Massachusetts Institute of Technology

Flight Transportation Laboratory

Report FTL-R79-5

A SYSTEMS ANALYSIS OF SCHEDULED AIR TRANSPORTATION NETWORKS

William M. Swan

June 1979

FLIGHT TRANSPORTATION LABORATORY
REPORT R79-5

A SYSTEM ANALYSIS OF SCHEDULED AIR TRANSPORTATION NETWORKS

BY: WILLIAM M. SWAN

ABSTRACT

This work establishes the conditions for airline system design building from submodels of smaller aspects of air transportation. The first three sections develop submodels which then are combined in extensive numerical studies of singles market services. The final section discusses the changes to this problem that occur due to network effects.

The first section develops a simple model of the cost of providing scheduled transportation on a link. The cost of aircraft of various capacities are divided into a per-frequency cost and a per-capacity cost for conventional subsonic turbojet designs. This cost structure implies that the more capacity provided in conjunction with a fixed schedule of departures the lower the average cost per seat. It is suggested that such aircraft scale economies create a trend toward monopoly or at least oligopoly services.

The second section develops a model for demand. The market for transportation is argued to be the city pair. Demand for scheduled service is expressed in terms of fare, frequency and load factor. Fare, frequency, and load factor are combined into total perceived price for the service. This price depends on the consumer's personal value of time.

With only a few competitors in such a market, only a few of the technically possible qualities of service will be offered. The services available will be suited better to some tastes than to others. Distributional effects influence the politics of regulation and have been neglected in the past. In this light it is shown that competitive firms are likely to design their services for the same value of time. Product matching increases costs without improving the distribution of benefits.

Chapter 4 develops in detail the statistical model used to estimate denied boarding rates from long run design load factors. The development raises doubts about the viability of competition in this dimension.

Chapter 5 develops the optimal service for a single carrier on a single city pair market. Optima defined by maximum traffic at zero loss show the importance of the flexibility in aircraft capacity for long run system design. Both algebraic solutions and extensive numerical studies suggest that optimal designs depend on traffic and distance. Changes in frequency and capacity are large; load factor and fare are more stable. Optima are shallow for U.S. domestic cost structures.

The final section brings to the discussion issues associated with networks of services. Most U.S. domestic city pairs have amounts of traffic of only modest size compared to the efficient aircraft capacities. Networks overcome these limitations by sharing vehicles among markets. This is done at the expense of extra departure costs. The network design tradeoff in its simplest form is shown to be between larger aircraft capacities and longer stage lengths. The corresponding routing patterns emphasize stops and connections or direct flights. Network design adds another degree of flexibility to the design of transport services: the number of intermediate stops per passenger trip. This affects both cost and service quality.

LIST OF FIGURES

List of Figures (continued)

LIST OF TABLES

List of Tables (continued)

List of Tables (continued)

TABLE OF CONTENTS

Table of Contents (continued)

Table of Contents (continued)

1 Introduction

1.1 Systems Analysis

What is a systems analysis? In the title of this work the term has two implications. The first is that the discussion attempts to describe a complex subject in a way which will provide useful insights on a number of issues of relevance and concern. Our concern is not just for profits, not solely for efficiency, and not for strict legal justice, but for a realistic combination of these three conflicting demands. The second implication of the term systems analysis is that the work focuses on the combination of several detailed sub models involved in overall airline system design. For this work those sub models are: (1) airline cost prediction, (2) consumer demand description in terms of quality, value, and distribution of benefits, (3) the statistical process of matching supply and demand, (4) the nature of optimal single link, single market services, and (5) the complicating effects of networks of services. Each of these gets a chapter in this work.

1.2 Air Transportation Networks

This work focuses on U.S. domestic trunk airlines. We deal exclusively with U.S. domestic air passenger transportation, although the concepts carry over to intercity freight modes. We specialize thus for two reasons: because of the increased clarity such a concrete example provides and because data are more often available for illustration for U.S. airline transportation than any other mode or place.

Much discussion in the transportation field in general (Mohring [20]) and in air transport in particular (Douglas and Miller [10]) has focused on the economics of operations with the assumption that markets were viably competitive. Large airline markets are able to support multiple schedules of service. However, the ability of such competition to provide levels of service one would desire is called into question in this document. Furthermore, there is no doubt that there are markets too small to support more that one firm. This work suggests that most of the U.S. domestic markets are of only modest size in comparison with the efficient choice of equipment and schedule for a single carrier. However, the severity of this constraint is considerably relieved by the operation of airline networks. This condition, however, argues to the advantage of firms of at least medium size.

At the outset it will be useful to state an assumption and two conclusions which should be recognized as prejudices of the author, however well they are documented by the succeeding developments. The assumption is that aircraft costs vary with capacity in such a way that big aircraft are cheaper per seat. The first conclusion is that there is a tendency for airline service to be dominated by large firms operating big networks of services. The second conclusion is that this tendency is not overwhelming. The system design is so flexible that regulation which attempts to overrule natural tendencies by specifying fares is likely to fail at its objective.

1.3 Relationship to Previous Work

In a work as broad as this it is impossible to review at the outset the list of all previous work addressing the separate issues. Nor would it be appropriate to do so, since the thrust of this work is to explore the interactions of existing detailed understandings rather than to refine any particular sub-model. However, in the field of air transportation there have been several predecessors who have taken the larger view. It behooves us to establish the relevance of our work to their field.

There are two approaches to analysis of the air transportation industry. The traditional approach is to look at the existing regulated industry and draw conclusions about what characteristics are caused by interaction of regulation with production, profits, and the marketplace. These works are usually called industry analyses. The second approach is to examine the technical possibilities of production and enquire what natural proclivities need to be counteracted. These works can be called technical analyses. We will pay our respects to the traditonal literature first.

<u>Industry Analyses</u>

The earliest industry analysis is by Wheatcroft [29] in 1956. While much of this book is dedicated to analysis of political and pricing problems for air transport in Europe, the opening chapters represent a systematic attempt to establish the underlying cost and demand characteristics for air transport. The understanding of cost dependency on aircraft capacity is particularly good. These and other costs are developed from a fundamental knowledge of aircraft design and operations. Wheatcroft devotes much attention to whether big airlines are more efficient than small ones. Unfortunately he does this without reference to the nature of the airlines' markets. In contrast, our analysis will focus at the market level.

At nearly the same time another economist (Cherington [20], 1956) was writing a book on U.S. domestic airline pricing. Like Wheatcroft, Cherington devotes his opening chapters to transport cost. Cherington advanced the case by bringing network variables into consideration. Specifically, he used measures of airline size, route turnover, length of haul, length of trip, and station strength. This list is irregular in that some of the variables describe the geography of demands served by the airline while others characterize the decisions taken in network design. Nonetheless the introduction of technical network issues as part of a systematic cost analysis was an important step. The rest of Cherington's book discussed specific pricing activity. Except for his open acceptance of price discrimination, the relevance to our work today is small.

Taken together, the books by Wheatcroft and Cherington are as yet unequalled in the depth and subtlety of their understanding of the major technical issues of air transportation supply. Both represent industry analyses of the type which must become somewhat dated 20 years later. Nonetheless their supply chapters share with this work a focus on the structural characteristics of the use of airplanes to provide transportation. Both books take the broader view of investigating what is possible in the way of service, given the technical characteristics of supply. Because the laws of engineering are only slightly more mutable than the laws of mathematics and physics on which they are based, many of the understandings in Wheatcroft and Cherington remain relevant today. Unfortunately, neither author had the inclination to formalize his concepts into simple mathematical expressions or to relate them to a sophisticated treatment of demand. What we have is a series of tantalizing verbal hints displaying the author's understanding of the importance of technical matters.

Caves [2] began a tradition of more systematic economic analysis which continues in the academic publications of today. Caves focused particularly on the interaction of regulation and existing airline service. The obligatory early chapter on cost is based less on causal understanding and more on data analysis: "One must . . . use techniques of simple regression." (1)

Abandoning analytical cost construction in favor of statistical hypothesis testing is unfortunate. Due to the small sample size, the difficulty of adjusting for quality of service and network characteristics, and the high degree of collinearity among observed measures of airline activity, the attempt to discover technical cost characteristics by regression analysis has achieved only mild and dubious success.

Caves did introduce two important concepts concerning the market for scheduled air transport. First, Caves considered entry and competition at the city pair level. The second important issue of the market place is the problem of "product competition" or quality of service. For regulated airlines, Caves' focus is on high frequency, low load factor services. For our models of a non-regulated system there are possibilities for quality to be too low.

Since Caves there have been a number of economists who have contributed to the examination of air transport. Of particular note to us are two: Jordan [29] showed that the costs of operations dedicated to a single high density market may differ from more representative airline costs. Eads [31] argued that low density networks could not be economically served with medium capacity aircraft. This latter point has been partially contended by this author in a separate work [42].

The latest and best contribution by economists is a work by Douglas and Miller [10] on economic regulation of domestic air transport. This time the customary early chapter on cost characteristics discussed a few technical issues, mentioned without elucidation two network measures, and proceeded to accept

---

(1) Caves [2], p64. The phrase is somewhat out of context, but telling nonetheless.

as sufficient a regression analysis largely independent of the initial discussion. The regression showed no economies of airline firm size for U.S. domestic carriers.

The real contribution of Douglas and Miller is a far more sophisticated treatment of demand. Previous articles (Grenau [16] and DeVany [33]) had developed the usefulness of the concept of value of time in evaluating trip quality. Borrowing from Simpson [43] and Gordon and de Neufville [14] in the engineering field, Douglas and Miller collapsed both frequency and load factor into time indices and then converted to money value. The use of explicit numerical forms allowed Douglas and Miller to go beyond conversational treatment of the interaction of costs and quality. The focus on determining the optimal quality/price combination from technically possible services and the explicit statement of demand values is essentially that taken in chapter 5 of this work. The major difference being that Douglas and Miller ignore the possibility of changing more or less continuously among aircraft capacities. Discussion of the technical differences is reserved for the later sections.

Technical Analyses

The second possible approach to examining the air transportation industry is to construct cost and behavior patterns as they might exist in the absence of regulation. Such costs and market behavior must be developed from logical or engineering models because little experience exists. This work maintains that not only the aircraft costs but also network characteristics are important to the development of costs. In the field of airline network analysis, there have been several technical works which go beyond the economics of Cherington.

We may group the work by Gordon and de Neufville [14] and the thesis by Greig [18] together since they are both dedicated to examining the European air network with an eye to "rationalization." From our larger view, both works suffer somewhat from a limited definition of optimal network performance. It is difficult to collapse a network into a single performance indicator, as both these works try to do. Furthermore, the work on single link demand by Douglas and Miller and the corresponding chapter here suggest that such a simplification will highlight some issues while ignoring other ones which are at least as important. In addition, restricting the problem to the use of a single aircraft capacity is not realistic in the long run, however well it may apply to airline fleet averages in the shorter run. By overcoming these objections we hope to progress beyond the particular technical focus of these two excellent works. (1)

At the other extreme, the general discussion of regulation

---

(1) In this section the author is referring to an early working paper by Gordon and de Neufville. The present reference [14] and also [50] have broader discussions which reach conclusions similar to ours on the basis of their simpler statement of the problem.

and economics by Kahn [44] has several chapters of interest. Kahn shares with Douglas and Miller and with this work an interest in the quality of service. Unfortunately, Kahn's experience with regulating public utilities appears to have given him a tendency to think in terms of a market with regular and repetitive consumer decisions. The situation in transportation where consumers can alternate between competing services causes some additional complication in the issue of quality regulation and in the types of competition possible. In particular load factor for an airline is somewhat different from load factor for a utility. (1) Nonetheless it is Kahn's approach that is closest to ours in that Kahn chooses to deduce behavior from specific fundamental characteristics of cost and of the marketplace, To his more general discussion of regulated companies we bring the specific case of air transport. The move is not without relevance, since Professor Kahn was recently head of the CAB, which regulated airlines in the U.S.

1.4 The Present Work

This work explores the consequences of designing scheduled air transportation services taking into consideration the ability to choose from a range of aircraft capacities. The consequences of this degree of freedom permeate the entire analysis, from demand and load factor considerations to the type of service which is optimal for a link or on a network. Airline networks themselves are a means of taking advantage of cost savings associated with larger aircraft capacities. In a sense it is these vehicle economies which motivate the use of networks at all.

This work is unique in its focus on the option of changing aircraft capacities and on its consequences in network design. It goes beyond the studies which have fixed aircraft capacity at one value or in steps. It also goes beyond the excellent short paper by Anderson [36] by tracing the consequences in the distribution of benefits and in network design. On the other hand this work does not solve the problem of optimal network designs. Indeed it seems that the problem may not have any solutions of general applicability. The major contribution of this work, it must be hoped, is to frame the appropriate questions and include the relevant issues.

The work divides itself up into five sections. The first three sections develop submodels which are combined in extensive numerical studies in the fourth. The result is a more sophisticated statement of the common carrier transportation problem for a single link than has been employed in the past. The final section discusses the changes to this problem that occur due to network effects.

---

(1) For utilities load factor and utilization are nearly synonomous. For a transportation company t'e terms refer to two different measures.

The first section develops a simple model of the cost of providing scheduled transportation. In wholesale form the product of a firm offering scheduled transportation between two points is stated to be a scheduled frequency of service and capacity. This product is not really transportation since nobody need be moved anywhere by it, but it turns out to be extremely useful to develop the cost of a schedule as an intermediate good. The operating cost of aircraft of various capacities is shown to have a fixed portion and a portion which rises with capacity. This relationship is developed first by detailing the causes of the fixed cost (crew costs, structural scale economies, engine scale economies), second by reference to design work by Simpson and Moore [25], and finally by statistical illustration using U.S. domestic airline data. The importance of changing aircraft capacities as part of the design of a transportation service is established in this way. Aircraft cost per flight is divided into a per-frequency cost and a per-capacity cost for conventional subsonic turbojet designs.

This cost structure implies that the more capacity provided in conjunction with a fixed schedule of departures the lower the per seat average cost. It is suggested that such aircraft capacity economies create a trend toward monopoly or at least oligopoly services.

The second section develops a model for demand which is relevant to the supply discussion of the previous chapter. The market for transportation is argued to be the city pair. Reference to legal definition of monopoly markets suggests that products in the same market must be functional substitutes for each other. Travel by different modes or qualities of service in the same city pair are products in the same market; travel to other destinations is not.

Demand for scheduled service is expressed in terms of that service's fare and two other aspects of the technical performance: average displacement time for the schedule and the probability of capacity being available. The latter two measures are shown to depend on frequency and load factor, respectively. The total service--fare, frequency, and load factor-- is combined into a single total perceived price for the trip. This price depends on the consumer's personal value of time.

With only a few competitors in such a market, only a few of the technically possible qualities of service will be offered. General statements by Chamberlin [4] are found to apply to this case. The services available will be better suited to some tastes than to others. The distribution of benefits for a particular service quality or small set of service qualities will not be even across values of time. The tradeoff in design is between multiple service levels at higher cost and a single level of service at lower average cost. Distributional effects have a profound influence on the politics of transport regulation and have been neglected in the past.

In this light it is shown by an argument analogous to Hotelling's classical ice cream salesmen on a beach [17] that competitive firms are likely to design their services for the same value of time. Product matching increases costs without improving the distribution of benefits.

The next section develops in detail the matching transportation supply in seats with demand in passengers. For the first time in the literature, the statistical model which underlies the intuitive load factor/service discussions [10,14] is stated fully and explicitly. The most broadly useful statement of the problem defines the demand distribution for a randomly selected departure in a schedule. The variability of such a distribution is part due to known cycles in demand against the more regular schedule of departures and part due to random variations in demand. The cyclic components of variability are shown to have a somewhat Gaussian distribution for available airline data. Random variations are estimated from a theoretical model. The random component of demand variability is shown to scale imperfectly with average load side.

Estimates of variance for a Gaussian demand distribution allows calculation of the fundamental service index. This is the probability of a denied boarding due to variations of demand exceeding available capacity. A numerical approximation of this index is made for airline load factors.

Competition in load factor alone shows that airlines who match the perceived price of their services as shown in the demand chapter will also have a tendency to match their load factors (if they can by design). Since de Neufville [8] and Nason [21] have already suggested that airlines match frequencies, the matching of perceived price may lead to a matching of fare, frequency, and load factor-- in fact a complete duplication of services. The proviso is that the competing firms employ the same technology. This may not be the case if they have different networks.

The next section develops the optimal service for a single carrier on a single city pair market. The work parallels recent efforts by Douglas and Miller [10] on the same lines. The major addition is the new dimension of changing aircraft capacities. Including this degree of freedom in the analysis permits frequency and load factor to be adjusted independently in a market. Optima defined by maximum traffic at zero loss show the importance of the adaptability of design aircraft capacity and the relative stability of optimal load factors. Apparently variations in optimal load factors in previous analysis (DeVany [33]) and numerical work (Douglas and Miller [10], and Gordon and de Neufville [14]) were caused in part by the coupling of frequency and capacity forced by fixing aircraft capacity at a single value.

Both algebraic partial solutions and extensive numerical studies suggest that optimal designs depend on traffic and distance. Frequency and capacity changes are important; load factor and fare are more stable than previous studies showed.

Optima are quite shallow for U.S. domestic cost structures. Constraining frequency, fare, or load factor is not sufficient to force the traffic far from the optimal levels if the two remaining dimensions can be readjusted. This inherent flexibility in design, it is argued, makes difficult or ineffective regulation aimed at changing the fundamental market equilibrium by setting fares.

The final section brings to the discussion issues associated with networks of services. It is shown that most U.S. domestic city pairs generate modest amounts of traffic in comparison with efficient aircraft capacities. Networks overcome these limitations by sharing frequency cost among several markets. They do so at the expense of extra departure costs. The network design tradeoff in its simplest form is shown to be between larger average aircraft capacities and longer average stage lengths. The corresponding routing patterns emphasize stops and connections or direct flights.

It is shown by illustration that the average cost for a network cannot be estimated without knowledge of distributions of flow densities and link frequencies. Thus previous network studies have not gone far enough to be able to capture the entire dimensions of network behavior. Illustrations also show the importance of network effects even in the largest airline markets. Aircraft capacities and service frequencies in use today are a result of firms operating large scale networks of services and not a result of isolated city pair market considerations.

Network design adds yet another degree of flexibility to the design of transport services. The new dimension is the number of intermediate stops per passenger trip. More stops mean more load building, which allows larger aircraft and cheaper costs of frequency.

In conclusion it is argued that thin airline markets are too small to need regulation of service or entry because the fullest exploitation must occur just to recover costs. Dense markets can support isolated competition, although it may be somewhat wasteful of resources. Medium markets may support only one carrier if served in isolation. Unfortunately, flexibility in service quality makes price regulation an ineffective curb on firm behavior. However, medium and even thin markets can be competitively served as part of large airline networks.

The growth of air service, then, should tend to be from single quality single carrier service to competitive multistop network services to nonstop network service, and finally to a broad selection of service/fare options. There is some doubt whether regulation is either necessary or even useful in achieving these ends.

2 Airline Costs

2.0 Introduction

To say an airline produces available seat miles is like saying a newspaper company produces pages of print. The statement is true but it is not a terribly useful improvement over ignorance. To say an airline produces passenger miles is like saying a newspaper produces informed readers. The statement is nearer the truth, but it represents so specialized a model of the situation that only one singular insight is available. In this chapter we will develop a slightly more complex model of what an airline does. Then it will be possible to say how much these activities cost in a simple and broadly applicable way.

This chapter is divided into three main sections. The first structures the problem of airline costs and introduces a series of useful definitions. The second estimates the cost coefficients for the cost structure developed in the first section. The third major section compares these estimates, both as a structure and as values, to other work.

Throughout this chapter and the whole of this work we will stress the structure of the cost models. As long as the numbers are within 10% to 20% of the truth, we shall be satisfied; we are interested in the consequences of the cost structure and not its current values. It is beyond the scope of this work to establish greater accuracy or detail. In the longer run, relative cost coefficient changes of 20% can be expected. A useful conceptual model should not be overthrown by such changes.

This cost structure and the cost coefficients are meant to represent the long run cost characteristics which face U.S. domestic airlines. In a loose sense, the structure and values should carry over to regional airlines and international operations, but later network discussions may not. In the broadest possible sense, we are talking about regular route, point to point, common carrier services. Further, we are talking about regularly scheduled repetitive services. Private transportation, where one user rents the entire vehicle and determines its path of travel, is not included in our discussion. Thus charter aircraft, irregular route trucking, unit trains, and oil tankers are all part of another class of transportation. Airlines, LTL (Less Than Truckload) trucking, boxcar rail, urban rail, and ocean freighter services are all systems of the type we will be discussing. Here the discussion will be entirely about passenger carriage by U.S. trunk airlines, but the parallels with surface modes occasionally occur and in some cases their involvement in the problem is unavoidable.

2.1 A Structure for Airline Activities

<u>Some Definitions</u>

We will classify any countable measure of what airlines do as an airline activity. The two broadest categories of airline activities are aircraft movements and passenger movements.

A direct aircraft movement between a city pair is an aircraft stage. When we discuss just the movement without concern whether there are any passengers on board we refer to the city pair as a link. A series of consecutive aircraft stages is called a route. A set of links is called a network. The terms network and network design refer to aircraft movements unless they are specially qualified to mean otherwise.

A direct passenger movement between a city pair is a hop. A series of passenger hops make up a path. Passenger movements differ from aircraft movements in that they have a purposeful origin and destination. When we discuss origin- destination passenger movements without concern with the path, we refer to the movement as a trip.

Wholesale/Retail Distinction

By separating airlines' activities into aircraft movements and passenger movements we create a useful division of airline activities. The set of aircraft movements describes what might be thought of as the wholesale form of the airline product. The costs associated with wholesale outputs in this case are largely determined by technical considerations, which are dominated by the operating costs for the aircraft themselves. In the airline industry the aircraft costs are specifically called direct operating costs (DOC's). The process of converting wholesale aircraft movements into retail consumption is a separate step in airline production.

The production of aircraft movements as thus defined for airlines is a technical process with clearly structured costs. It is these costs and their structure which we will be exploring in this chapter and throughout this work. The cost of retailing this production is secondary in a conceptual sense because the structure of the retailing costs does not add to the design tradeoffs and problems established by the structure of the wholesale costs. The retailing process is still of great interest. We will focus on aspects of this process in chapters 3 and 4.

Engineering Cost Methodology

The essence of the engineering approach to cost construction used here is that physical activities can be counted and a cost applied to such measures. The art is to choose the activities so the cost coefficients are constant over a broad range of activity levels. This choice is guided by knowledge of the details of the activities and their physical relationships. The process of finding the activity measures which produce the simplest cost coefficients is analogous to finding the transformation of variables which best simplifies the formula for a geometrical shape.

Because the complexity is contained in the structure rather than in the cost coefficients, the cost functions here are different from cost functions established by statistical analysis of aggregated inputs and outputs. Statistical analyses of industry activities are almost always compromised in form to allow for calibration.

Measures of Wholesale Activities

We have generally described wholesale production by the activity of aircraft movements. Engineering cost functions for wholesale production will be simpler if we divide aircraft movements further into departure cycles and cruise miles. There is one departure cycle per stage. The number of cruise miles is defined as the great circle intercity distance for the stage in question. To reach a structure with constant cost coefficients, it is necessary to divide these categories each in two. The first part is a cost per vehicle, independent of capacity. The second part is a cost per seat. In making this distinction we have divided an aircraft conceptually into a vehicle part and a capacity part. We measure aircraft size or capacity in seats. We measure aircraft frequency in vehicle hops per link. We now have four measures of aircraft activity:

number of vehicle departure cycles (vehicle departures)
number of vehicle cruise miles (vehicle miles)
number of seat departure cycles (seat departures)
number of seat cruise miles (seat miles)

Total activities in each of these categories for any airline network will be sufficient to determine the cost of wholesale production. It will not matter whether the wholesale product is consumed for us to determine its cost. We will estimate the cost coefficients in section 2.2.

Measures of Retail Activities

We have generally described the retail part of production by the activity measure of passenger movements. Engineering cost functions for retail production will be simplest if we divide the passenger movements further into passenger trips, passenger miles, and passenger departures. Passenger miles are the total great circle distance of all the hops in the path used for the trip. Passenger departures are equal to the number of hops used for the trip. (1) We now have three measures of passenger activity:

passenger trips
passenger miles
passenger departures

Total activities in these three categories will be sufficient to determine the retail cost of airline activities. The cost coefficients for these three measures are developed from a combination of our work and previous authors' in appendix A. Because the structure of these costs tends to reinforce the structure already established by the dominant wholesale costs, our discussion of retail costs is less thorough than the

---

(1) We do not make a distinction between a connection and a through flight for consecutive hops. To do so is the first improvement we would suggest in our work.

discussion of wholesale costs.

Service Measures

At the reported, accounting level airline activities are divided into aircraft hours which generate direct costs (DOC's) and the rest of airline activities which cause indirect operating costs (IOC's). At our engineering level of wholesale and retail production activities we redivided activities into passenger, vehicle, and seat departures and miles, and also passenger trips. We now introduce a third frame of reference for looking at airline activities, the schedule.

The airline schedule describes the airline wholesale products (aircraft movements) in the form in which they are presented to the public by the retailing process.

A schedule is a set of aircraft movements useful for trips in a market. For nonstop service, the schedule is the link frequency and the seats available. We implicitly assume the timing of vehicle departures and the distribution of seats among those depatures are reasonable. For multistop service, the schedule includes not only nonstop flights but also any multistop flights or connections which are useful paths. The schedule frequency will be generally considered in terms of nonstop link frequency or its equivalent in convenience. (Equivalence will be obtained by compensation in price or in higher frequencies. These issues will be discussed when they are needed in the network chapter.)

The schedule is as far as we can go in describing airline service from the airline side alone. The quality of service depends on the traffic because average load factor is important. The price of service also depends on the interactions between the airline and demand.

This completes our discussion of airline cost structure. We now move on to the matter of calculating the cost coefficients for our engineering cost frame of reference.

2.2 Estimates of Airline Cost Coefficients

Conversion from Reported Cost Measures to Engineering Cost Coefficients

Accounting terminology divides airline costs into direct (DOC) and indirect (IOC). DOC is the cost of aircraft hours and it comprises 60% of airline costs. It is this cost which gives the structure to wholesale activities. Indirect costs apply against both wholesale and retail activity measures. We view indirect costs as a secondary addition to direct costs.

DOC is reported on the basis of block hours of aircraft use. We will use the term DOC in the sense of aircraft operating cost per hour, unless we explicitly state otherwise. Block hours correspond closely to hours with the engine on, so block hours include both time spent in cruise miles and time spent in departure cycles. We will convert DOC per hour to a per aircraft departure and per aircraft mile basis using cruise speeds and departure times. Then we will subdivide the per aircraft numbers into per vehicle and per seat coefficients. Finally we will add indirect costs to arrive at the final coefficients for wholesale activity measures of vehicle and seat departures and miles. The cost coefficients for the retail activities of passenger miles, departures, and hops will also be determined from IOC's.

The cost of the activity measure aircraft block hours is manipulated in the next four sections to obtain the major fraction of the cost coefficients for aircraft movements. In developing values for these coefficients we must demonstrate that the structure of these costs is valid. That is, we must show that each coefficient is significant and that it is reasonably constant over a range of activity levels. We begin the discussion by covering several influences which might affect the tendency of the cost coefficients to be constant.

DOC's for an Aircraft Type

There are a number of issues which can be best disposed of by discussion of the costs of a fleet of aircraft all of the same capacity and design. We refer to the DOC of such a fleet as the DOC of the aircraft type. The issues which can be usefully discussed in this context are those influences on DOC which we do not include in our engineering activity measures, i.e. issues of what aircraft costs do not depend on in any significant way.

DOC for an aircraft type is not stongly dependent on how many aircraft of the type are operated in the airline's fleet. Jordan [28] in particular has investigated the DOC for different fleet sizes and reached this conclusion. The size of the fleet or analogously the size of the firm has been a variable of interest for economists seeking to discover economies of scale for the industry. We shall try to show in later sections that economies of scale when phrased in this manner may not be relevant to issues of social concern.

Also, aircraft DOC is nearly independent of where the aircraft is operated. Aircraft are machines and they work in very similar fashion in different geographical places. Independence of costs from geographical variation is a particularly useful simplification for network discussions.

One sign that costs are independent of where the aircraft is operated is the existence of markets for hours of aircraft use. Airlines rent aircraft by the (block) hour to each other and to charter operators. In this way aircraft hours have come to be viewed as a particularly useful intermediate measure in the process of manufacturing transportation services. DOC's reported for the same aircraft type by different firms provide one of the few relevant comparisons of interfirm performance. Restrictions necessary for a valid comparison are developed later in this chapter. For now the important thing is that aircraft cost should be independent of locale.

Another item that is absent from the list of factors having a significant impact on DOC is load. Capacity is a dominant characteristic for determining aircraft costs; how much of that capacity is full is not. Physical laws allow us to put some bounds on the accuracy of this statement. The major load related cost is the expenditure of energy to carry more weight. Lift associated drag for aircraft rises linearly with aircraft weight. The effect of this on operating expenses can be estimated. Fuel costs for airlines are 30% of DOC. Lift absorbs roughly half of this energy in cruise, and payload is less than a quarter of total weight. Combining these estimates, there can be at most a 3.8% difference in DOC between operating dead empty and chock full. The difference as a percent of total costs is even less.

There is one final and most important independence for DOC's. The block hours against which DOC's are charged include both cruise and departure time. If DOC were substantially different in cruise than during terminal area maneuvering, there would be different DOC's per hour for stages of different length. This would not change the validity of our engineering cost structure, but it would make the coefficients harder to estimate from available DOC data. Fortunately, DOC does not vary significantly with stage.

It is possible to illustrate a lack of dependence of DOC on stage length. Figure 2.2.1 shows a scatter of block hour costs plotted against average stage length for the most common aircraft type. A least squares linear regression of these data points produces an hourly cost formula with little dependence on stage length. (R^2=.14)

The ATA formula provides further support for the concept of a departure cost proportional to the cost of the same amount of time spent in cruise. This formula [1] is the industry standard for comparing new aircraft designs' average per mile costs at different stages. Within the formula the fundamental measure is hourly costs. Average cost per mile is found by multiplying stage time by DOC (per hour) and dividing by stage length. The only significant correction for departure cycles is the 21 minute addition to stage time for air and ground maneuvers.

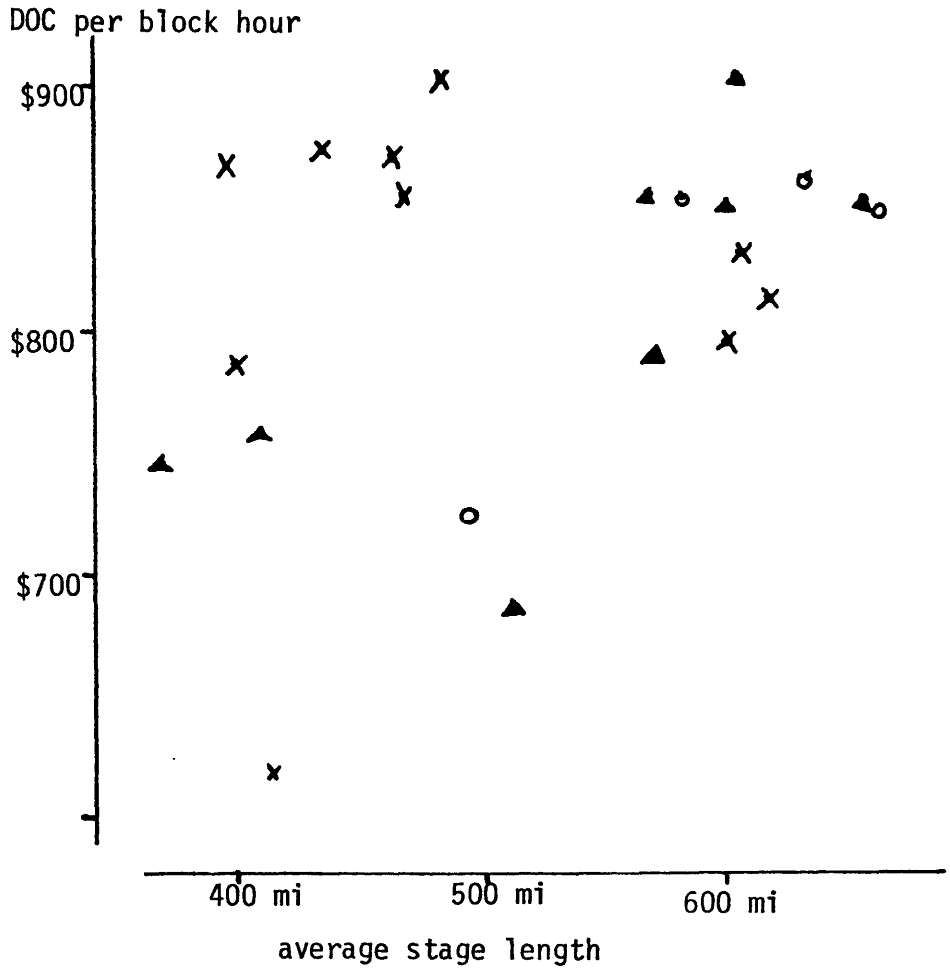
Having stated that we wish to think of aircraft hourly costs as only marginally dependent on where the aircraft is operated, how far it is going, or how full it is, we can now address the issue of what aircraft costs do depend on.

DOC's for a Family of Aircraft

The next sections discuss what aircraft as a class of vehicles cost to operate. The structure of aircraft costs has not changed with the change from piston prop aircraft to jet aircraft to turbofan designs. The same structure applies to new designs, such as STOL fixed wing and VTOL rotary wing designs [32]. Families of air transport vehicles of similar design performance (such as runway requirements, speed, range, and seating density) display hourly costs which are nearly linear with vehicle capacity. The same statement is true of trucks [37] and trains [25], within the bounds of their operations.

U.S. domestic airlines all use aircraft from one design family. Currently these aircraft are pressurized turbofan fixed wing vehicles with cruise altitudes near 30,000 feet, cruise speeds near mach 0.8, and runway requirements of a mile or so. DOC for these aircraft varies with their capacity. At present the available capacities range between 70 and 500 seats. At any given point in time only a handful of aircraft types with

Figure 2.2.1: DOC vs Stage Length for 727 Aircraft
U.S. Domestic Airline Fleet Averages, by airline



Data on 727-200 (x), 727-100 (▲), and 727-100C/QC (o) aircraft reported operating cost less depreciation and rentals, all airlines
Source: CAB <u>Aircraft Operating Cost and Performance Report</u>, Vol X, 1976.

different capacities are actually represented, but in a long run technical sense, the entire continuum is relevant. The single characteristic on which all later discussion rests is that aircraft DOC's display economies of scale with respect to capacity. That is, larger aircraft cost less per seat. The very existence of larger aircraft would seem to prove this point, since smaller aircraft are more convenient from every other aspect. Wherever there is a combining of loads onto a public transportation system, it seems safe to assume that there is some cost advantage in doing so. However, it will not do to rely on such logical arguments when observed data is available to illustrate the point.

Figure 2.2.2 plots hourly costs for four aircraft of identical design in terms of range, field length, seating density, and technical state of the art. These designs were performed by Simpson and Moore [26]. The line in figure 2.2.2 illustrates two significant points. First, there is a large cost per vehicle (the intercept). And second, the cost per seat (the slope) is constant over a broad range.

Observing such lines in real world data is difficult. Although airline aircraft are all from a similar family, there are small differences in technical age, (1) range, required runway lengths, and seating density. These differences average out among the family in the long run, but they distort the few data points we have. Furthermore some of the variations which are relevant from technical considerations have not been exercised in practice and some have. For instance there are few very large short range designs and few small long range designs.

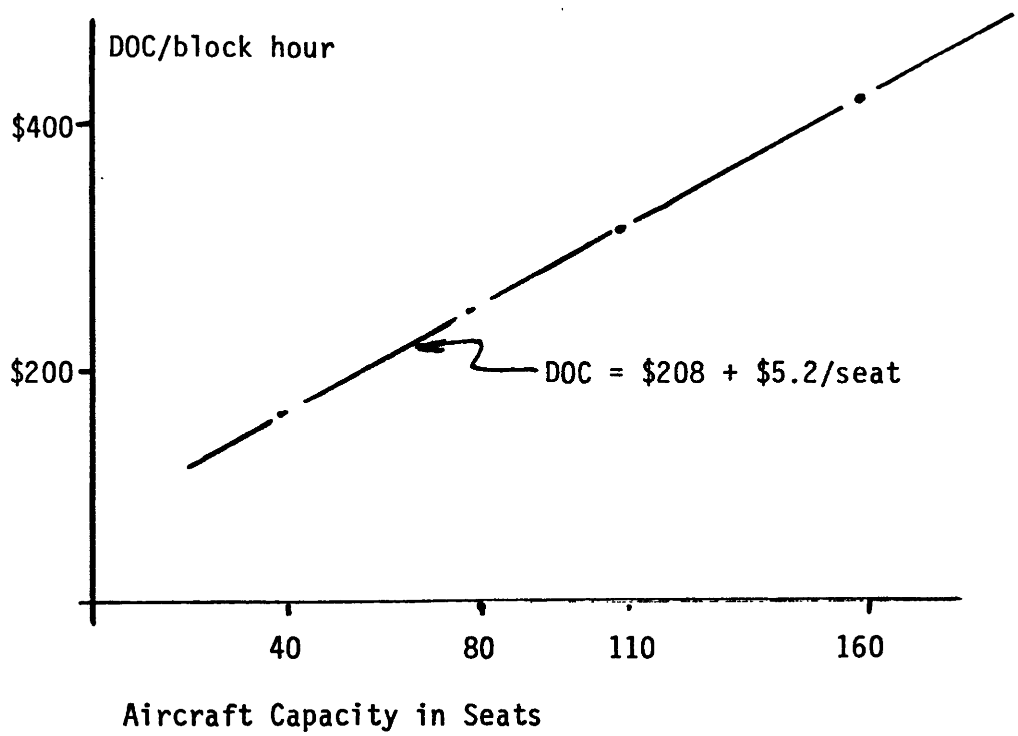
In spite of these difficulties, we can illustrate the design trend using current data if we make some adjustments. We start with the list of aircraft in table 2.2.1. This is every aircraft type with over 20 vehicles in the U.S. domestic fleet in 1976. Where airline accounting practice allocates costs arbitrarily between model variations, we have used a weighted average. Thus the DC-9 is a composite of the -10 and the -30 variants heavily favoring the -30 due to its greater numbers. We make only one adjustment to these numbers. The newer of these types were designed to operate at slightly higher cruise speeds than earlier designs. We are fortunate in having a very careful analysis by Sercer [24] to use in estimating the practical cruise speeds for these aircraft. The estimates of departure times by Sercer are open to a number of questions of comparability of detail, but the cruise speed estimates should be very good. We use these cruise speeds to adjust the block hour costs for the aircraft types to costs for a set of aircraft which were designed to cruise at the same speed. The speed we use for the standard is 507 mph, so the adjustment in block hour costs is to multiply the reported costs by 507 and divide by the cruise speed deduced by Sercer.

The result of our study of adjusted DOC's is presented in

---

(1) The age of an aircraft is really the year of its design or first manufacture. Advances in engineering and materials continually improve the state of the art in commercial aircraft.

Figure 2.2.2: DOC vs. Capacity for Comparable Designs



Original Designs by Wesley Moore at MIT, 1972 (ref 26)

1970 technology
Costs converted to 1976 dollars
design range is 2000 miles
DOC by 1967 ATA formula, modified

Table 2.2.1: Data for Aircraft Operating Costs

Aircraft type	Capacity(3)	Speed(4)	Adjusted Block Hour Costs (5)
DC-9(1)	87	463	$852
737-200	97	450	1226
727(2)	115	504	1049
707-100	132	504	1197
DC-8-50	135	510	1338
DC-8-51	194	508	1562
DC-10-10	237	510	1891
L-1011	247	526	2205
B747	357	534	2773

(1) The DC-9 is a weighted average of the -10 and -30

(2) The 727 is a weighted average of the -100 and -200

(3) Capacity from CAB ref 8 (Seats)

(4) Speed in practice from Sercer ref 24 (mph)

(5) Adjusted Block hour costs are hourly costs from CAB ref 8 multiplied by (507mph/speed)

figure 2.2.3. A nearly straight line appears to fit the data and the intercept is still positive and significant.

The reasons for the economies of scale in aircraft capacity are technical, which is why they have persisted over the ages. To quote Wheatcroft in 1956 speaking of the first generation of commercial transports:

"(a) Larger scale gives aerodynamic advantages and lower proportionate drag.
(b) The structural weight is a lower percentage of the total weight. [due to changes in proportions.]
(c) The operating weight (crew, radio, navigational equipment, etc.) is a smaller percentage of total weight.
(d) Larger engines . . . are more efficient . . . per pound . . . and per gallon.
(e) The costs of aircrew salaries and expenses remains more or less fixed and becomes a smaller percentage of the total costs."

---Wheatcroft [29], p33.

The important observation from figure 2.2.3 is that the cost per seat mile is a constant (the slope) and the cost per vehicle departure is about 50 times that constant (the intercept).

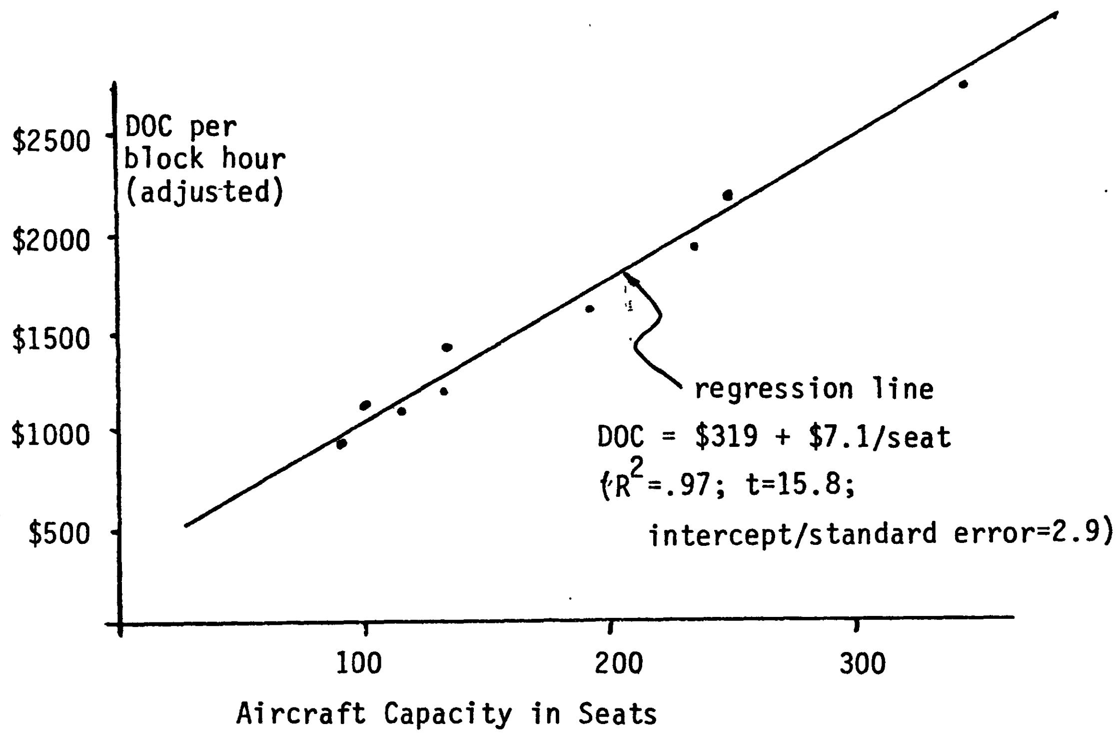
<u>Adjustment of Costs for Return on Investment</u>

Depreciation has been included in the DOC figures of the preceding section. Depreciation is only part of the cost of owning aircraft. An aircraft is a physical plant with capital value. Even idle, it incurs expense due to the cost of capital, the rate of technical obsolescence, the risks accruing to ownership, and some mechanical deterioriation. These ownership expenses depend on clock time, not use. Most of this expense is reported on the books as depreciation or as accounting profits. In this section we estimate the total ownership costs including both depreciation and return to capital. We then remove the depreciation expense from the DOC figures and replace it with our newly calculated ownership costs. This is merely a technical adjustment of costs and does not affect the main flow of our developments.

