quantitative models for seat inventory control. As important as the database management system being used is the quality of the data being collected, and improvements in this respect can require changes to airline procedures and reservations systems. Both of these tasks represent medium-term objectives for airlines that hope to improve their revenue management systems. The most complex and longest term objective involves the concurrent reconfiguration of reservations systems to manage seat inventories by passenger itinerary and ticket revenue, and the development of mathematical decision models to forecast future demand and to maximize system revenues in the context of these reservations systems.

8.2 Directions for Further Research

The need for further theoretical and empirical research related to airline seat inventory control and revenue management depends to a large extent on the future evolution of the "problem", as defined by the practical constraints discussed in this dissertation. Most important of these constraints is the specific technical approach to seat inventory control that is adopted by airlines, the existing approach being fare class/flight leg control. The current status of seat inventory control practices and the planned development of more sophisticated approaches to revenue management suggest two distinct yet interrelated areas for further research.

The mathematical models required to determine optimal booking limits for future flights represent the first research area of significance to the future development of seat inventory control systems. In this thesis, the problem presented by the nested fare class inventories currently used by most airlines to control bookings by flight leg was addressed. The EMSR framework was developed for application specifically to the problem as defined by this approach to seat inventory control.

The development of virtual nesting approaches to control seat inventories by fare level and travel itinerary introduces complications that will require extensions to the EMSR formulations or the development of new techniques for determining optimal seat availability. Specifically, the problems associated with the incorporation of passenger upgrade potential and overbooking factors into a virtual nesting approach will have to be addressed. The EMSR decision approach provides a basis for the development of optimization models for virtually nested systems. The potential revenue losses from too great an emphasis on high-revenue through passengers in such systems, however, will require the addition of an algorithm that will identify situations in which multiple single-leg itineraries will generate a higher total revenue than a single multiple-leg itinerary.

The network formulations reviewed in Chapter Four provide directions for the development of optimization models that can take into account the true impacts on airline system revenues of requests for different fare products and flight itineraries. Further work on these approaches might reduce the size of the formulations involved and the computing resources required. Concurrent improvements to computing capabilities could lead to the implementation of system-wide revenue maximization models, in which case the basic approach to controlling seat inventories might have to be changed once again.

The practical application of large network formulations to airline seat inventory control depends ultimately on the defined scope of the "problem". If usable network-based revenue maximization routines are to be developed for dynamic applications, the scope of the problem might have to be limited to, for example, a single set of connecting flights operating through an airline's hub airport. Expanding the scope of the network problem to larger portions of the airline's route system or to longer time periods would once again make the size of the formulations required a practical limitation on the use of network models.

No matter what optimization models are applied to leg-based or virtually nested ...ns for seat inventory control, the importance of accurate data inputs is likely to increase. The need for better estimates of revenue and, especially, demand for future flight departures represents the second major area requiring further research. The research areas of most immediate importance to both existing and planned seat inventory control systems are discussed briefly below.

The estimation of future demand for flights and/or O-D itineraries requires the development of forecasting models to predict forthcoming requests at various points in time prior to departure. Extensive empirical analysis of airline reservations and boarding data is necessary to develop such models and to answer the following questions:

1. What historical data sample is most representative of a particular time period, market, and flight departure for use in forecasting demand for a future flight?
2. How can seasonal variation from past years as well as current traffic growth trends be incorporated into estimates of demand by fare class for future flights?
3. How do estimates of future demand for a particular flight change as actual bookings are accepted for that flight?

This last question will require analysis of conditional probabilities and densities of demand for homogeneous subsets of the total demand in a market. As mentioned earlier, there is a practical need for generalizable forecasting models that can be applied to a variety of markets and time periods. Given the limitations of existing airline demand data, deriving generalizable conclusions from empirical analysis will be difficult.

An important component of any estimate of future demand for a flight is the inclusion of the potential of vertical and horizontal choice shifts, as well as booking losses when requests are refused. Again, conditional probabilities of each of these events given that a requested seat is not available must be estimated. The most immediate need of both the existing leg-based and planned virtual nesting approaches to seat inventory control is the development of models to estimate the probabilities of vertical choice shifts (i.e., fare class upgrades).