Ownership costs actually depend on clock time, not block hours. However, in transportation industries depreciation is almost always reported against hours or miles of use. For ground modes most of depreciation is wear, which is truly a per mile cost. The situation is different for airplanes. Up to the present day, changes in the value of airplanes have been caused by their design growing out of date rather than by their mechanical life. Thus aircraft depreciation depends on clock time, not block time, and is therefore an ownership cost. Airplane designs may change more slowly in the future, but at the present aircraft depreciation like return to capital is a daily ownership expense, independent of use.

Capital costs are proportional to market value. Market value is usually estimated from original price and depreciation. Both original cost and the depreciation period depend on the age

Figure 2.2.3: DOC vs. Capacity, Reported Figures



DOC points are U.S. domestic fleet averages for all 9 aircraft types with more than 20 in the fleet. (ref 8, 1976)
DOC adjusted to constant cruise speed, see text.

of the design. (See table 2.2.2.) Older designs are worth less (new or used) and depreciate more quickly. The original price estimates of figure 2.2.4 give an indication of long term trends of price for designs which are neither the oldest nor the newest. Applying a 14 year mortgage at 8% (1) to these figures produces aircraft ownership expense of $964 per vehicle per day plus $13.93 per seat per day. Aircraft utilization is roughly 8 hours per day, which gives an ownership cost of $120.5 per vehicle per block hour plus $1.75 per seat per block hour. For a 125 seat aircraft, this becomes $339 per hour. (2)

We may now substitute these values for the depreciation fraction of the DOC values of figure 2.2.3. (This depreciation was 15% of DOC.) The new total DOC per hour becomes $392 per vehicle and $7.79 per seat.

A better estimate of ownership costs is difficult and not germane to the issues at hand. The important points are that daily costs do exist, that they depend on aircraft capacity, and that they are usually allocated against hours of use by assuming a use pattern. (3)

<u>Discussion of Indirect Costs</u>

One is on firm ground characterizing the cost of vehicles (DOC) because they are machines subject to physical laws. For IOC the case is more difficult. The laws of performance which apply to human labor in a service industry are not so well defined. A great deal of non-specialized service labor goes into transportation-- loading and unloading, sorting, dispatching, scheduling, selling, and managing to name a few categories. Most of these activities do not involve much physical plant, so the economies of scale typical of industrial production are reached at very modest levels of activity. That is, the process of boarding 30 aircraft loads is pretty much a tenfold repetition of the process for 3.

Appendix A compares several assessments of airline IOC's including one done by this author several years ago. We have updated and selected from among the several values, but we admit the treatment is crude. Nonetheless, casual conversations with airline staffs over the years have tended to bear out these

---

(1) 8% is chosen to provide a 4% return after taxes. All figures are in constant dollars. Aircraft appreciate on the books due to inflation (that is they sell for more than their depreciated value), so some extra cost might be added to pay taxes on these apparent capital gains.

(2) This is on the order of twice normally reported depreciation figures. Such an estimate is supported by aircraft lease costs reported on CAB form 41.

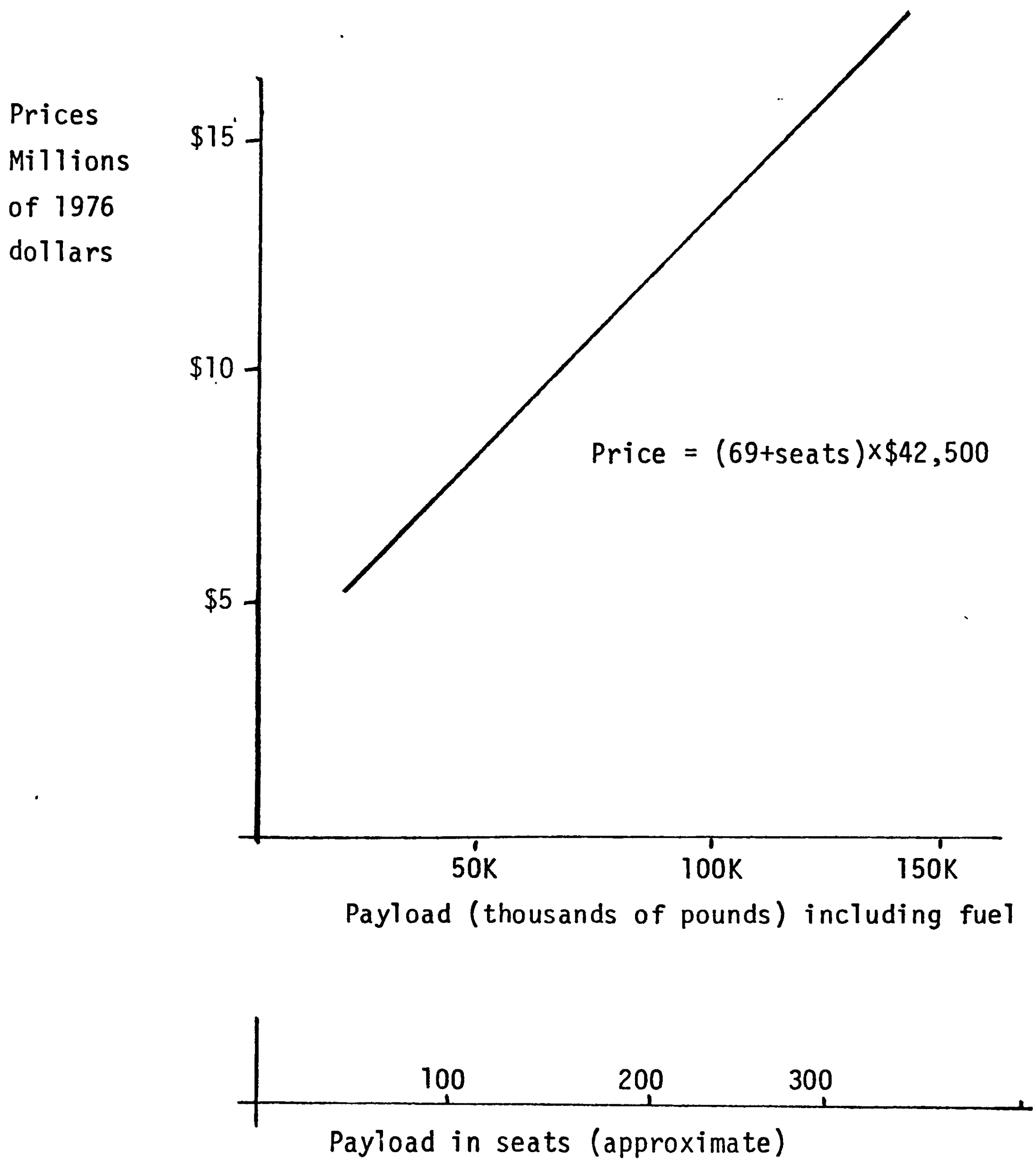
(3) Ownership costs should not be allocated against use if the amount of demand peaking or the average stage length is expected to change much during the analysis. This matter is explored in section 2.3.

Table 2.2.2: Depreciation Rates Accepted by the CAB

class of aircraft	Example	approximate design year	depreciation rate	life (years)
turboprop (old designs)	FH227	mixed	8.5-9.5%	10.5-12
turbojet (first jets)	DC-8	1958	7.9-9.5%	10.5-12.5
turbofan (modern jets)	B-727	1965	7.0%	14
wide body (latest designs)	L-1011	1972	5.6%	18

Source: CAB DPFI, ref 5.

Figure 2.2.4: Aircraft Prices



Source: Simpson, MIT NASA Workshop R72-7, ref. 26
1972 dollars converted to 1976
operating weight empty converted to payload or seats
Original data: Prices from Flight Magazine, 20 April 1972
Weights from Jane's All the Worlds' Aircraft
Seats from CAB Aircraft Operating Cost and Performance, ref. 8 and 39.

estimates, so we have more confidence in them than the methodology alone inspires.

Cost Coefficients for Aircraft Miles

With the above DOC discussion in hand, estimation of costs per cruise mile becomes straightforward. DOC per aircraft hour has been found to be

$$\text{DOC/hr} = \$7.79/\text{seat} + \$392/\text{vehicle} \qquad (2.2.1)$$

These figures have been adjusted to represent a family of aircraft with a cruise speed of 507 mph. Dividing by this speed we get

$$\$/\text{mi} = \$.015/\text{seat} + \$.77/\text{vehicle} \qquad (2.2.2)$$

To these must be added indirect costs from appendix A. These costs are an overhead of 19% on all but the ownership expense. We introduce three variable names:

c1 = cost per aircraft mile
c11 = cost per seat mile
c12 = cost per vehicle mile

The last two of these are our engineering cost coefficients. The cost per aircraft mile becomes

$$c1 = c11 \text{ SEATS} + c21$$

or

$$c1 = \$.0176 \text{ SEATS} + \$.816 \qquad (2.2.3)$$

Cost Coefficients for Aircraft Departures

The cost for a departure depends on the amount of the time the departure cycle takes. For uncongested airports, these cycle times are determined by local ground and air speed limits and traffic distances which do not vary for different aircraft types. This throws the variations in Sercer's data (presented in table 2.2.3) into question. However the average time, 22 minutes, is very close to that used by the ATA [1]. We use this time to calculate ddeparture cycle costs for aircraft. Thus 22/60ths of a block hour's cost gives a departure cost of

$$\$/\text{departure}=\$2.84/\text{seat}+\$144/\text{vehicle} \qquad (2.2.4)$$

To these must be added $183 per vehicle and a 19% overhead from the indirect costs developed in appendix A. (The overhead does not apply against the ownership cost fraction of the figures in equation (2.2.4).) We introduce three more variables:

c0 = cost per aircraft departure
c10 = cost per seat departure

Table 2.2.3. Departure Cycle Times for Aircraft

Aircraft type	
DC-9 (avg)	16 min.
737-200	16 min.
727 (avg)	20 min.
L1011	25 min.
DC-10-10	25 min.
DC-8-61	22 min.
DC-8-50	21 min.
707-100	23 min.
747	32 min.
Average	22 min.

Source: Sercer, ref 24.

c20 = cost per vehicle departure

The last two of these are two more of the engineering cost coefficients. The cost per aircraft departure becomes

$$c0 = c10\ SEATS + c20$$

or

$$c0 = \$3.27\ SEATS + \$378.9 \qquad (2.2.5)$$

As one might expect, there are considerably more indirect costs involved in an aircraft departure than a cruise mile.

Cost Coefficients for Passenger Movements

The indirect cost estimates of appendix A provide a cost per passenger boarding and a cost per passenger mile which come to $12.64/boarding and $.008/mile including overheads. Our engineering cost structure does not include the term passenger boarding. A passenger boarding is one passenger trip plus one passenger hop. We make a division of boarding costs between the two parts. Our estimate for the hop cost reflects the passenger time involved. The remainder of boarding cost is the per trip cost, which is the cost of initial boarding and ticketing. The division we employ is entirely arbitrary, but it appears reasonable. We get for the three retail cost coefficients

c30 = $6.64 per passenger hop
c32 = $6.00 per passenger boarding
c31 = $.008 per passenger mile

These three coefficients complete the list of engineering cost coefficients and bring to an end this calibration section.

2.3 Discussion of Costs and Cost Structure

Cost of a Schedule and Schedule Competition

Section 2.1 defined a schedule for a nonstop market as the link freuency and the seats available. We can now state that for a market with intercity (great circle) distance D and a schedule with frequency FQ and capacity SEATS the cost is

$$COST=(\$3.27+\$.0176\cdot D)\cdot SEATS+(\$379+\$.816\cdot D)\cdot FQ \qquad (2.3.1)$$

This formula may be approximated by a simpler one, which is useful as a rule of thumb although it is not accurate enough for cost calculations:

$$COST/FQ \cong (D+200)\cdot(CAP+50)\cdot\$.0176 \qquad (2.3.2)$$

This formula gives the per stage cost as a function of aircraft capacity (CAP). (1) This rule of thumb formula will prove a useful way to think of costs when it comes to matters of network design.

Notice that both (2.3.1) and (2.3.2) imply that the costs of capacity (SEATS) and of frequency (FQ) for a schedule are mathematically separable. This occurs because we have considered a range of aircraft capacities (CAP). In the long run, airlines can alter the average aircraft capacity for their fleet and they can also change the distribution of capacities. In the short run, airlines cannot alter the average capacity although they can rearrange the capacities used for particular movements. Global changes in aircraft capacities in use are part of long run cost options.

The degree of freedom of changing fleet capacities is often ignored in analyses of the airline cost function, with misleading results. The consequences of this problem are discussed in chapter 5.

This form of the cost function shows that for a fixed schedule (fixed FQ) the larger the capacity (SEATS) provided the lower the average cost per seat. Because of economies of larger aircraft, competition among many airlines is difficult. If each carrier offers its own independent schedule, there is duplication of frequency costs at great expense. (2) Thus the minimum cost service on a link is with only one carrier, and there is little likelihood of more than a few carriers serving any one isolated market.

Of course we really can not reach any conclusions about competition from the discussion as it has progressed so far. The question of matching supply in seats to demand and the process of competition in the dimensions of frequency and load factor will be the focus of the next chapters. However, the cost discussion so far does allow us to focus on service by a severely limited number of firms and very likely by a single carrier.

Cruise Speed

The cost derivation above is far from formal or rigorous. In a sense the particular form of the cost calibration, a linear dependence on capacity with a sizable positive intercept at zero, is an assumption of our work and not a result. The assumption is that the technical trends do manifest themselves in practice. The discussion above is intended as an approximate calibration or perhaps just a "sizing" of that assumption. Still it is of value to discuss the differences between our numbers and those produced by other analyses.

The most outstanding disparity between the numbers above and conventional values is in aircraft cruise speed. Aircraft design cruise speeds are in the neighborhood of 585 mph, not 507 mph as

---

(1) Notice that SEATS is the total schedule capacity while CAP is the per aircraft capacity.

(2) Appendix C points out that collusive scheduling is rare in present practice.

above. There are two reasons why the cruise speeds Sercer obtained from airline schedules differ from theory. The first is circuity. The speed we use, 507 mph, is based on travel time vs. great circle distance between the two points. Aircraft do not travel such straight lines, there being between 3% and 10% circuity as a matter of plan. (1) Thus the miles per hour measured in great circle miles is lower than that measured along the aircraft track.

In addition to circuity, airplanes suffer from winds. Winds average out on round trips, but aircraft spend more time flying against them than with them, so average speed is reduced.

Departure Time

The other figure which differs from numbers this author has found prevalant in industry is the departure cycle time. We use 21 minutes; airlines will tend to use 30. We are content to use Sercer's value, obtained from regression of airline schedules. This produces a lower departure cycle cost for our study than might be usual. It turns out that the estimate of departure costs on the indirect side which we use from appendix A is higher than most. Consequently our final cost figures will not be far from the mark.

In general there is little controversy about the existence of economies of aircraft capacity. The general difficulty has been that the consequences of this type of cost dependence have not been explored. This will be the major purpose of our work.

Singularity of 727 Costs

There has been some confusion about the economies of aircraft capacity due to an anomaly in costs which has developed in the 1970's. The anomaly is that the 727-200 aircraft, which seats 125, compares favorably in reported seat-hour costs to the L-1011 and 747, which seat 250 and 360 respectively.

The reasons for this are several. First and foremost, the 727 has smaller seats and higher seating density. It also has less galley space and less cargo ability. Second, the 727 is a medium range design, which makes it cheaper per seat than long range designs with the same capacity. Third, the 747 and L-1011 bear the costs of a substantial environmental constraint on noise, while most of the 727 fleet does not. Fourth, the 727 is a uniquely high production aircraft and is uniquly cheap as a consequence. And fifth, reported depreciation for 727 fleets is very low because some aircraft have been fully depreciated and others are valued in uninflated dollars. None of these advantages is overwhelming, but taken together they add up so that unadjusted hourly costs for the 727 are exceptionally low. This need not concern us here since our discussion is concerned with long term trends.

---

(1) Circuity for trucks and rail can be 20% and 40% respectively, which shows the necessity of converting to great circle distance before making comparisons.

Block Hour Utilization

Daily ownership costs per aircraft were discussed in section 2.2. In the course of that discussion the point was made that such costs are traditionally allocated against hours of aircraft use. Since aircraft operating costs per hour for aircraft of the same type or even of the same capacity can be used to compare firms, it is important to know what assumptions about ownership costs are necessary for such comparisons. (1)

First of all, "utilization" in the vocabulary of the transportation industry does not include time spent loading or unloading the aircraft. "Utilization" refers to hours with the engine on. Let us assume as in figure 2.3.1 that an aircraft can be in use up to 14 hours a day. Ten night hours have little or no traffic and are reserved for maintenance. With normal loading times the maximum possible "utilization" can range from 7 to 12 hours for practical differences in average stage length. Apparently for a valid comparison of DOC's, stage lengths must be similar.

Stage length is not the only influence. The amount of traffic and traffic peaking affects actual aircraft "utilization". There is generally insufficient traffic to allow use of an aircraft (including loading) for more than 14 to 16 hours a day; few people travel at night. Further, use at full intensity may not be the best economic decision at off peak hours in the middle of the day. Therefore, airlines will not seek to achieve the same fleet utilization unless they face similar hour to hour traffic patterns. Thus the actual utilization plotted as points on figure 2.3.1 may all be efficient use of resources, even though they inconsistently fall short of the maximum established either by our formula or the ATA's [1]. (2)

From these considerations it should be clear that the conditions for interfirm comparison of aircraft expenses are similar stage lengths and similar traffic peaking patterns. (3) These issues are specifically excluded from our considerations. We assume all time of day issues have been solved as a separate part of the design process.

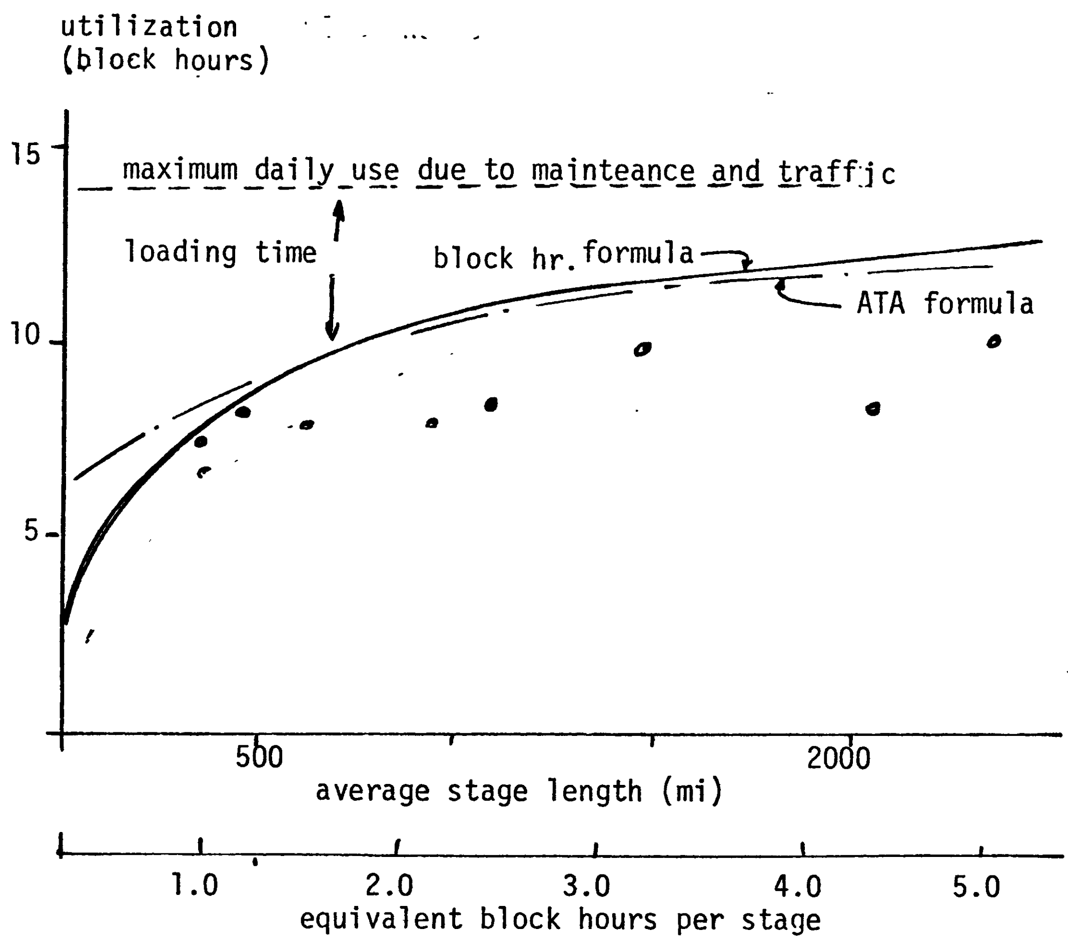
---

(1) Ownership costs cannot be eliminated before comparison because some firms may use low price and high operating cost aircraft. Their costs for aircraft stages of the same capacity should be comparable with cost for new, expensive but efficient aircraft.

(2) In the summer of 1978 United Airlines increased the utilization of parts of its fleet by 20% in one single network schedule revision. This accompanied a change in discount travel fares and activity. Thus off peak idleness was avoidable.

(3) Comparing firms with similar utilization and different peaking patterns accepts the different degrees of response to the peaking as both efficient.

Figure 2.3.1 Aircraft Utilization



Block hour formula from considerations of a 14 hour day and a 40-min. loading time:

Block Hrs = (STAGE/507)/(1.0+STAGE/507) x 14

Data points (•) are selected industry reported figures for all aircraft of one type in the airline's fleet

Log-Linear Cost Stuctures

The discipline of economics as practiced in modern times emphasizes cost formulas which can be calibrated using linear least squares regressions and which reveal the cost or scale elasticities in the calibration parameters. (1) This trend has produced formulas for aircraft DOC like:

$$\$/ASM \sim STAGE^{-.07} \ CAP^{-.16} \qquad (2.3.3)$$

Where ASM are airline system wide available seat miles. STAGE is fleet average stage length, and CAP is fleet average aircraft capacity. Exponents are from Douglas and Miller [10], p 16. The negative exponents in this formula indicate economies of scale with respect to aircraft capacity and stage length.

In order to see if the log-linear form can describe our cost structure, we have calibrated a similar expression using data points generated by exercising our aircraft cost structure and numbers from section 2.2. The DOC per mile formula we used is included on table 2.3.1. The 30 points presented in table 2.3.1 were used to arrive at the following results:

$$\$/ASM \sim STAGE^{-.29} \ CAP^{-.22} \qquad (2.3.4)$$

This represents the results that the log-linear regression techniques could achieve with a set of ideally comparable data points. We call this the regression of "perfect" data. (2) Figure 2.3.2 plots the results of this regression against the original data line. The log-linear expression appears able to follow the original line over the range of calibration (up to 1200 miles), but fails to predict accurately at practical stages outside the data points. The Douglas and Miller regression when scaled to intercept the data at 600 miles also suffers at longer stages.

The more interesting question is whether aircraft capacity dependence can be tracked by the log-linear form. The answer is that both the regression of "perfect" data and the Douglas and Miller regression track quite well at all but the smallest capacities, as seen in figure 2.3.3. It is regrettable Douglas and Miller did not make use of the information contained in their cost function.

Industry Cost Presentations

In order to show that the engineering cost stucture can explain the cost curves the airline industry is accustomed to seeing, we have plotted DOC lines from the costs developed in section 2.2 in the representation which is traditional for the airline industry. Figure 2.3.4 illustrates that the cost "taper"

---

(1) Adam Smith based his cost arguments on engineering considerations. The technique has since fallen into disuse in academic circles.

(2) Statistical performance is listed on table 2.3.1.

Table 2.3.1: Data Points For Regression Calibration

"perfect" data generated by exercising the formula

$$\$/ASM = .0153 + .774/CAP + 144/(CAP\cdot STAGE) + 2.85/STAGE$$

for the values of CAP and MI indicated by the table below:

STAGE \ CAP	100	125	150	200	250	300
200	x	x	x	x		
400	x	x	x	x	x	
600	x	x	x	x	x	x
800	x	x	x	x	x	x
1000		x	x	x	x	x
1200			x	x	x	x

mean STAGE=700 miles

mean CAP=186 seats

correlation: 35%

REGRESSION RESULTS:

$$\$/ASM = .523\cdot STAGE^{-.29}\ CAP^{-.22} \qquad (R^2=.977)$$

STAGE exponent significant at 99% level
CAP exponent significan at 99% level
Standard Error = 3.3% (because of log form, error is a %)

Figure 2.3.2: Log-linear Representation of DOC per Stage

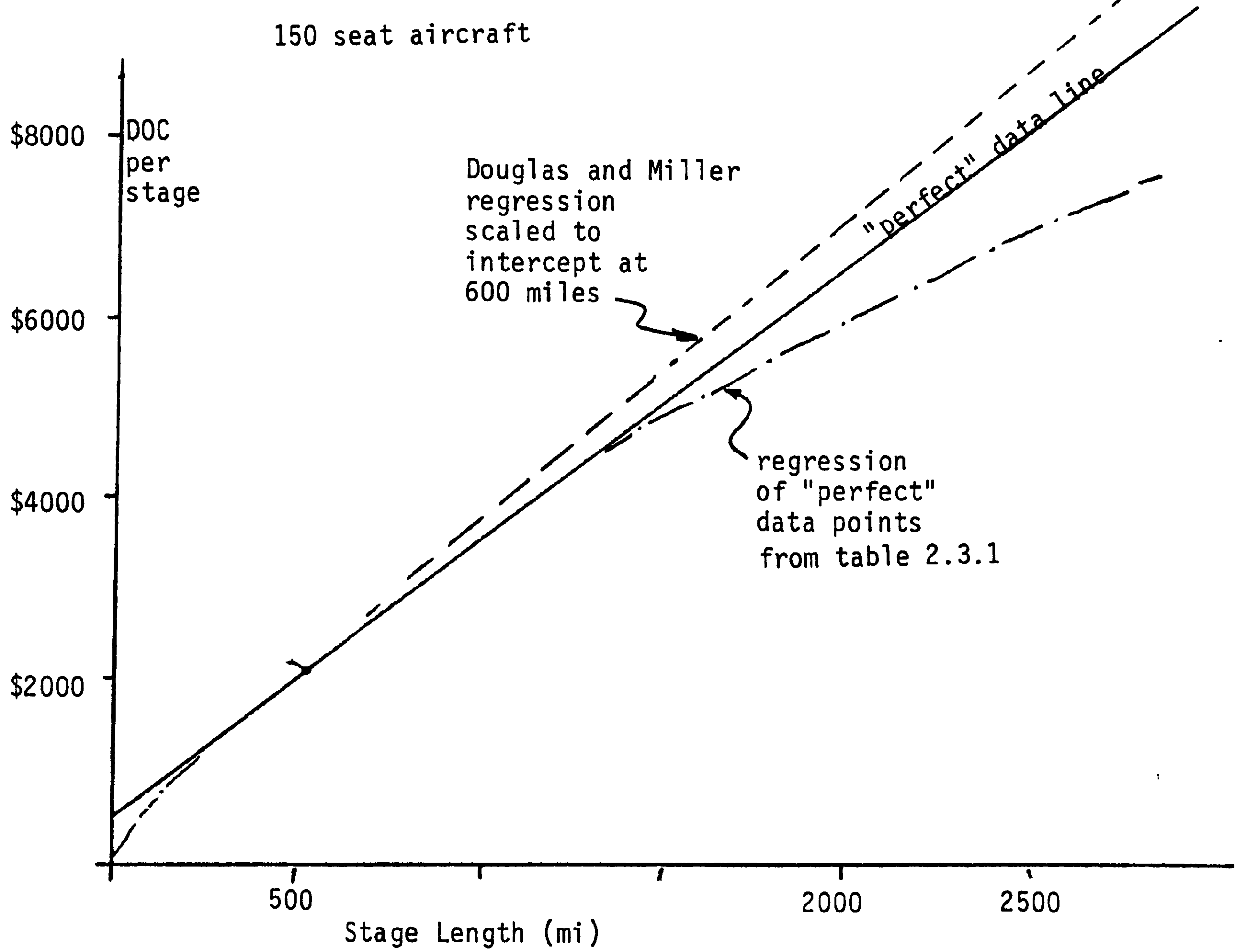


Figure 2.3.2

Figure 2.3.3: Log-linear Representation of DOC vs. Aircraft Capacity

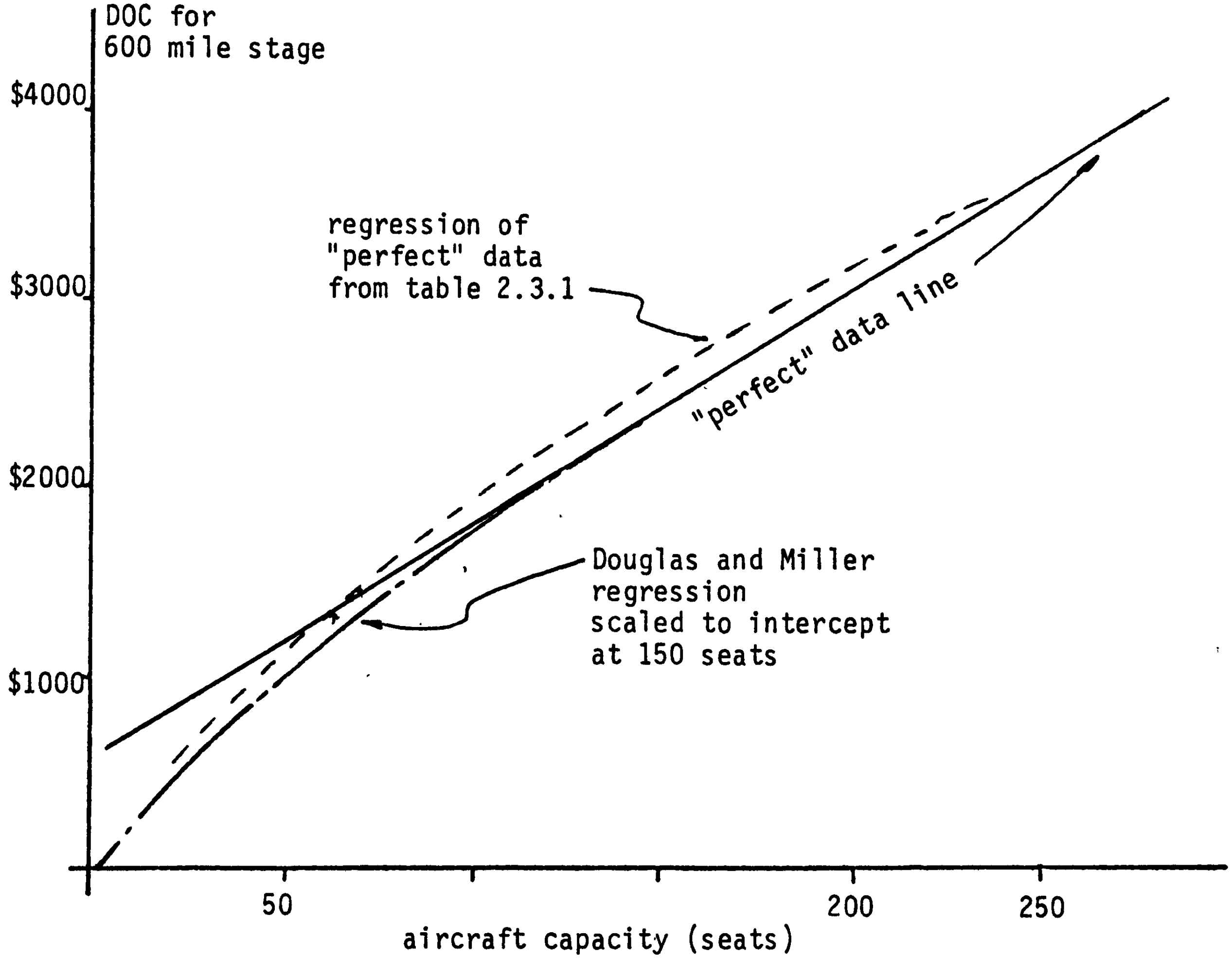


Figure 2.3.3

Figure 2.3.4: DOC in Traditional Seat Mile Format

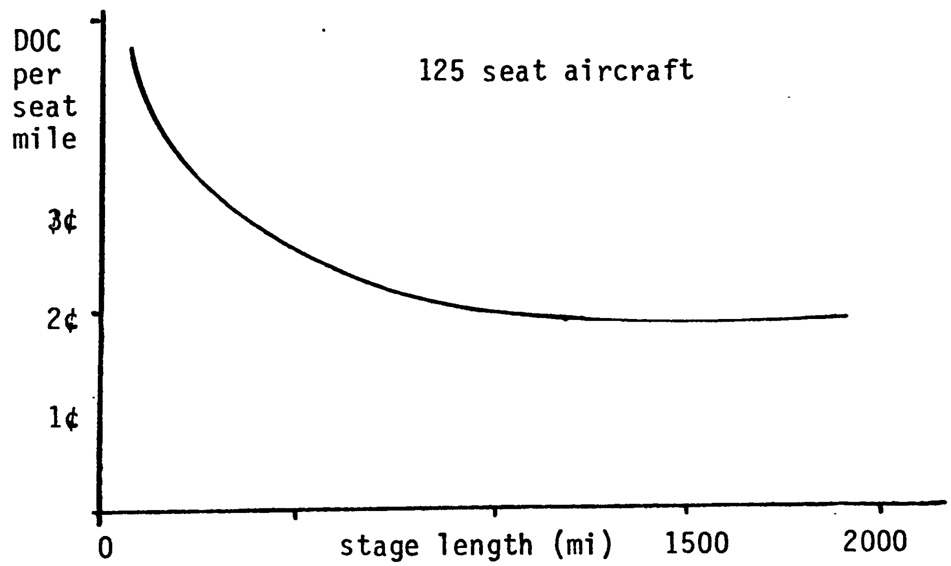
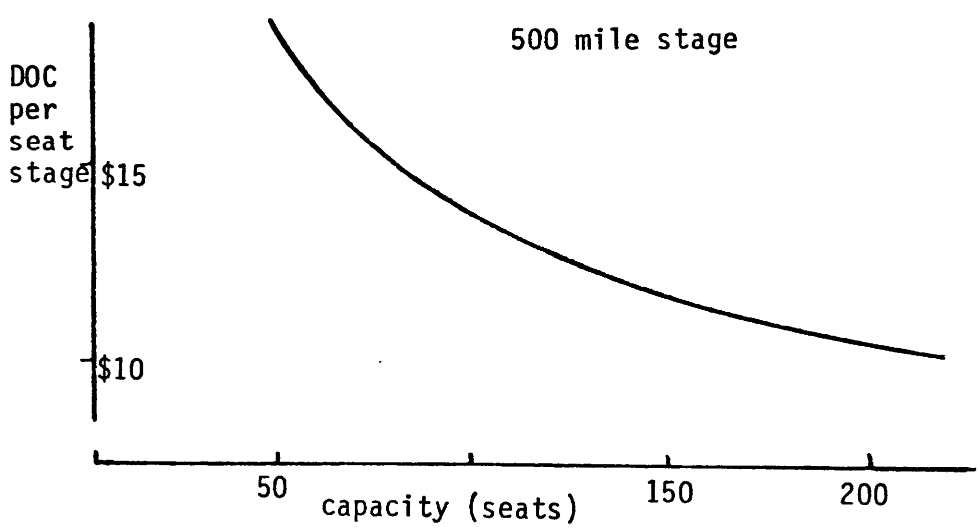


Figure 2.3.5: DOC in Traditional Aircraft Capacity Format



with distance can be explained by our simple assumptions. In the range of normal average stage lengths (200-1200 miles), stage length is an important influence on costs.

In figure 2.3.5 the effect of aircraft capacity is seen. In one sense this is an economy of scale, in that producing more capacity in the same place (i.e. on the same departure) reduces average capacity cost. Above 200 to 300 seats, savings begin to diminish. Aircraft in general use are approaching these ranges, but their economies are somewhat compromised by network effects. We will not be able to explain this more fully until chapter 6.

2.4 Conclusion

It seems a little early to reach conclusions. We have not introduced a definition of what is a market or how to measure the qualtiy of service. Nonetheless, if firms compete for the traffic in a city pair by offering their own schedule of departures and capacity, it is clear that the ability to use larger aircraft to serve larger traffic levels will reduce average passenger costs for firms serving larger fractions of the total market. Thus we antipate that service at the city pair market level will be characterized by one or only a few competitors. This limitation motivates the demand model discussion of the next chapter.

3 Characterization of Demand

3.0 Introduction

It is natural and perhaps forgivable for us to think of demand for travel as more or less fixed and to focus on the cost of supply. But in a larger sense the value of goods or services deserves as much attention as their cost. This is particularly the case with a transportation service where the consumer may find only one or two carriers and only one service type available. This section develops a simplified model for the aggregation of individual decisions which we call demand. (1) The focus is particularly on the distribution of values for different levels of service. In chapter 2 we discussed the supply side costs without discussing the market equilibrating process through which the supply is used. In this chapter we follow an analogous course by discussing demand with the supply side being a postulate. There is one difference, though. We already suspect from the cost structure that monopoly or oligopoly service will be the norm in a city pair market. More than one or two competitors would fractionalize the traffic to the point where aircraft too small and too expensive are necessary. Our demand model will take a particular look at two-firm competition.

The developments in this chapter pull together disparate discussions of consumer behavior and air travel demand models. In the general field of demand description we have elegant mathematical derivations from utility theory and cruder statistical formulas from forecasting. Utility theory models consumer behavior while forecasting models demand. The distinction is one of emphasis, but the practical differences between the two approaches should not be underestimated. Between these two camps there is much fertile ground. In order to discuss the tradeoffs in airline service design, we need the concept of value of quality from the theoreticians, but we approach the issue using the terminology of the forecasting formulas, because it should be more familiar to most of our readers. The model we end up with represents the smallest revision of the conventional transportation demand formulations which will capture all the relevant issues. The arena thus entered is coming to be called behavioral demand modeling.

In the discussion below the central variable is the total perceived price for a trip. This price includes both the fare paid and travel time. As an expected value, travel time includes delays which depend on frequency and load factor. Different people value this time at different levels. It is by appealing to subsets of the distribution of these values that firms or modes with different service/price offerings divide a market.

---

(1) Demand and consumption as used here are not synonomous. Demand refers to the entire function which describes the aggregated values of all potential concumers. The actual number of consumers is called the traffic.

But there may be a tendency to match services' qualities. Such behavior will neglect to provide appealing options to some parts of the demand population. These are the issues we hope to make clear.

3.1 Definition of a market

It will be important for clarity of later discussions that the market for scheduled transportation be carefully defined. We choose to describe services which are practical alternatives for each other as belonging to the same market and services which can not replace each other in a functional sense as being in different markets. This definition of a market does not specify the mode(s) involved. The test that services which can be used for the same purpose by the consumer are in the same market is the same test which the Supreme Court employs in determining markets for monopoly cases. (1)

The substitutability criterion implies that transportation in different geographical places represents different markets: travel from New York to Miami can not substitute for travel from Pittsburgh to Chicago. Thus while vehicles are not sensitive to where they move, people are. The city pair is not sufficient to define a transportation market. The type of service must also be specified before the substitutability test can be applied: passenger travel in the Pittsburgh- Chicago city pair is a different market from freight carriage between the same cities. In general we will be discussing passenger travel, although our model often applies to freight as well.

A market is usually sufficiently defined by specifying the origin- destination pair and the class of service (e.g. bulk freight, package, or people). There will be grey areas at this level of distinction, but the imprecisions will be more theoretical than practical. There is one interesting aberration. In air travel, substitutability may not always imply the exact origin- destination pair. There are cases where the destination is a region larger than a city. In some contexts it makes sense to discuss the market for travel from Boston to Europe or Boston to the West Coast. Usually these broader descriptions apply to personal pleasure travel. The extreme case of this is a market for pleasure travel from Boston to anywhere pleasant. But for the purposes of this discussion, markets are defined by city pairs. All travellers are included, but no other destinations. The broader use of the term is mentioned only as an exception to the rule that substitutability of service defines a market as a city pair.

The implications of this definition should not be neglected. Because goods can be transported, the market for physical products is national. But transportation itself is a service. The service takes place in one geographical market. A firm which provides the only practical means of transport in a city pair has

---

(1) Nason [21], p14 footnote.

a monopoly in that market. The consumer wishing to travel in that city pair has only one option. This point has escaped a good many air transportation analysts in recent times. From a practical point of view the issues of cost, competition, and monopoly power apply at the market and not the industry level. Exploitation or excess profits can occur in a single market.

If the market for transportation is specified by the city pair, then airlines which provide services all over the country operate in many different markets. They are able to do this because aircraft can be easily converted from production in one market to production in another. In fact, the same aircraft movements may produce service in several markets at once.

A result of airlines operating in many markets is that the consumer perceives the firm as if it were a retailer carrying an assortment of loosely related goods. (1) A person goes to a hardware store for tools and to an airline for trips. From the consumer's viewpoint the relevant issue is that if only one firm provides service to the destination city, that firm will carry all the traffic and may operate as a monopolist.

3.2 The Quality of Service Within a Market

If there is more than one practical way to get from here to there, the consumer's choice of mode or firms within a mode will depend on the differing qualities of service offered. In addition to the price, the consumer considers several other aspects of service. It is convenient to express these as additions to the money price of travel. (2)

The first of these additional costs of consumption is the amount of time it takes to make the trip. Travel time in particular has a cost. A passenger's travel time may be thought of as the amount of his own labor necessary to produce the trip. In some cases, the cost of travel time can dominate the total perceived price of the trip.

A second part of travel cost is the cost in time and money of access and egress. This is really a small transport problem within the larger transport problem. The custom of ignoring the sub problem of access to consumption will be followed throughout this analysis.

One cannot ignore the dimension of quality of service associated with the timeliness of that service. The customary index of the timeliness of service is its frequency. For repetitive consumption or as an expected value for a single trip, a high frequency of service improves the convenience of the

---

(1) The author is entirely indebted to unpublished comments by R.W. Simpson for this concept. Much of this discussion has grown out of work with Prof. Simpson.

(2) The use of heuristic prices for service qualities is quite well established. Recent examples are Grenou [16] for air and Roberts [23] for freight.

scheduling, i.e. it saves time. For a scheduled service, the consumer can read the schedule and avoid long waits. However, some time inconvenience is still involved for low frequencies. For the purposes of quantifying this time we shall use the expected displacement time as an index of the time inconvenience associated with the schedule frequency. This is the standard assumption and has been used by Douglas and Miller [10], Gordon and de Neufville [14], Eriksen [11], and others. Displacement time is the difference either forward or backward between the desired time and the nearest scheduled departure. It will be convenient to consider this as part of the time cost of a trip.

As far as the consumer is concerned, displacement time is a more fundamental measure of the quality of service than frequency. However, for a well designed schedule even with considerable peaking in the desired departure times, expected displacement time is proportional to the reciprocal of frequency. This relationship applies both in theory [22] and in practice [41].

The final part of travel time is time lost due to high load factors. In a normal scheduled transportation system, most of the consumers are served on their first choice of departures, even though demand varies from flight to flight. In order to serve most of the people most of the time, it is necessary to provide more capacity than the average amount of demand. But only absurdly large capacities serve all the people all the time. Sometimes space is unavailable due to an unusually large surge in demand. The ratio of average load on a regularly scheduled service to capacity is the average load factor. Load factor is an important aspect of service quality not because it has an overwhelming impact on non-money cost to the consumer but because it has a powerful influence on average passenger trip cost and thus on the price charged for the service.

The probability of a denied seat (due to load factor) can be quantified, and an estimate is developed in chapter 4. The penalty for such a happening is that the customer must wait one entire headway period for the next scheduled departure and take a chance on that being full. (1) Thus, the time associated with load factor can be thought of as a probability of denied service which goes up with average load factor times a headway which drops with the reciprocal of frequency. We go deeper into these matters in chapter 4. The practical gist of the matter is, the higher the load factor, the more chance of a full departure, while the greater the frequency, the smaller the extra wait time.

---

(1) Strictly speaking the customer may move to the second most convenient service, the displacement being from his original desired departure time rather than the time of his first choice of departures. Only third and fourth choice departures must be an average of one headway further from the desired time. However, displacement of passengers to much earlier times may not be possible, in which case the suggested formulation is correct. In any case it is the conventional approach and seems intuitively reasonable.

Up to now we have discussed dimensions of service quality which can be converted into time penalties. We now mention a few other quality dimensions.

Comfort, perceived safety, and reliability are all aspects of service quality which can be crudely quantified. At least in theory, a price can be put on them which adds to travel cost. In air travel, however, these dimensions of quality add little to the cost of the trip and they can be ignored, with great reduction in the complexity of the demand models.

More often ignored but with less justice is the cost of finding information on available services and of arranging to purchase the service itself. In the freight modes there is manipulation of the cost of finding out about a service. This is done in order to influence demand. Unprofitable services may not be admitted to exist upon casual inquiry with shipping agents. (1) Profitable services are marketed door to door. The situation is similar for passenger travel. In the passenger business, rail and bus schedules are not practically available from travel agents and discount air travel information is sometimes withheld. (2) There are so many products available in the transportation arena that the services and prices for a specific market are often hard to obtain. This adds to the cost, at least for some consumers, and often influences the choice of mode.

3.3 Fare, Perceived Price, and Demand

The previous section implied that the total perceived price to the consumer of a transportation service could be expressed as:

$$PP = F + v \cdot T + h \cdot q \qquad (3.1)$$

Where PP is the total perceived price of the service
F is the money price or Fare.
T is the total travel time defined below.
v is the value of time for the consumer.
q is a variable representing other quality dimensions in quantized measures. Dimensions include availability of information as well as comfort and safety.
h are a series of values, i.e. implied (hedonic) prices, for the various quality dimensions q.

Total travel time is defined as:

---

(1) Documentary evidence is not available to prove these points, since such practices are illegal.

(2) Studies have been made to disprove this point for air travel. The results do not coincide with thi… author's private experience.

$$T = tb + f1(FQ) + f2(LF,FQ) \qquad (3.1a)$$

Where FQ is the schedule frequency.
LF is the average load factor
tb is the physical travel t (block time)
f1 is the displacement time function. In appendix c this is shown to be $f1=5.7/FQ$
f2 is the expected value of delays due to load factor. This function is developed in 4.1 as $f2=57 \cdot LF^{9}/FQ$.

Although the functional forms of f1 and f2 often differ, this definition of total travel time is the generally accepted one (cf. [10],[14]).

Real people making choices may not trade off matters of fare, time, and quality in this simple linear form. This formulation corresponds to the first term of the Taylor series of each of these fundamental tradeoffs. It will be absolutely necessary to address these tradeoffs in the discussion of transportation systems, but it will not be necessary to go beyond the relative slopes v and h in the price expression. Interactions between F, T, and q and curvatures of the derivatives of PP with respect to these variables are not necessary to our later developments.

There are two very nice aspects of this linear formulation for perceived price. First of all, the numbers v and h are dimensional quantities which have reasonable boundaries and which may be obtained by asking people for their values on consumer surveys. They can also be estimated from non-transport related data (e.g. v should be associated with the wage rate; safety values can be taken from tradeoffs in other activities). As a last resort, they can be imputed from transport choices in the usual way, by statistical inference.

The second advantage of the linear expression for total perceived price is that consumer surplus can be expressed in a simple closed form for the case of constant demand elasticity with respect to total perceived price. The simplest demand function for our use will be:

$$D = k_1 \cdot PP^{\alpha} \qquad (3.2)$$

Where D is the total Demand (1)
PP is the Perceived Price
k_1 is a market density constant
α is the (negative) elasticity

---

(1) Since the model is eventually calibrated from observed behavior in individual markets, the demand curve is neither income nor time compensated. In a less theoretical sense, we have already chosen to ignore second order effects in the use of equation (3.1) for perceived price. We do so because such refinements do not affect our general conclusions, and because demand model calibrations are an order of magnitude away from being accurate enough to detect such effects.

This is not a general form at all. It has been chosen strictly on the basis of algebraic convenience. Since available calibrations of forecasting demand models only provide first order derivatives and since our analysis does not require higher order terms, this form will do as well as any other. (1) Using this form, consumer surplus may be obtained as in figure 3.3.1. With actual perceived price defined as Pa and the resulting actual traffic as Da, surplus S becomes:

$$S = \int_{Da}^{\infty} PP\, dD - Pa \cdot Da$$

or, reversing the axis of integration:

$$S = \int_{Pa}^{\infty} D\, dPP$$

The second form is easier to integrate. By substituting $D = k_1 \cdot PP^{\alpha}$ and performing the integral we get:

$$S = \frac{k_1 \cdot Pa^{\alpha+1}}{\alpha+1} - \frac{k_1\, \infty^{\alpha+1}}{\alpha+1}$$

The second term is zero for $\alpha < -1$ and a large constant for $0 < \alpha < -1$. For the usual case where $-3 < \alpha < -1$ the second term vanishes. The first term can be reexpressed using (3.2) as:

$$S = -Da \cdot Pa / (\alpha + 1) \qquad (3.3)$$

This hybrid expression is convenient to remember. Surplus is proportional to the product of actual traffic and actual perceived price.

<u>Discussion of Demand Model Form</u>

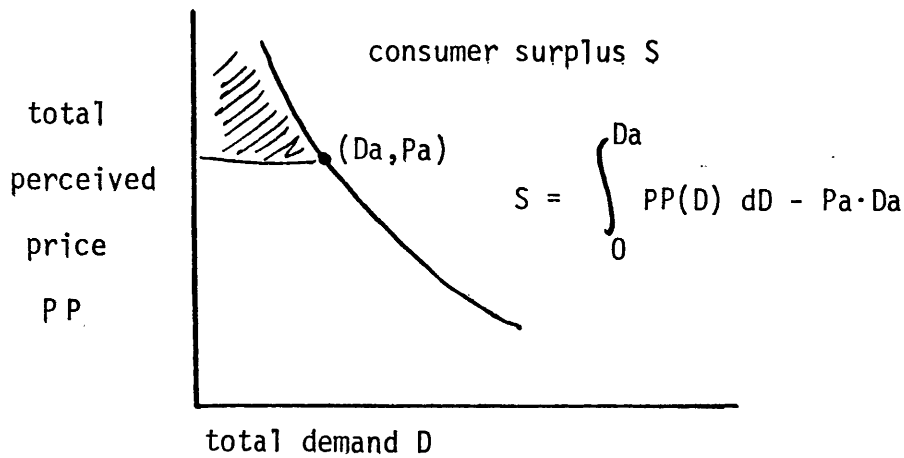
There is an interesting consequence of the mathematical form of the demand model. While fare elasticities as normally defined are known to change with income, the elasticity with respect to perceived price (α in (3.2)) may be nearly constant across income groups. If α were constant across values of time, our model implies that consumers with high value of time will be insensitive to changes in money price (fare F) and sensitive to travel time changes (T), while people with low values of time will be inelastic with respect to time and elastic with respect to money price. This seems intuitively correct. Whether α is nearly constant across consumers of differing values of time is a matter for future statistical evaluation, but our assumption that α is constant does not violate currently available information and is convenient conceptually.

The approach to demand modelling which uses explicit values of time (v) and quality (h) is intellectually clearer and mathematically more tractable (with respect to consumer surplus) than the form most often used in air transportation analyses today. The more traditional form ([3],[11]) is

---

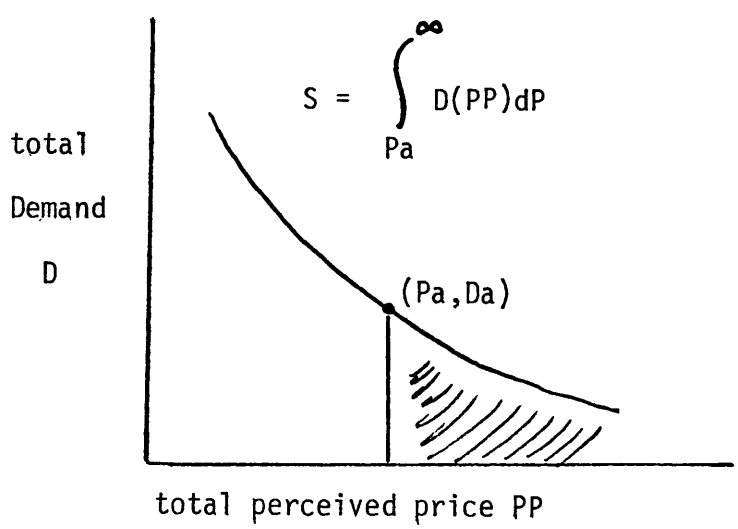
(1) Sadly as it may need improvement, it is not our place at this time to advance the state of the art in demand modelling.

Figure 3.3.1: Consumer Surplus



consumer surplus transformed for easier integration

(independent variable on horizontal axis)



$$D = k \cdot F^{b} \cdot T^{g} \cdot q^{u} \qquad (3.4)$$

In this formula the value of time is Fb/Tg and the value of quality is Fb/qu. (1) Sometimes when these expressions are evaluated they produce numbers for value of time or service quality outside of reasonable ranges. Values v and h (not g and u) are the derivatives in the form which influence the technical design, so they should be as accurate as possible.

It is unusual to combine the concept of hedonic prices (v,h) with the more primitive formulation of consumer surplus. However, the resulting model is mathematically straightforward and conceptually simple once it is accepted that the price for valuing and for paying for consumption includes both money changing hands and labor input (T) on the part of the consumer. An unusual consequence of this treatment is that figures like figure 3.3.1 do not allow one to deduce the actual fare F, which represents revenue to the producer. The process of matching fare to cost and supply to demand in an air transportation market is not as simple as it is in a market for hard goods. We will not even specify the conditions which make a match until chapter 5.

Calibration

The two formulas, ours (3.2) and the traditional one (3.4) can be made to agree in value and slopes at any given point. This is what we have done to evaluate aa in our equation from the best available prediction for b and g in the traditional form. The best available estimates for b and g are from a two equation regression of air travel markets by Eriksen [11]. Appendix B details the transformation which produced our estimate of $\alpha=-1.5$ from Eriksen's results.

3.4 Disaggregation of Demand

Although F, the money price paid for a transportation service, may be the same for all consumers, the total perceived price will vary according to the consumer's value of time and value of quality. For instance, value of time might be predicted by the consumer's wage rate. In this case the trip is more expensive for a rich man than a poor one. (2) There will be differences in value of time for the same person depending on the purpose of the trip. Pleasure trips may involve a value of time equal to the marginal wage less the marginal taxes while a business trip will involve the marginal wage plus employment taxes and overheads. In any case, there are different implicit prices for different people and different trips. Thus the

---

(1) These values are derived from dF/dT=(dD/dT)/(dD/dF) and dF/dq=(dD/dq)/(dD/dF). The symbol d is used to indicate a partial derivative.

(2) For ease in terminology we treat income and wealth as perfectly correlated.

perceived price of a trip varies among the population.

The distribution of perceived prices among the consuming population is not an academic point. We saw in chapter 2 that small aircraft are very expensive, so division of a market into many small pieces is uneconomic. From this we suspect that only a limited number of types of service are likely to be available. Services tailored for one type of consumer will be expensive for another type. For instance when only one type of service is offered, the choice of service affects who gets the greatest benefits. The important social issue is not so much technical efficiency as the distribution of benefits among consumers with different values.

An arbritrary example will help to illustrate this. The numbers employed will be more fully explained in chapter 5. Let us consider three classes of people: those with a value of time of $1 per hour, those with a value of time of $10 per hour, and those with a value of time of $50 per hour. At this point in our development we must take as an assumption that it is possible to offer scheduled passenger service in an 800 mile city pair market by air at a fare of $59, a frequency of 5.2 flights per day, and a load factor of 67% average. Travel time including flight time, displacement time, and the expected value of delays due to load factor is 3.32 hours. The perceived price of such a service is $62, $92, and $225 for our $1/hr, $10/hr and $50/hr people, respectively.

For this market it is also technically possible to offer service at a fare of $47, using a larger aircraft twice a day at a 77% load factor. Travel time is 7.36 hours when delays are included because of large amounts of displacement time and frequent unavailability of service due to full flights. The respective perceived prices of such a service are $54, $120, and $415. Although fare is 20% less, the two wealthier classes of travellers experience a very large increase in the perceived price of the service. On the other hand, the $1/hr people notice improvement.

Table 3.4.1 describes the situation for three technically feasible (zero profit) options of air service for this example. Each option favors one of the types of customers over the other two. Within the available service types, the "charter" service is best for the $1/hr people, the "standard" service is best for the $10/hr people, and the "premium" service is the best for the $50/hr people.

Operations at the "charter" level of service represent great inconveniences of time for the $10/hr and $50/hr people. Operations at the "standard" level represents a wastefully high quality of service as far as the $1/hr people are concerned. To first order however, the operator of the service is content to operate at any of these points, since all three are technically

---

(1) A risk averse operator might prefer the higher value of time customers because they are fare inelastic and he can raise his fares if his costs turn out higher than expected.

Table 3.4.1: Perceived Prices at Different Levels of Service

services:	charter	standard	premium
fare:	$46.88	$58.58	$75.94
frequency:	2.0	5.2	11.0
load factor:	76.5%	66.5%	61.0%
travel time:	7.36 hrs	3.32 hrs	2.53 hrs

All services are technically feasible at 400 passengers a day and zero loss. Numbers from chapter 5.

perceived prices:

services:	charter	standard	premium
@ $1/hr	$ 54	$ 62	$ 78
@ $10/hr	$120	$92	$101
@ $50/hr	$415	$225	$202

Relative Traffic:

services:	charter	standard	premium
@ $1/hr	1.00	0.91	0.64
@$10/hr	0.66	1.00	0.86
@$50/hr	0.34	0.86	1.00
total:	2.00	2.77	2.50

Relative Surplus:

service:	charter	standard	premium
@ $1/hr	0.59	0.55	0.53
@ $10/hr	0.87	1.00	0.95
@ $50/hr	1.54	2.10	2.21
totals	3.00	3.65	3.69

relative consumer surplus takes into accound both perceived price and relative traffic.

possible at zero loss. (1)

3.5 Market Split

Distribution of Values Among Consumers

For the set of all potential consumers or trips, the distribution of implicit prices for time and other quality dimensions may be thought of as a probability density function defined over an n-dimensional space. Each dimension would be one aspect of quality. Figure 3.5.1 is a contour map for the two dimensional case. Here the two dimensions of tastes are value of time and value of reliability. Reliability is measured in percent but not otherwise defined in this example. From the contour map one can deduce a peak of demand in the $6/hr and $30/% range and a correlation in the demand population between value of time and of reliability.

Splitting the Market by Value of Time

In figure 3.5.1 there is a dashed line which represents the watershed between those consumers who prefer a service which takes 4 hours, is 99% reliable and costs $50 (service A) over a service which takes 8 hours, is 91% reliable and costs $18 (service B). Given the concept of a distribution of tastes among consumers, services which compete in the same market with different cost/quality combinations split the market by such a watershed line. With more than two service choices, the demand space is divided into multiple regions. If the number of services is no more than one greater that the number of quality dimensions, all firms can compete against all other firms to varying degrees.

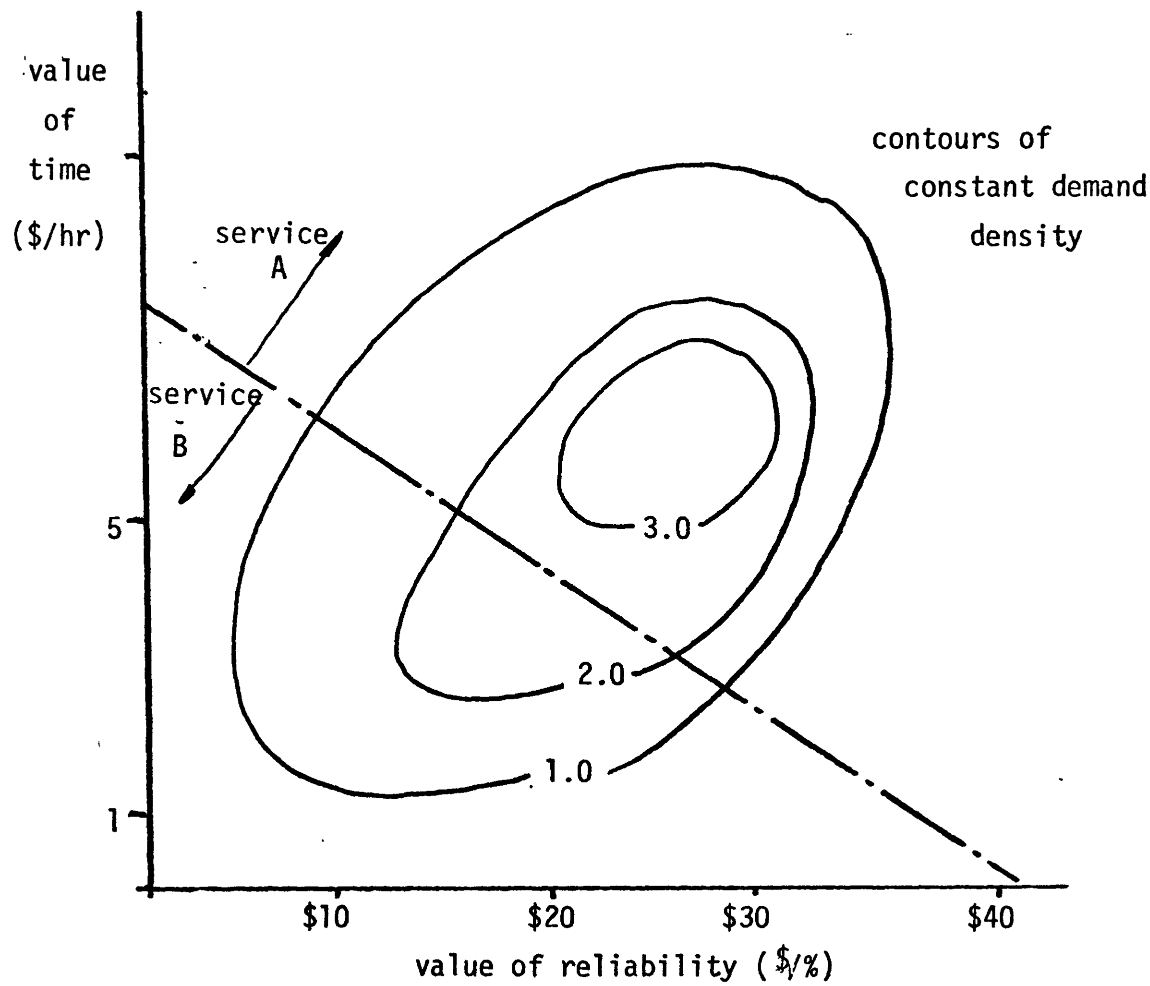
Figure 3.5.1 is the last time we shall use more than one dimension of quality to illustrate this point. Usually we will assume that value of time is the dominant factor and other quality dimensions q and their implicit prices h will be droppped from the discussion. Figure 3.5.2 illustrates the case for a single dimension of quality. Instead of a contour map, the vertical axis can be used to display the demand density. Figure 3.5.2 shows the demand density and figure 3.5.3 shows the watershed line between two services.

In practice the watershed line between the two services in figure 3.5.3 will be somewhat muddied. The travel time is a distribution imperfectly represented by its expected value. In particular there are variations in access and displacement times. Some# people will live near an airport and want to leave just when a flight departs and others will live across town and have other time of day plans. Thus the distribution of travel times will muddy a division of traffic made according to average travel times. (1) This situation will lead to a watershed line which

---

(1) Further lack of clarity will occur due to inaccurate perceptions of time or modal choice decisions which are not strictly rational.

Figure 3.5.1: Hypothetical Density Function for Demand in a Market with two quality dimensions



Contours are marked in units of passengers per time $ and reliability $

Figure 3.5.2: Distribution of Value of Time

Demand Density Function for the Single Quality of Service Measure, Time

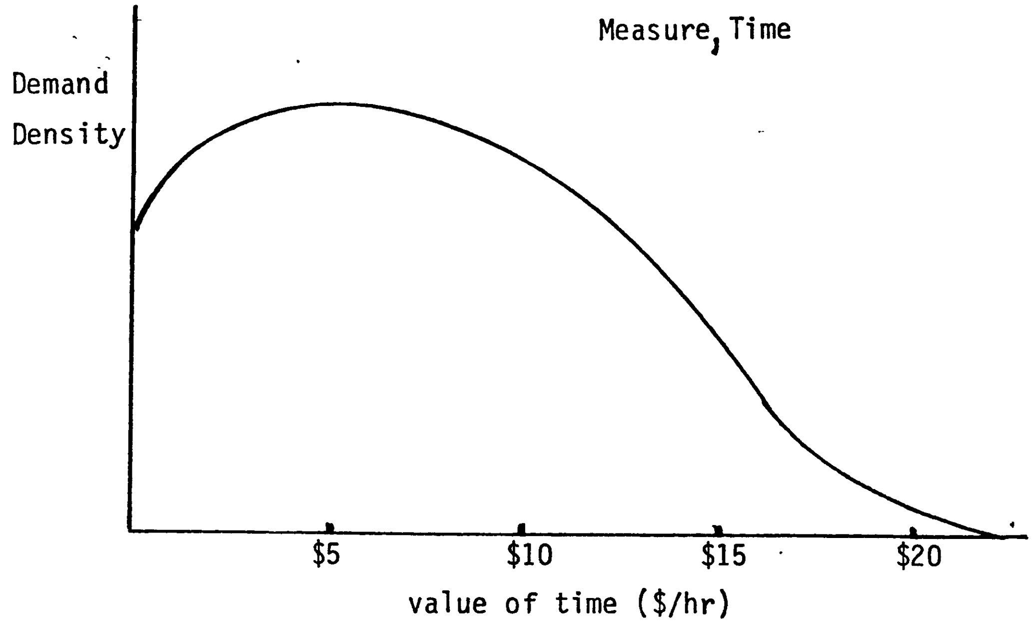
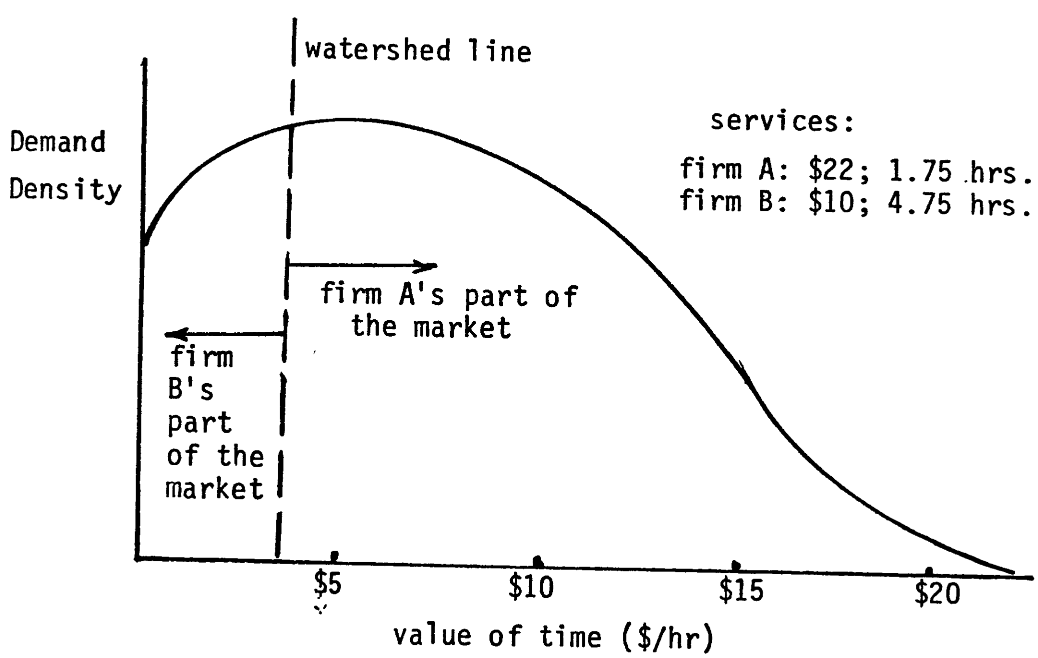


Figure 3.5.3: Market Split Between Firms by Value of Time



may look like the one in figure 3.5.4.

When two services are identical except in detail, the model as we have drawn it fails to provide any useful insights into the form market split will take. The market split line would become horizontal rather than vertical, dividing the market according to some criterion not presented in our figures. Such competition is very much a possibility in air travel. An example of one of the forms it may take is presented in section 3 of chapter 4 and another is explored in appendix C. But the larger issues of the amount and distribution of benefits are apparent from the pictures we have drawn, so we will not develop the demand model further at this time.

Our presentation of the problem puts the demand modelling process in a very untraditional form: modal split is made with the full disaggregation of demand in the dimension of values of service quality (figure 3.5.3), and there is a probabilistic choice of modes or services only for a small part of the population for which it is a close decision. We see that only a group of people near the watershed line are involved in the process of competition between services.

As long as the muddying of the service choice decision is symmetrical, i.e. there are as many people below as above the average values, the presentation in figure 3.5.4 with the curved watershed line is not necessary. For clarity, we shall revert to the use of a distinct, straight watershed for the rest of this discussion. (1)

<u>Distribution of Demand Stimulation</u>

We now wish to combine two concepts involving the distribution of values of time. We let the demand distribution line of figures 3.5.2 through 3.5.4 represent the <u>maximum</u> traffic at each value of time. By maximum we mean the traffic if the best mode offered technically efficient service most suited to just that value of time. Thus for any single service level offered, actual traffic will only equal the demand density line at one place. The situation is illustated in figure 3.5.5. The shaded section on the left represents demand for which the available service quality is unduly expensive when compared to service more closely tailored to its needs. The shaded section to the right similarly represents demand for which the service is a mismatch, only this time the quality is too low.

<u>Market Split between Competitors</u>

With this concept in mind, we can now define two philosophically different ways of altering market share between firms (market split). We consider two firms splitting the market as in figure 3.5.6. Notice that at the watershed line, both firms satisfy the same fraction of the potential demand.

The first type of change in market split is the kind which could come from technological improvement of one firm. Thus if

(1) For a discussion of a calibration of such a demand model, see Blumer and Swan [3].

Figure 3.5.4: Market Split Watershed Muddied by Distributions of Travel TIme

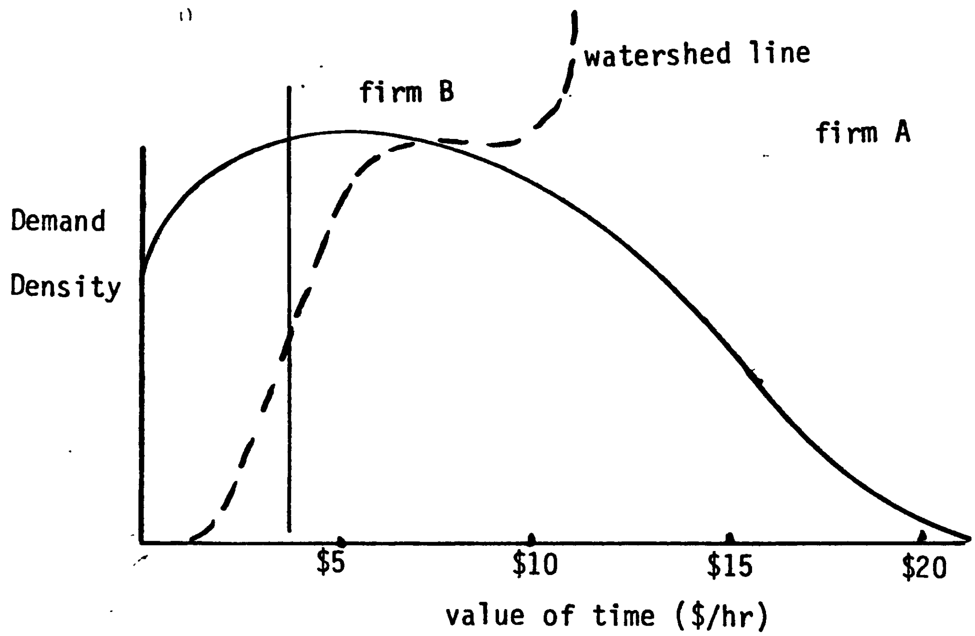


Figure 3.5.5: Ideal Service Quality at only one Value of Time

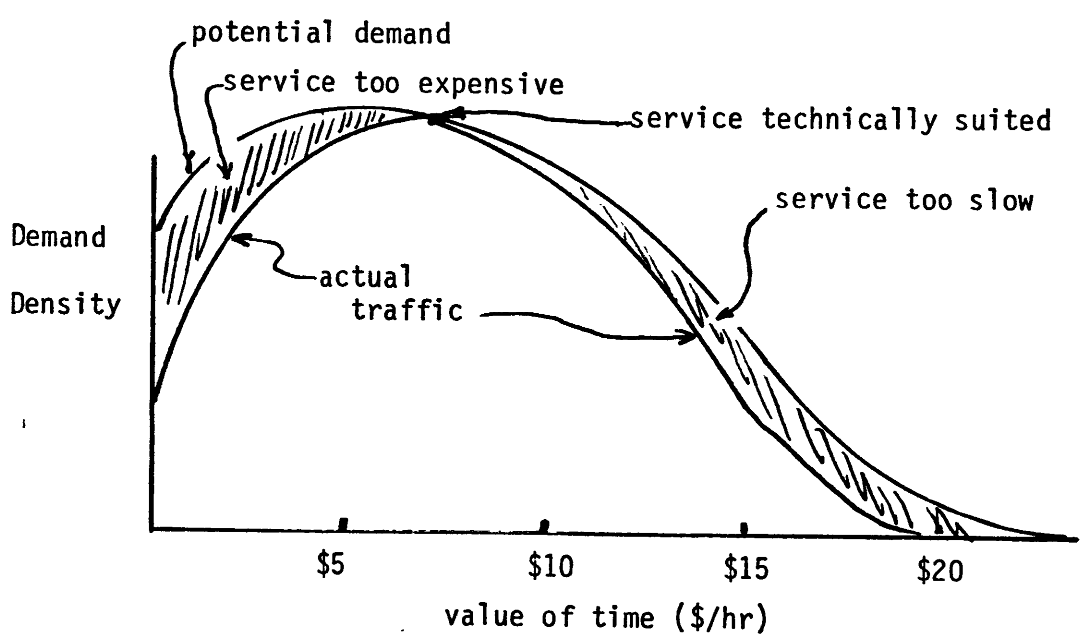


Figure 3.5.6: Market Split and Demand Satisfaction Shown Together

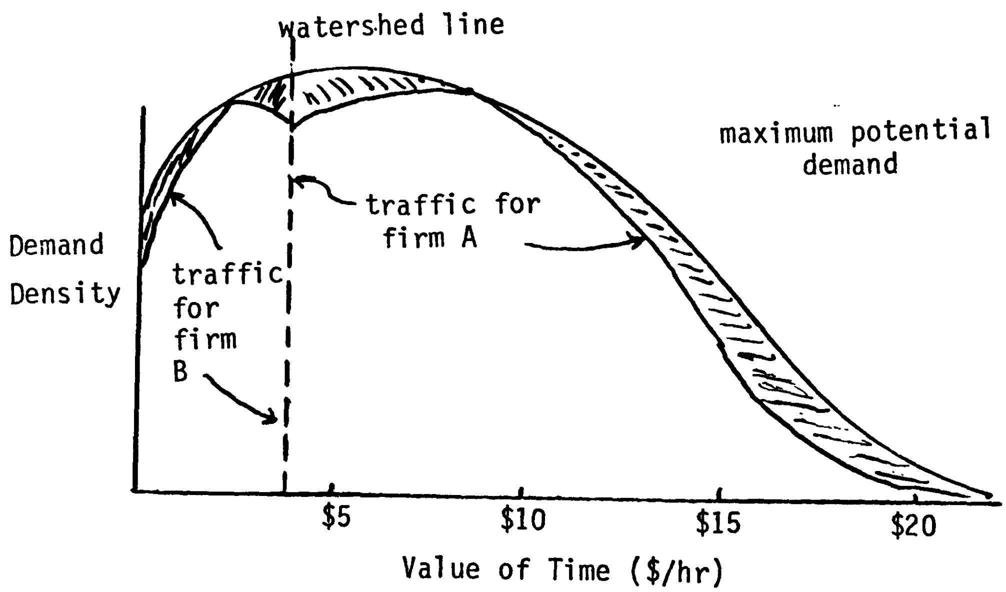
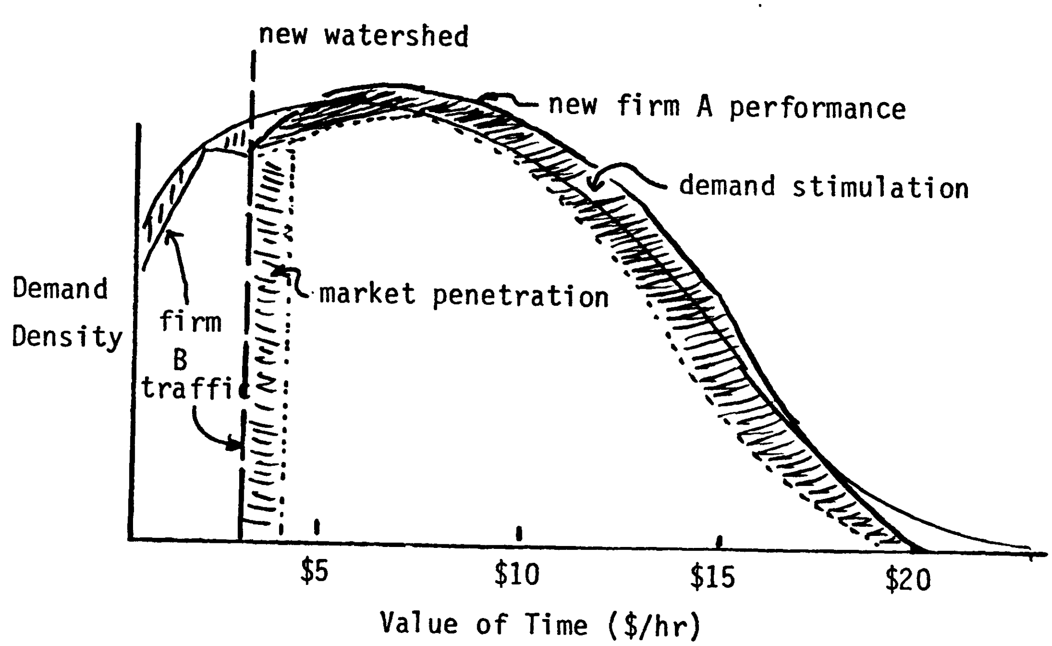


Figure 3.5.7: Market Split Change by Technical Advance



firm A suddenly became faster and cheaper, but remained designed for the $7/hr demand, the new market split would involve an increase in satisfied demand at the higher value of time levels, but little penetration of firm B's market. This situation is illustrated in figure 3.5.7. There is a broad increase in traffic indicated by the dark shaded band and a narrow penetration of firm B's market indicated by the column lying next to the new watershed line.

The other type of market split change occurs when firm A using its old technology merely reconfigures itself to favor the low value of time market served by firm B. This is illustrated in figure 3.5.8. Here firm A is reconfigured to offer service at a lower fare and frequency and higher load factor. Comparison with figure 3.5.6 shows a large amount of traffic removed from firm B, an increase in the fraction of demand satisfied in the $1/hr range, and a decrease in the amount of satisfaction for all demand above $5/hr.

This model of demand including a quality dimension (in this case, time) sheds some interesting light on the process of competition irrespective of the exact nature of the transportation technologies involved. First we notice a predilection for competing firms to match their price/quality offerings. This is what Nason [21] has shown to happen for airlines with identical technologies. (1) If we assume that the total traffic accruing to a firm is either a direct objective of management or instrumental in achieving management's objectives, then premium firms will tend to degrade their services (and drop prices) in order to capture the lower end of the market. They will do this as long as they are not losing too much traffic at the upper end of their market due to increasing the mismatch of service. Similarly, low price/quality firms will attempt to raise their quality (and price) until constrained by the loss of the poorest of their customers. (This type of competition was originally discussed by Hotelling [17] in the context of ice cream salesmen positioning themselves on a beach.)

If the two competitors have different technologies, their service/price packages may be held apart by the limitations of their operations. In addition, if demand stimulation dominates the market response, firms with identical technology may be held apart by the desire to satisfy different market segments. On the other hand we cannot ignore the predeliction of technically similar firms to offer services as nearly indistinguishable from each other as possible.

This process of product matching leaves the extreme top and bottom of the demand relatively unsatisfied. Market entry may be easiest for a firm directing its attentions to the top or the bottom of the market. In airline parlance, for a Freddie Laker Skytrain or a supersonic Concorde.

The demand model alone has less to say about three, four, or more firm equilibrium than it does in the case of two firms.

---

(1) Nason's development of the argument focuses on the single quality dimension of displacement time.

Figure 3.5.8: Market Split Change by Change in Service/Price Option

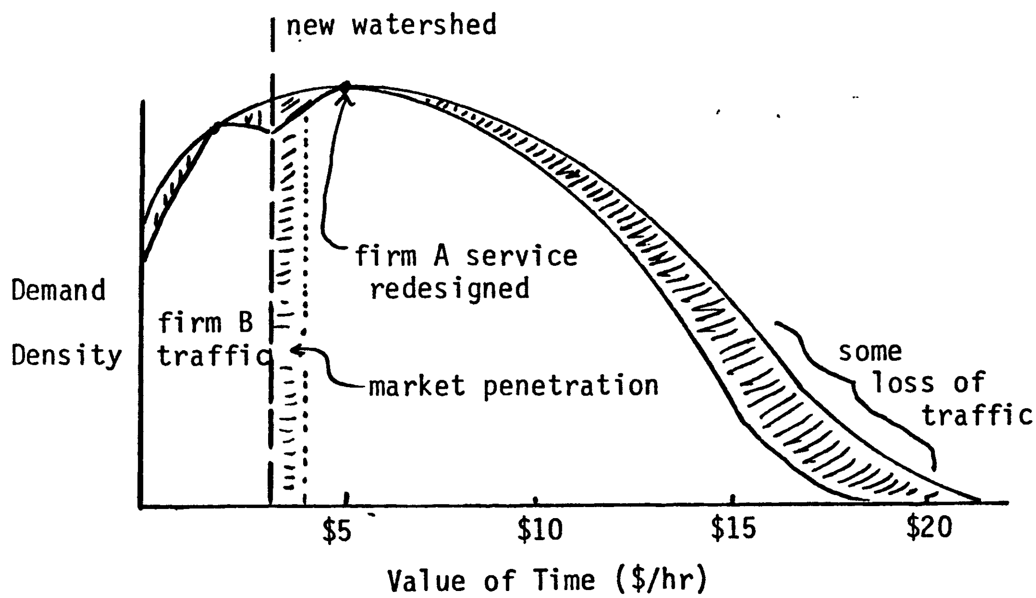
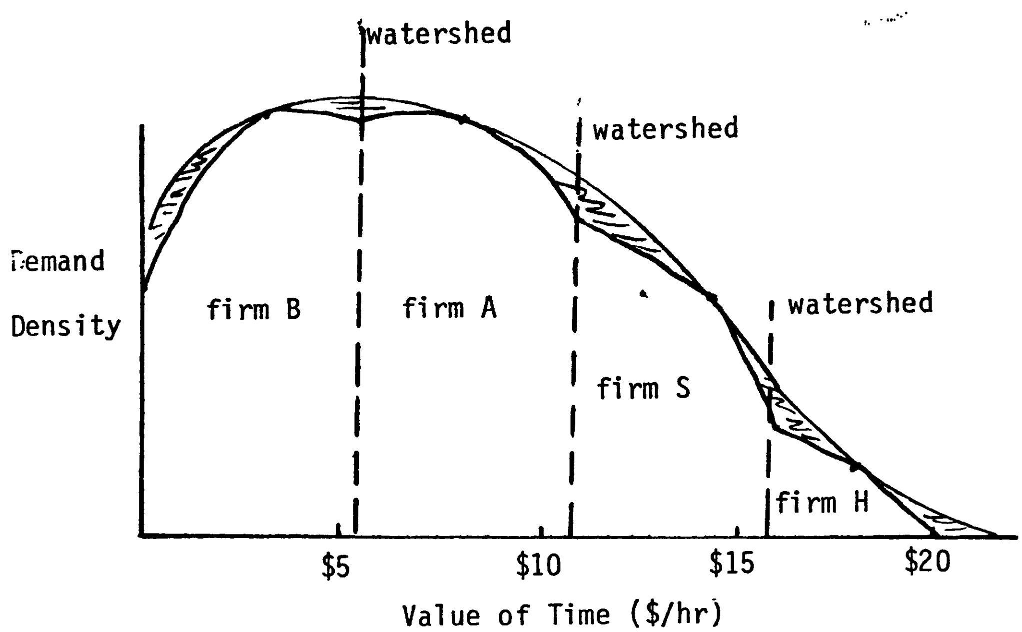


Figure 3.6.1: Multiple Service Levels



However, as the number of firms grows, losses and gains due to market penetration become smaller and demand stimulation effects become stronger. A greater diversity of services can be expected.

The above discussion has taken the view that maximum traffic is instrumental in achieving the objectives of the firm. The objective of maximum profits may be more readily achieved by isolating a segment of the demand from even remotely useful alternative services and then exploiting the local monopoly in that segment of demand to its fullest. Thus we notice that consumers lying near a modal split watershed line are unlikely to be unduly exploited by either firm, while consumers at the extremes may be offered unsuitable services and also suffer higher price markups whenever they can be singled out in the marketplace. In air freight we have an example of this in Federal Express. Federal Express serves the very high value of time traffic without direct competition in performance. They have been able to raise prices to cover costs. (Cf. [15].)

3.6 Multiplication of Service Levels

We have already introduced the situation in which a market is served by two different service levels. In figure 3.5.8 two firms satisfy more demand than the one in 3.5.6. An interesting possibility to contemplate is a situation such as figure 3.6.1 where many levels of service are available and almost all the demand is suitably served. This is the case with the greatest possible consumer benefits.

Unfortunately the provision of service by each new firm is accompanied by non-capacity costs. The per vehicle cost in chapter 2 is one such cost. Vehicle costs recur every day that a firm offers a frequency, but they do not rise with traffic. Thus vehicle frequency cost is like a fixed cost. (This will become more obvious in chapter 5 when explicit matching of demand and supply is treated in detail.) Each additional firm's service represents an extra fixed cost. Finding the optimal number of firms is a matter of trading off the extra fixed costs for each service type against the extra benefits in terms of demand satisfaction because of a better matching of service level to desires. This problem has been discussed theoretically for a general industry by Chamberlin [4].

Joint Service Offerings

Throughout this discussion we have implicitly assumed that each firm offers only one level of service. To some extent multiple service levels can be offered by the same airline by altering space availabilty and timing for different types of fare. (1) In this way multiple price/quality combinations can be

---

(1) When information about actual rather than probable loads on a service is available to the consumer, differentiation becomes difficult. See chapter 4.

offered without increasing fixed costs. However, a single firm is not competing with itself when it offers several quality options and its pricing and service choices reflect the fact that cross elasticities refer to traffic which is still on that firm's service. Although multiple service choices may increase demand satisfaction, the firm is in a position to capture much of the extra value as revenue. We continue to consider the one firm, one service case, avoiding these secondary complications.

Trends for Optimal Service Distributions

In the general case a small market probably should receive only one level of service while larger markets generate larger total benefits and thus justify multiple service levels and more fixed costs. At least this would be the case in a technically efficient system design. Market growth should be from one choice of service to many.

The distribution of demand according to its value of time allows us to perceive the dual objectives for efficient system design. The first objective is reduced cost or technically efficient provision of service levels. The second objective is an appropriate distribution of service level options. The service levels should be suited to the distribution of consumer values and should offer several options, especially in dense markets. Whether this will happen in a competitive situation is a good question.