Estimating fare class upgrade potential will require empirical analysis of historical reservations data, to determine aggregate upgrade behavior. Once again, the limitations of the data collected from existing reservations systems will make any such estimates subject to numerous qualifications. Disaggregate data on fare class upgrades would have to be collected from surveys of individual passengers and/or reservations transactions. The key to deriving accurate estimates will be analysis of passenger behavior under different conditions and in different markets, with the objective of developing a usable model of passenger choice.

Estimates of horizontal choice shifts will become more critical as the scope of the seat inventory control problem addressed by optimization models is expanded. The time dimension of the seat inventory control problem has yet to be addressed in the context of system revenue maximization. With the growth of "mega-airlines" that offer a wide variety of departure times each day in a single market on flight itineraries through several connecting hubs, the revenue and traffic implications of "horizontal" choice shifts by passengers to alternate flights and times can be substantial. To further complicate the problem, the times, fare products offerings, and seat availability of competitors' flights in the same O-D market all affect the revenue maximization objectives of each individual airline.

Passenger behavior is also important with respect to cancellations of bookings and passenger no-shows. The relationship between optimal fare class booking limits and optimal overbooking limits was established in this thesis, and incorporated into the EMSR formulations. The approach employed, however, assumed a given optimal overbooking factor by fare class. Overbooking models need to be developed to determine optimal overbooking factors by fare class as a function of time, based on reliable estimates of no-show and cancellation behavior by fare class.

Empirical analysis of historical reservations and boarding data is required to determine the extent to which no-show and cancellation patterns differ by fare class, time period, and market. Furthermore, the potential revenue impacts of significantly different rates, particularly those associated with non-refundable and partially refundable low-priced fare products must be assessed. This information should be incorporated into the revenue-based decision approach of the EMSR framework, such that overbooking limits are determined in conjunction with the optimal fare class booking limits.

The estimation and forecasting of passenger demand by O-D market, flight itinerary and fare class represents an extremely large area of research that can lead to significant contributions to airline revenue management, as well as to many other airline and transportation applications which depend on accurate demand estimates. Modelling passenger behavior presents the greatest challenge in this area, especially in light of the limited data availability and inherent instability of the airline markets that generate the required data. Research into demand and passenger behavior is nonetheless essential if the inputs required by even existing optimization methods are to result in reliable optimal solutions.

At this point in time, the need for better demand modelling is greater than the need for more sophisticated optimization methods. Any further development of optimization methods should be directed toward the immediate needs of existing leg-based seat inventory control systems and the imminent needs of virtually nested systems. Concurrent development of improved optimization models and demand estimation techniques would be desirable, so that the requirements of the former can be addressed by the latter. Above all, the emphasis in further research must continue to be on the limitations posed by reservations system capabilities, data availability, and the rapidly changing nature of airline competitive practices.

Recent developments in airline seat inventory control have focused on the personnel and database management systems required. Potentially the largest gains from more sophisticated seat inventory control techniques are yet to be realized, as reservations systems are revised and linked together, innovations to allow O-D seat inventory control are implemented, and the use of decision models is expanded. The potential benefits to both the airline industry and the traveling public are great. Airlines can look forward to better management of their available capacity, higher load factors and, higher total revenues. The traveling public, in turn, can look forward to greater access to available seats and a wide variety of fare products, as airlines gain greater control over their pricing and distribution practices in a highly competitive marketplace.

References

1. This research, conducted over a ten-year period, is summarized in "Boeing Promotional Fare Management System: Analysis and Research Findings", Sales Technology Department, Boeing Commercial Airplane Company, Seattle WA, July 1982.

2. "Discount Fares and the Potential for Profit or Loss", Strategic Planning Department, Boeing Commercial Airplane Company, Seattle WA, October 1980.

3. William M. Swan, "A Systems Analysis of Scheduled Air Transportation Networks", Flight Transportation Laboratory Report FTL-R79-5, Massachusetts Institute of Technology, June 1979, pp 83-84.

4. DOT data for 1985 show the *average* airline expense per passenger to be $13.31 for passenger service and $9.59 for passenger handling [Source: *Air Transport World,* September 1986]. The *marginal* expense per passenger could well be lower than the sum of these two figures.

5. Alfred Kahn, *The Economics of Regulation,* Volume 1, John Wiley Publishers, New York, 1970, p. 142.

6. "Some Considerations in the Review of Domestic and Transborder Air Fares", Canadian Transport Commission, Ottawa-Hull, 1984.