3.7 Conclusion

At this point we leave our general discussion of demand behavior and turn our attention to the fundamentals of transport technology in serving a single isolated market. Throughout the rest of the discussion we shall refer to the conceptual model implied by the figures in this chapter and especially figure 3.5.8 as our demand model. The understanding is that equation (3.2) is to be exercised at each value of time.

In the next chapter we examine in detail a part of the process of matching demand and supply, that is the issue of load factor. In chapter 5 we do the matching for a single market and firm, discussing the levels of firm and the distribution of benefits which should occur.

4 Load Factor

4.0 Introduction

Chapter 2 presented the cost of providing transportation capacity on a regular scheduled basis. Chapter 3 discussed the nature of the demand for those seats, made up as it is by the sum of a number of different personal decisions to travel. The next step is to examine in detail a simple single link matching of these two numbers.

Load factor for a regularly scheduled service has been defined in Chapter 3 as the ratio of the mean traffic to the (constant) capacity. As such, it was used as an index of quality of service and become part of the travel time component of perceived price. The fundamental index was the rate of space denial. At that time detailed discussion of how load factor affects denial rate was deferred until this chapter. Section 4.1 develops a model which predicts the rate of space denial from a careful consideration of statistical matters.

The numbers for demand as it was discussed in chapter 3 were totals of the demand for all the departures in a daily schedule. Because we were discussing long term design, this daily total was an average for numbers of daily repetitions of the schedule. The schedule is operated for a number of months. For the calculation of mean denial rates not only the mean demand but also the variability about the mean is important. Variations in demand come from both regular cycles and unpredictable randomness. Section 2 discusses the components of total variability in demand and their estimates. theoretical development.

Finally, the nature of the reservations and space availability process permits competition favoring what may be undesirably high load factors and correspondingly poor service. This point is raised in Section 3. Detailed treatment of load factor problems in this chapter allow for single-market, single-link optima to be developed in the next by matching demand, capacity, cost, and service. General conclusions may then be applied to framing the network design problem in the following chapter.

Although the concept that load factor influences space availability is widely used, the problem has yet to be correctly addressed as a probabilistic model. The treatment in this chapter corrects the best efforts to date (1) and brings to light several new issues. (2) In addition the issue of competitive practices with respect to load factor is seldom addressed even by those employing essentially correct space availability models in

---

(1) Douglas and Miller [10], Chapter 3, which contains two analytical errors, but which is correct in intent.

(2) most especially the problems of correlation of the variability of demand with the mean.

competitive situations. (1)

4.1 Load Factor and Denied Boardings

High average load factors for a transportation service cause some fraction of users to be unable to take the departure they would prefer. This section explores two mathematical descriptions of the process by which some but not all of the departures in a series come to be over-subscribed and a number of measures which come out of such models. While the fundamental results have often been correctly grasped, (2) the formal treatment which follows has not been presented in the literature.

We start by specifying with greater detail the phenomenon we wish to model. A transportation schedule can be once each day, once each hour, four times weekly, or any other arrangement of departures which is useful. For our theoretical model we insist that each departure be performed by a aircraft of the same capacity. Although changes in capacity will not be discussed, the model will apply in cases when demand for departures with different capacities scales with the capacities. (3)

We describe the demand for a randomly selected departure in a schedule as a random variable d. We assume that all we know about d is contained in a probability distribution O(d). We do not distinguish among departures in the group; they all have the same O(d). In practice there may be information about the demand for consecutive departures in the schedule, but departure to departure correlations will not be captured by our O(d). The known cyclic variations in demand are, of course, included in O, as are random elements of variability. As defined, O(d) could apply to almost any grouping of departures in a network.

Not all of the demand is carried on their preferred departure The demand variable d includes passengers who cannot get on the aircraft because there is insufficient capacity.

The probability distribution O has a mean μ and a standard deviation σ. The variability may be caused by such cyclic changes in demand as weekend peaks against a daily departure schedule or morning and evening peaks against an hourly schedule of departures. Appendix E shows that this effect is quite large. In addition, some of the variability of demand is caused by random fluctuations in the desire for travel. It does not matter to the model at this point what is the source of the demand variability; we take it that the distribution O(d) is known.

Three demand distributions are presented in figure 4.1.1. Figure 4.1.1A shows the demand distribution for a group of

---

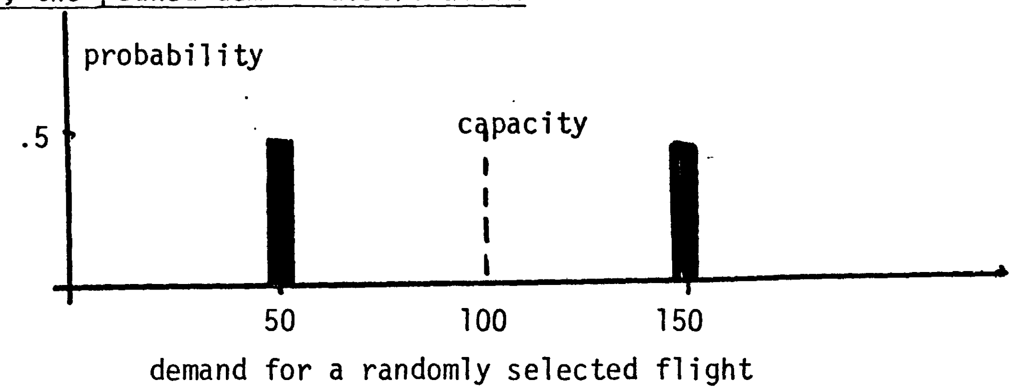
(1) Notably Dorman [9].

(2) See especially Gordon and DeNeufville [14].

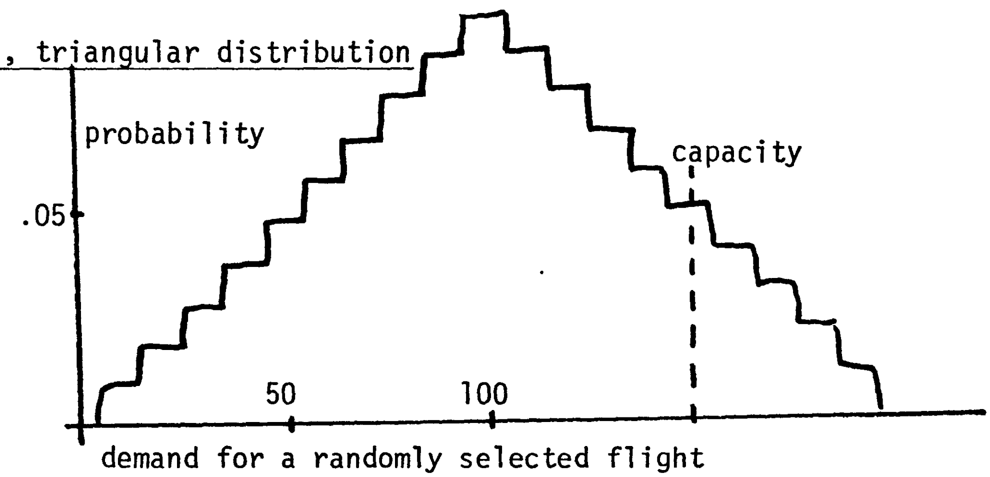
(3) Both the mean and distribution of demand must maintain a constant ratio to aircraft capacity.

Figure 4.1.1: Three Example Demand Distributions

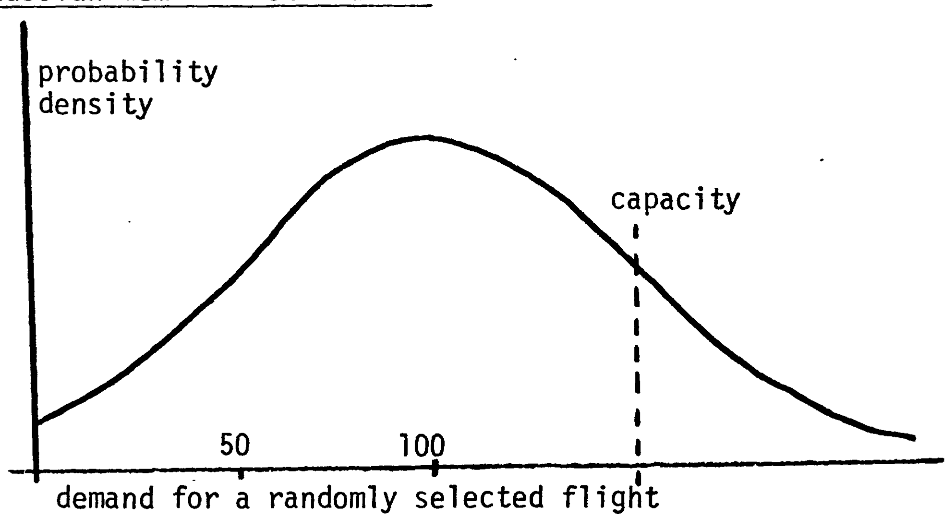
Case A, two peaked demand distribution



Case B, triangular distribution



Case C, Gaussian demand distribution



departures with a strong peaking pattern. For half the departures, the demand is 150 passengers; for the other half the demand is only 50. This produces the two-peaked distribution for a randomly selected departure. Figure 4.1.1B displays a nearly triangular distribution with a broader spread of demand levels. The probability distribution O(d) for Figure 4.1.1B is still discrete Figure 4.1.1C shows a Gaussian probability density function O(d) as a continuous distribution. This last case is the one we shall ultimately use to describe air travel.

If we define the capacity of each departure in the group as c, we can write formulas for several measures, each of relevance in a different circumstance. The model now states that the demand is distributed as O(d) and the capacity is always c. For those times where demand d is less than capacity c, all the demand is carried. When demand exceeds capacity some fraction of the demand is denied space.

Before we start we must discuss what happens to those people who are denied space. Do they go home and not make the trip? Do they get first priority on the next departure? Do they join the lottery for seats on the next departure? Associated with this is the question of correlation between the demand for a departure and the demand for the departure succeeding it.

Our discussion will focus on two of these possibilities. We make the distinction that O(d) is the distribution of demand initially preferring a particular departure. We define Os(d) as the demand for a randomly selected departure including any passenqers not accomodated on previous departures in the schedule who do spill to the next departure.

If passengers for whom space is unavailable on their preferred departure turn away entirely and do not attempt to board the following departure, then Os=O. To distinguish this case we will refer to it as the turnaway model and we will call its distribution Ot.

If all the passengers for whom space is unavailable on their preferred departure are spilled to later departures, then Os will have a higher mean and standard deviation than O. We will discuss the case where a turned away passenger enters the lottery for the next departure without priority. This approximates airline reservations practice. If we maintain the same preferred distribution for each departure and add to that the distribution of spills, the process forms a Markov chain. We will refer to the demand distribution including spills as Om(d). This we call the spill model.

For a Gaussian O(d) the distribution Om(d) can be calculated. We have done so numerically for the cases in this chapter. We assumed a capacity of 20 and mean demands in the range 10 to 20. We divided the Gaussian probability density function O(d) into a distribution by integrating from -.5 to .5, from .5 to 1.5, and so on up to 49.5. This gives a probability vector Po(d) for the number of people who prefer a given separture. (We lump all the probabilities in the negative tail of the Gaussian distribution in the first element of Po, Po(d=0).) From this vector we can generate the transition matrix T(i,j) where T(i,j) is the probability of having a demand j for a departure given there was a demand i for the preceding departure and any excess of i over capacity spilled to the next. We can say

$$T(i,j) = Po(j) \text{ for } i<c$$
$$T(i,j) = Po(j-i+c) \text{ for } i>c$$

If we define Pm(j) as the distribution for the mth departure in the series and if we know the demand for the first departure was k, then

$$Pm(j) = Pk(i) \cdot T(i,j)^m$$

Where Pk(i) is zero except Pk(k)=1. It is a property of $T(i,j)^m$ that for m sufficiently large, each row is identical and independent of m. This is called the steady state transition matrix for the Markov chain. We raised T(i,j) to sufficiently high powers to achieve this result in order to obtain the numerical values for this section. The resulting Pm is the discrete form of Om. (1)

Conceptually, the spill distribution Om is the original distribution O scaled up and spread out by the spill rate. This is illustrated in figure 4.1.2.

Neither the turnaway nor the spill model represent all we know about a series of departures in a schedule. Because peaks and valleys of demand cycles are known, some correlation between the demands of consecutive departures can be measured. With a significant positive correlation between consecutive departures, even the spill model understates the ultimate probability of demand exceeding capacity. Nonetheless it is a suitable approximation for modest spill rates.

Using first the simple turnaway model and then the spill model we define several measures for the cases of figure 4.1.1, e.g. the three cases illustrated. Our primary focus is on the Gaussian distribution, case C, but the simpler distributions make good illustrations.

<u>Nominal Load Factor (ln)</u>

The simplest number to define is the nominal load factor. By this term we mean the ratio of the mean of the demand to the capacity. Mathematically we may define the nominal load factor as:

$$\mu/c = ln = \frac{1}{c}\int_{-\infty}^{\infty} d \cdot O(d)\ dd \qquad (4.1)$$

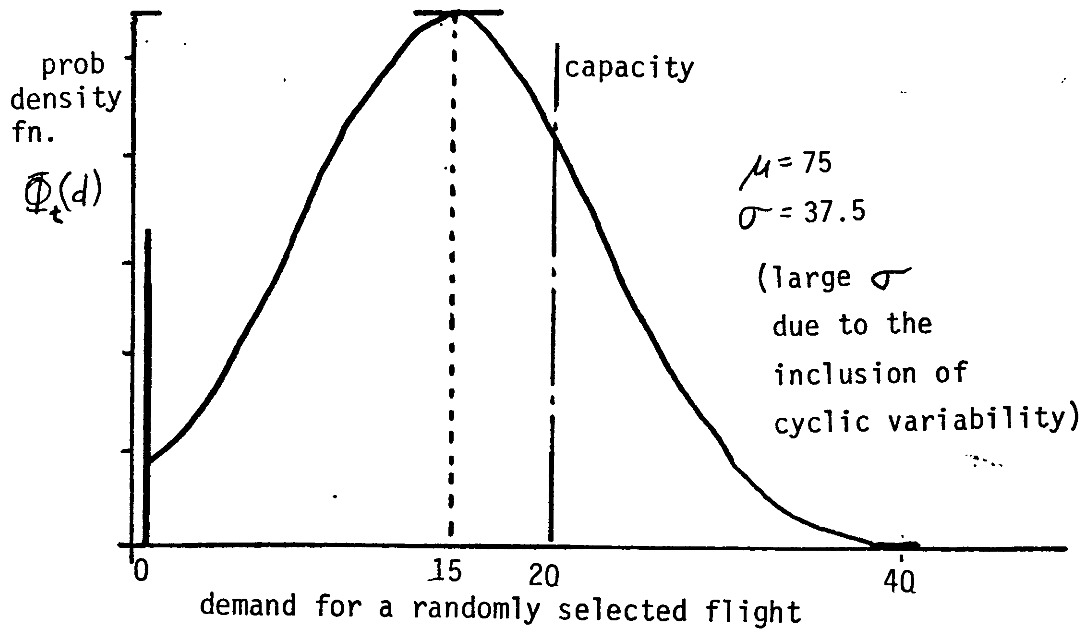
The nominal load factor describes only the basic demand. Whether it spills to succeeding departures or not is not an issue, so the measure is the same for both turnaway and spill models. The capacities assumed for cases A, B, and C were 100,

---

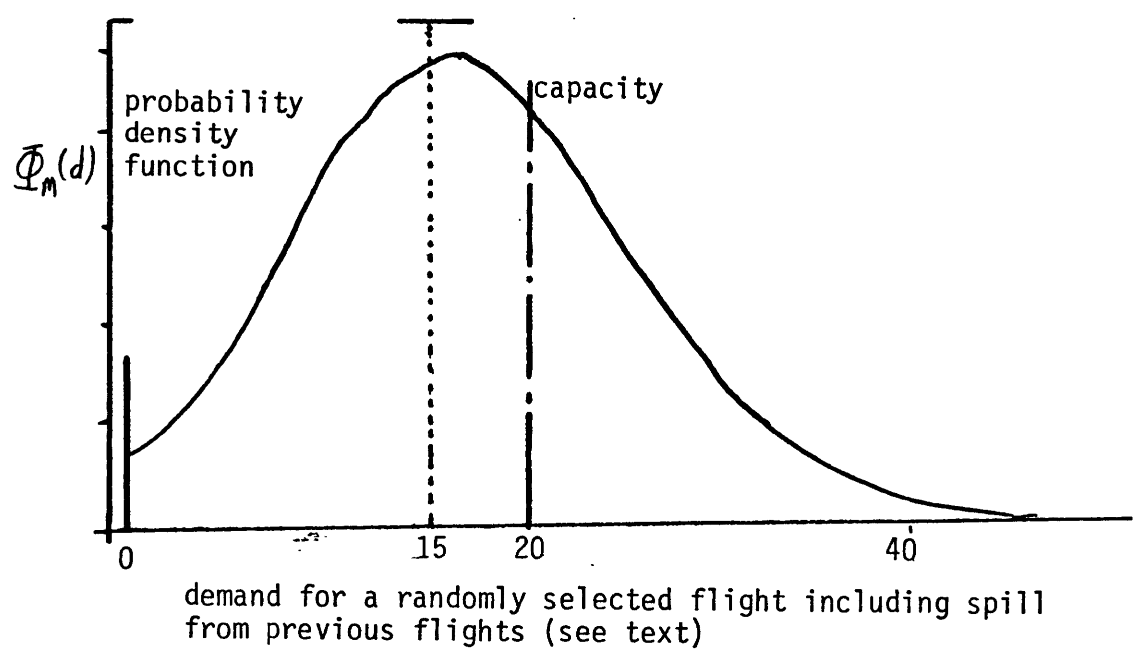
(1) Note: although there are errors in the Douglas and Miller [10] treatment of this topic, Table 6-5 appears to be numerically correct.

Figure 4.1.2: Comparison of single turnaway and spill model demand distributions

A: Gaussian distribution for single turnaway model



B: Corresponding spill model distribution



150, and 150 seats and μ is in every case 100. Therefore ln has a value of 1.00, 0.67, and 0.67 for cases A, B, and C respectively (see Table 4.1.1). From case A it is clear that this load factor (ln = 1.0) will not be achieved in practice. In case A, half the departures have a demand of 50 over the capacity. The actual load factor for these departures is 100%. From a practical standpoint, nominal load factors near or above 1.00 will not be handled by our approximations, although they may occur in fact.

Probability of a Full Departure, P100

A crude index of level of service is the fraction of departures which are filled. We call the probability of a full departure P100. This is merely the integral of the distribution O above the capacity:

$$P100 = \int_{c}^{\infty} Os(d)\ dd \qquad (4.2)$$

In the turnaway model Os=Ot=O. Table 4.1.1 lists the values for P100 for the three example cases from figure 4.1.1. For case A, P100 is 50%. For case B, the integral is 15%. For case C the integral is the shaded area of Figure 4.1.3, e.g. the cumulative normal above the capacity. The value is 15.9% in case C.

For the spill case, Os=Om. Needless to say, the probability of a full departure is higher when customers can spill to later departures. We have calculated Om only for the Gaussian case. For the spill model of curve C, the probability of a full departure is 19%.

Reported load factors will most closely correspond to the spill model. Because of no shows and standbys, reported load factors above 95% but below 100% probably include a large fraction of cases which are described as 100% load factors in our numbers For the purpose of comparison with observed load factors, the probability of a 95% or higher load factor according to the spill model is presented for several load factors and demands in table 4.1.2.

Flight availability Av is sometimes defined as

$$Av = 1 - P100$$

The name flight availability is misleading since it is not the probability of a person being turned away from a departure. We will discuss that measure next. It is possible to have half the departures full (Av =0.5) and have no one turned away. This would be the case if the two-peaked demand of figure 4.1.1A were served by a 150 seat aircraft. The flight availability is only an informal index of service levels.

Space Availability Pd

The best measure of the quality of scheduled service is the probability of a person being denied space on a departure. We call this probability the space availability, Pd. This number takes the situations in which anybody is turned away (Os(d>c)) and weights them by the number turned away (d-c). For the

Table 4.1.1: Measures Using Simple Turnaway Model for Cases Illustrated in figure 4.1.1

	case A	case B	case C
Demand mean	100	100	100
Demand standard deviation	50	41	50
Assumed Capacity	100	150	150
(1) Probability of a full departure	0.5	.15	.159
(2) Probability of denied boarding	.25	.02	.04
(3) Nominal Load Factor	1.00	0.67	0.67
(4) Actual measured load factor	0.75	0.65	0.64
(5) Passenger observed load factor	.833	.749	.764

Table 4.1.2: Flight Unavailability at 95% load factor for comparison with observed data;

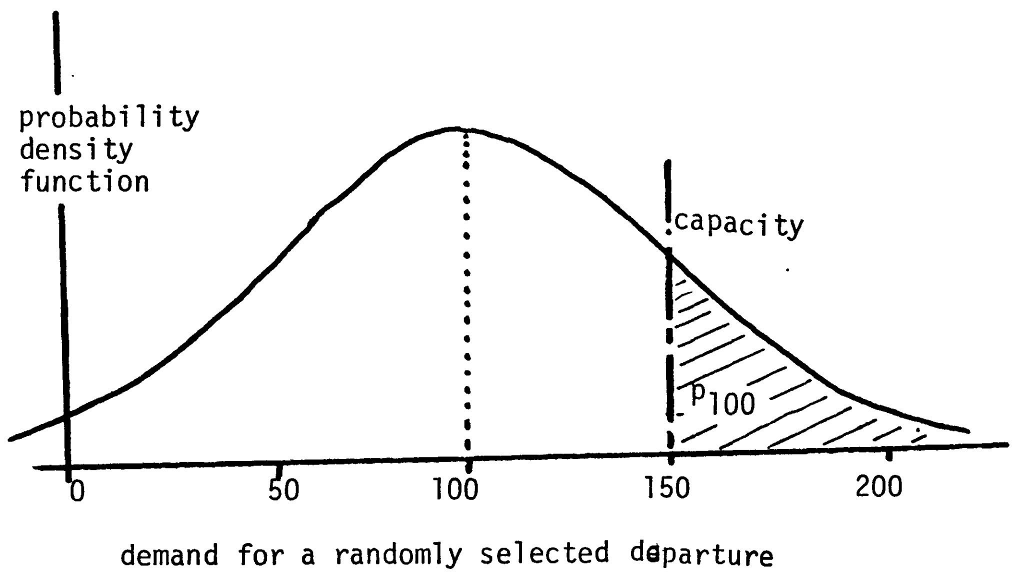
Values on table are the probability of a randomly selected departure having over 95% load factor
Spill model Øm

average load factor	Gaussian Demand Ø μ/σ = 1.0	2.0	3.0
.50	.212	.043	.005
.60	.383	.138	.042
.70	.524	.302	.081
.80	.621	.494	.373

numerical calculations from a 50 x 50 Markov chain steady state transition matrix. See text.

Figure 4.1.3: Probability of a full departure

Case C, Gaussian Demand Distribution with μ=100, σ=50



turnaway model

$$Pd = 1/\mu \int_{c}^{\infty} O(d) \ (d-c) \ dd \qquad (4.3)$$

This number is presented in table 4.1.3 for several Gaussian O's.

For the spill model the space availability is similarly:

$$Pd = 1/\mu \int_{c}^{\infty} Om(d) \ (d-c) \ dd \qquad (4.4)$$

Here Om is the distribution of demand for a randomly selected departure including customers spilled from other departures. Pd in this case is presented in table 4.1.4 for several Gaussian O's.

<u>Ultimate Denial Rate, Pu</u>

Space availability Pd is the denial rate for a single departure. On the other hand, the demand model in Chapter 3 used for its measure the chances of waiting exactly one extra headway. For the spill model, there is the additional chance of second, third, and fourth denials. The space availability Pd is only an approximation of the measure we need for the demand model. A better approximation of the desired measure would allow turned away customers to enter the lottery for the next departure with the same odds as before, and would count the cumulative number of turnaways. If the spill rate is small, the spill model approximates the multiple spill behavior fairly well. From this formulation we get the number of denials per customer as the power series sum:

$$Pu = Pd^{1} + Pd^{2} + Pd^{3} + Pd^{4} + \dots \qquad (4.5)$$

Where Pd^{n} approximates the probability of being turned away from n consecutive departures. We call Pu the ultimate denial rate. Mathematically Pu is still an underestimate of the number of denials per boarding because the knowledge that there are spilled customers is not included in the calculation of the two-spill probability (≠Pd2), but such an approximation should be valid over normal load factor ranges. This spill model ultimate denial rate is the number we shall use in the demand model as the probability of enduring one extra headway of wait. We will see in section 4.2 that the case of a Gaussian distribution with u/o =2 describes airline demand fairly well. In this light the values for Pu in table 4.1.5 will be used in the demand model (3 2) when it is exercised in chapter 5. For algebraic convenience we will employ the approximation

$$Pu = 2.5 \ln^{9} \qquad (4.6)$$

This approximation was obtained by curve fits of the numbers in table 4.1.5. Quite a few more complicated functions and other exponents on load factor were tried, but this approximation seemed best. Over the critical range of load factor, 0.6 to 0.8, the accuracy is acceptable.

Table 4.1.3: Space Unavailability for the lost turnaway model

nominal load factor	Gaussian Nominal Demand Distribution μ/σ 1.0	1.5	2.0	2.5	3.0	10.0
.50	.083	.020	.004	.001	.000	.000
.55	.116	.035	.011	.003	.001	.000
.60	.151	.056	.021	.008	.003	.000
.65	.186	.079	.036	.017	.007	.000
.70	.221	.105	.054	.029	.016	.000
.75	.254	.132	.076	.045	.028	.000
.80	.286	.159	.099	.065	.044	.000
.85	.317	.187	.124	.087	.063	.002
.90	.346	.214	.149	.110	.085	.007
.95	.373	.240	.174	.134	.108	.019
1.00	---	.265	.199	.160	.133	.040

Table 4.1.4: Space Unavailability for the spilled customers model

nominal load factor	Gaussian Nominal Demand Distribution μ/σ 1.0	1.5	2.0	2.5	3.0	10.0
.50	.108	.022	.004	.001	.000	.000
.55	.198	.046	.012	.003	.001	.000
.60	.284	.084	.026	.009	.003	.000
.65	.374	.138	.051	.020	.008	.000
.70	.463	.215	.092	.042	.020	.000
.75	.531	.310	.155	.078	.042	.000
.80	.578	.416	.248	.141	.082	.000
.85	.605	.519	.370	.244	.158	.002
.90	.614	.607	.508	.396	.297	.009
.95	.710	.674	.640	.579	.513	.047

Table 4.1.5: Ultimate Denial Rate
Gaussian Nominal Demand with $\mu/\sigma = 2.0$

nominal load factor ln	ultimate denial rate	Approximation $2.5 \times ln^9$	Space unavailability Pd
.50	.004	.005	.004
.55	.012	.012	.012
.60	.027	.025	.026
.65	.054	.052	.051
.70	.101	.101	.092
.75	.183	.188	.155
.80	.329	.336	.248
.85	.587	.579	.370
.90	1.033	.969	.508
.95	1.780	1.576	.640

Crucial range of ln is .55 to .70

Implications of the Demand Model for Turnaway and Spill

It turns out the combination of the demand model and the ultimate denial rate has implictions about the rate at which customers turned away from their preferred departure are retained by the system. It behooves us to explore the behavior implied by our models to see if it is reasonable.

We do so by a case study for a market with service defined by a fare of $75., a frequency of 5.7, a block time of four hours, a value of time of $10./hr. and a capacity of 100 seats total. We use a Gaussian demand distribution with $\mu/\sigma = 2$, and the demand equation from chapter 3 (equation (3.2)). Looking at Table 4.1.6 we see a series of a priori demands in column (1). These are not the actual traffic that occurs with capacity set at 100 but rather they are the mean demand u that would occur if capacity were so large there were no denied boardings. Column (2) shows the turnaway model space availability for the Gaussian O(d) and a capacity of 100. This we take as our a priori estimate of the rate of denial for the preferred departure. The traffic according to our demand model reacts to this denied boarding rate, and μ is reduced. Iterating between estimates of traffic (adjusted for denial rate) and predictions of ultimate denial rates (using the adjusted traffic) we get a new observed traffic μo in column (3). This is what our demand model calls the demand. With capacity at 100 and load factor at μo, the ultimate denial rate was Pu from column (4).

We can now compare the losses as estimated by the demand model and the number of turnaways. The demand model losses are the difference between columns (1) and (3):

$$\text{losses} = (\mu - \mu o)$$

The turnaways are those losses plus the ultimate denial rate for the traffic which remains:

$$\text{turnaways} = (\mu - \mu o) + Pu \cdot \mu o$$

The ratio of losses to turnaways is the rate of loss implied by th demand model. The rate is shown in column (5) of table 4.1.6.

The combination of the demand model and our more detailed load factor submodel appears to cantain no unhappy surprises. Column (5) shows a constant fraction of lost turnaways with load factor. The losses appear to be reasonable. For instance, at a measured load factor uo of 67.6%, we see that there could have been 70 customers (u) had space been available in the peaks. Of the people turned away, 33% failed to take a later departure.

Notice that the problem now has three means. The first mean is the mean of the Gaussian single departure demand O. The second mean is μo, the mean of the traffic as predicted by the demand model with iterations for the denial rate. Finally there was another mean which was the mean of the distribution of spilled demands, of Om. This μm was above μ (see figure 4.1.2). In fact for all reasonable distributions, μm is above the a priori μ of column (1). Of these three means, the only observable number is μo, mean onboard load.

Table 4.1.6: Comparison of Denied Boardings Lost for Demand Model under Typical Circumstances

Capacity = 100; $\mu/\sigma = 2.0$

(1) original nominal demand μ	(2) turnaway model space availability Pd	(3) ultimate mean demand μ_o	(4) ultimate denial rate Pu	(5) Fraction of denials lost % lost
60	.017	59.3	.023	.34
65	.032	63.6	.043	.34
70	.050	67.6	.074	.33
75	.071	71.0	.115	.33
80	.095	74.0	.166	.33
85	.119	76.6	.227	.33
90	.145	78.7	.290	.33
95	.170	80.6	.359	.33
100	.195	82.3	.433	.33

(1) $\mu = (75+10\cdot(5+4\cdot Pu^9))^{-1.5}$ for Pu=0. (equation 3.2)

(2) Pd for Gaussian demand with $\mu/\sigma = 2.0$ as in lost turnaway model on table 4.1.3

(3) $\mu_o = (75+10\cdot(5+4\cdot Pu^9))^{-1.5}$; Pu from (4).

(4) $Pu = 2.5\cdot\mu_o^9$; load factor for capacity = 100.

(5) $\% = (\mu-\mu_o)/(\mu-\mu_o+Pu\cdot\mu_o)$

Table 4.1.7: Buffer Space for Fixed Service Standards
for explanation, please see text

buffer for a Pd of 2% according to spill model

μ/σ	load factor	buffer in σ's
1.5	.49	1.56
2.0	.58	1.45
2.5	.65	1.35
3.0	.70	1.29

buffer for Pd of 10% according to spill model

μ/σ	load factor	buffer in σ's
1.0	.48	1.08
1.5	.61	0.96
2.0	.70	0.86
2.5	.76	0.79
3.0	.81	0.70

This case study has served to clarify some fine points involved in the combination of the demand and denial rate models. The two models are not perfectly mated by nature, and some care must be taken to use the most appropriate measure of denial rate from the load factor submodel for the demand model.

Buffer for Constant Denial Rate

We think of the difference between the capacity and the mean demand as the buffer space. As one would expect, the buffer space necessary to achieve a standard of Pd goes down with the standard deviation of the demand, σ. This can be seen from the first and second columns of table 4.1.7. The first column shows that as σ goes down (or μ/σ goes up), the load factor allowable for a specific denial rate goes up. What is surprising is that the buffer space necessary goes down even when it is nondimensionalized by the standard deviation. The third column of table 4.1.7 shows a decrease in buffer space for a fixed denial rate beyond linear adjustments for changes in σ. This occurs because the tail of the Gaussian distribution represents different fractions of μ for different values of σ. (1)

The point may seem unnecessarily technical, but it is natural to assume that a standard of service in terms of denied boarding can be specified by stating a buffer space measured in standard deviations of demand. Indeed both this author and Douglas and Miller made just that assumption at first. It proves inadequate

4.2 Estimating Variability of Demand

The preceding section discussed in detail the theory of demand distributions for a series of departures and the consequent ultimate denial rate. The fundamental input information for the process was O(d), the distribution of the demand d preferring a randomly selected departure. The statistic we can measure which is closest to O(d) is the distribution of measured load factors over a set of departures, O(l).

As we can see from figure 4.2.1 in the case of airline passenger transportation the distribution of load factors O(l) takes on a nearly Gaussian shape with μ/σ near 2.0. The Gaussian shape occurs because the variability of demand is composed of a number of independent influences.

Random and Cyclic Variability

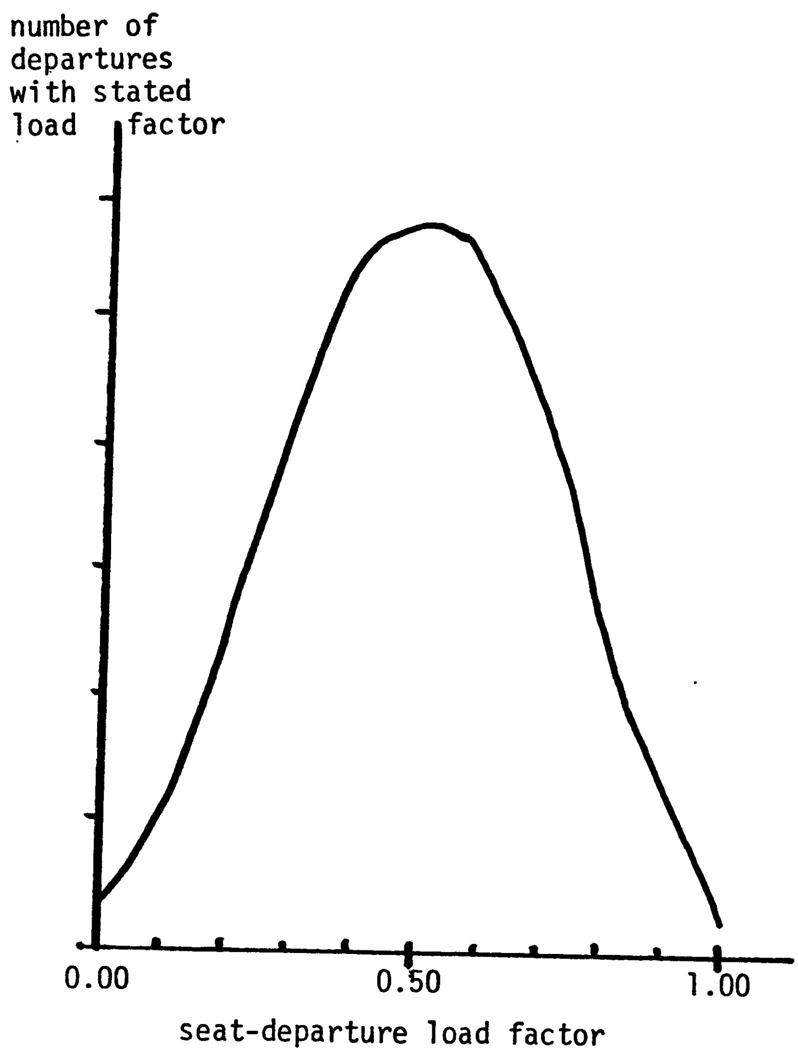
Our analysis of demand variability makes the distinction between cyclic and random demand variations. Cyclic variations could be more closely matched if the supply system found it economic to do so. Random variability is entirely outside of design efforts to reduce.

Cyclic variations are those variations which can be observed

---

(1) For a rigorous mathematical demonstration of this, see appendix D.

Figure 4.2.1: Frequency distribution of load factors from service segment data on trunks and locals, Feb 1976 (ref 40)



by adding up the loads from large numbers of departures by category. For instance weekly cycles are observable by comparing Monday's total traffic to Tuesday's and so on. Cyclic variability is part of O only because we have defined O to apply to all departures in a schedule without entering into time of day issues. An important consequence of this definition of cyclic variability is that it should be proportional to the mean demand. If we call the contribution to the standard deviation of demand by cyclic variability σc and the mean of demand μ, then $\sigma c/\mu$ should not change with μ.

This is not the case with the random variability of demand. We will call the contribution to the standard deviation of demand due to variabilities which do not scale with the mean σr. Random variability σr averages out whenever more than one observation is added together. In fact, random variability is the variability which exists when the cyclic variability is removed. Cycles may be used to predict the expected value of the demand for a departure for a Tuesday in January at 2:00 p.m. Then the variation among the four Tuesdays is random. Random variability is very nearly the variability which averages out whenever you try to observe it.

This is the important aspect of random variabilities: they average out. They average out not only in observations but also in practice as load sizes get bigger. As we shall develop shortly, random variabilities do not scale as the mean demand u. They scale with its square root. This is an interesting design influence and motivates much of the effort spent determining how much of observed load factor variability is cyclic. Table 4.2.1 shows that a market of 10 passengers a departure may have variabilities both cyclic and random requiring a 50% load factor while a 100 passenger market can survive a 65% load factor for the same standard of service.

<u>Summary of Issues</u>

We can summarize the situation for scheduled air transportation as we have drawn it so far: We can use mean load and load factor to deduce rates of denied boardings. The deduction of denied boarding rates depends on assuming a Gaussian distribution of demand for a departure with a known mean and standard deviation per departure. The mean demand is approximately the average load. The standard deviation must be deduced from system averages.

The variability of demand can be classified as part cyclic and part random. Cyclic variabilities are those which scale proportionally to the mean demand. They are created by observable seasonal, daily, and hourly cycles. These cycles could be matched by cyclic changes in supply or in price, but there are compelling technical and institutional reasons why this is imperfectly done. To the extent that cyclic variations are not matched by changing departure capacities or prices, they become part of the standard deviation of demand in the load factor and denied boarding rate model. Because cyclic variability scales with load size, when cyclic demand variabilities dominate, small and large airplanes need the same

Table 4.2.1: Numerical Example of Load Factor Changes
for constant denied boarding rates

demand per departure μ	cyclic variance $\sigma_c = .354\mu$	random variance $\sigma_r = 1.71\sqrt{\mu}$	total variance $\sigma = \sqrt{\sigma_r^2 + \sigma_c^2}$	design load factor for 2% denial rate
10	3.5	5.4	6.44	.50
25	8.8	8.5	12.2	.58
50	17.7	12.1	21.4	.63
100	35.4	17.1	39.5	.65
200	70.7	24.1	74.7	.67
400	141.	34.1	145.	.68
700	247.	45.2	251.	.68
1000	354.	54.0	358.	.68

design load factor obtained from interpolated values of μ/σ at a denial rate of 2% from table 4.1.4. See also table 4.1.7.

load factors to achieve the same denied boarding standards.

In addition to cyclic patterns of variability there are variations of demand which have no correlation pattern among departures. Completely random variations in demand are hard to observe because aggregation of markets or departures will tend to average them out. No adjustment in scheduled capacity can anticipate these variabilities because they are predictable only as a distribution. According to accepted probability theory for groupings of independent statistical events, random variabilities scale as the square root (1) of the mean demand, so as mean demand grows, random variations as a percent shrinks. This allows larger aircraft to tolerate higher load factors without increases in the denied boarding rate.

The relatively greater difficulty of attaining high service levels at small load sizes is reflected in the load factor standards for small and large aircraft airlines in the U.S. The curve in Figure 4.2.2 has two distinct points representing load factor standards for 80 and 140 seat aircraft, taken as representative for the two classes of carriers. This curve then suggests what similar standards of denied boarding rates would imply at higher and lower aircraft sizes. The curve shows the effect of the different relative growth of cyclic and random variabilities in demand. The numerical values are only for the purposes of illustration.

<u>Estimates of Variability</u>

We can examine in greater detail the components of cyclic and random variabilities in demand. To do so will bring to light several points of interest with respect to reported load factor statistics. Appendix E estimates the size of cyclic variations of demand from the variations in traffic caused by the combination of monthly, daily, and hourly cycles. The estimate from appendix E is that cyclic variations are in the neighborhood of 40% of the mean demand. (2) We now move on to discuss and estimate the uncorrelated or random variability.

<u>Random Variability</u>

Random variability (σ_r) in demand is whatever variability remains after cyclic variations have been removed. We shall show here that random variabilities do not scale linearly with mean demand.

<u>Bernoulli Trial Model</u>

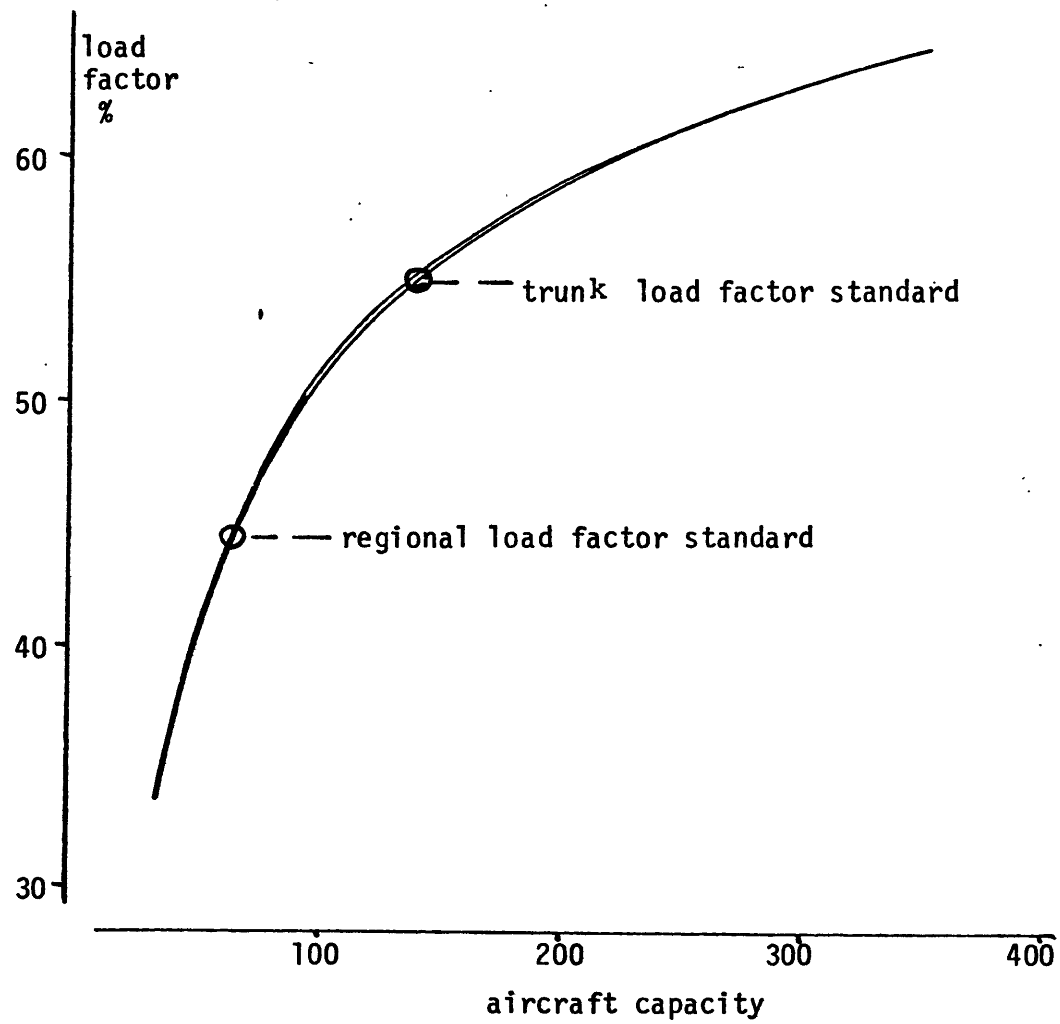
Part of the random variability σ_r is due to the process of a large group of people independently deciding whether or not to

---

(1) We imply here a specific model which is presented in detail below.

(2) This estimate has been substantiated in private discussions with American Airlines and by data from the Federal Register [40], Lankford [19], and preliminary estimates as part of a Ph.D. thesis by Tim Carmody at George Washington U.

Figure 4.2.2 Load Factors for constant flight availability



Load Factor standards from CAB DPFI, ref. 5

Ratio of cyclic to random variations determined so that the curve passes through the two circles drawn and also obeys the appropriate growth rules. (see text) take that trip at that time. A model of such behavior allows this contribution to or to be predicted. We assume that there is a very small probability P that each customer at an origin will want to use the departure in question. For a large number of people n we predict a mean of P*n customers for the departure. If the departures form an hourly schedule every day of the week and month of the year, the probability P would be predicted from cycles like those presented in appendix E. The result of n of these decisions to travel or not to travel is a well known probability distribution, called the binomial distribution. For large numbers n and small probabilities P the distribution is Gaussian with mean and standard deviation: (1)

$$\mu = P\ n$$
$$\sigma r = \sqrt{\mu}$$

The decision process is called a Bernoulli trial, so we refer to this as the Bernoulli trial model.

If one agrees with the logic that strangers' decisions to travel do not influence each other but that the mean probability is influenced by hourly, daily, or monthly cycles, then the prediction of σr equaling $\sqrt{P*n}$ is valid. This estimate falls directly out of the Bernoulli trial model without benefit of statistical calibration. This is just as well because observed load or load factor data almost always represent averages (or totals) over several cycles. Randomness will tend to average out, being less than 3% for any aggregation to over 1000 passengers.

Logistics Loss

The Bernoulli trial nature of demand is not the only component of variability which does not scale linearly with load size. We observe the variability of load factor distribution rather than of loads, so we must consider variations in capacity as well as demand. The concept of variability in capacity may seem odd at first, but it will be a particularly useful way to treat an otherwise awkward problem. The problem is that no airline has all possible aircraft capacities in its fleet. Most fleets are limited to between 1 and 6 aircraft types. For instance, American, TWA, Eastern, and United each have fleets with only 5 distinguishable capacities. This means that the airlines are unable to match capacities to mean loads exactly. It is possible to model this as a further fractional adjustment of load factor, one which is correlated with load size. One way to think of it is as a number of extra empty seats which randomly show up for a scheduled departure. We name this effect of missing target capacities the logistics loss. It is really caused by aircraft "integerization" effects.

One can get a feeling for the size of this effect for

---

(1) For the case where people travel in groups of m,

$$\sigma r = \sqrt{P \cdot n \cdot m}$$

airlines by examining the extra capacities necessary to meet a given design requirement. Figure 4.2.3 plots the next available capacity against load requirements for one airline. In the first line of figure 4.2.4, the shaded areas of figure 4.2.3 are reproduced on the horizontal axis. These shaded areas are the extra seats which must be scheduled in meeting the load requirements shown on the horizontal axis. Figure 4.2.4 shows that for the four trunks the logistics loss drops as a fraction of larger required load sizes. The treatment is not at all rigorous, but the variability of load factor due to logistics loss does contribute to σr.

We define the ratio of the required seats to the next available capacity as the seat factor. Figure 4.2.5 is the seat factor distribution found by sampling at 105, 115, 125,......,295 over all four of the lines in figure 4.2.4. By itself, this effect produces a standard deviation which is 10% of the mean at an average capacity of 200.

For illustration we will treat logistics loss as if it scaled in the same way as the Bernoulli trial losses. Thus when Table 4.2.1 assumed a σr 70% higher than that suggested by the Bernoulli trial estimates σ of the preceding discussion, that 7 % came from logistics loss.

One warning must be added. The name "logistics loss" might have been applied to the phenomenon variously referred to as deadheading, empty backhaul, or vehicle repositioning. While there is a loss in systemwide average load factor associated with such movements, it will be convenient to separate this irregular and local phenomenon from the considerations above. The discussion up to this point has involved the difficulty of obtaining high load factors in any market or link irrespective of its location in a network. Network losses associated with directional flows and backhauls or with sparse demands and weak links do not occur regularly over all markets or systems. Surprising as it may seem, mathematical assignments of aircraft to move networks of demands can achieve target load factors on nearly all links. (1) Such phenomena are not included in what we have called "logistics loss."

<u>Random Variability Conclusions</u>

For the modelling in later chapters, the problem of variability of demand which does not scale with the mean will be ignored. The concepts presented here are new and unproven. Present documentation is more on the order of evidence than proof In order not to jeopardize later conclusions by involving untested theories, the issue has been bypassed and a constant ratio of standard deviation to mean assumed for all load

---

(1) This occurs with symmetrical flows and aircraft capacities variable down to small capacities, as is the case for air travel. The author's experience in modelling U.S. domestic airline networks confirms this. Backhaul and deadhead links do occur in networks designed to serve sparse networks of demands. Cf. chapter 6, section 3.

Figure 4.2.3: Logistics loss illustrated

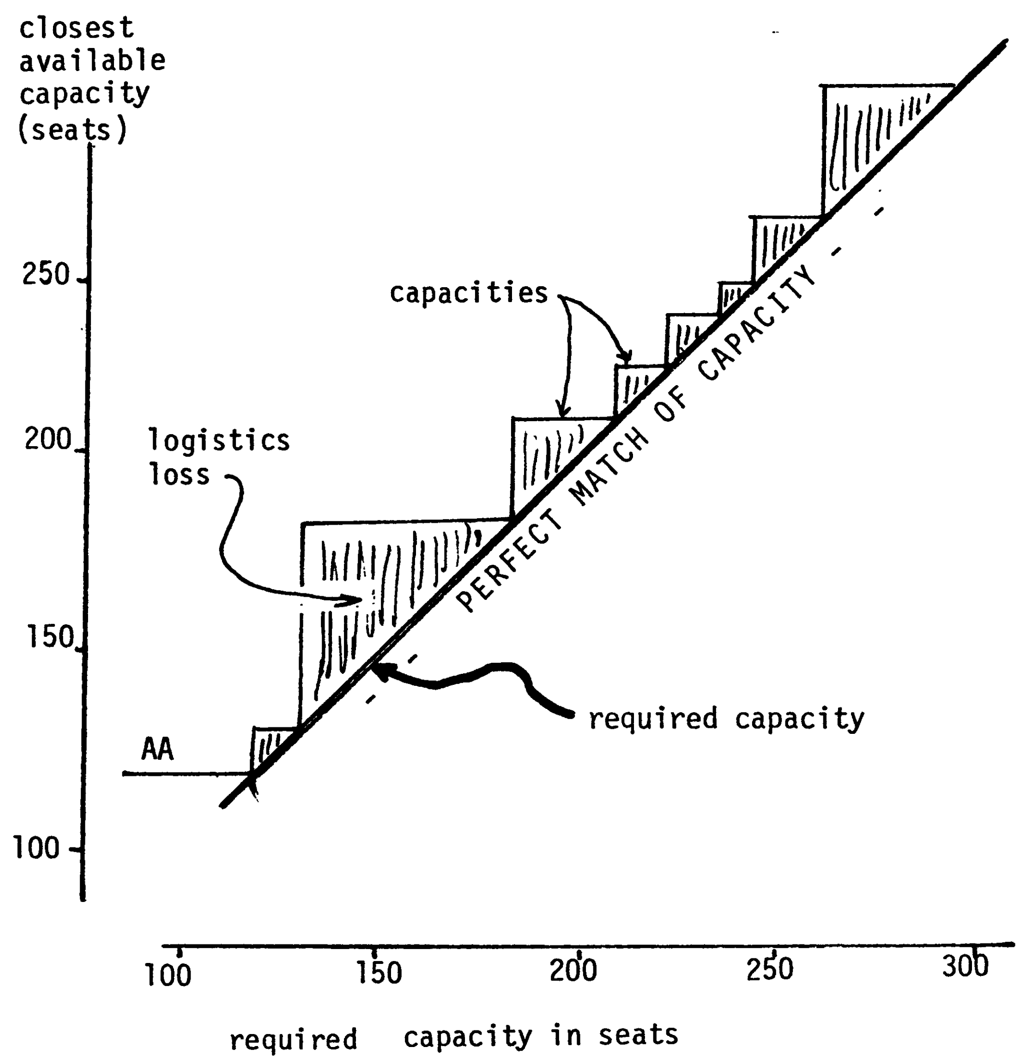


Figure 4.2.4 . Logistics Loss for four trunk airlines

1 or 2 aircraft capacities combined to equal or exceed required capacity

logistics loss is defined to be the extra seats beyond requirements

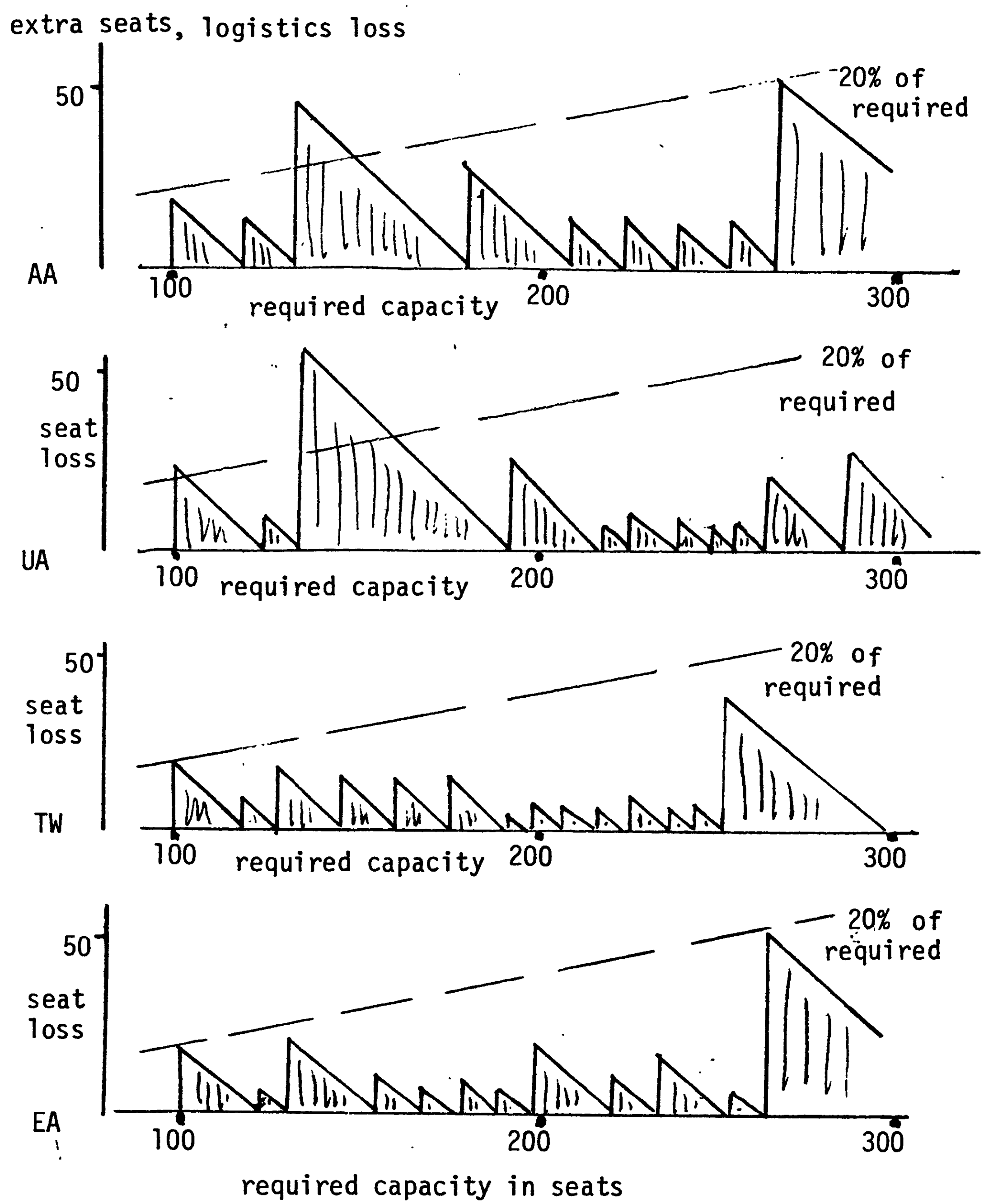
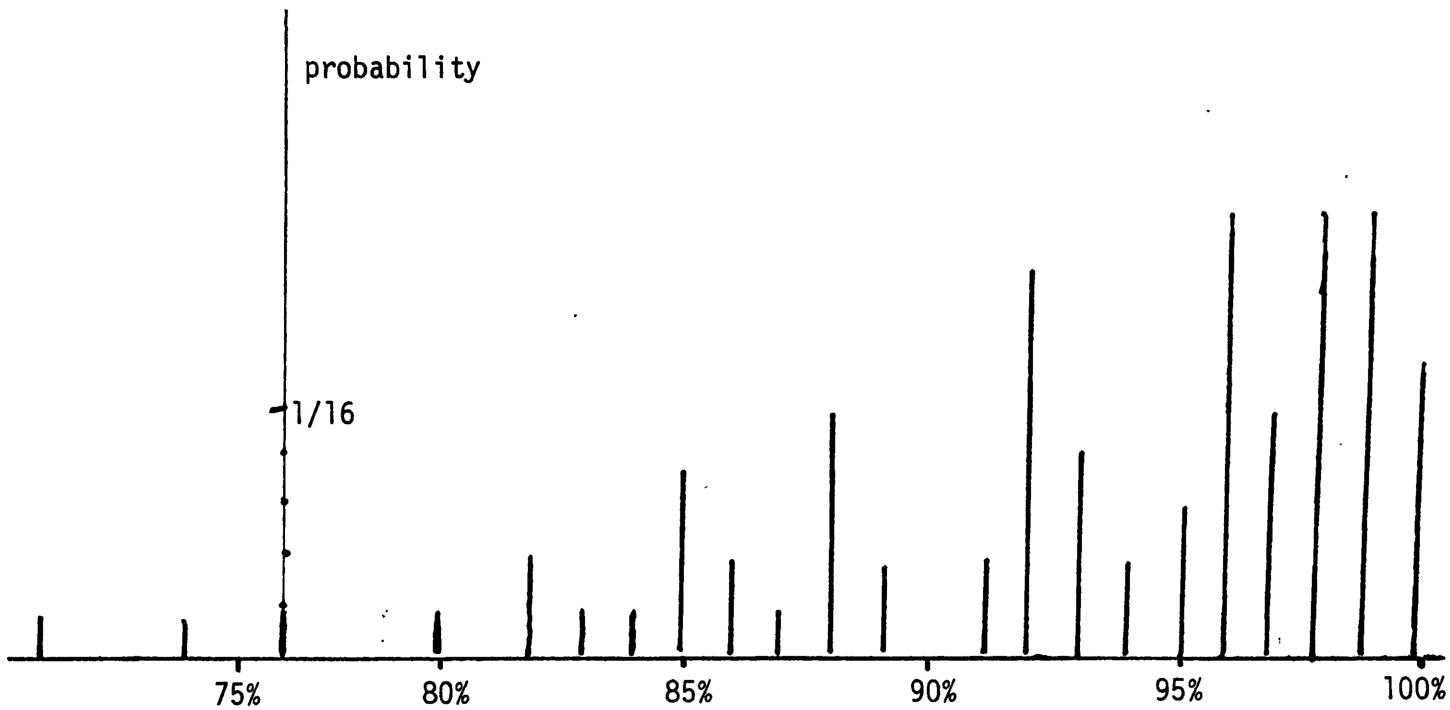


Figure 4.2.5

Figure 4.2.5: Distribution of Logistic Load Factors

(for detailed explanation, see text)



RMS σ is 10.4% of mean.

logistics load factor is demand size devided by next largest one or two plane capacity
Demand size sampled from 105 to 295 by tens. Capacities from 1975 fleets of
American, United, Eastern, and TWA (see figure 4.2.4)

capacities. That constant is $\mu/\sigma=2.0$, reflecting a slight upward adjustment of cyclic variability for the contribution from random effects. If scaling with size does behave as predicted, it serves to heighten the cost edge already obtained by larger aircraft capacities and thus the cost/service dominance of the largest competitor in the market.

So far, we have developed a detailed understanding of the load factor problem and how it is handled by the demand model. But as a practical matter the demand model does not fit the facts in the case of two similar services competing at slightly different load factors and prices. This problem of modelling competition by load factor is discussed in the next section.

4 3 Load Factor and Competition

Load factors influence the quality and cost of a transportation service. They also influence in an unusual way the form competition can take. This occurs because of the unusual nature of the reservations process as we have come to know it in air transportation. The phenomenon is named "cream skimming" although it is more nearly analogous to taking milk off the bottom.

Consider two firms offering nearly identical transportation services. Firm A offers a low price and operates at a high load factor. At least that is its intention. Assume that our expression from chapter 3 (equation (3.1)) for perceived price including delay time establishes a price of $100. (1) Firm B tries to design a competitive service with the same perceived price, but with a different arrangement of fare and load factor. B's service is priced higher, but has a greater space availability (lower load factor). At the same value of time, B's service is also perceived to price at $100. According to theory so far, A and B should split the market. In practice, a consumer making an individual decision to travel can always attempt to employ A (at low fare) and go to B only when A is full.

If the consumer has knowledge of actual seat availability instead of merely average denial rates, he can do better in the long run than either of the design perceived prices of A or B. If all consumers use departure by departure information, B will be unable to get any traffic as long as A has a seat to sell. In this case B faces a different demand distribution than A. B's distribution becomes the tail of the total market distribution, which A does not serve. We see in figures 4.3.1 and 4.3.2 what happens with a particular discrete distribution of demand. Firm A carries the majority of the market at a high load factor and correspondingly low mean cost (Figure 4.3.2A). Firm B gets only the people who cannot fit on A on big demand days. (For this illustration we have assumed these people are all willing to pay the premium to ride on B. Even then, B does not get half the traffic, although his service is of equal value according to the

---

(1) For some value of time, say $10/hr.

Figure 4.3.1: Sharing of a distribution of demand by day between two services of unequal price

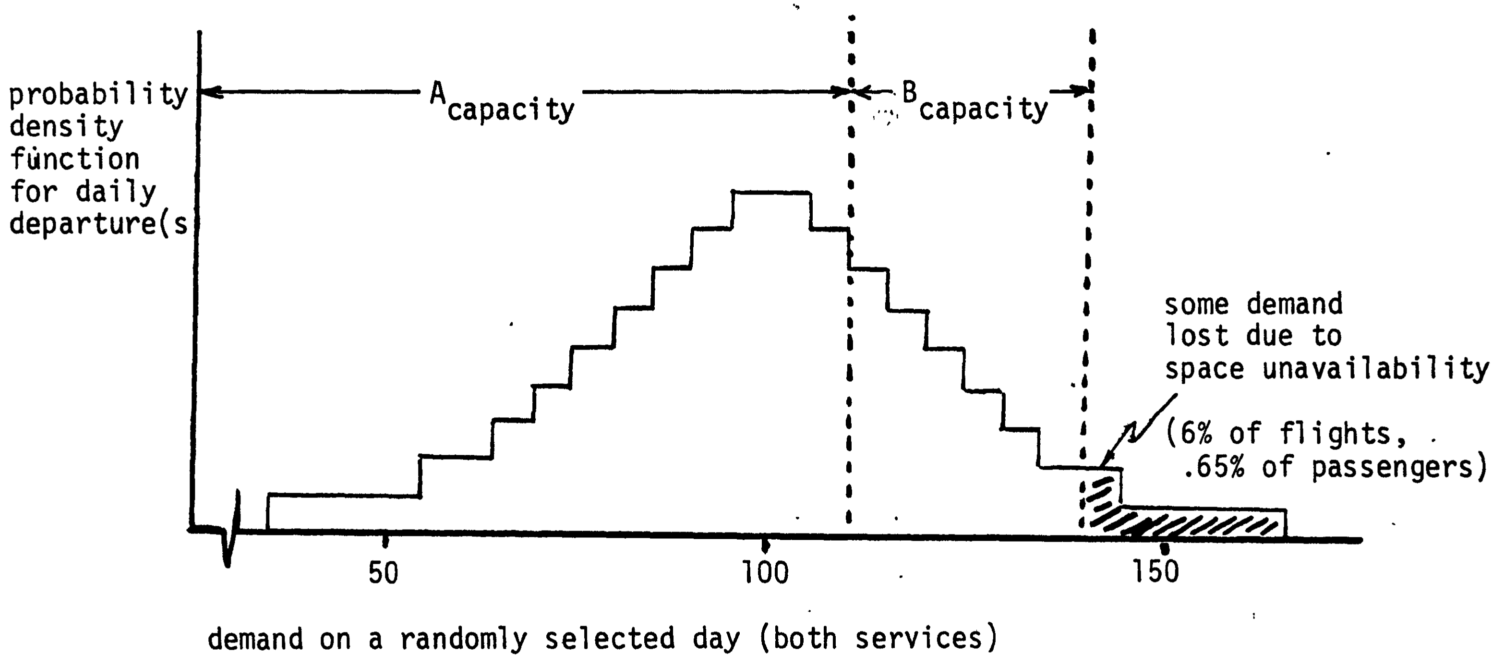
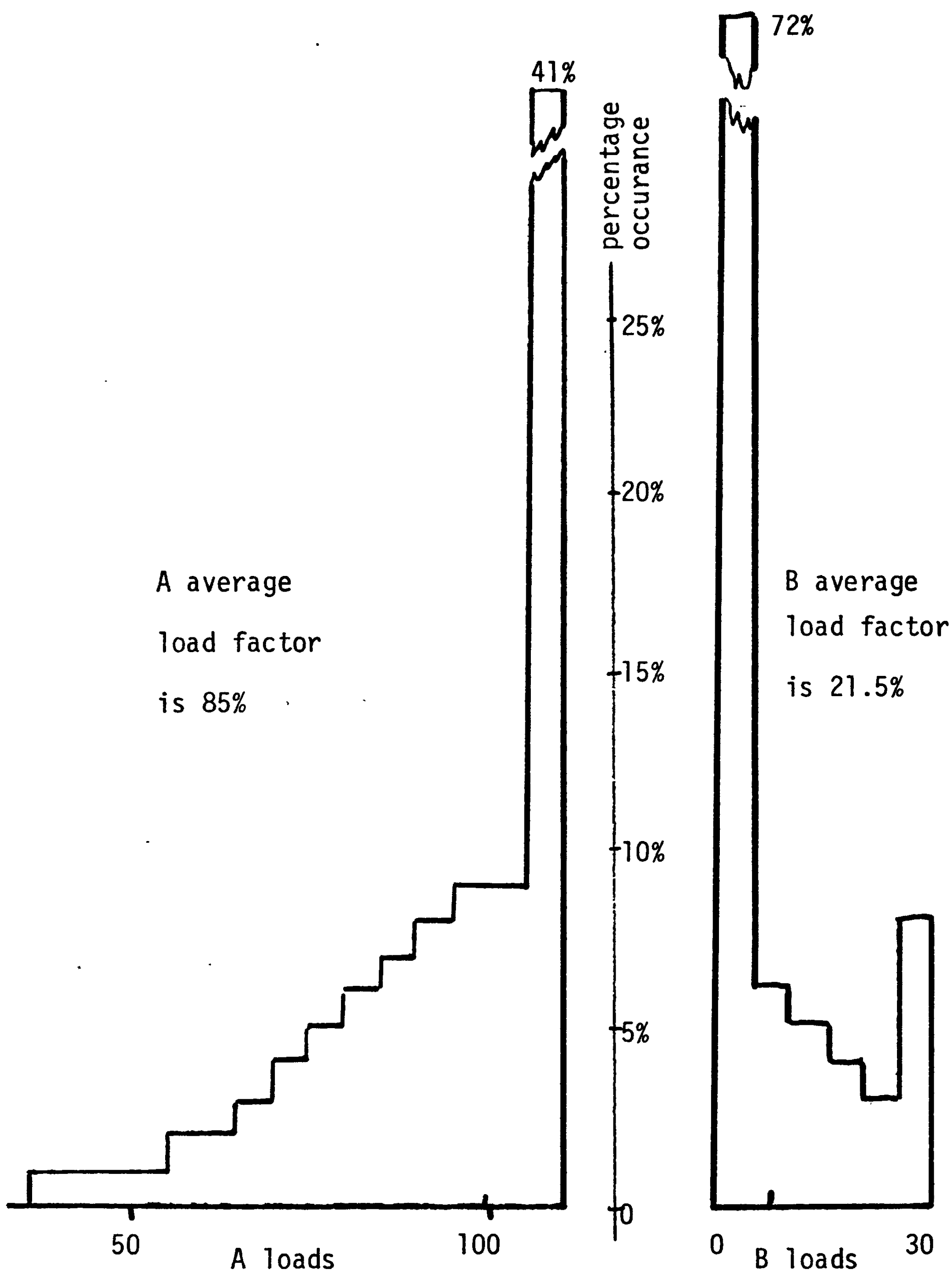


Figure 4.3.2: Distribution of observed loads for A and B for competitive case as illustrated on figure 4.3.1



demand model.)

This type of market behavior is possible because the transportation services are scheduled well in advance at a level of knowledge about demands which reflects statistical averages while the consumer's decision is made in a shorter time frame with greater knowledge of supply and demand for a departure. With greater knowledge, the consumer can get the better of the bargain The situation is nearly unique to air travel and is caused by the reservations process, which is an unusual part of the buying and selling transaction.

The result of such a situation is a marketplace tending toward high load factors and high rates of denied boardings. Moves by one competitor to adjust load factor up and price down must be met by the other competitor or he will be caught in the cream skimming effect illustated in 4.3.2. On the other hand, moves to slightly improve quality at higher price are foolish. (1)

Airlines find it difficult to overcome this effect of passengers using premium services only as backup. Round trip ticketing on one airline can influence passenger behavior by exposing them to a doubled probability of not being accommodated on the high load factor carrier. Other attempts to make customers experience the mean reliability instead of the specific seat availability are difficult.

Other modes have greater success in overcoming this problem. For package freight movements, reservation information is generally not available. Repetitive shippers learn to anticipate expected values of capacity availability. Thus the demand model as stated applies. In cases where knowledge of capacity availability is possible, consumers can be tied to the high quality services for all their shipments by preferential treatment (2) of regular customers during capacity crunches and by contractual commitments for regular carriage.

In the most open move to overcome this problem, ocean freight companies have attempted to freeze out high load factor/low price competition (tramps) by forming shipping conferences with service contracts which insist on the use of member firms. Such a contract is illegal in domestic freight (regulated by English common law), and furthermore such agreements are hard to imagine for individually ticketed

---

(1) Of course an airline could intentionally offer the backup service like B's at a very high price. A service with a 22% load factor, such as B's, would not have a fare which would allow a perceived price of $100, so some of the overflow would be lost. The possibility strikes this author as only an theoretical one.

(2) Such preferences are nominally illegal, but the author is assured by reliable sources that they are exercised.

(3) Ted Keeler at Berkeley suggests a "club dues" scheme for binding repetitive customers to an airline. This adjustment is

passenger travel. (3)

4 4 Load Factor Summary

The development in sections 1 and 2 of this chapter states that level of service measured as the probability of a customer being able to obtain a seat on his choice of departures depends not only on the mean load factor for the service in question, but also on the magnitude of the mean load on that service. The suggestion that percent variations of demand go down as mean demand rises is a new and radical one. The conventional assumption is that load factor alone is necessary for predicting service. Since no data is available for estimating the extent of this phenomenon and since statistical proof or even calibration is not possible as part of this work, the level of service measures used in chapters 5 and 6 do not make any adjustment for load size For these chapters the approximation in table 4.1.5 will be assumed to apply. This data has been summarized in equation (4.5)

---

outside of the rules of the marketplace as we are discussing them. It also may not be practical.

5 Isolated Single Link Service

5.0 Introduction and Problem Statement

Service between two cities is largely defined by the fare, frequency of departures, and load factor. (1) Each of these characteristics influences the utility to consumers, their benefits, and the amount of demand. These three characteristics also determine the costs and revenues for the operator and thus his profits or losses. The discussion below focuses on differing combinations of fare, frequency, and load factor for single link, single carrier air passenger service. We seek those arrangements which are in some sense optimal.

In the next three parts of this chapter we will discuss three classes of optima. First, we discuss what we believe to be the best representation of desirable market conditions, that of traffic maximizing at zero losses. We shall refer to this as the maximum traffic case. Second, we compare maximum profit optima to see if they too could be a reasonable description of existing conditions. Finally, we examine the condition of optimal service with subsidy, to see what might be achieved in that direction. This will be called the maximum surplus case.

Throughout this chapter we will specify optima in terms of frequency, fare, and load factor. Objective functions will deal in terms of profits, consumer surplus, and traffic. The arena for this discussion is a single isolated city pair market. We imagine a single airline serving origin-destination passengers nonstop between two cities. There are no network effects from other markets. A single schedule of service is offered at a single price. The demand is described by an aggregate demand function. We characterize the market by the distance between the cities and the density of demand for travel.

What is the nature of optimal services for markets of different distances and densities? Are optimal load factors constant across markets of different range or density? Can optimal frequencies be designated as a simple function of market density? Should fares change dramatically with market density? These are the questions we would like to answer before we start to explore the nature of service in a network.

<u>Supply Summary</u>

We may summarize what chapter 2 said about the cost of providing a schedule of service on a single air transportation link by the formula

$$COST=(c10+c11\cdot d)SEATS+(c20+c21\cdot d)FQ+(c30+c31\cdot d)PAX \qquad (5.1)$$

---

(1) load factor is an independently adjusted degree of freedom because capacity can vary by choosing different aircraft sizes. The number of intermediate stops is also important, but not in the single llink case here.

Where the following definitions apply:

COST is the total expense of operating the day's schedule of services

SEATS is the total capacity per day. (SEATS/FQ is the average capacity per departure.)

FQ is the daily frequency

PAX is the daily average traffic

d is the intercity distance

Chapter 2 derived values for the constants for 1976 U.S. domestic air travel:

$$\left.\begin{array}{l} c10=\$3.27 \\ c11=\$0.0176 \end{array}\right\} \quad c10+c11\cdot d=c1$$

$$\left.\begin{array}{l} c20=\$379.8 \\ c21=\$0.816 \end{array}\right\} \quad c20+c21\cdot d=c2$$

$$\left.\begin{array}{l} c30=\$12.64 \\ c31=\$0.008 \end{array}\right\} \quad c30+c31\cdot d=c3$$

This simple formula (5.1) recalls several interesting results.

First, total capacity (SEATS) and frequency (FQ) can be adjusted independently. Changes in aircraft capacity are allowed over a continuum. There may be upper and lower bounds on the vehicle capacity, but between those bounds adjustments of frequency and capacity can be made independently.

Second, the average cost per seat falls wth the number of seats. This is because the cost of frequency (c2) is spread over more seats. Frequency takes on the aspects of a fixed cost in this respect; it does not depend on capacity or traffic. The concept if a cost of frequency is fundamental to our discussion. The cost of frequency is the cost of the vehicle trips in the schedule but not the cost of the seat trips or passenger trips. We are reminded that vehicle costs are that part of aircraft costs not associated with capaciaty.

Third, we are reminded that this is the cost for a single scheduled transportation service over a link assuming the time of day distributions of PAX, FQ, and SEATS have been worked out and vehicle utilization is known.

Demand Summary

We may summarize our demand formula from equations (3.1) and (3.2) in chapter 3. For city pair market:

$$PAX = k1\cdot(FA+v(tb+(k2+k3\cdot LF^{q})/FQ) \qquad (5.2)$$

The following definitions apply:

FA is the fare

LF is the market average load factor

k1 is the market demand density

k2/FQ is the expected value of the time lost due to denied boardings. k2=5.7 in the calibration discussed in appendix C.

k3 LF^9/FQ is the expected value of the time lost due to denied boardings. k3=57 is derived from the approximation that the denial rate is $2.5 \cdot LF^9$ (from equation (4.5) in chapter 4) times the headway, which is 22.8/FQ.

α is the market elasticity with respect to total perceived price. Appendix B calibrated α=-1.5.

tb is the physical travel time or block time. From chapter 2, tb=.37+d/507

We recall from chapter 3 that we must explore this formula for a spectrum of different travellers' tastes, represented by their value of time, v.

In the expression above, frequency affects not only the schedule displacement time but also the penalty for denied boardings which result from too high an average load factor.

Load Factor Summary

Aside from its ramifications in the demand model, the defintion of load factor is a separate condition in the matching of demand and supply. This provides us with a third equation:

$$LF=PAX/SEATS \tag{5.3}$$

It may seem surprising to go to the trouble to explicitly state the definition of load factor. However, doing so points up one of the major issues involved in the matching of supply and demand. Is load factor a result of the matching of supply and demand or is it part of the matching process? In our formulation, the matching of SEATS to PAX is not an incidental fallout of the supply and demand interraction. Load factor is a design parameter addressed by the supplier of the schedule. As a measure, it is more fundamental than the measure of SEATS. That is, solutions to the single link optimum design will state load factor and passengers and leave capacity to be deduced.

Objective Functions

The matching of supply and demand in a single carrier single link transportation market is not the simple process of setting PAX equal to SEATS. First one must translate the many and conflicting objectives of society and the actors involved into a mathematical objective function. Then one must find the optimum conditions subject to the three equations above (5.1-5.3) and whatever additional constraints the statement of operating and optimal conditions adds.

We need a precise mathematical definition for the conditions for optimality. Unfortunately, the objectives of society are not strictly the maximum benefits with respect to the production and consumption of goods, nor are the objectives of the firm strictly the maximization of profits. Consumer benifits and corporate profits are only leading members of a list of objectives.

Our discussion will cover three distinct objective functions. One is a mixed objective, and the other two are more strict statements of producer and societal benefits.

The first and most important objective we will discuss is the maximization of traffic (1) at conditions of zero loss. (Zero loss and zero profit points are the same, since we have included a fair return on investment in the costs.) This statement of the objective is as good an approximation of today's compromises as is possible to state mathematically. The condition of zero loss is one of justice to the producers. For air travel it is a practical constraint except in a few low density markets. The condition of maximum traffic is as good as any other in achieving benefits to consumers.

The second objective we will examine is the one of maximum profits for the monopoly producer. Since this case produces huge profits, it is of little practical relevance.

The third objective we will examine is the maximization of total surpluses. That is, the maximization of the sum of consumer surplus and the offsetting losses that the producer might experience. This is the classical societal optimum for an industry, but since it turns out to involve substantial losses for the carrier, it is of little relevance to U.S. industry conditions.

It may seem strange that we do not discuss equilibrium in competitive markets. We avoid doing so because a profit maximizing treatment of competition will tend toward monopoly. We deduce this from the economies of scale in the cost of producing the intermediate good of a transportation schedule (as defined in chapter 2) and the unstable competitive situations discussed in chapters 3 and 4. In the presence of such economies of scale and product matching tendencies, any model of stable competition in an isolated market must rely on assumptions of some form of "polite" or constrained behavior among a few competitors. Such oligopolistic competition might well exist in an unregulated transportation environment, but it is outside of our scope to predict its nature here. We will reserve discussion of the viability of competition in general until after the development of the final arguement in favor of single large operators, the network effects.

5.1 Maximum Traffic Optima

Although differences of opinion exist about the objective(s) of the managements of large industries, there is a consensus that the avoidance of losses is at the very least a constraint on behavior. In the presence of such a constraint, one objective which should be instrumental in achieving both the desires of managment and welfare and justice for society is the maximizing of total passengers carried. This is the scenario we shall explore first. It seems the best point of reference for later

---

(1) We make the careful distinction in terms between demand, which may not all be satisfied, and traffic, which is the passengers actually carried.

cases. It is also the closest to realistic conditions, competitive, regulated, or otherwise. We consider isolated city pair markets. We employ the cost structure developed for the air transportation example in chapter 2 and echoed in equation (5.1) above We employ the demand model developed in chapter 3 and summarized in equation (5.2). This predicts the traffic that occurs given the fare, frequency, and load factor of the service. We require for this section that the fare exactly cover average costs including a fair return on investment. This produces zero losses. Our objective is to inquire after the technical performance which maximizes traffic.

The constraint that fare equal average cost produces zero (excess) profits. This constraint provides and expression for the fare:

$$FA=COST/PAX=c1/LF+c3+c2\cdot FQ/PAX \qquad (5.4)$$

The first two terms of this expression are the cost of capacity and the cost of passenger handling. We will see these two terms in exactly this form as part of the fare for all our objective functions. These terms are the cost that can be definitely associated with each passenger. The final term is the cost of the frequency of service, averaged over traffic.

For the case where all customers have the same value of time, maximizing traffic also maximizes consumer surplus, as discussed in chapter 3. For the case where customers have a range of values of time, maximizing total traffic is not the same as maximizing total consumer surplus. The two optima are close to each other, but they do not coincide. Maximizing consumer surplus in a market with mixed values of time favors high value of time customers more heavily that low value of time customers. It is not the purpose of this work to comment on the distribution of wealth or benefits in society, nor does the calibration of our demand model permit us to assess these effects with any pretense of accuracy. Therefore we examine the condition of maximum traffic with the understanding that it is a simplification of the condition of maximum surplus at zero loss. This is not to endorse the implications for the distribution of welfare among the population. We have already pointed out in chapter 3 that a distribution which is unfair in one sense or another must inevitably exist.

Before we state the mathematics of the problem, it will be useful to define two further variables which can serve as notation for the rest of this chapter. The first we wish to define is the total travel time:

$$TT=tb+(k2+k3\cdot LF^{q})/FQ \qquad (5.5)$$

The second is total perceived price:

$$PP=FA+v\cdot TT \qquad (5.6)$$

The problem now becomes:

Maximize PAX(FA,LF,FQ)

Subject to:

$$PAX=k1 \cdot PP^{\alpha} \quad (5.2)$$

$$LF=PAX/SEATS \quad (5.3)$$

$$FA=c1/LF+c3+c2 \cdot FQ/PAX \quad (5.4)$$

It will be illuminating to eliminate the varible FA by using (5.4), but further simplification produces no great insights.

We substitute (5.4) into (5.2) and take the derivative dPAX/dFQ=0 (1) in order to find the conditions for optimum PAX. This produces:

$$c2/PAX-(dPAX/dFQ) \cdot c2 \cdot FQ/PAX \; -v(k2+k3 \cdot LF^{9})/FQ=0$$

But dPAX/dFQ=0, so this can be rearranged to establish frequency:

$$FQ=PAX^{0.5} \; (v(5.7+57 \cdot LF^{9})/c2)^{0.5} \quad (5.7)$$

In this final step we have substituted the values for the k's. From (5.7) we can see that optimal frequency should rise as the square root of traffic (PAX), provided the load factor is constant.

This is as far as we can profitably go analytically. Further algebra results in an equation which must be solved numerically. Instead, we have done a more intuitive numerical search for the optima in the next section.

However, insights can be gained into the nature of the load factor optima by a similar derivation. The condition dPAX/dLF=0 produces from equation (5.2):

$$-c1/LF^{2}+v \cdot k3 \cdot 9 \cdot LF^{8}/FQ=0$$

The real roots of this are all

$$LF=(c1 \cdot FQ/9k3 \cdot v)^{0.1}$$

Substituting for FQ from (5.7) we get:

$$LF=PAX^{0.05} \cdot (c1^{2} \cdot (5.7+57LF^{9})/263169 \cdot v \cdot c2)^{0.05} \quad (5.8)$$

Once again explicit values have been used for the k's. We note that if LF on the right hand side does not change much, LF on the left hand side rises only as the twentieth root of traffic.

Values for COST, SEATS, PP, TT, and FA may be deduced from the values of PAX, FQ, and LF.

---

(1) Our typeface does not include the partial derivative sign. Within the text, we rely on the symbol "d" for partial derivatives.

Numerical Studies of Maximum Traffic Case

One specification perpetually causes confusion both comceptually and in practical expositions. The constant k1 in the demand equation (5.2) specifies the market density or potential demand. The actual traffic in terms of passengers can only be stated in conjunction with a level of service in terms of fare, frequency, and load factor. The problem is, how to specify k1? The most convenient approach is to adjust k1 so that the maximum traffic optimum just obtains the stated market size. With this definition, market size is the actual traffic instead of some theoretical conditions. A 400 passenger market is understood to have 400 passengers at the maximum traffic optimum. The k1 would be the smallest k1 which could produce 400 passengers. We can use this simplification of terminology because minimizing k1 for constant PAX is the same as maximizing PAX for constant k1 . From (5.2) $PAX/k1=PP^{\alpha}$. Optima for either are derived from dPP/dFQ=0 and dPP/dLF=0. An 800 passenger market would have 800 passengers at its maximum traffic optimum, but k1 would not be exactly double the k1 for a 400 passenger market because the service at the optima would be different. For convenience we shall follow this approach throughout this chapter and the next. Market size will be specified in terms of the passengers carried at the maximum traffic optimum.

We take as the central case for our numerical studies a 400 passenger, 800 mile market. For a start we will use a value of time of $10/hr. (1) This is a fairly dense air market of typical length for U.S. domestic trunk operations. For this market 400 passengers can be obtained with the smallest k1 (k1=352441) at a fare of $59, a frequency of 5.2, and a load factor of .665. (2) Aircraft size works out to be 116 seats. These values were obtained by numerical search for the maximum traffic optimum.

We are not concerned that the frequency is a fractional number. First of all, rounding off to 5.0 flights a day (and adjusting the load factor and fare to reoptimize) causes only a 0 025% (one-fortieth of 1%) drop in traffic. This is unimportant. Secondly, we do not want answers to be influenced by integerization effects because such influence may obscure long term trends. 5.2 flights per day may be interpreted as an annual average, which may certainly be fractional.

We are also not concerned that the optimal load factor (.665) is higher than what has been normal for such a market in

---

(1) Value of time was shown by DeVany [35] to be near $13 (1976$) This was a single value for markets involving air travellers participating at 1968 prices and service levels. A lower figure is used here as representative of modern markets with a greater number of low value of time trips being considered.

(2) Cost and demand constants (c's and k's) have been given values as indicated in equations (5.1) and (5.2)

recent U.S. domestic trunk operations. (1) Neither our demand model nor our cost model are accurate enough to permit us to draw conclusions from numerical comparison with real world operations. Nor do our single-carrier, zero-profit, maximum-traffic conditions necessarily represent existing conditions. What will be meaningful are changes from this optimum in different scenarios. We will soon see that load factors at optimum climb with market size. This will be a valid observation which should have implications in the real world. That they climb above some number like 75% is not significant.

Changes in Optimum With Market Density

There are three ways to accomodate increases in traffic in a market. Frequency can grow, load factor can rise, or aircraft capacity can increase. Table 5.1.1 shows that for our conditions of cost and value, all three degrees of freedom come into play at significant levels. Optimal frequency, fare, and load factor all change with market density. This is one of the conclusions of our work.

If one variable can be ignored, it should be changes in load factor. Load factor changes only as $PAX^{.06}$. (2) This result was anticipated in equation (5.8) in the analytical section. Frequency is the most important adaptation. Frequency rises at just over the square root of traffic ($PAX^{0.58}$), bearing out the analysis in equation (5.7). Vehicle capacity is nearly as important. Capacity changes as the cubic root of traffic ($PAX^{0.36}$). The possibility of capacity changes has been ignored in much recent literatere on the subject. (3) Changes in capacity coupled with modest changes in load factor produce a small reduction in fare. Fare drops as $PAX^{-0.12}$.

These relationships are approximate and depend on our specific cost and demand assumptions. The significant conclusions are (1) the optima do change broadly with market density, (2) all three of frequency, capacity, and load factor rise together, and (3) load factor changes the least. In additon, increases in load factor and in vehicle size reduce average cost and thus fare. Fare falls as traffic rises. This favors production by a single carrier in a market, as we have noted.

---

(1) As this is written, load factors appear to be arriving at this range, but the last fifteen years of regulated service has seen most load factors in the .50-.55 range.