7. Steven Shaw, *Air Transport: A Marketing Perspective,* Pitman Books, London, 1982.

8. George W. Douglas and James C. Miller, *Economic Regulation of Domestic Air Transport,* The Brookings Institution, Washington D.C., 1974, pp 68-69.

9. Nawal K. Taneja, *The Commercial Airline Industry,* D.C. Heath and Company, Lexington MA, 1976, pp 66-68.

10. John R. Meyer *et al, Airline Deregulation: The Early Experience,* Auburn House, Boston MA, 1981, pp 257-259.

11. Alexander T. Wells, *Air Transportation: A Management Perspective,* Wadsworth Publishing, Belmont CA, 1984, pp 299-301, 321-322.

12. "Boeing Promotional Fare Management System: Analysis and Research Findings", *op cit.*

13. *Aviation Week and Space Technology,* July 14, 1986.

14. R.W. Dodd, "An Assessment of the Effectiveness of the Restrictions on Low-Priced Air Fares", Research Branch Report 40-80-12, Canadian Transport Commission, Ottawa-Hull, September 1980, p. 34.

15. *ibid,* pp 101-125.

16. Joseph P. Schweiterman, "Fare is Fair in Airline Deregulation", *AEI Journal on Government and Society,* July-August 1985, pp 32-38.

17. *ibid,* p. 38.

18. R.W. Dodd, *op cit.*

19. Joseph P. Schweiterman, *op cit.*

20. R.W. Dodd, *op cit.*

21. Peter P. Belobaba, "TWA Reservations Analysis: Demand Distribution Patterns", Flight Transportation Laboratory, MIT, Cambridge MA, April 1985.

22. Fares available with no advance purchase for travel on Trans World Airlines, August 30, 1986.

23. Walter Nicholson, *Microeconomic Theory: Basic Principles and Extensions,* 2nd Edition, The Dryden Press, 1978, Chapter 3.

24. Moshe Ben-Akiva and Steven R. Lerman, *Travel Behavior: Theories, Models and Prediction Methods,* Chapter 5.

25. H. Simon, "A Behavioral Model of Rational Choice", in *Models of Man,* John Wiley Publishers, New York, 1957.

26. R.D. Luce and P. Suppes, "Preference, Utility, and Subjective Probability", in *Handbook of Mathematical Psychology,* Vol. 3, Wiley, New York, 1965, Chapter 19.

27. Marvin L. Manheim, *Fundamentals of Transportation Systems Analysis: Basic Concepts,* MIT Press, Cambridge MA, 1979.

28. K.J. Lancaster, "A New Approach to Consumer Theory", *Journal of Political Economy,* Vol. 84, 1966, pp 132-157.

29. Michael D. Meyer and Eric J. Miller, *Urban Transportation Planning: A Decision-Oriented Approach,* McGraw-Hill, New York, 1984, pp 257-258.

30. Barry C. Smith, "Reservation Recapture Rate Analysis", American Airlines, Presentation to AGIFORS Reservations and Traffic Control Study Group, New York City, April 14, 1986.

31. *ibid,* p. 2.

32. H.R. Thompson, "Statistical Problems in Airline Reservation Control", *Operational Research Quarterly,* Vol. 12, pp 167-185.

33. Marvin Rothstein, "OR and the Airline Overbooking Problem", *Operations Research,* Vol. 33, No. 2, March-April 1985, pp 237-248.

34. Marvin Rothstein, "Stochastic Models for Airline Booking Policies", Ph.D. Thesis in the Graduate School of Engineering and Science, New York University, 1968.

35. Robert Crandall, Chairman — American Airlines, quoted in "Life Under Deregulation The American Way", *Travel Weekly,* May 31, 1985.

36. *ibid.*

37. T.G. Plaskett, Senior Vice-President Marketing — American Airlines, Presentation to Shearson, Lehman Brothers Airline Industry Seminar, Key Largo FL, February 14, 1986.

38. "Airline Observer", *Aviation Week and Space Technology,* April 1, 1985, p. 35.

39. "Yield Managers Now Control Tactical Marketing", *Lloyd's Aviation Economist,* May 1985, pp12-13.