(2) Exponents in this and the following numerical studies were obtained by solving lnLF=K+x·lnPAX for x using two widely spaced values for LF and PAX from the appropriate table. Curve fit was plotted and seen to be reasonable in all cases. This method was convenient and suitable given the qualitative nature of the conclusions.

(3) Specifically Douglas and Miller [10], Dorman [9], but <u>not</u> Anderson [36].

Table 5.1.1: Maximum traffic Case
Optima at Different Market Sizes

800 mile market, value of time = $10/hr, optima at π = 0.

traffic	fare	frequency	load factor	capacity
50	$85	1.8	.60	47
100	72	2.4	.615	67
200	65	3.6	.64	87
400	59	5.2	.665	116
800	54	8.0	.70	147
1600	51	11.8	.725	187
3200	48	17.6	.75	243

1400 mile market, value of time = $10/hr, optima at π = 0.

traffic	fare	frequency	load factor	capacity
50	$112	1.4	.615	58
100	98	2.0	.635	79
200	89	3.0	.66	101
400	81	4.4	.69	132
800	75	6.6	.715	170
1600	71	10.2	.75	209
3200	67	16.0	.785	255

Changes in Optima With Distance

At longer distances the wait time and denied boarding penalties are a smaller fraction of total time and cost. This means more can be saved by using larger aircraft and higher load factors than is lost due to increased displacement times. Thus at the optima, frequency should fall and load factor rise as distance increases.

Table 5.1.2 details our case study for a value of time of $10/hr and a traffic of 400 passengers. As expected, frequency falls and load factor rises with distance. Corresponding to a fall in frequency, capacity rises. The net result is that fare is not proportional to average seat cost for some standard size vehicle, but has a lower slope and a higher intercept, as shown in figure 5.1.1.

The desirability of lower frequency and higher load factors at longer distances escapes most analysis. One source of error is assumimg constant vehicle size. Fixing vehicle size nearly freezes frequency with distance (for constant market size). Such vehicle size constraints have been employed by economists ([9],[10],[13]). Another viewpoint was expressed by the CAB in the domestic passenger fare investigation [5]. In that study, load factor was fixed at all ranges, but the fare formula was not linear with distance. The combination implied changing to larger aircraft at longer distances.

In our study of the effect of market distance, the important variations are in fare and frequency. Fare is the dominant parameter in this case, rising as the square root of distance ($d^{0.55}$). Frequency falls with distance ($d^{-0.29}$), and capacity correspondingly rises ($d^{0.23}$). Load factor is once again nearly stable ($d^{0.06}$).

Conclusions from Market Density and Distance Studies

A number of authors (Gordon and de Neufville [14], Douglas and Miller [10], Dorman [9]) have analysed air travel markets without considering the airline's ability to adjust aircraft size almost continuously. This treatment creates an inappropriate algebraic coupling between load factor and frequency. (The product of load factor, frequency, and capacity equals demand.) The cumulative effect of our load factor/delay model, our demand model, our cost structure, and the definition of optimal in this section suggests that load factor is the most stable of the variables. The most volatile is frequency, which in the absence of constraints on aircraft size translates almost directly into changes in capacity per flight. In the full formulation of the problem, capacity changes strongly and load factor is nearly constant. (1)

Distributional Effects

We stated at the start that the demand model must be

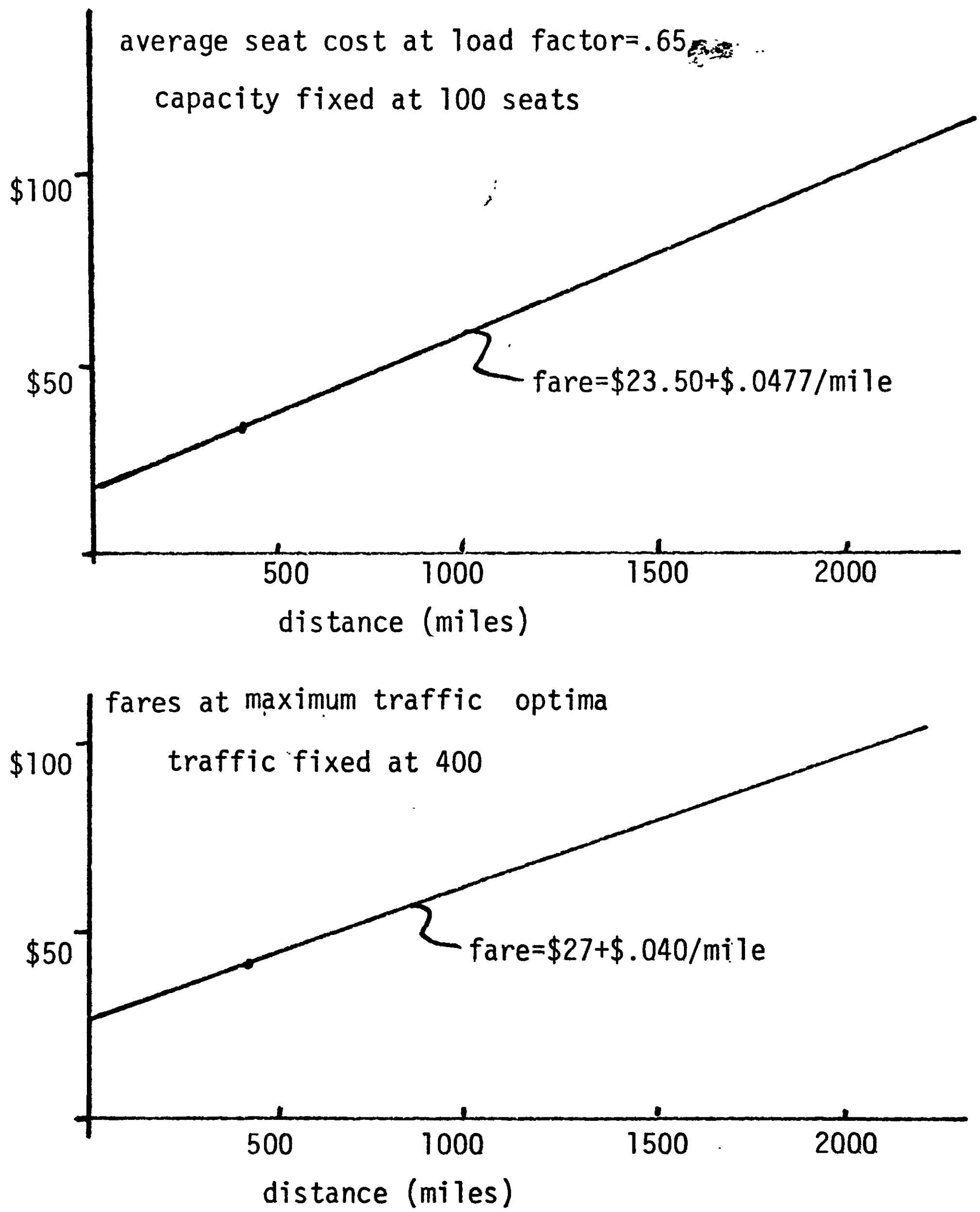
---

(1) Interestingly, the analysis by Gordon and de Neufville could be reformulated for constant load factor and variable aircraft size with only minor changes.

Table 5.1.2: Maximum Traffic Case,
Optimar at Different Distances
Traffic = 400 passengers, Value of time = $10/hr

distance	fare	frequency	load factor	capacity
100	$30	7.4	.610	89
200	$35	7.0	.625	92
400	$43	6.2	.640	101
800	$59	5.2	.665	116
1000	$67	5.0	.675	119
1400	$81	4.4	.690	132
1800	$96	4.2	.700	138
2200	$110	3.8	.710	148
2600	$126	3.6	.715	156
3000	$139	3.5	.725	158

Figure 5.1.1; Fares vs Distance at optimum for 400 passengers



evaluated over a range of values of time. Low value of time consumers prefer low fares even at the expense of high load factors or low frequency. High value of time consumers prefer higher prices if service levels are correspondingly raised. Thus the best service for one class of customers is not the best for another. Table 5.1.3 illustrates this for our central case. Optimum fares and services rise for higher values of time as the optimum adapts to minimize the total perceived price including the cost of time delays. (1)

Table 5.1.4 shows what might happen in a mixed market with the seven values of time between $1/hr and $19/hr represented equally (2) as seven disaggregate groups. Each column of the table lists the passengers in each value of time group for the fare, frequency, and load factor combination indicated by reference to the appropriate row in table 5.1.3. Except for the execrable service designed for the $1/hr people, the total demand is very nearly the same for all the different types of service. In other words, a large error in the assumed value of time produces a small result in terms of total demand when a distribution of values of time exists in the market. A maximum for total demand exists near the service designed for the average value of time, but it is a very shallow maximum. Apparently any of a broad range of service levels produces nearly the same total demand.

A slightly more suitable presentation of this point is given in table 5.1.5. Here frequency and load factor are as before, but fares have been adjusted to remove losses. This causes a small additional erosion of demand. Now every column is a zero loss condition for the market. This makes comparison more just. As before, the demand peak is shallow. Only the very lowest quality of service (column 1) is inappropriate.

Although the total traffic is stable, the distribution of traffic changes from column to column. In the last column of table 5.1.5 the $1/hr consumers will be complaining that service is over-priced. Demand is off 30% in this group (from 57.2 to 39.9 passengers). In the second column, on the other hand, the $19/hr consumers will find flights infrequent and often full. Our demand model states that the trip is more important to $19/hr

---

(1) Table 5.1.3 also states that the per trip consumer surplus as defined in Chapter 3 is low for low value of time consumers and high for high value of time consumers. This is a characteristic of the demand model as defined in chapter 3; it is not a conscious part of our ground rules. Whether the consumer surplus of $1 per hour and $30 per hour travellers should be compared in this way depends on one's definition of social welfare in general. In this work we will be careful not to draw conclusions from such relative values generated by the demand model.

(2) By equally we mean traffic is 57.14 for each at its own best service level. Computational inaccuracies in the original determination of the k1's cause 57.14 to come out as 57.1 to 57.3 in the diagonal of table 5.1.4.

Table 5.1 3: Optimum Service Levels for different values of Time
Maximum Traffic optima , 400 passenger markets, 800 mile distance

Row no.	value of time $/hour	fare	frequency	load factor	aircraft capacity	consumer surplus (000's)
(1)	$1	$47	2.0	.765	262	44
(2)	5	54	3.8	.69	153	59
(3)	7.5	57	4.6	.675	129	67
(4)	10	59	5.2	.665	116	78
(5)	12.5	60	5.8	.66	104	80
(6)	15	62	6.2	.65	99	86
(7)	19	64	7.0	.64	89	96
(8)	30	69	8.1	.635	72	120
(9)	50	76	11.0	.61	60	162
(10)	100	88	15.4	.59	44	259

Table 5.1.4: Demand at Service Levels Designed for Other Values of Time

market distance=800 miles.
service offered at a small loss

traffic by value of time category	Service in terms of Fare, Frequency, and Load Factor defined by row number and value of time for table 5.1.3 row: (1) v: $1	(2) $5	(3) $7.5	(4) $10	(5) $12.5	(6) $15	(7) $19
$ 1	57.3	51.8	49.0	47.0	45.2	43.8	41.5
$ 5	47.5	57.1	56.6	55.7	54.7	53.6	51.9
$ 7.5	42.2	56.6	57.2	57.0	56.4	55.7	54.4
$10	38.1	55.4	56.9	57.2	57.1	56.7	55.8
$12.5	34.9	53.9	56.2	56.9	57.1	57.0	56.5
$15	32.4	52.6	55.4	56.5	57.0	57.1	56.9
$19	29.3	50.6	54.1	55.6	56.5	56.9	57.1
total traffic:	282	378	385	386	384	381	374
loss	$614	$202	$164	$179	$228	$276	$444

Each Column in this table represents the passengers in each value of time category that would accrue to a service designed specifically for the value of time at the top of the column.

The diagonal elements are the traffic from the value of time category for which the service is optimal.

For further explanation, please consult text.

Table 5.1.5: Demand for Service Levels Designed for specific values of time

market distance=800 miles

service offered without a loss by adjusting fares upward from those used in table 5.1.3

traffic by value of time category	Service in terms of Frequency and Load Factor defined by row number and design value of time for table 5.1.3: row: (1) v: $1	(2) $5	(3) $7.5	(4) $10	(5) $12.5	(6) $15	(7) $19
$ 1	53.6	50.9	48.3	46.3	44.3	42.7	39.9
$ 5	45.6	56.4	56.0	55.0	53.8	52.5	50.2
$ 7.5	40.7	55.9	56.7	56.4	55.6	54.7	52.8
$10	37.0	54.8	56.4	56.7	56.3	55.7	54.2
$12.5	34.0	53.4	55.7	56.4	56.4	56.1	55.0
$15	31.6	52.1	55.0	56.0	56.3	56.2	55.5
$19	28.7	50.2	53.7	55.2	55.9	55.5	55.9
total traffic	271	373	381	382	379	373	363
fare:	$49	$55	$57	%59	$61	$65	$66

Each column in this table represents the passengers in each value of time category that would accrue to a service designed specifically for the value of time at the top of the column. Minor adjustment of fare from [illegible] the value that was optimal on table 5.1.3 corrects each column to a zero loss operating condition.

For further explanation, please see text

Table 5.1.6: Perceived Prices of Different Levels of Service

Perception for people with value of time of:	Service in terms of Fare, Frequency, and Load Factor defined in table 5.1.3 as optimal for row and $/hr: row: (1) v: $1	(2) $5	(4) $10	(6) $15	(7) $19	(9) $30	(10) $50
$ 1	$54	$58	$62	$65	$67	$72	$78
$ 5	84	74	75	77	79	83	89
$10	120	94	92	92	93	96	101
$15	157	114	108	108	108	110	114
$19	187	130	122	120	120	120	124
$30	268	173	158	154	152	150	152
$50	415	253	225	210	210	207	202

Diagonal elements are the services with the minimum perceived price for the value of time listed at the left. These are the values of time for which the service is designed.

For further explanation, please see text.

people (they pay a higher total perceived price and have higher surplus per trip), so traffic is off only 12%. The relative sizes of traffic erosion in different value of time categories is a result of the specific calibration of the demand model. The existence of modest distributional effects is a general conclusion which will survive any recalibration.

Another way to look at the relative evaluatons of service levels is to state the perceived price of each class of service by each class of customers. Reverting to the fare, frequency, and load factor conditions of table 5.1.3, we see the ranges of perceived prices in table 5.1.6. Data from this table was used in the example in chapter 3. The numbers themselves depend on the calibraton of the demand model. The important qualitative point is that there is a distribution of opinions about any given service level.

The overall consclusion from table 5.1.3 is that different values of time require different services and fares at optimum. However, tables 5.1.4 and 5.1.5 illustrate that only the extremes of these ranges have a significant impact on the total demand for a market with a range of values of time. There are two real world consequences of this. First, one cannot say from a single parameter such as fare, frequency, load factor, or vehicle size that a service is far from optimum. (However, demonstrably bad combinations of these parameters exist.) And second, some people in the market will always feel the service is not well designed for their needs.

We shall take advantage of the observation that the value of time for which a system is optimized has little effect on the total demand. From now on the numerical examples will employ a single value of time ($10/hr) to predict the total traffic in the market. This simplification has the added advantage of avoiding the philosophical problems of adding consumer surplus from different classes of consumers.

Constrained Operations

The preceding section made the point that the optimum at zero losses was very shallow. We can illustrate this point by exploring the shape of the optimum with one or another of the variables constrained. For instance, in figure 5.1.2 the conditions have been found for maximum traffic and zero losses with the fare fixed at values ranging from $52 to $70. If the fare is $54 (the optimum was $59), an adjustment of frequency from 5.2 to 3.5 and load factor from .665 to .64 produces a zero loss serice level which earns 95% of the optimum traffic. Thus with the degree of freedom of fare removed, the system can still adjust to a condition not far from optimum. Figure 5.1.2 is for a single value of time, with a range of values of time in the market the optimum can be even more shallow.

A similar story can be told in the load factor demension. For instance in figure 5.1.3, if the load factor changed to 76% from the optimum 66.5%, and increase in frequency from 5.2 to 6.3 and an adjustment in fare can nearly make up for the increased denial rate. The change in aircraft capacity is important here; the change is from 116 to 79 seats. and only a 5% erosion of

Figure 5.1.2: Range of Possible Fares

Frequency and load factor adjusted to maximuze passengers at stated fare and zero losses. 800 mile, 400 passenger market at $10/hr.

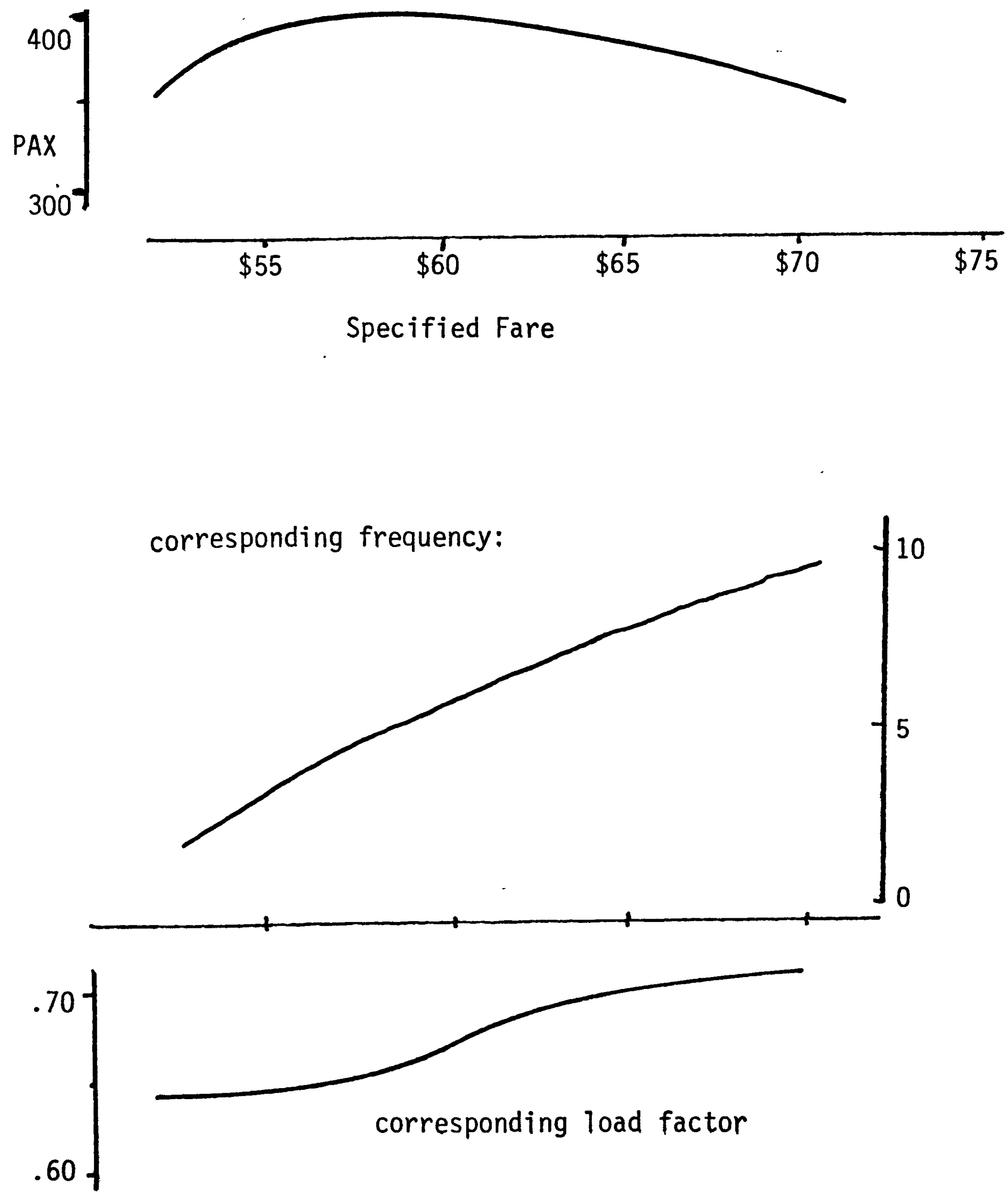


Figure 5.1.3: Range of possible load factors

Fare and frequency adjusted for maximum passengers at stated load factor and zero losses. 800 mile, 400 passenger market at $10/hr.

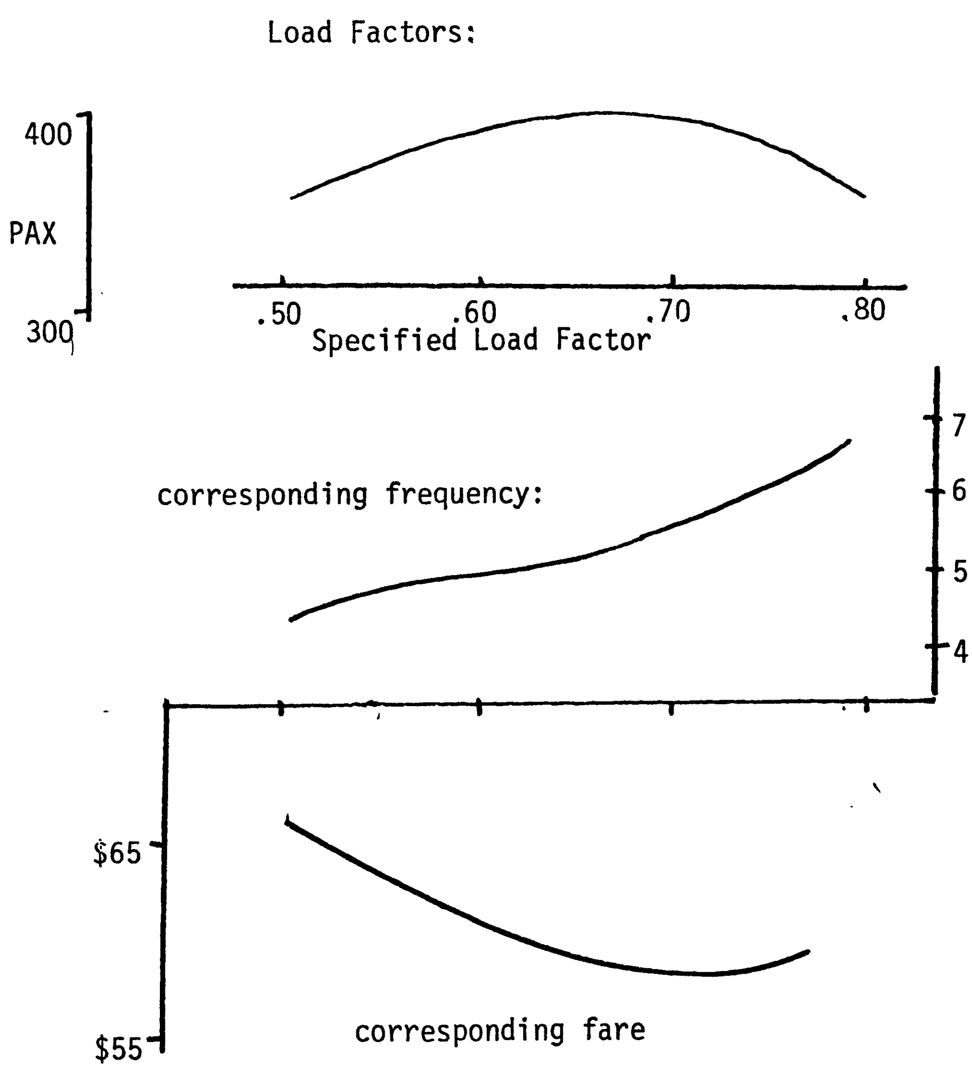
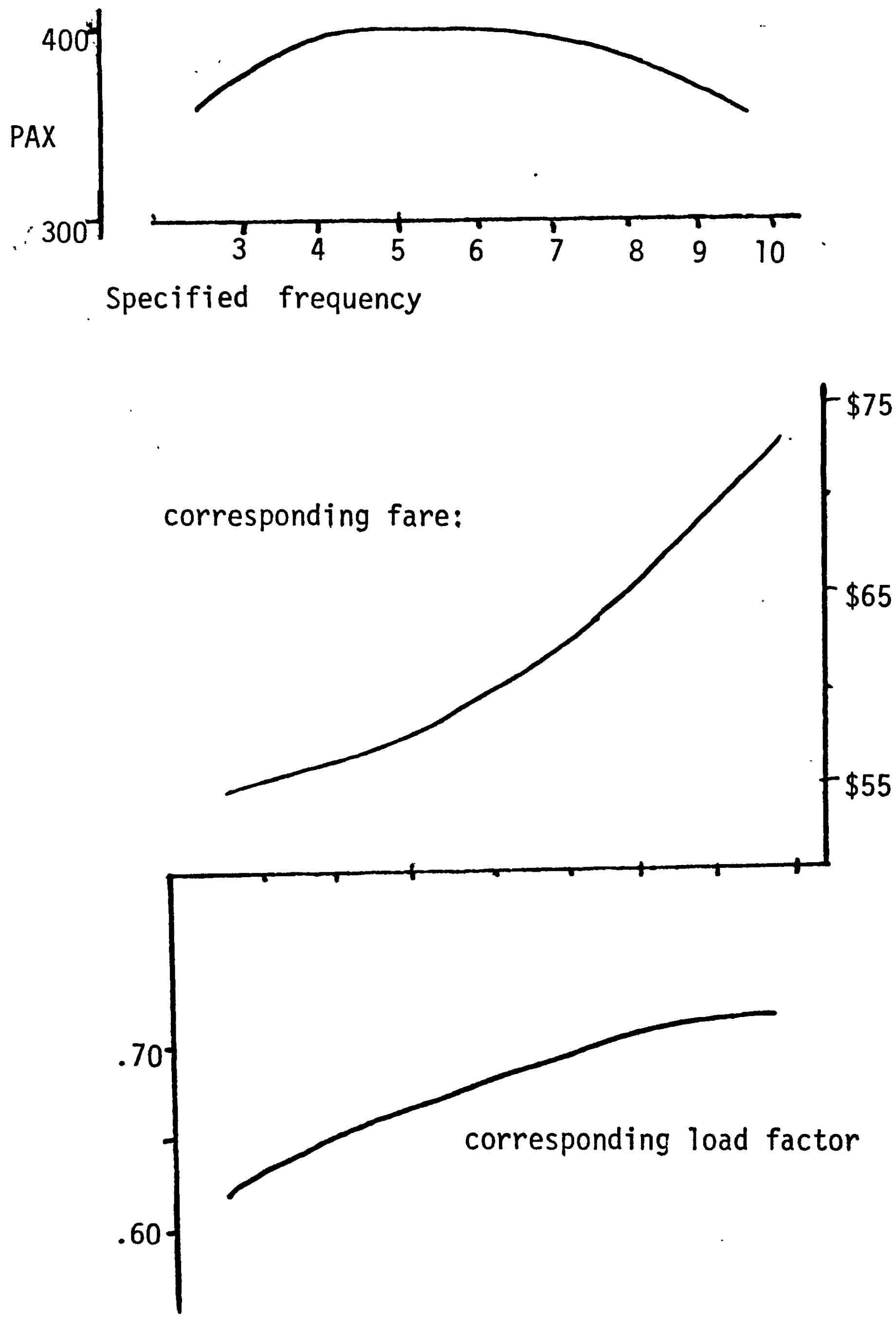


Figure 5.1.4: Range of possible frequencies

Fare and load factor adjusted for maximum passengers at stated frequency and zero losses. 800 mile, 400 passenger market at $10/hr.



note: data points are same as figure 5.1.2, Range of possible fares case.

Figure 5.1.5: Range of possible aircraft sizes
Data corresonds to Figures 5.1.2 and 5.1.3

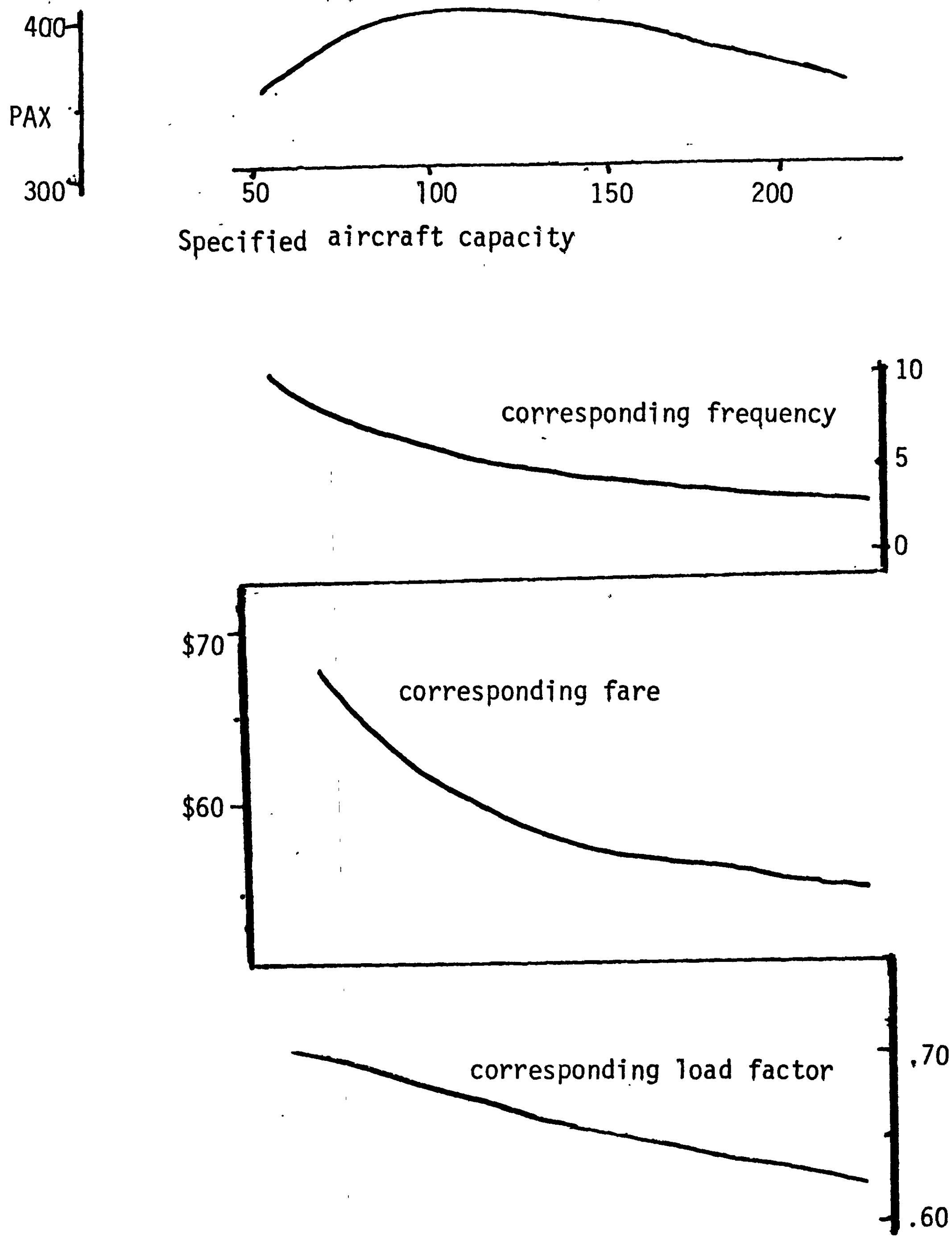


Table 5.1.7: Ranges of Service Levels near Maximum Traffic Optimum
800 mile, 400 passenger market with value of time = $10/hr

Ranges for a 5% erosion of demand:

	optimum	range	% ranges
fare	59	54-66	92-113
frequency	5.2	3.3-8.0	63-154
load factor	.665	.56-.76	84-114
vehicle size	116	68-180	59-155

Ranges for a 10% erosion of demand:

	optimum	range	% ranges
fare	59	52-71	89-121
frequency	5.2	2.6-9.4	50-181
load factor	.665	.51-.79	77-119
vehicle size	116	54-219	47-189

traffic need occur.

The range of reasonable operating frequencies is correspondingly broad. Figure 5.1.4 shows that anywhere from 2.6 to 9.4 flights per day can be flown with no more than a 10% loss of traffic. We are using only a single value of time instead of a range of values. With a range of values of time, the optimum tends to be even more shallow.

The ranges of fare, frequency, and load factor possible are presented in table 5.1.7 for a 5% and 10% erosion of traffic. Of course it is not possible to take all parameters to their extremes simultaneously. Low fare operations must be at high load factor and low frequency, and vice- versa.

From a technical standpoint it is interesting to note that a mistake in aircraft size can be adjusted for by changing fare, frequency, and load factor. Figure 5.1.5 shows the nature of the capacity optimum.

A consequence of this flexibility in service design is that regulation which places restrictions on one aspect of service will have only modest influence on the value of the optimum. This makes regulation of service standards within reasonable limits relatively harmless and regualtion aimed at simulating a change to a different opjective criterion ineffective.

<u>Regression Analysis</u>

Formulas for predicting optimal fare, frequency, load factor, and capacity as a functon of market distance and traffic at the optimum can be created by using data points like those presented in tables 5.1.1 and 5.1.2 to calibrate log-linear expressions. Such formulas are only valid as an interpolation among data points. (1) Nonetheless, the formulas will prove useful for the qualitative feel provided by the exponents and later in chapter 6 for interpolating.

The formulas take the form of:

$$\text{VARIABLE}=k\cdot \text{PAX}^{n}\cdot d^{g}$$

Values for k, g, and n for each variable of fare, frequency, load factor, and capacity are listed in table 5.1.8. These values were obtained by least squares regression of the logarithmic version of the formula above using the data points listed in table 5.1.9. The values of 5.1.8 agree with those reported earlier in this section. Differences in detail are due to the greater and more careefully distributed range of data points used to calibrate the formula.

<u>Competition in U.S. Domestic Markets</u>

Chapters 3 and 4 suggested that competitive services between airlines might well take the form of product matching. With product quality in a city pair market measured in terms of

---

(1) See Bard [49], chapter 1, for a discussion of the distinction between an interpolation formula and one that may be extrapolated.

Table 5.1.8 Values for regression formula

$$\text{VARIABLE} = k \cdot PAX^{\eta} \cdot d^{\vartheta}$$

	k	ϑ	η	R^2
fare (FA)	4.202	0.516	-.118	.995
frequency (FQ)	0.904	-.252	0.570	.979
load factor (LF)	0.338	.0523	.0566	.993
capacity (SEATS/ FQ)	3.274	0.199	0.373	.988

Values used in local one-dimensional curve fits in earlier parts of thes section:

	ϑ	η
fare (FA)	.55	-.12
frequency (FQ)	-.29	.58
load factor (LF)	.06	.06
capacity (SEATS/FQ)	.23	.36

Table 5.1.9 Data used to calibrate regression formulas

distance	traffic	fare	frequency	load factor	capacity
200	2318	28.3	18.2	.69	185
	1043	30.6	11.6	.66	136
	466	33.8	7.6	.635	97
	213	38.1	5.0	.61	70
	106	43.2	3.4	.585	53
500	2022	39.3	14.6	.71	195
	926	42.2	9.4	.685	144
	423	46.3	6.0	.65	109
	203	51.9	4.2	.635	76
	108	58.3	3.0	.61	59
1000	1845	56.6	12.0	.755	208
	861	60.8	7.8	.71	155
	405	66.2	5.0	.675	120
	202	72.7	3.4	.65	91
	113	81.0	2.6	.63	83
1500	1781	73.4	10.6	.755	223
	841	78.5	6.8	.725	171
	402	85.0	4.4	.69	132
	205	92.5	3.0	.665	103
	119	99.7	2.2	.65	83
2000	1751	89.7	10.0	.775	226
	833	95.7	6.2	.735	183
	403	103.3	4.0	.70	144
	209	112.3	2.8	.675	111
	124	119.2	2.0	.66	94
2500	1738	106.0	9.2	.78	242
	832	112.5	5.8	.75	191
	410	121.1	3.8	.715	151
	215	130.2	2.6	.69	120
	128	140.2	2.0	.675	95
3000	1733	121.9	8.8	.79	249
	827	129.7	5.4	.75	204
	413	138.5	3.6	.725	158
	217	148.2	2.4	.695	130
	131	158.0	1.8	.675	108

frequency and load factor, how viable is single link competition at representative U.S. domestic demand volumes? Table 5.1.10 shows a list of markets in the U.S. ranked by traffic flows. Optimal service quality from table 5.1.8 allows fares to be predicted for a single carrier operating at the same traffic and distance. These fares are listed in the lower half of table 5.1.10. For the situation in which competitors duplicate frequencies, the average cost is listed in the column headed "2 competitor fare". For all U.S. domestic markets, marginal passenger cost is noticeably below average cost for service levels fixed at optimum. (1)

Of course competition may take the form of sharing the total optimal frequency among competitors. Unfortunately, the game theoretic approach to such scheduling leads to matching of schedule times, as shown by de Neufville [38]. The condition of interleaved or coordinated schedules occuring in competitive markets appears to be an accident unless the arrangement is collusive. (2)

Sensitivity Studies

How sensitive is the zero profit optimum to changes in the cost or demand parameters? In view of the conclusion that load factors vary least, the first condition to explore is the importance of the value for the turnaway rate in the demand model.

Let us double the probability of a space denial for a given load factor. The new optima for the standard 400 passenger, 800 mile, $10/hr market show the same ranges of fare and frequecy as before. Table 5.1.11 can be compared with table 5.1.1 and 5.1.2. Optimal load factors are down by 4 points, but they too show the same kinds of variation as before. The conclusions, which were based on the ranges of these parameters and not their absolute values, are unaffected.

The opposite is true if the economies of scale with respect to aircraft size are removed. We may cut these in half by allocating half the cost per vehicle mile and per vehicle departure to seat miles and seat departures respectively in such a way that the total costs at 125 seats remain constant. (3) The new optima are at roughly double the frequency, and the new load factors are 5 points above the old. Service has changed a lot, although fares are more stable. This can be seen by comparing table 5.1.12 here with tables 5.1.1 and 5.1.2 in the earlier part of this section. In general, the size of the economies of scale in vehicle capacity is quite important to the optimum service

---

(1) The usual definition of marginal cost for a service does not fix both the frequency of service and the load factor because aircraft size is not usually allowed to vary.

(2) But network effects change this conclusion, as real data suggest. Cf. appendix C.

(3) New numbers are c_{10}=4.79, c_{11}=.021, c_{20}=189.9, c_{21}=.408

Table 5.1.10 Sizes of U.S. Domestic City Pair Markets

rank in passengers boarded	1976 traffic per day one way	distance (miles)	competitors with over 10% of traffic	airlines providing service in 1976	cumulative traffic to rank
1	2308	721	3	12	1%
50	402	407	2	7	21%
100	252	196	2	4	29%
200	159	185	2	5	40%
400	86	604	2	4	52%
600	60	627	1	2	60%*
1000	33	527	1	2	69%

*40% of the domestic trips were in markets of less than 60 pax/day

---

rank	city names	optimal fare	optimal frequency	Marginal cost %	2 competitor fare %
1	Chicago-NYC	$46	14.2	87%	113%
50	ChicagO-Kansas City	$43	6.1	75%	125%
100	Miami-Orlando	$38	5.9	67%	133%
200	Dallas-Oklahoma City	$40	4.4	63%	137%
400	New Orleans-St Louis	$63	2.3	64%	136%
600	Denver-Tucson	$69	1.8	60%	140%
1000	Detroit-Norfolk VA	$72	1.3	54%	146%

Optimal fare and load factor are optimal single link specifications as on table 5.1.8 . Fare deduced.

Marginal cost is the cost of one more passenger at the stated load factor and frequency, expressed as percent of the average cost, which is the fare.

2 competitor fare is average cost for double the frequency and the same capacity and traffic. This implies frequency and load factor duplication and is not intended to represent actual competition.

Table 5.1.11: Maximum Traffic Optima with Doubled Denied Boardings

varying demand density at 800 miles, value of time = $10

traffic	fare	frequency	load factor	capacity
50	$87	1.8	.565	49
100	76	2.6	.58	66
200	67	3.6	.60	93
400	61	5.4	.625	119
800	56	8.0	.65	154
1600	52	11.8	.675	201
3200	49	18.4	.71	246

varying distances with traffic at 400 passengers, value of time = $10

distance	fare	frequency	load factor	capacity
200	$35	7.0	.585	98
400	44	6.2	.60	107
800	61	5.4	.625	119
1000	69	5.0	.635	126
1400	85	4.6	.645	135
1800	100	4.2	.65	147
2200	115	4.0	.655	150
2600	131	3.8	.665	158
3000	145	3.6	.675	165

level for a market.

This point should be emphasized. If there are no economies of scale in this respect, the optimal solution is at infinite frequency, 100% load factor, and a vehicle size of zero. In order to avoid this conclusion, those analysts who ignore scale economies are forced to fix aircraft size. (1) By adding the constraint of fixed capacity, the extreme solution is avoided. But if there are really no economies of scale with aircraft size, a smaller aircraft will always produce a yet better optimum. The logical extension of this is not a public common carrier transport system, but individually targeted vehicles such as in road passenger transport. As we argued at the start of chapter 2, that common carriers exist at all is a sign that there are economies to be gained from using larger vehicle sizes.

While the economies of scale affect the location of the optima, they do not change the amount of variation with market size or distance. Table 5.1.13 shows that the exponents of the dependency of fare on market size or distance are high for the standard case and both sensitivity studies. (The exponents in tables 5.1.13 were derived from curve fits of the data in tables 5.1.1, 5.1.2, 5.1.11, and 5.1.12.) The exponent for load factor is small in every case, showing the least dependence for this variable. Capacity is important in all cases. Table 5.1.14 shows the ranges of these variables for each study. The ranges for capacity are certainly not negligible.

These results suggest that our conclusions from the market size and distance studies are largely independent of the numerical values of the cost and demand parameters (within normal ranges). The conclusions do rely on the demand model structure, but this structure differs only in technical detail (2) from the models accepted by the best work to date. (3)

Maximum Revenues at Zero Profits

A second form of sensitivity study is sensitivity to the exact definition of the objective function. What if we had maximized passenger revenues instead of passengers? The constraint that fare equal average cost should be maintained in this section. However it is often assumed that managements maximize their revenues instead of the traffic carried. These are not the same objectives, even for a single value of time market. The maximization of revenues will favor service of higher quality and higher price than the maximum traffic optimum. For the standard case of an 800 mile market with 400 passengers

---

(1) This is done by Douglas and Miller [10], Gordon and de Neufville [14], Dorman [9], but not Anderson [36].

(2) The functional dependence of seat availability on load factor and displacement time on frequency are different due to different approximation schemes. Consult chapters 3 and 4.

(3) Douglas and Miller [10], Gordon and de Neufville [14], Dorman [9], and others.

Table 5.1.12: Maximum Traffic Optima with Frequency Cost Halved

varying demand density at 800 miles, value of time = $10

traffic	fare	frequency	load factor	capacity
50	$80	2.6	.64	30
100	71	3.6	.655	43
200	65	5.4	.68	55
400	60	8.0	.71	70
800	56	12.2	.74	89
1600	53	18.8	.775	110
3200	50	30.2	.815	130

varying distances with traffic = 400, value of time = $10/hr

distance	fare	frequency	load factor	capacity
200	$35	10.2	.665	59
400	43	9.4	.69	62
600	52	8.6	.70	67
800	60	8.0	.71	70
1000	66	7.6	.725	73
1400	84	7.0	.735	78
1800	99	6.6	.75	81
2200	114	6.2	.76	85
2600	129	5.8	.765	90
3000	144	5.6	.77	93

Table 5.1.13:

Comparison of exponents for sensitivity studies

dependence on market size: variable $\sim PAX^{\vartheta}$

variable name	ϑ standard case	ϑ load factor sensitivity	ϑ frequency sensitivity
fare	-.12	-.13	-.104
frequency	.58	.55	.60
load factor	.06	.055	.055
capacity	.36	.40	.34

dependence on distance: variable $\sim d^{\eta}$

variable name	η standard case	η load factor sensitivity	η frequency sensitivity
fare	.55	.55	.55
frequency	-.29	-.26	-.24
load factor	.061	.053	.055
capacity	.23	.21	.18

Table 5.1.14:

Ranges of variables for sensitivity studies

ranges for PAX=50 to 3200; distance = 800 miles.

variable name	standard case	load factor sensitivity	frequency sensitivity
fare ($)	48-85	49-87	50-80
frequency	1.8-17.6	1.8-18.4	2.6-30.2
load factor	.60-.75	.565-.71	.64-.815
capacity	47-243	49-246	30-130

ranges for distance = 200 to 2600 miles; traffic = 400 passengers

variable name	standard case	load factor sensitivity	frequency sensitivity
fare ($)	35-126	35-131	35-129
frequency	3.6-7.0	3.8-7.0	5.8-10.2
load factor	.625-.715	.585-.665	.665-.765
capscity	89-156	98-158	59-90

with a $10/hr value for their time, the maximum revenue service is at fare=$81, frequency=10.4, and load factor=.61. (The maximum traffic service was at fare=$59, frequency=5.2, and load factor=.665.) Traffic for the maximum revenues case is only 321 passengers for the same demand which could produce 400 passengers at the maximum traffic point. Fare is up 38%, demand is off 20%. Service is twice as frequent at lower load factors. Consumer surplus is off 12%.

Ti might be claimed that these conditiond represent what happened in the last 10 years in the U.S. domestic markets. Not only is demand suppressed, but such services are not at all suited to the low value of time customers. Thus while the total benefits are only slightly reduced, the distribution of these benefits is skewed. The poltical forces which fostered deregulation and the actions by the airlines since seem to be evidence in support of such claims.

The general observation is that the maximum revenue otimum is not near the maximum passengers optimum. Whatever its merits as a discription of management objectives, the maximum revenue, zero profits case is a poor surrogate for broader social objectives such as maximizing consumer surplus.

5 2 Maximum Profits Optima

We now examine the case of a monopolist maximizing profits. By profits we mean accounting profits above the normal return on investment. Normal cost of capital was included in the cost function. The general conclusion is that such profits can be large, although this depends on the market size and elasticity. The profit optimum case is of particular interest because the greater the traffic on a single firm's service, the lower the average cost per passenger and the higher the level of service at zero profits. We suspect that a single firm might come to dominate such a market.

Total profits from operating the transportation service are simply the revenues less the costs:

$$\pi = FA \cdot PAX - c1 \cdot PAX/LF - c2 \cdot FQ - c3 \cdot PAX \qquad (5.9)$$

The variable COST was removed from the problem by the use of its definition in equation (5.1). The variable SEATS has been removed by the use of the load factor definition of equation (5.3). The variable PAX remains in the formula as a useful notation for the expression of equation (5.2), but it is understood that the three independent variables are FA, FQ, and LF. The conditions $d\pi/dFA=0$, $d\pi/dFQ=0$, and $d\pi/dLF=0$ will be employed to determine the unknowns at the optimum profit point. These conditions will locate the maximum profit point when there is one.

The condition $d\pi/dFA=0$ determines the appropriate fare. We apply the condition to (5.9) and simplify to get:

$$FA=c1/LF+c3-1/\alpha\cdot PP \quad (5.10)$$

First of all, we notice in (5.10) that the expression (c1/LF+c3) is the sum of the SEAT-related costs and the PAX-related costs expressed as a cost per passenger. We saw this term in the previous section and we will see it again. It represents all costs of service except the cost of frequency.

For comparison, equation (5.4) shows the average cost per passenger of frequency as (c2·FQ/PAX). The maximumprofit point will be reasonable only when the third term in (5.10) exceeds this cost, i.e. when profit is postive:

$$-1/\alpha > c2\cdot FQ/PAX \quad (5.11)$$

Using (5.2), $PAX=k1\cdot PP^{\alpha}$, we get:

$$-1/\alpha\cdot PP > c2\cdot FQ/(k1\cdot PP^{\alpha})$$

Rearranging obtains:

$$k1 > c2\cdot FQ\ (-\alpha)\ PP^{-\alpha-1} \quad (5.12)$$

Thus there is a minimum market size (k1) which can be made profitable. With demand below this minimum, no single fare service can attract enough traffic to recove its costs. (1) This comes as no surprize.

But we can say more. PP, the perceived price in (5.10), contains FA plus time related terms, which we summarized as TT in equation (5.5). Thus (5.10) becomes:

$$FA = c1/LF+c3-1/\alpha\cdot(FA+TT\cdot v)$$

Rearranging obtains:

$$FA = (c1/LF+c3-1/\alpha\cdot TT\cdot v)/(1+1/\alpha) \quad (5.13)$$

As the market becomes inelastic with respect to perceived price (α moves from -1.5 toward -1.0), the maximum profit fare approaches infinity. Of course this is really far outside the range of validity for the demand model, but the indication is that profits can be a large fraction of revenues. In any case there is an important dependency of profits on elasticity, as one would expect.

For our value of α=-1.5, the maximum profit fare is

$$FA=3\cdot(c1/LF+c3+0.667*TT*v)$$

Since the first two terms of this expression are usually 75% of the total costs, fares at maximum profit points are generally well above average costs. In short the maximum profit point for an isolated monopoly market can be relatively profitable.

What about service levels? Frequency can be determined from

---

(1) We have not allowed price discrimination in the market at this time.

$d\pi/dFQ=0$. This produces:

$$0 = dPAX/dFQ \cdot (FA-c1/LF-c3)-c2 \qquad (5.14)$$

Using (5.6), we may state

$$dPAX/dFQ = dPP/dFQ \cdot (PAX \cdot \alpha/PP)$$

Substituting this into (5.14) and rearranging obtains

$$1 = (FA-c2/LF-c3)(PAX \cdot \alpha/PP \cdot c2)(dPP/dFQ) \qquad (5.15)$$

Using (5.5) and (5.6), we may take dPP/dFQ as:

$$dPP/dFQ = -v(5.7+57LF^{q})/FQ$$

Substituting this into (5.15) and moving FQ to the left hand side produces

$$FQ=PAX \cdot (FA-c1/LF-c3) \cdot (-\alpha v(5.7+57LF^{q})/PP \cdot c2) \qquad (5.16)$$

The first thing to notice is that as before frequency at optimum rises as the square root of traffic. Secondly, the excess of revenues (FA) over non-frequency costs (c1/LF+c3) in the second term of the product is the per capita profit. The per capita profit rises as the square of frequency. Since frequency costs are linear with frequency, profits as a fraction of costs grow with market size. Finally, as the market goes to long hauls the numerator of the last term of (5.16) becomes small with respect to the denominator. The numerator is the displacement and scedule delay component of the total perceived price PP. Thus, frequencies should drop at longer hauls.

No particular insights into optimum load factors are available by manipulating the equations above. The best thing to do is to examine some numerically located optima for the conditions we have defined above. We have done this by guessing FQ, LF, and FA for markets of different densities and distances and searching for the optima. The results are presented next.

<u>Numerical Case Studies for Maximum Profit Optimum</u>

We will specify the market size not by k explicitly but by the number of passengers which would occur in the maximum traffic optimum case for the same market. Thus each market size statement will contain the information for comparing traffic. The standard market is our 800 mile, 400 passenger market with $10/hr as the value of time. the maximum traffic point was at fare=$59, frequency=5.2, and load factor=.665.

For this standard market the profit optimum conditions are far different from the maximum traffic optimum. The maximum profit traffic is only 69 passengers. The fare is over four times the maximum traffic fare. Consumer surplus is half, and the sum of consumer surplus and producer surplus (profits) is 2/3 of the maximum traffic values (see table 5.2.1). As we suspected from (5.16), the maximum profit point is outside the

Table 5.2.1: Maximum Profit Points

Value of time = $1.00 per hour:

	max π	max pax	max π at fare=$47
fare	$155	$47	$47
frequency	0.8	2.0	0.8
load factor	.70	.765	.84
traffic (pax)	74	400	268
profits π	$7400	$ 0	$1100
surplus	24800	43500	38100
total	32200	43500	39200

Value of time = $10 per hour:

	max π	max pax	max π at fare=$59
fare	$245	$59	$59
frequency	2.0	5.2	2.1
load factor	.61	.665	.72
traffic (pax)	69	400	270
profits π	$12000	$ 0	$ 2000
surplus	38000	74000	65000
total	50000	74000	67000

Value of time = $50 per hour:

	max π	max pax	max π at fare=$76
fare	$479	$76	$76
frequency	4.5	11.0	4.1
load factor	.56	.61	.66
traffic (pax)	71	400	276
profits π	$25600	$ 0	$4200
surplus	90800	161900	143000
total	116400	161900	147200

bounds of reasonable operations. Profits are 2/3 of revenues.

It is also possible to maximize profits at fixed fare. For instance, we might fix the fare at the maximum traffic optimum of \$59. This represents the sort of behavior which is often assumed relevant to a price-regulated market. Maximizing profits at this fare produces service and traffic quite far from the best obtainable. Passengers are only 2/3 and frequency and load factor are both degraded from optimum at zero profits. Excess profits in this case are still over 10% of revenues. These conclusions depend on the calibration values for cost and demand elastisticity. The general observation is that there is room for maneuvering to off optimum conditions even when the fare is held at the maximum traffic optimum. (1)

Table 5.2.1 compares maximum traffic, maximum profit at fixed fare, and maximum profit at any fare for several values of time in our standard maximum. The non-competitive maximum profit point is so far beyond the range of reasonable operations that the condition of unconstrained profit-maximizing will be of little further interest to us. Both equation (5.16) and our numerical studies put the maximum profit point at profit levels unlikely to be tolerated by society. The single and important exception to this in the case of very thin markets where the fullest exploitation of monopoly power is necessary to recover the cost of frequency.

5 3 Maximum Surplus Optima

In the maximum traffic, zero profits case the marginal cost of adding one more passenger to the system is

$$dCOST/dPAX = c1/LF+c3 \qquad (5.17)$$

We must be very careful about what this marginal cost means. It is the cost at fixed frequency and fixed load factor of one more passenger. Thus it is the marginal cost at constant service quality To fix both FQ and LF implies that the marginal passenger is accomodated by adding 1/LF seats to the capacity, i.e. by changing aircraft sizes. Taking marginal cost while maintaining exactly the same service levels (in FQ and LF) has been addressed only once before (by Anderson [33]).

The fare in the maximum traffic zero profit case was above (5.17) by an amount necessary to recover the cost of frequency:

$$FA = c1/LF+c3+c2 \cdot FQ/PAX \qquad (5.4)$$

Under these conditions it seems desirable to expand the service to those people willing to pay at least the marginal cost of their carriage. The objective function which will set fares (and

(1) In the light of our previous comments, such maneuvering in the U.S. domestic operations may have been toward maximum revenue operations instead of maximum profits.

frequency and load factor) to take advantage of this is the maximization of total surplus. With fare at that indicated by (5.17) and costs as indicated in (5.4), profits will be negative.

In practice this rule of allowing losses to exist in order to serve the marginal users at the marginal price is difficult to implement Society seems willing to cover the costs of physical fixed facilities such as airport land and runways out of tax revenues, but the cost of frequency of service is expected to be recovered from the traffic. This can be done by price discrimination, but price discrimination among customers in the air mode is difficult, both as a legal and a practical matter. The marginal customer is hard to identify by an enforceable rule. Nonetheless, the condition which establishes marginal cost pricing is of interest for comparison with the average cost pricing which was part of the maximum traffic case developed above

Profits were defined in equation (5.9):

$$= PAX\ (FA-c1/LF-c3)-FQ\cdot c2 \qquad (5.9)$$

Consumer surplus was defined in chapter 3 as:

$$S = PAX\cdot PP\cdot(-1/(\alpha+1)) \qquad (5.18)$$

The objective function is the sum of the two:

$$MAx(S+\pi) = \text{objective} \qquad (5.19)$$

We differentiate with respect to FA and use PAX and PP as notation for the terms involving the instrumental variables in the optimization process, FA, FQ, and LF:

$$PAX\ dPP/dFA\cdot(1+\alpha(FA-c1/LF-c3)/PP)+PAX=0 \qquad (5.20)$$

We notice that PAX can be eliminated, so the optimum fare in the maximum surplus case is nearly independent of the market density. Also dPP/dFA, the change in perceived price with a change in fare, is $1, so (5.20) simplifies to:

$$FA = c1/LF+c3 \qquad (5.21)$$

This is the condition that price equal marginal cost (equation (5.17)), as anticipated in the introduction to this objective function. Notice that this would produce a loss. A major observation is that fares do not cover the cost of frequency of service Frequency becomes in essence a public good. This is the same conclusion reached by Anderson [33] from a similar problem statement. Anderson's contribution is too little recognized.

The frequency can be set by differentiating (5.19) with respect to FQ and using (5.21) to simplify. The solution is <u>not</u> independent of market size:

$$FQ = PAX^{0.5} (v \cdot (5.7+57LF^{9})/c2) \qquad (5.22)$$

This is the same formula as for the maximum traffic case (equation (5.7)), except that the term PAX in this case will reflect a lower fare and a different frequency than the maximum traffic optimum. Notice that optimum frequency does depend on the market size, rising as its square root.

The optimum load factor is nearly independent of market size. We differentiate (5.19) with respect to LF and simplify using (5.21):

$$LF = FQ^{0.1} (c1/9 \cdot k3 \cdot v)^{0.1} \qquad (5.23)$$

Once again this is the same expression as for the maximum traffic zero profits case, except that the frequency FQ will be higher than in that case. Substitution of the frequency formula (5.22) produces a dependence of optimal load factor on the twentieth root of traffic:

$$LF = PAX^{0.05} (c1^{2} \cdot (5.7+57LF^{9})/263169v \cdot c2)^{0.05} \qquad (5.24)$$

This is the same as equation (5.8).

To summarize, the maximum surplus solution adjusts the fare downward from the maximum traffic solution. The costs of frequency are not covered by the new fare. The fare does reflect the cost of load factor, i.e. the marginal seat cost divided by the average occupancy. With the fare set at this new low level, optimum frequency and load factor should be adjusted to have the same dependence on observed traffic as in the maximum traffic case. At lower fares traffic will be higher than in the maximum traffic case, so frequency and load factor will also be higher.

The two major conclusions of the maximum surplus case are (1) losses occur because fare is below average cost and (2) the optimum frequency and load factor bear the same relationship to traffic as in the maximum traffic zero profits cases. Further insight will be available from the numerical studies which come next

Just as in the maximum profits case, there is a minimum market size (minimum k1) at which the surplus is positive. With the possibility of subsidy, this market size is smaller than that for zero profit operations. The net surplus from equations (5.9), (5.18), and (5.19) with substitution of (5.5) for PAX is:

$$k1 \cdot PP^{\alpha+1} (-1/\alpha+1) + k1 \cdot PP^{\alpha} \cdot (FA-c1/LF-c3) - FQ \cdot c2 > 0 \qquad (5.25)$$

We have required surplus to be positive. We know that (FA=c1)/LF-c3) from (5.21) so (5.25) can be simplified to:

$$k1 > FQ\ c2\ (-(\alpha+1)) \cdot PP^{-\alpha-1} \qquad (5.26)$$

Comparison with equation (5.12) which expresses the minimum market k1 for a profitable market shows that the net surplus minimum is smaller. For $\alpha=-1.5$, the minimum underlying demand may be one third or less of that for a profitable market.

<u>Numerical Study of Maximum Surplus Case</u>

Table 5.3.1 shows the maximum surplus optima for different distances for market with 400 passenger maximum traffic optima. Although fares are 25% less, traffic is 33% higher and losses amount to 20% of costs. The net change in surplus is only 1%. Thus for major markets at least, the failure of society to find a way to subsidize the public good of frequency of service causes very little welfare loss. This conclusion is based on our calibration numbers and will not necessarily carry over to other modes. However the conclusion is fairly convincing for air. Given the accuracy of our numbers, a 10% change would have seemed small

The situation is slightly different at very low demand densities. Table 5.3.2 shows that traffic can nearly double in low density markets if their frequency is subsidized. Measurable gains of surplus (on the order of 5%) are possible for subsidized low density services. In table 5.3.2 demand densities below 50 are represented by their relative k1 values because such low density markets do not have comparable zero-profit traffic. Subsidy in low density cases can be as much as 50% of costs.

Current airline practice in the U.S. is to subsidize markets of 25 to 50 passenger per day (1) so that their fares equal average costs for the denser part of the system and thus are near marginal costs for the markets in question. This practice would seem to have some theoretical justification.

<u>Price Discrimination</u>

The maximum surplus case derived a price covering the marginal cost of a passenger but not covering the cost of frequency. One way the cost of frequency could be recovered would be by price discrimination among consumers of the service. For instance, if the high value of time customers (who are in our model relatively fare inelastic) paid more and the low value of time customers paid less, a more efficient optimum could exist at zero losses. It would be possible to do better still by offering the high value of time customers a service suited to their optimum (low load factor, high frequency) and low value of time customers service suited to theirs (higher load factors, some flights not available). Technically, both services can be combined on the same flights. The fixed costs of frequency could be shared. The sharing need not be "fair". As long as each group pays the costs of its seats (adjusted for load factor), who pays for the frequency is not important, philosophically. The political and legal implications of such an arrangement are less clear cut.

Price discrimination without service differences is practical in freight operations. Value of service pricing uses a commodity's own value as an index of its price elasticity. Expensive goods are charged more per ton than cheap ones. Air

---

(1) actually <u>points</u> with 25 to 50 origins per day are subsidized. This gets us into network effects. Here a single satellite to hub feeder service is treated as an origin- destination market.

Table 5.3.1: Maximum Surplus Optima over Distances
value of time = $10/hr
demands set so PAX = 400 at maximum traffic optimum

distance	passengers[1]	fare[1]	freq	load factor	capacity	loss[2]	surplus[3]
100	149%	71%	9.2	.625	103	25%	102.1%
200	144%	72%	8.4	.635	108	24%	101.8%
400	138%	73%	7.0	.65	120	22%	101.5%
800	133%	76%	6.2	.675	127	21%	101.1%
1000	131%	76%	5.8	.69	131	21%	101.0%
1400	128%	79%	5.2	.70	141	20%	100.9%
2200	124%	80%	4.4	.72	156	18%	100.7%
2600	123%	81%	4.2	.725	162	17%	100.6%

Table 5.3.2: Maximum Surplus Optima over Demand Density
distance = 800 miles; minimum frequency is 1.0

nominal passnegers	actual pax[4]	fare[1]	freq	load factor	capacity	loss[2]	surplus[3]
3200	115%	88%	19.4	.76	249	12%	100.3%
1600	119%	84%	13.0	.735	199	14%	100.4%
800	125%	76%	8.8	.70	162	17%	100.7%
400	133%	70%	6.2	.675	127	21%	101.1%
200	144%	64%	4.4	.66	99	26%	101.8%
100	159%	55%	3.2	.635	78	31%	102.9%
50	183%	50%	2.4	.62	62	36%	104.4%
50	92	$47	2.4	.62	62	36%	n.a.
(1/2)	35	$48	1.4	.59	42	46%	n.a.
(1/4)	14	$50	1.0	.565	25	60%	n.a.
(1/8)	7	$50	1.0	.565	12	75%	n.a.

[1] fare and passengers as a % of maximum traffic case
[2] loss as a % of total costs
[3] surplus is consumer surplus less loss as a % of max traffic case
[4] actual pax as a % of nominal (except for last four lines)

freight price discrimination is limited by competition from trucks and between regular commodity service and containerized (non-discriminatory) service. A good can be overcharged for its carriage only if the mode is markedly superior in perceived price (logistically for freight) to other modes. That is the same thing as saying that the goods for which a mode is technically best suited are captive and can be charged above marginal costs. Goods for which air and truck service are equally appealing can only be charged the marginal cost of their carriage when competition is allowed. That is to say that goods whose value of time falls on the watershed lines of the demand figures in chapter 3 will usually pay the lowest prices for the service.

In passenger transportation, the use of indicators such as age, sex, occupation, or marital status to determine elastic segments of demand is legally difficult. However, simultaneous service differentiation and price discrimination can occur. Excursion fares represent services aimed at lower value of time optima. Excursion fares probably do not cover the average cost per customer of frequency in the same way as regular fares do. If frequency costs are not equally shared by both fare groups, some price discrimination is occuring.

One conclusion from the maximum surplus case is that only low density markets seem worthy of consideration for subsidy. In these cases, price discrimination within a market can take the place of cross subsidy from other markets or other activities. Price discrimination could be allowed or even encouraged in low density markets. In higher density markets, stricter standards of justice among consumers may be enforced, although multiple choices of service level are still to be encouraged. For freight this would mean that commodity pricing might exist in low density markets, but only service based differentials would be appropriate in bigger markets. In passenger transport, low density markets might have larger price differentials for different service qualities than would denser markets.

In network studies we will find that the cost of frequency cannot even be allocated to a given market, so the range for discrimination among customers extends itself geographically as well as across values of time. Those markets which are not captive to the network (i.e. have alternative service in the city pair) will probably pay close to their marginal cost. We will discuss these network matters in greater detail in their proper places.

5.4 Summary and Conclusions

This chapter has demonstrated that the optimum design for a single link transportation service involves differing frequency, vehicle size, load factor, and fare depending on the market density, distance, the nature of the objective function, and details of the distribution of values among the consumer population. An objective of maximum profits produces severely constrained services and unrealistically high fares. An objective of maximum total surplus produces substantial losses.

Objectives of maximum traffic or maximum consumer surplus at zero loss produce shallow optima which depend on the distribution of values of time among the population. Prediction of optimal or even likely operating conditions can not be made without accurate demand data. There is an absence of simple, fundamental, global answers.

The absence of peaked optima means that in any real situation details of short term practicality may determine operating conditions. It may also be that the particular choice of conditions is a matter of chance. This makes regulatory control difficult if not impossible.

There are three degrees of freedom in designing a service for a city pair: fare, frequency, and load factor. From the consumer's viewpoint these add up to total perceived price. From the producer's viewpoint they determine profits. There may be some simple natural equilibrium between consumers and a producer or producers. If there is, regulations fixing only one degree of freedom are unlikely to displace the equilibrium perceived price or profits by much. There is sufficient flexibility in the system to adjust. Furthermore, the optima are by nature broad and flexible. Regulation of the product itself is unlikely to alter the fundamental efficiency or inefficiency of the market. (Regulation of the rules of the market place including the number of producers and the pricing shemes allowable may have considerable influence.)

The lack of peaked optima influences the network design problem. Apparently a broad range of total seats and frequency may be provided in a market; cost related adjustments in fare can produce a service close enough to optimum. Therefore the definition of optimum network design is extremely difficult. Networks add yet another degree of freedom to the perceived price and cost functions. This degree of freedom is the number of intermediate stops per passenger trip on the service side, which translates into stage length on the cost side. Even without network adjustments, optima cannot be generally defined with any relevance. By examining networks designed to provide capacity and frequency in markets at minimum cost, we shall see the optima in this dimension can be shallow too. This is the task of the next chapter.

6 Networks

6.0 Introduction

Chapter 2 showed that a schedule of service could be most cheaply provided by a single carrier. Chapters 3 and 4 explored the consequences on the demand side of having service provided by one or only a few carriers. Chapter 5 developed in detail the technical possibilities for the cheapest and best (single carrier) service for markets ranging from 50 to 1600 passengers per day. All these discussions examined a single city pair market served by a dedicated nonstop vehicle link. But with few exceptions airlines operate <u>networks</u> of services involving hundreds of city pairs. Networks were not forced on existing airlines because new airlines could not be certificated. Nor is network service entirely due to the market and managerial forces which favor national corporations. Airlines operate networks because there are technical cost savings gained in so doing. These savings come from combining passengers from several markets onto a single large aircraft. Larger aircraft are cheaper per seat, so combining loads produces savings. In a sense, airline networks are a direct consequence of vehicle economies of scale.

This chapter focuses on some very simple network concepts which are of practical relevance to today's airline system. While formal statement of the network design problem has seen much study (cf. Wong [41] and his references), little has been done to translate the consequences into meaningful terms. Work by Gordon and de Neufville [14] on the airline network problem has broader relevance for the sparser European networks with largely connecting traffic. Regression analysis by Fruhan [12] using network measures suffered from over simplification and perhaps misstatement of the problem. The discussion below, then, is largely new work.

We will not be able to develop any closed form mathematical statements about networks. Indeed, part of the contribution of this effort is to illustrate that the minimum level of detail for such a solution goes far beyond the simplified measures used in the past. In the course of the discussions we will attempt to develop a level of understanding which will preclude further misuse of global activity measures of the kind influenced by network considerations.

Transportation services are operated in networks because most markets are too small to be served alone. Networks are a powerful means of mitigating the problem of small aircraft capacities. In the domestic U.S. there are only 50 city pairs with over 400 passengers a day each way. (1) For comparison, there are 25 airlines operating jet equipment of the kind described in chapter 2. Airlines operate at the frequencies and aircraft capacities which were discussed in chapter 5 because they operate networks of services where each flight is part of trips taken in a dozen or more city pair markets.

The fundamental technical gain in network design is in the direction of increased aircraft capacity. These scale economies

---

(1) 1976 CAB ticket survey data [7]. See also table 5.1.10.

are partially lost because load building means more expense in stops and shorter stage lengths. Designs for minimum cost involve balancing the use of large aircraft on short hops against longer hauls but smaller aircraft capacities. Thus vehicle cost (approximately equal to the cost of 50 seats) is traded off against departure costs (approximately equal to the cost of 200 miles of cruise). These costs are evaluated not for a single link or market, but as a total for the entire complex of services in the network.

6.1 The Network Design Problem

Networks: Supply Side

In its simplest form a supply network is a set of city pair links served at various frequencies by aircraft of various capacities. The cost is determined from the sum of the single link costs determined in chapter 2 and employed extensively in chapter 5. The only difference from chapter 5 is the cost for traffic which connects from one link to another. (For the moment we will not care whether the passengers travel on a through flight or make a physical change of planes.) For air travel this cost is a minor correction in expense. The early illustrations in this chapter will treat all movements through intermediate cities as passenger connections and will consider the expense to be roughly $6, or half the cost of an original boarding. This figure is discussed with the rest of the costs, in section 2.2 of chapter 2.

Networks are sometimes measured in terms of seat-miles or ton-miles. These are measures of network extent and tell nothing about the cost or characteristics of the service. A two link network of a million seat-miles is very different from a 50 link network of a million seat-miles. The same comments apply to measures of seat-departures and aircraft departures. For assessing costs or general viability of networks, measures of the density of network services are more useful indices of the state of the network. Such measures include aircraft capacity averaged over miles, departures, or block hours, stage length averaged over aircraft or seats, link frequencies or link departures averaged over links or link miles, links per city, and departures per city.

Attempts have been made (cf Greig [18]) to characterize air transport networks by simple measures, but no such general descriptors can be used for design because the distributions of stage lengths, capacities, and frequencies and even the correlation among these distributions affect general cost and service levels. This is why regressions using network measures such as the ones performed by Fruhan [12] have met with mixed results.

It is not determined what is the minimum level of detail that needs to be specified to characterize network average costs, but examples in this chapter will illustrate that distributions of capacity, stage length, and frequencies are necessary, though they may not be sufficient. (1)

Networks: Demand Side

Strictly speaking demand does not exist as a network. Demand applies to travel in a city pair. Nonetheless, groupings of city pair demands for the purpose of serving them with a supply network may be referred to as a network of demands. A network of demands is merely a list of city pair markets with enough common terminals to make joint services possible. Problems arise when demand networks are specified by a list of cities, with the list of city pairs by implication including all viable markets. It turns out the definition of viable is critical. For example, changing the minimum cutoff from 25 to 20 passengers a day will generally increase the number of markets by a large amount without affecting total network traffic significantly. Thus, while measures of the extent of networks of demand such as cities, passengers, or passenger miles (2) are stable, measures of demand network density such as passengers per market, markets per city, and markets per potential city pair (3) are not broadly useful because the number of viable markets is sensitive to the definition of viable. Perhaps the only stable density measure available to characterize a network of demands is average trip length.

Without contradicting the above statements, it is still useful to make a point about one aspect of network density. Usually a few markets dominate the total demand in a network. Among N cities there can be $N(N-1)/2=M$ city pair markets. M rises as N^2. Let the top m markets comprise 90% of the demand. Then m will tend to be between N and $N^{1.3}$ in the real world situations. The statement cannot be proven; it is not true in any general mathematical sense. But in the author's experience it is a useful rule of thumb. (4) Demand in a network is seldom uniformly distributed among markets. When the number of markets is near N, a thin or sparse network exists. Under these circumstances the flexibility in routing and aircraft capacity disappears and integerization and time of day effects dominate the technical design problem. This is the case in most international networks, but it is rare in U.S. domestic trunk networks, which is the problem we are addressing. (For an eloquent exploration of the problems of sparse networks, see Rosenberg [51], chapter 3.)

---

(1) Here we use the term "average cost" intuitively to indicate general cost performance. Measures such as $/seat mile have meaning only in the context of fixed demands and service levels.

(2) When specifying demand one should specify the fare and level of service for each market.

(3) The number of potential city pairs is developed in the next paragraph.

(4) The reasoning is somewhat circular, since it was the author's experience that established the rule of thumb.

This information points up the great difficulties in specifying network densities. A very important index for a network is the number of city pair markets per terminal. In fact an airline which dominates the service at a city is in a powerful position not only at that city but in all markets that may reasonably be routed through that city. But the markets per terminal rise non-linearly with the number of terminals included in the case. Thus network design responds to three partially interrelated measures: number of terminals, number of markets, and markets per terminal.

Networks: Load Factor

Load factor in a network is usually presented as an average over seat departures. Load factor is also sometimes presented as an average for a segment or link. Neither of these measures is appropriate. When studying a network design, load factor is best defined by city pair markets, regardless of which links are used in the path between the origin and destination points. (1) System average load factors per departure, per mile, or per block hour are like the other network density measures; the distribution about the average contains much necessary information.

Using a definition of load factor per market implies space is blocked on each link of any path used by a market. A given link will carry several markets, and sections of that link's capacity will be allocated at different load factors. There may even be seats on a link which are not reserved for any market and which are part of unassignable network overheads. (2) Under this formulation, the conventionally reported link load factors are averages among several markets and they may in theory include some deadhead seats. Designing a network using the traditional measure of link load factors creates a mathematical problem of complication and elegance. Using the distinction of market load factors eliminates the problem entirely. (3)

Networks: Intermediate Stops

The process of matching supply (in seats) and demand (in passengers) for a single link was complicated by the degree of freedom called load factor. Load factor affected both service level and cost. In networks there is yet another degree of freedom: the number of intermediate stops experienced by the traveller. Like load factor, this dimension affects both service and cost.

A city pair's demand can be served in a network either directly nonstop or by connecting or through service using one or

---

(1) The set of links used to make a multistop trip in a market is called a path.

(2) Such deadhead seats seldom occur in the relatively well-connected U.S. domestic networks.

(3) If variability in demands is partially random, as developed in chapter 4, then combined loads will have greater than expected space availability. However, this is a second order effect.

more intermediate stops. Our terminology calls each way a passenger might travel through a network a demand path for the market. Both stops and extra miles have significant costs, so practical options for a path neither stop too often nor go too far around. This limits the number of potential paths for any given market to perhaps a dozen options. (This limit makes possible the linear programming network designs like the one done later in this chapter.)

The matching of demand and supply for each market in a network is defined by the fare, load factor, frequency, and number of intermediate stops for that market. Cases often occur where more than one path serves the market. In this case some index such as a weighted average should be used to define stops and total frequency for the market. As a practical matter, the examples in this chapter will discuss market service in terms of equivalent non-stop service. The equivalence will be obtained by one of two methods. At first we shall maintain the single link perceived price by reducing fare just enough to compensate for the time spent in the intermediate stop(s). (1) (For this purpose 1 hour of time or $10 of value represents an intermediate stop.) We shall incorporate this $10 penalty into the network cost structure and then we shall look at minimum cost designs. This method is used in the simplified network examples. The second approach follows more correctly the methodology of Eriksen [11] in calibrating the market demand model. In the final network example, the frequency on multistop paths will be raised so that the expected total travel time including both stops and schedule delay is the same as for the design single link service. Neither of these compromises is ideal; the design should be able to explore either reduced cost or increased frequency as an option in compensating for intermediate stops.

In U.S. domestic practice, air passengers make an intermediate stop on roughly a third of their trips. The number of stops per trip averaged over a network of demands is an interesting index of the state of development of the network. For air use, the network average ranges between 1 and 2 stops per passenger trip. Truck and rail networks can range much higher than this.

We now begin to explore network designs using this compromise approach. Our design criterion is minimum cost for providing certain service levels to the network of demands. Because the $10 of passenger experienced cost is included, the design considers service quality in the dimension which is uniquely relevant to networks-- the number of intermediate stops. The other quality dimensions, load factor and frequency, are fixed. This approach highlights the special network issues.

---

(1) We assume that people passing through an intermediate airport are on through flights or are making a close and convenient connection. This is the case for U.S. domestic trunk systems. European systems involve a second displacement time at the intermediate airport, as developed by Gordon and de Neufville [14].

Optimal network designs in the sense of optimal used in chapter 5 would consider all degrees of freedom in design.

Network Example: Simple Linear Network

Network relationships may be illustrated using the smallest possible network. To simplify the design problem, we will require four departures daily for each market and load factors of 67%. This fixes the service quality in the dimensions not unique to network design. Figure 6.1.1 describes a simple linear network of four cities and five city pair demands. In figure 6.1.2 three service patterns are represented: the spanning tree, an intermediate solution, and an all direct service pattern. In this case the direct service pattern also characterizes the network of demands because the average seat stage length is the average passenger trip length and there are no intermediate stops per passenger trip.

The three service patterns may be compared for fairly typical airline conditions. With distances of 500 miles and demands of 200 passengers per market per day, the three service patterns are tallied in table 6.1.1. For this network of demands, the spanning tree service pattern is unduly expensive because too many long haul passengers are being cycled through too many intermediate stops. This costs both money to the airline and time to the passengers. The direct service pattern, on the other hand, is too expensive because the aircraft capacity is small and the expense of vehicle frequencies is excessive when averaged over the on board demand. The intermediate network design represents a balance between eliminating intermediate stops and combining markets to achieve densities suited to larger aircraft. As the number of vehicle departures grows from left to right in table 6.1.1, the average stage length grows (a saving) but the average aircraft capacity drops (a loss).

What kind of network is appropriate depends on the size of the demand in comparison to the per vehicle costs. The example above was carefully chosen to lead to an intermediate solution to the network problem. In our the vehicle costs are approximately equal to the cost of 50 seats. (1) Table 6.1.2 displays the costs of the three network designs for different demands. At the top of the table the low density markets are best served with many intermediate stops and short stage lengths (the tree network). In the tree network maximum load building is achieved, spreading vehicle costs as much as possible. Stage length has been sacrificed in order to use increased aircraft capacities. At the high densities at the bottom of the table, most of the economies of increased vehicle capacity have been attained, and maximum stage length is reached by doing without load buildingnd using aircraft capacities below the absolute largest achievable. Stage lengths and aircraft capacities are also presented in table 6.1.2.

The other influence on network design is the relative

---

(1) We recall that our engineering cost structure divides aircraft costs into per vehicle and per seat costs. Vehicles provide frequency. Seats provide capacity.

Figure 6.1.1: Simple Linear Network Problem:

A ←500 mi→ B ←500 mi→ C ←500 mi→ D

demands:	AB - 200	passengers	per day
	BC - 0	"	"
	CD - 200	"	"
	AC - 200	"	"
	BD - 200	"	"
	AD - 200	"	"

Assumed service levels and costs:

all markets must be served four times daily

load factor is 67% by design

Costs from Chapter 2:

$12.64 per passenger boarding

$0.008 per passenger mile

$379.8 per aircraft departure

$3.27 per seat departure

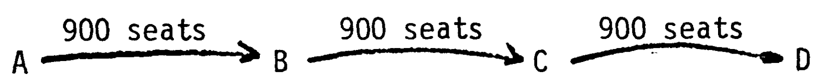
$0.816 per aircraft mile

$0.0176 per seat mile.

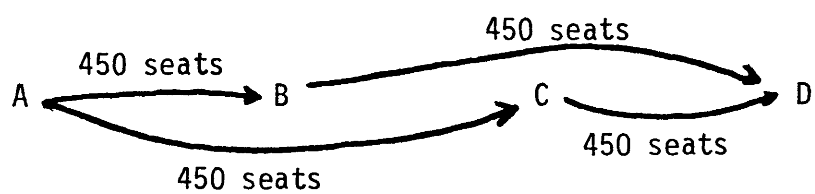
A cost of $6 fnr handling and $10 for passenger's time is added for each intermediate passenger stop.

Figure 6.1.2: Linear Network Service Patterns

Minimum Tree:



Intermediate Design:



Direct:

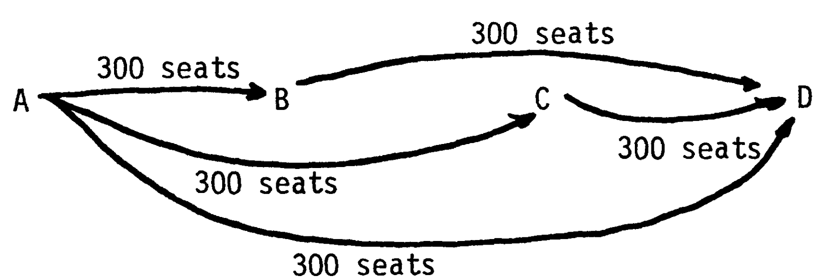


Table 6.1.1: Data for Simple Linear Network and three service designs for distances of 500 miles and demands of 200 passengers Service four times daily at 67% load factor.

	service designs		
measure	tree	intermediate	direct
vehicle departures	12	16	20
vehicle miles	6000	12000	18000
seat departures	2700	1800	1500
seat miles (000's)	1350	1350	1350
Costs ($000's)	$75	$69	$71
Average Vehicle Stage Length	500mi	750mi	900mi
Average Passenger Trip Length	900mi	900mi	900mi
Average Passenger Hop Length	500mi	750mi	900mi
Average Stops per Trip	1.8	1.2	1.0
Average Vehicle Capacity (per mi)	225	113	75
Average Capacity per Departure	225	113	75
System Passenger Handling $	$19,840	$19,840	$19,840
System Minimum Seat Costs[1]	$28,665	$28,665	$28,665
System Frequency Costs	$ 9,752	$15,869	$22,284
Seat costs above minimum[2]	$16,724	$ 4,181	0
total costs	$74,683	$68,555	$70,789

[1]Minimum Seat Costs are the cost of direct milage and one departure for all passenger traffic at the stated load factor.

[2]Seat costs above minimum are the costs of intermediate hops for seats at design load factor. This includes $6 handling costs, $10 of passenger time costs, and $3.27 per seat departure cycle.

Table 6.1.2: Linear Network at different Densities

market densities	System costs in $1000 for each design: tree		intermediate		direct
50	$25.8		$29.0		$34.4
100	42.1	or	42.2		46.6
200	74.7		68.6		70.8
300	107.3		94.9	or	95.0
400	139.9		121.2		119.3

optima are underlined

---

market densities	Aircraft Size in Seats for Network Design: tree		intermediate		direct
50	66		33		19
100	113	or	66		38
200	225		113		75
300	338		169	or	113
400	450		225		150
stage length:	500mi		750mi		900mi

Table 6.1.3: Linear Network System Costs at 1000 miles

costs in $1000

market densities	System Costs in $1000 for each Network Design: tree	intermediate	direct
100	$62.4	$67.5	$76.7
181	101.3 or	101.3	108.9
200	110.5	109.3	116.4
544	275.7	253.1 or	253.1
800	398.8	360.1	340.4

cost minima are underlined

Table 6.1.4: Linear Network System Costs at 100 miles

market densities	System Costs in $1000 for each Design: tree	intermediate	direct
25	$10.6	$11.5	$13.5
40	13.6 or	13.6	15.3
100	25.7	22.0	22.4
120	29.1	24.8 or	24.8
200	45.9	35.9	34.3

cost minima are underlined

importance of departure and cruise costs. In our example each departure costs roughly the same as 250 cruise miles. For a network with 1000 miles between cities instead of 500, departure costs are a smaller fraction of the total. Under these conditions, the tree network becomes more suitable because the reduced importance of the cost of intermediate stops favors the network design with more such stops (see table 6.1.3). The stage length of the tree network (1000 miles) is long enough to nearly exploit the achievable economies in this dimension, so longer stage length network designs are not needed.

Table 6.1.4 shows the case for a short haul network with distances of only 100 miles between cities. Here the cost of extra passenger departure and landing cycles is dominant and the network with the least of them and the greatest stage length is the best. Thus the direct flight service dominates all but the smallest demand levels.

Not all the qualitative conclusions from this small network example carry over to all networks in all cases, as we shall see in our next example. In particular the regular growth of stage length with increased number of links sometimes fails. Such occurrences make general laws about network design impossible. However, the trends apply to most transport networks in most cases. Increased demands favor more direct service and longer stage lengths. Increased departure costs do the same.

Algebraic Statement: Simplified Trends

Even with the simple example above, we notice several network trends. Larger aircraft are used with shorter stage lengths, more stops for passengers already on board, and generally less direct service. Smaller aircraft are used in more direct service, longer stage lengths, and less stops for passengers on board. We may crudely formalize some relationships. Let us define

D the demand per city pair market (system average)
T the demand average trip length
M the number of city pair markets
N the number of cities

We can characterize the demand network using these variables, although they are not sufficient to determine the network design. (1) Corresponding to the demand network descriptors, there are a number of useful variables for characterizing the supply network design:

l the number of links served
f the frequency per link (network average)
c the aircraft capacity (system average)
s the average stage length (either per vehicle or per seat)

---

(1) Distributions of densities and stage lengths are also necessary to characterize a network for design. Even these may not be sufficient.

n the number of stops per passenger trip (system average)

For a design without slack capacity and without passenger circuity:

$$n \sim T/s \tag{6.1}$$

This says that passenger departures and seat departures are closely correlated when load factor is constant. The number of intermediate stops per passenger trip (n) tends to rise as the average stage length (s) falls. (1)
We can also say

$$D \cdot M = f \cdot l \cdot c = \text{constant} \tag{6.2}$$

This states that traffic and seats must match when load factor is constant (=1.00) and circuity is negligible. As frequency (f) rises, capacity (c) and number of links served (l) tend to fall. Finally we can indicate a third trend:

$$n \sim M/l \tag{6.3}$$

This says that networks with fewer links route more traffic over the links which do exist. As the number of links falls, the number of intermediate stops tends to rise.

These relations say that for a fixed network of demands one can have (1) frequency at the expense of extra intermediate stops and smaller aircraft capacities, or (2) direct service at the expense of frequency and aircraft capacities, or (3) low cost service (large aircraft capacities) at the expense of frequency. (2) Mathematically combining equations (6.2) and (6.3) we get:

$$f \cdot c/n \sim D \tag{6.4}$$

From this we see most clearly the new degree of freedom offered by network design. For a single link, frequency times capacity must equal demand (when demand has been adjusted for load factor). In the single link case there is only one market (M=1) and one link (l=1). The number of stops (n) must be fixed (n=1). In a fully developed network, n ranges from 1 upward. For airline practice n is between 1 and 2, as we have said.

In network design the correspondences in equations (6.1) to (6.4) are not strict proportionalities. The trends can even be reversed over some ranges of network design options. Much depends on the distribution of demand per market and frequency

---

(1) The equation is true when only one aircraft capacity is employed. Otherwise it is merely a trend.

(2) For an excellent discussion of just these tradeoffs from a slightly different perspective, see Gordon and de Neufville [14], and Rosenberg [51], chapter 8.

and capacity per link about their averages and on the spatial layout of the network. However, for efficient solution to networks of demands distributed over regular geographical areas, the tendencies hold true, as we hope to illustrate further.

6.2 Small Hub and Spoke Case Study

Network design involves simultaneous tradeoffs among vehicle departure cycles, vehicle miles, capacity departure cycles, capacity miles, and the cost of intermediate stops to the traffic itself. In order to explore these considerations, it will be best to start with a stylized network rather than a real list of cities and markets. With the symmetry and simplicity of a stylized example, the measures and tradeoffs will be more obvious at the first. Our next step will be to examine more realistic situations.