40. "The Optimal Partitioning of an Airplane's Seating Capacity", Sales Technology Dept., Boeing Commercial Airplane Company, Seattle, WA, February 1982.

41. Oli B.G. Madsen, "Booking Policy for Flights with Two Passenger Types", *European Journal of Operations Research,* forthcoming.

42. "The Optimal Partitioning of an Airplane's Seating Capacity", *op cit.*

43. Dennis F.X. Mathaisel and Herve Duchesne de Lamotte, "Optimization of Air Transport Yields and Capacity", Memorandum M83-1, Flight Transportation Laboratory, MIT, Cambridge MA, April 1983.

44. Dennis F.X. Mathaisel, "Optimization of Air Transport Yields and Capacity: A Tactical Model", Memorandum M83-3, Flight Transportation Laboratory, MIT, Cambridge MA, May 1983.

45. Fred Glover, Randy Glover, Joe Lorenzo and Claude McMillan, "The Passenger-Mix Problem in the Scheduled Airlines", *Interfaces,* Vol. 12, No. 3, June 1982, pp 73-79.

46. "The Optimal Partitioning of An Airplane's Seating Capacity", *op cit.*

47. Douglas Aircraft Company, "Electronic Bulkhead Seat Allocation", McDonnell-Douglas Corporation, Long Beach CA.

48. Ewalds D'Sylva, "O and D Seat Assignment to Maximize Expected Revenue", Sales and Technology Dept., Boeing Commercial Airplane Company, 1982.

49. Richard D. Wollmer, "An Airline Reservation Model for Opening and Closing Fare Classes", Unpublished Internal Report, McDonnell-Douglas Corporation, Long Beach CA, 1985.

50. Richard D. Wollmer, "A Hub-Spoke Seat Management Model", Douglas Aircraft Company Unpublished Report, McDonnell-Douglas Corporation, Long Beach CA, 1986.

51. Kenneth Littlewood,"Forecasting and Control of Passenger Bookings", AGIFORS Proceedings 1972, pp 95-117.

52. A.V. Bhatia and S.C. Parekh, "Optimal Allocation of Seats by Fare", Presentation by Trans World Airlines to AGIFORS Reservations Study Group, 1973.

53. Helmut Richter, "The Differential Revenue Method To Determine Optimal Seat Allotments By Fare Type", AGIFORS Symposium Proceedings 1982, pp 339-362.

54. Michael Mayer, "Seat Allocation, or A Simple Model of Seat Allocation Via Sophisticated Ones", AGIFORS Symposium Proceedings 1976, pp 103-125.

55. Bernard Titze and Raimund Griesshaber, "Realistic Passenger Booking Behavior and the Simple Low-Fare/High-Fare Seat Allotment Model", AGIFORS Symposium Proceedings 1983, pp 197-223.

56. Jorn Buhr, "Optimal Sales Limits for 2-Sector Flights", AGIFORS Symposium Proceedings 1982, pp 291-303.

57. Ken Wang, "Optimum Seat Allocation for Multi-Leg Flights With Multiple Fare Types", AGIFORS Symposium Proceedings 1983, pp 225-237.

58. "Boeing Promotional Fare Management System: Analysis and Research Findings", Sales Technique Department, Boeing Commercial Airplane Company, Seattle, WA July 1982.

59. J.J. Beckman, "Decision and Team Problems in Airline Reservations", *Econometrica,* Vol. 26, 1958, pp 134-145.

60. M. Rothstein, "An Airline Overbooking Model", *Transportation Science,* Vol. 5, 1971, pp 180-192.

61. "Boeing Promotional Fare Management System", *op cit.*

62. Peter P. Belobaba, "TWA Reservations Analysis: Demand Distribution Patterns", *op cit.*

63. Joao Sa, "Reservations Forecasting in Airline Yield Management", Master's Thesis in Flight Transportation, Massachusetts Institute of Technology, February 1987.

64. Tables of Z values of the cumulative normal distribution can be found in most statistical textbooks, including, A.M. Mood, *Introduction to the Theory of Statistics,* McGraw-Hill, 1950.

65. John M. Chambers, William S. Cleveland, Beat Kleiner, and Paul A. Tukey, *Graphical Methods for Data Analysis,* Chapter 6: Assessing Distributional Assumptions About Data, Wadsworth International, Belmont CA, 1983.