We start with the geography of figure 6.2.1. There are two major hub cities and each hub has three nearby satellite cities. Inter-hub demand is large. Hub to satellite demand is modest, and inter- satellite demand is small. To keep the network design tradeoffs clear, we will fix frequency and load factor. Load factor for all markets will be required to be 65%; daily intercity frequency will be required to be two flights. There will be a $10 expense of passenger time for each passenger who passes through an intermediate terminal. Otherwise the costs from chapter 2 can be used:

Cost per vehicle stage=$379.8
Cost per vehicle mile=$.816
Cost per seat stage=$3.27
Cost per seat mile=$.0176
Cost per passenger hop=$6.00
Cost per passenger trip=$6.64
Cost per passenger mile=$.008

For the stylized case of this study, any aircraft capacity can be used on any link. There will be no deadhead or backhaul seats. Consequently, passenger and seat costs can be combined using the sytemwide design load factor in the manner analogous to equation (5.21) in chapter 5. Also at a design frequency of two, link costs can be stated explicity by rearrangement of the previous numbers:

Cost per link=$759.6
Cost per link mile=$1.632
Cost per passenger hop=$21.03
Cost per passenger mile=$0.0351
Originating passenger correction=-$3.34

The network design which minimizes the total <u>link</u> costs is the minimum spanning tree with 7 links. This is shown in figure 6.2.2. Also shown in figure 6.2.2 is a second 7 link network of almost identical total link costs, but greatly reduced number of passenger transfers. In our experience network designs often have a natural minimum network very close to the mathematically defined minimum spanning tree but with the most heinous circuities corrected. In general it is best to start with this natural minimum network when designing by adding links.

Links can be added to the tree network until all 28 city pairs are served nonstop. The nonstop solution maximizes link costs, but minimizes total passenger related costs. In this process the first links to add for the case study are hub bypass links hooking the southwestern satellites to the eastern hub (or vice-versa). The first such link removes one hub transfer for all 130 passengers headed from the western satellite to the east. Bypass routes from the eastern satellites to the west are just as good in the symmetrical case. With six bypass links added, 560 transfers are eliminated. The system has half the possible links (13 out of 28) and almost all the reasonable savings. The number of transfers and a few miles of passenger circuity can be reduced by adding some short local intersatellite links and then the longer east-west intersatellite links, but the rewards are few.

This link addition process is traced from the start in tables 6.2.1 through 6.2.4. Table 6.2.1 lists the number of links and the totals of the activities necessary to calculate costs. Physical system costs (without the $10 of passenger experienced transfer costs) are listed in the last column. As links and link miles are steadily added, passenger departures are steadily reduced. Passenger miles are also slightly reduced by the removal of circuity. The minimum cost to the operator of meeting the design service and capacity levels occurs with 9 links.

The design tradeoffs are highlighted in table 6.2.2. Here links and link miles above the minimum and passenger departures and miles above the all-nonstop case are reported. These are the "extra" links and "extra" passenger departures listed in the first four columns. One sees at the start the addition of one to six hub bypass links adds modestly to link costs while saving large numbers of departures. There is a cost minimum at six "extra" links for the case where both operator physical costs and passenger time costs are included. Additional links serve smaller and smaller markets or groups of markets. Beyond link 14, a few passenger stops are removed at great expense in terms of link numbers and link miles.

Notice that the optimal system from a minimum physical cost view has 9 links and the optimum design from the view of minimum cost including the cost of passenger inconvenience has 13 links. There is a tradeoff of cost and quality beyond the tradeoffs already made in determining frequency and load factor in chapter 5. That tradeoff is between cost and the number of intermediate stops.

Notice also that in a full optimization shared costs of frequency would be reflected in the specifications of per market frequency. The instrumental variable in network design is aircraft route frequency, involving several links, stops, and markets. Aircraft route frequencies are justified by the benefits to all the markets involved. In our discussion we do not optimize design over a network; we merely explore the cost advantages involved.

Measures which correspond more closely to single link or system average transportation parameters are presented in table 6.2.3. In a sense the all nonstop service for a network corresponds to single link service as an average among the markets. Starting with this nonstop network at the bottom of columns 2 and 3 and moving upward, we delete links which are expensive and thinly used. There is a steady increase in the average traffic density on the remaining links. Notice that density per departure and per mile are different due to distributions of aircraft capacities over links. Higher link densities mean lower average per seat operating costs because the cost of frequency (vehicle cost) is averaged over more seats. Costs do not fall in direct proportion to the economies of aircraft capacity because intermediate stops (column 6) are being experienced by passengers, causing both operating and traffic-borne expense.

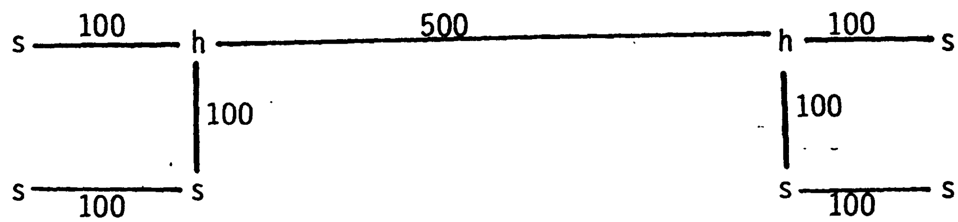
From this table we can notice two irregularities which apply in the real world. First, vehicle stage lengths do not directly correspond to passenger hop lengths. Vehicle stage lengths do not even change monotonically with link addition because of the distribution of aircraft capacities over link lengths (column 4).

The second point that is typical of network designs is seen in column 7. Passenger circuity for a network of demands is the ratio of path miles to the direct market distances. Because mileage cost is significant, minimum cost designs tend to use the shortest path the network of existing links provides. Thus passenger circuity is never very large. This is almost always the case for airline network designs, even though circuity can be substantial in mathematical network designs with lower costs per capacity mile.

Table 6.2.2 showed a cost minimum at 6 links added to the minimum network (or 15 links removed from the all-direct network). This minimum changes depending on the problem. If daily frequency requirements are set at three flights per day instead of two, as the second to last column of table 6.2.4 (column (a)) shows fewer links are cheaper. If passenger demand is quadrupled, more direct service (more links) is called for, as shown in column (c) of table 6.2.4. Figure 6.2.3 shows which links are involved in the three designs of columns (a), (b), and (c) of table 6.2.4.

It would be valuable to define general laws for tne minimum cost number of stops, trip length, number of links, and so forth for networks in general. Indeed it was once the hope of this research to be able to do so. However, optimal network design depends very much on the distribution of demand densities over the network as well as on the relative departure and mileage, vehicle and capacity costs. Furthermore, as we can see in this simple case study, the optimum is quite shallow. In the 13 link optimum network there are several equal or very nearly equal paths for most of the multistop demands to take. What is more, the addition or deletion of a link has modest effects. Removal of a link generally diverts the traffic to a path of only slightly greater length and the same number of stops. For

Figure 6.2.1: Geography of Small Hub and Spoke Network



distances in miles

h - hub terminals

s - satellite terminals

all angles 90^{o} as shown, Intercity distances determined by plane geometry where not explicitly stated

Demands:

hub to hub - 1000 pax per day (1 market)

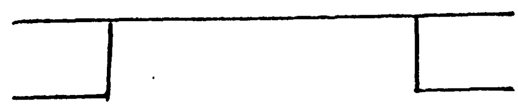
satellite to hub - 100 pax per day (12 markets)

satellite to satellite - 10 pax per day (15 markets)

total of 28 markets

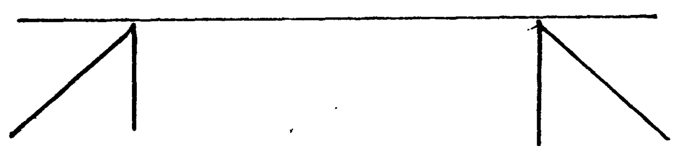
Figure 6.2.2 Seven LInk Designs for Hub and Spoke Example

Minimum spanning tree:



link costs=$3556;   Passenger departures=3650

"Natural" minimum network:



link costs=$3623; Passenger departures=3190

Table 6.2.1: Small Hub and Spoke Example: Activities

number of links	total link mi	Passenger departures	Passenger mi (000)	System cost (000)
7*	1100	3650	1041	$96.2
7	1182	3190	1016	93.6
8	1692	3060	1004	93.4
9	2202	2940	993	93.3
10	2809	2820	988	93.5
11	3416	2710	985	94.0
12	4016	2610	985	94.6
13	4616	2500	985	95.1
17	5016	2460	980	98.2
20	5798	2440	976	101.4
24	8212	2390	976	107.8
28	11024	2350	976	115.0

*minimum spanning tree

Table 6.2.2: Small Hub and Spoke Network Example: Extra Costs

extra links	extra link-mi	extra pax-dep	extra pax mi (000)	extra physical cost ($000)	extra total cost ($000)
0	0	840	40	$10.5	$19.0
1	510	710	28	10.4	17.5
2	1020	590	17	10.3	16.2
3	1627	470	12	10.5	15.2
4	2234	360	9	11.0	14.6
5	2834	260	9	11.6	14.2
6	3434	150	9	12.1	13.6
10	3834	110	4	15.2	16.3
13	4616	90	0	18.4	19.3
17	7030	4	0	24.8	25.2
21	9842	0	0	32.0	32.0
+7	+1182	+2350	+976	+83.0	+83.0

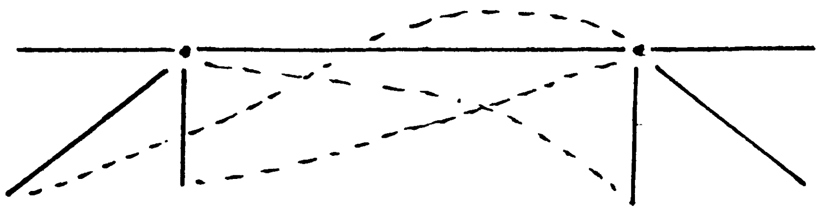
The figures on the bottom line with the + signs are the minimum costs to which the "extra" costs are to be added

Table 6.2.3: Small Hub and Spoke Example: Indices

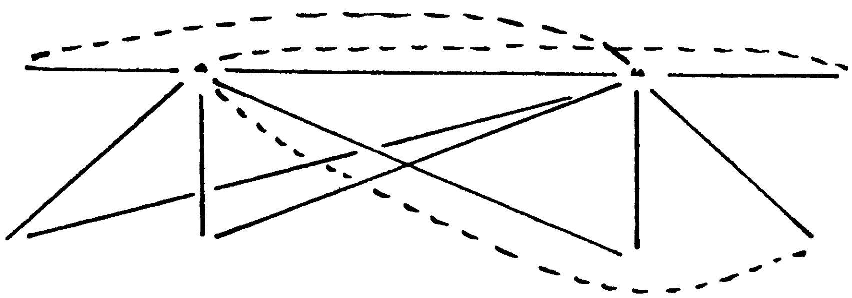
(1) total network links	(2) link densities: PAX-mi / Link-mi	(3) PAX-dep / Link-dep	(4) stage: link-mi / links	(5) hop: PAX-mi / PAX-dep	(6) stops per PAX trip	(7) circuity PAX-mi / demand-mi
7	859	456	169	318	1.36	1.041
8	593	383	212	328	1.30	1.029
9	451	327	245	338	1.25	1.018
10	352	282	281	351	1.20	1.010
11	288	246	311	364	1.15	1.010
12	245	218	335	378	1.11	1.010
13	196	192	355	394	1.06	1.001
17	195	145	295	398	1.05	1.000
20	168	145	290	400	1.04	1.000
24	119	122	342	408	1.02	1.000
28	89	84	394	415	1.00	1.000

Figure 6.2.3 Links Added to Tree Network for cases of Table 6.2.4

(a) First Three links for frequency=3 case:



(b) 3 more links for Standard case:



(c) Four more links for demand quadrupled case:

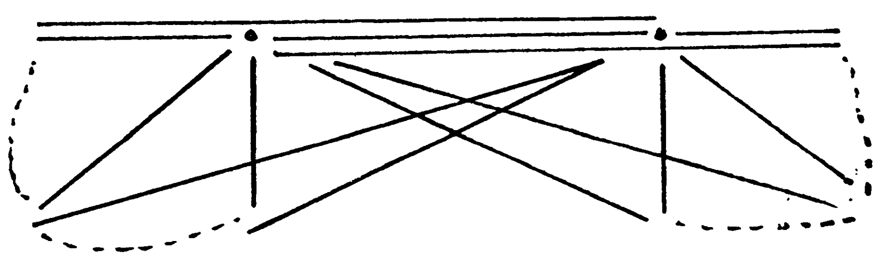


Table 6.2.4: Small Hub an d Spoke Example; Different Sizes

number of links in network	(b) Cost Standard case	(a) cost frequency=3 case	(c) cost pax·4 case
7	$102.4	$105.7	$386.8
8	100.5	104.9	375.8
9	99.2	104.4	365.8
10	98.3	104.3	356.7
11	97.5	104.6	348.8
12	97.2	105.1	342.1
13	96.7	105.4	334.6
17	99.3	109.9	334.2
20	102.3	114.7	335.5
24	108.3	124.1	338.3
28	115.0	134.7	342.6

costs in (000's) include $10 per passeger connection to compensate for passenger's time.

instance, adding the 6th new link increases the average stage length by 8%. This would seem to reduce average cost per mile. But at the same time aircraft capacity is reduced by 15%, which increases average cost per mile. Also, distributions of densities and stage lengths change. The net result is a very small change in average cost per passenger mile.

It is true in practice that stage length and aircraft capacity economies tend to share any gains made in redesigning a network for greater demand densities. As a rough rule of thumb, they share these gains equally. But this is merely a tendency, and cannot be reduced to a general rule.

Another thing we can see from this example is that the optimum is so shallow that it could easily be displaced by reasonable readjustments of costs or parameters. In a sense the broad and indefinite nature of the optimum network design is analogous to the broad optimum found for single link service design. Network design adds a new degree of freedom to a process of optimizing technical performance which is already quite adaptable. Network design has so much freedom to adapt that coupled with market by market flexibility in fare, frequency, and load factor there is no way to define optimality with any hope of generality.

6.3 Cost Allocation in a Network

We saw in chapter 5 that the cost of frequency was not included in the marginal cost of traffic in an isolated city pair market. At fixed frequency and load factor, marginal cost per passenger did not cover vehicle costs. Similarly, if we fix the frequency throughout a network, the marginal cost for traffic in any single sity pair market is the cost of capacity (adjusted for load factor for that market) and the cost of handling for the miles and departures experienced on the <u>path</u> used for that market. This is the answer we get if we require all markets and service levels be fixed, and that our derivative be the cost of adding one more passenger to one market. The costs of vehicle frequency are again excluded.

The cost of intermediate stops is included in marginal cost for traffic in a network, as is the cost of any miles of circuity. Because of stops and circuity, the marginal cost for traffic served in a market is equal or above the marginal cost for the same market isolated. This occurs even though the total network cost is below the summed costs of its markets served in isolation. Also with marginal cost depending on stops and circuity, marginal cost depends on the network design, which in turn depends on traffic in other markets in the network and the design optimality criterion.

The picture we now have is that the costs of all capacity movements and departures show up as marginal costs in one market or another and can be allocated to the traffic in those markets. (1) The cost of frequency cannot be assigned to a specific

---

(1) We have assumed no empty backhaul or deadhead seats.

market with any degree of certainty. Thus we cannot express average costs for a market, even though that is what we did to set fares in the single link case. With any degree of overlap of paths, frequency costs can seldom be allocated among any subgroup of markets with clear justice.

In addition, when it comes to allocation of costs to markets rather than mathematical definitions of marginal cost for traffic, the cost of intermediate stops and circuity for a market cannot with fairness be said to be caused by (the traffic in) that market. (1) With the market alone, those stops and that circuity would not exist. With the market combined in a different network, the stops might not exist. Intermediate stops in one market are incurred in part to reduce frequency cost which would result for some other market if it were served alone.

As we saw in the maximum surplus case in chapter 5, marginal cost as an issue is a simplification of maximum surplus considerations. But maximum surplus on a network should allow the degree of freedom of changing service levels in the other markets of the network, whereas marginal cost is assessed with all network performance fixed. (The author has illustrated inter-market traffic and service adjustments due to marginal traffic in a single market in [34].) With service and traffic adjustments possible, the usefulness of strict marginal cost definitions becomes less obvious.

With this in mind, the only expense that can without argument be allocated to a market is the cost of its own capacity direct without stops. The rest of network cost in frequency, intermediate stops, and circuity must be allocated around the network and among the markets. How it is allocated is almost entirely arbitrary. However, a reasonable upper bound is the cost of the service the market would get if isolated. For reasons of efficiency, the lower bound must be the marginal cost for traffic with stops and circuity included. In most cases this lower bound is too low and too close to the minimum marginal capacity cost to be an issue. For practical purposes cost allocation in a network by markets is arbitrary between the average and the marginal traffic cost of the market served alone.

Price Discrimination Among Markets in a Network

The rule above opens up opportunities for price discrimination among markets to cover non- allocated costs. (In the single link case, discrimination could occur only among customers within a market.) The problem is, how is discrimination among markets in a network defined? Is the non-discriminatory norm a constant markup over marginal traffic costs in a market? Is it a fixed sum addition per market for frequency? Is it a fixed sum addition per passenger for frequency? Is it a constant discount from costs for the market served in isolation? Is it related to the extra cost of adding

(1) The concept of extra cost per market has only secondary relevance to economic efficiency, but it is the first measure sought in examining justice of cost allocations. As a culture we intuitively seem to prefer to pay costs within their own markets.

the market to the network?

Legally speaking, inter-market discrimination even in clear theoretical network examples is probably not provable for prices between the average cost for service in isolation and the marginal capacity costs. Practically speaking, some exploitation by discrimination does occur unless one of the rules for pricing listed above is consistently applied across all markets. As long as prices are below isolated market minimums, society may be willing to let exploitation be limited by the usual combination of oligopolistic competition, the threat of entry, and the mixed non-profit-maximizing objectives of managements of large corporations.

In terms of economic efficiency alone the ideal situation would be to mark up maginal costs in proportion to the traffic's inelasticity. That is, discriminate among customers in each market according to their elasticities and among markets according to the general changes of elasticities short to long haul. In practical terms, this means tax business and long haul travel to recover fixed costs, and ignore market density. This appears a fair description of what has happened in the past. In a future competitive environment, dense markets may not support heavy markups.

Fares in the Hub and Spoke Network

The stylized case study in section 6.2 illustrated the basic cost tradeoffs in network design. Service requirements were oversimplified in order to highlight the design cost tradeoffs. Nonetheless, it is interesting to examine the fares that might exist in such a network. Table 6.3.1 shows the bounds suggested by our discussion above. The lower bound fares are the marginal cost of traffic in direct service for a market of such distance and service attributes. Because design load factors and frequencies do not change with traffic in our example, these costs are the same for 10 passenger or 100 passenger markets. These lower bound fares are the fares which would be suggested as the maximum net surplus fares for the traffic and service provided if provided in isolation.

The upper bound fares in table 6.3.1 are the average costs of serving the traffic as on a single link. Notice that the upper bound for the markets with traffic 10 is absurdly high. If served in isolation, a market with demand near 10 (1) can only be served with subsidy: the market is below the minimum for profitable service when served alone.

Finally in table 6.3.1 there is a column of possible fares. These are the fares for 100 passenger markets applied universally, except where the upper bound is cheaper. The income from this arrangement covers the costs. (The $10 connection penalty has not been included in any of the figures of table 6.3.1.)

It appears from this example that the gap between the upper and lower bounds on fares drops with traffic. The effect is

(1) We mean a demand of 10 at reasonable fare and service level. In this stylized example we have traffic of 10 for any fare.

Table 6.3.1: Bounds for Fares on Small Hub and Spoke Network

service levels in all markets at 65% load factor
frequency of 2.0

Market passengers	Market distance	lower bound	upper bound	number of markets	possible fare
10	100	$21.20	$112.21	4	$30.43
10	141	22.64	121.61	2	32.54
10	~700	42.26	232.46	9	61.28
100	100	21.20	30.43	4	30.43
100	141	22.64	32.54	2	32.54
100	~600	38.75	56.14	6	56.14
1000	500	35.24	36.30	1	36.30
incomes:		$76,602	$116,506		$96,047

physical costs at 9 links = $93,300
" at 13 links = $95,100

Figure 6.3.1: Fare Bounds for 800 Mile Market in a Network

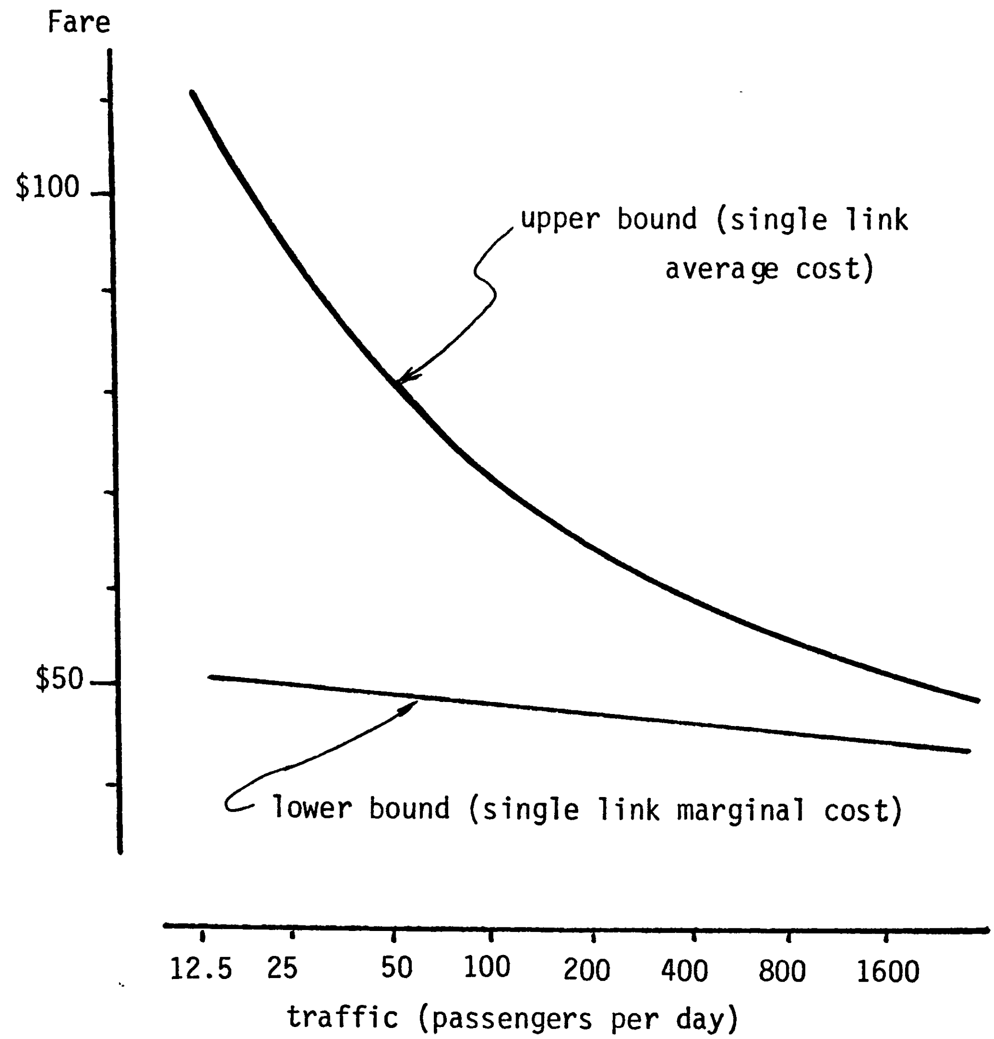


Table 6.3.2: Fare Bounds at Different Densities

200 mile distance:

PAX	frequency	load factor	minimum fare	maximum fare
12.5	1.0	.52	$27	$71
25	1.5	.54	27	59
50	2.2	.56	26	50
100	3.3	.58	26	44
200	4.9	.60	26	39
400	7.2	.62	25	35
800	10.7	.65	25	32

800 mile distance:

PAX	frequency	load factor	minimum fare	maximum fare
12.5	0.7	.56	$50	$108
25	1.1	.58	49	92
50	1.6	.61	48	80
100	2.3	.63	47	71
200	3.4	.65	46	63
400	5.1	.68	45	58
800	7.6	.70	44	54

3200 mile distance:

PAX	frequency	load factor	minimum fare	maximum fare
12.5	0.5	.61	$136	$255
25	0.7	.63	133	221
50	1.1	.65	129	195
100	1.6	.68	126	175
200	2.4	.70	123	159
400	3.6	.73	120	147
800	5.3	.76	117	137

somewhat diminished if frequency and load factor are allowed to adjust to their single link optima at different market sizes, but the point still holds true. Table 6.3.2 presents the service and fares for markets at 200, 800, and 2600 miles and different densities according to the single link maximum traffic optima. The optimal single link fares are also the upper bound for network fares. The final column of table 6.3.1 gives the maximum surplus fare for these traffic levels, which we have taken as the lower bound on network fares. Figure 6.3.1 plots the two fare bounds for the 800 mile market, showing graphically how the bounds narrow as market density grows. These are the bounds with load factor and frequency continuously adjusted to the maximum traffic optimum as market size changes.

6.4 Dense U.S. Domestic Network Example

We now turn our attention to a final network example. The network of demands is all 120 city pairs among the 16 largest traffic hubs in the U.S.. This provides an example of a network of markets so dense that isolated single link service is viable for most city pairs. A map and a list of the cities is presented in figure 6.4.1. The distances, approximate traffic in 1974, and the service at maximum traffic for each market served in isolation are presented in table 6.4.1. (1)

Initially we discuss the network in the same stylized terms as we did the previous hub and spoke example: we insist on service at 65% load factor and at a frequency of 8 per day for each market. We charge $6 in operating costs and $10 in passenger time costs for each passenger who passes through an intermediate city. This is in addition to the vehicle and mileage costs that the stop incurs. As before we examine the network by the process of postulating certain link configurations and calculating their costs.

Specifically we examine the 8 networks described in figure 6.4.2a-d. Each of the 8 solutions was meant to represent a distinct design case. None of the 8 choices is necessarily optimal in any sense at all. However, the 8 solutions do represent reasonably good networks. (2)

The smallest network contains only 16 links and is very nearly the minimum spanning tree. (A false node exists at Salt Lake City in this network.) This design is inappropriate for air and is included only to illustrate the extreme point in minimum link network design. Actually this 16 link network would be suitable only if the cost of adding links were huge and the cost

---

(1) Maximum traffic service in terms of frequency and load factor for each market was predicted according to the interpolation formulas from table 5.1.8 for distance and traffic as stated. Fare is average cost under these conditions.

(2) These and other stylized network designs in this chapter were explored by trial and error link addition and deletion using an interactive computer program to evaluate the networks.

Table 6.4.1: City Pair List

City Pair	distance	traffic	frequency	load factor	fare
NYC PHL	84	184	5.8	.570	$ 36
WAS PHL	133	186	5.2	.586	37
NYC BOS	191	2562	21.1	.686	28
WAS PIT	193	297	6.2	.613	36
PIT DET	198	238	5.4	.607	38
NYC WAS	215	2562	20.5	.691	29
ORD STL	256	866	10.6	.659	34
DET ORD	238	940	11.3	.659	33
PHL PIT	273	479	7.4	.641	37
BOS PHL	274	588	8.3	.648	36
NYC PIT	329	919	10.3	.670	36
LAX SFO	355	1489	13.2	.691	35
WAS DET	391	1108	10.9	.684	38
PIT ORD	404	556	7.3	.661	41
BOS WAS	406	1003	10.2	.682	38
ATL TPA	409	386	5.9	.649	43
MSY DFW	423	403	6.0	.651	43
ATL MSY	425	301	5.1	.642	45
DET STL	451	178	3.7	.626	50
PHL DET	452	295	4.9	.643	46
STL ATL	484	251	4.4	.640	49
NYC DET	489	1006	9.8	.689	41
TPA MSY	495	97	2.6	.610	57
BOS PIT	496	244	4.3	.640	49
PIT ATL	526	237	4.2	.641	51
STL DFW	537	969	9.3	.691	43
WAS ATL	540	518	6.5	.669	46
PIT STL	554	127	2.9	.623	57
WAS ORD	591	919	8.8	.693	45
ORD ATL	597	612	7.0	.679	48
DET ATL	602	275	4.4	.651	53
STL MSY	604	133	2.9	.627	59
BOS DET	623	263	4.3	.651	54
DFW DEN	664	455	5.8	.672	52
S-O SEA	671	582	6.6	.682	51
PHL ATL	672	321	4.7	.661	54
PHL ORD	675	645	7.0	.685	50
WAS STL	707	286	4.4	.659	57
ATL DFW	707	537	6.2	.681	52
NYC ORD	721	2645	15.4	.741	46

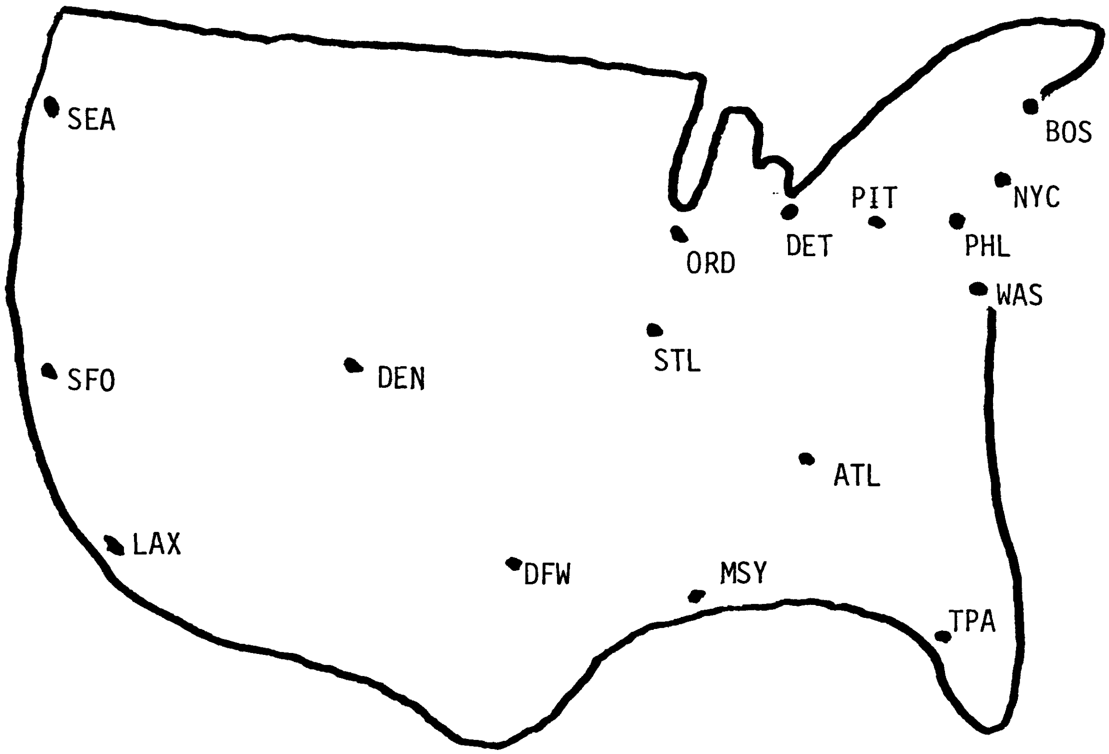
Table 6.4.1: City Pair List (continued)

City Pair	distance	traffic	frequency	load factor	fare
NYC ATL	755	1007	8.8	.706	$ 51
STL DEN	781	215	3.6	.653	62
ORD DFW	790	678	6.9	.693	54
WAS TPA	810	151	2.9	.642	67
PHL STL	820	138	2.8	.639	68
ORD MSY	831	322	4.5	.669	61
DEN LAX	839	655	6.7	.695	56
BOS ORD	860	689	6.8	.697	57
STL TPA	873	108	2.4	.634	73
PIT TPA	873	163	3.0	.647	68
NYC STL	882	516	5.8	.688	59
ORD DEN	907	1089	8.7	.716	56
TPA DFW	911	121	2.5	.639	73
PIT MSY	918	67	1.8	.620	81
PHL TPA	922	100	2.2	.633	76
DET MSY	936	78	1.9	.625	80
BOS ATL	946	216	3.4	.660	69
DEN SFO	956	392	4.8	.681	64
LAX SEA	959	647	6.4	.699	61
WAS MSY	962	136	2.6	.645	74
DET DFW	983	164	2.9	.652	73
DET TPA	991	296	4.1	.673	68
NYC TPA	1003	645	6.3	.701	62
ORD TPA	1006	460	5.2	.689	65
DEN SEA	1020	247	3.6	.667	71
BOS STL	1046	127	2.5	.646	79
PIT DFW	1049	107	2.2	.640	81
DEN MSY	1067	115	2.3	.643	81
PHL MSY	1094	100	2.1	.639	84
DET DEN	1144	164	2.8	.658	80
WAS DFW	1161	321	4.1	.681	74
NYC MSY	1177	388	4.5	.689	73
BOS TPA	1182	207	3.2	.667	79
ATL DEN	1208	537	5.4	.702	71
DFW LAX	1248	667	6.1	.711	71
PHL DFW	1289	186	2.9	.666	84
PIT DEN	1302	90	1.9	.642	95
NYC DFW	1363	664	6.0	.714	75
BOS MSY	1367	77	1.7	.638	100
WAS DEN	1476	237	3.2	.680	89

Table 6.4.1: City Pair List (second continuation)

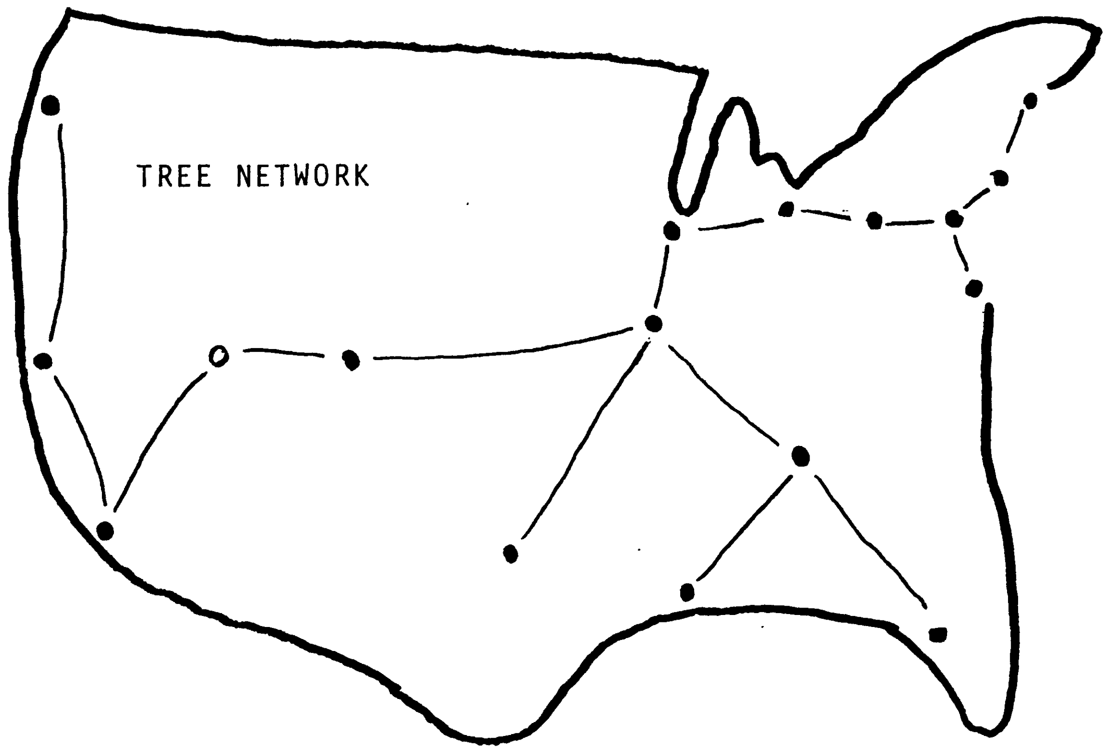
City Pair	distance	traffic	frequency	load factor	fare
DFW SFO	1493	348	4.0	.694	$ 86
TPA DEN	1520	59	1.5	.633	112
BOS DFW	1543	162	2.6	.668	97
PHL DEN	1575	116	2.1	.658	103
STL LAX	1581	367	4.1	.698	88
NYC DEN	1627	470	4.7	.709	88
MSY LAX	1658	211	2.9	.680	98
DFW SEA	1681	132	2.2	.664	105
STL SEA	1710	78	1.7	.647	116
ORD SEA	1730	356	3.9	.701	94
STL SFO	1736	192	2.8	.679	102
ORD LAX	1740	1245	8.0	.749	83
BOS DEN	1766	147	2.4	.670	107
ORD SFO	1853	744	5.9	.731	91
MSY SFO	1915	133	2.2	.669	115
DET SEA	1932	45	1.2	.633	138
ATL LAX	1934	295	3.4	.698	104
DET LAX	1977	386	4.0	.709	103
DET SFO	2086	207	2.8	.688	115
MSY SEA	2087	25	0.8	.616	163
PIT SEA	2124	21	0.7	.611	172
PIT LAX	2125	199	2.7	.688	117
ATL SFO	2141	210	2.8	.690	117
TPA LAX	2153	107	1.9	.666	129
ATL SEA	2182	74	1.5	.654	138
PIT SFO	1155	114	1.9	.670	132
WAS LAX	1188	525	4.6	.727	110
WAS SEA	2317	127	2.0	.675	133
PHL SEA	2383	55	1.2	.647	155
PHL LAX	2396	326	3.4	.711	120
TPA SFO	2403	51	1.2	.645	158
NYC SEA	2408	30	0.9	.628	174
WAS SFO	2430	359	3.6	.715	120
NYC LAX	2453	1722	8.8	.776	104
BOS SEA	2495	66	1.4	.655	155
PHL SFO	2526	215	2.7	.697	132
TPA SEA	2527	18	0.6	.612	200
NYC SFO	2574	1160	7.0	.762	112
BOS LAX	2600	429	3.9	.724	124
BOS SFO	2706	315	3.3	.714	132

Figure 6.4.1: Map and City Names for Domestic Network



NYC - New York City
BOS - Boston
WAS - Washington, D.C.
PHL - Philidelphia
PIT - Pittsburg
DET - Detroit
ORD - Chicago
STL - St. Louis
MSY - New Orleans
DFW - Dallas/Ft. Worth
DEN - Denver
LAX - Los Angelos
SFO - San Francisco
SEA - Seattle
ATL - Atlanta
TPA - Tampa

Figure 6.4.2a : Domestic Networks



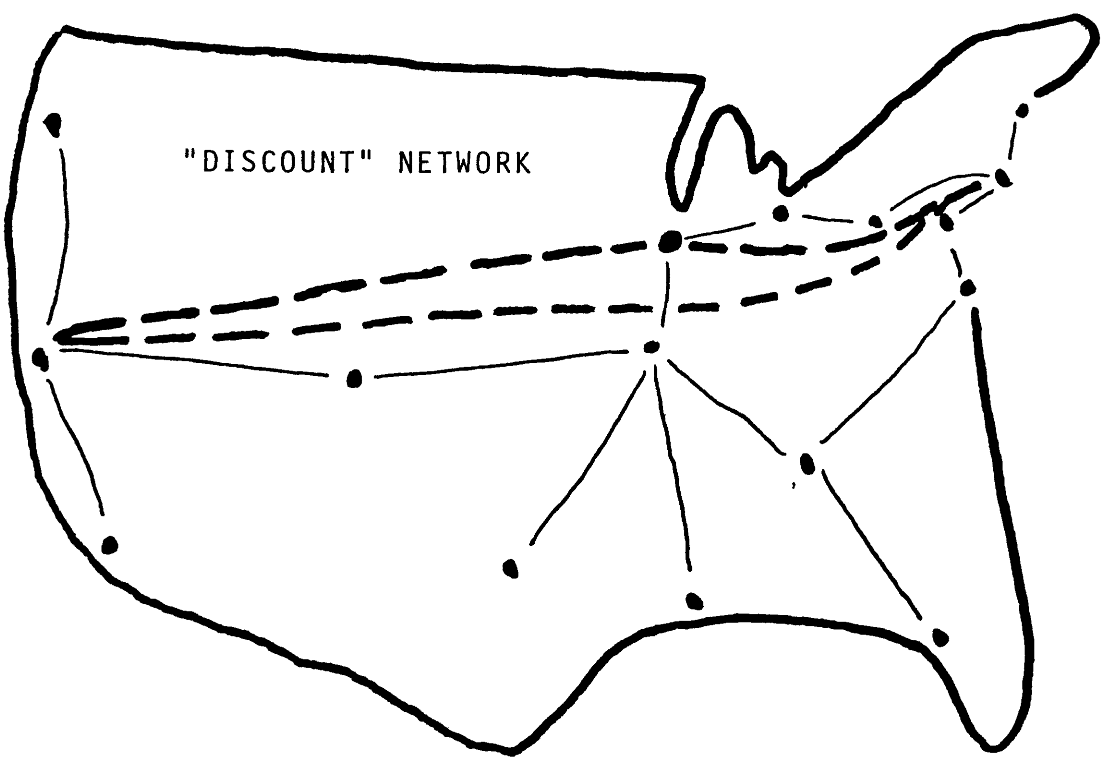
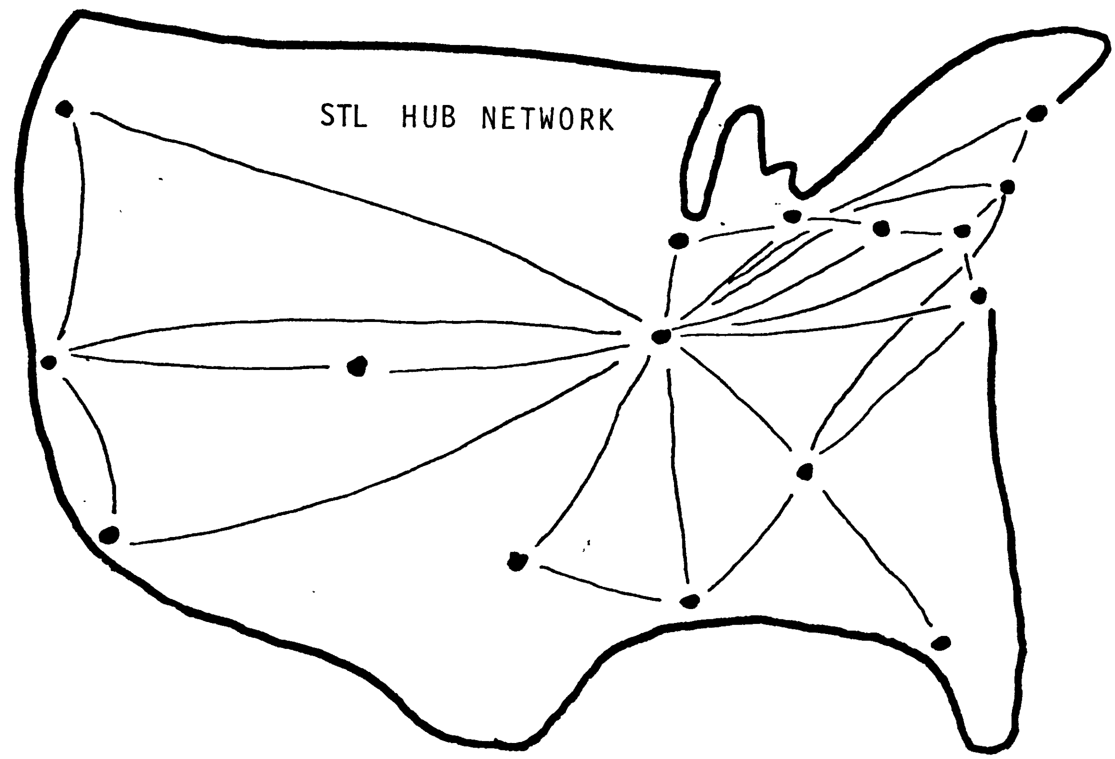


Figure 6.4.2b: Domestic Networks



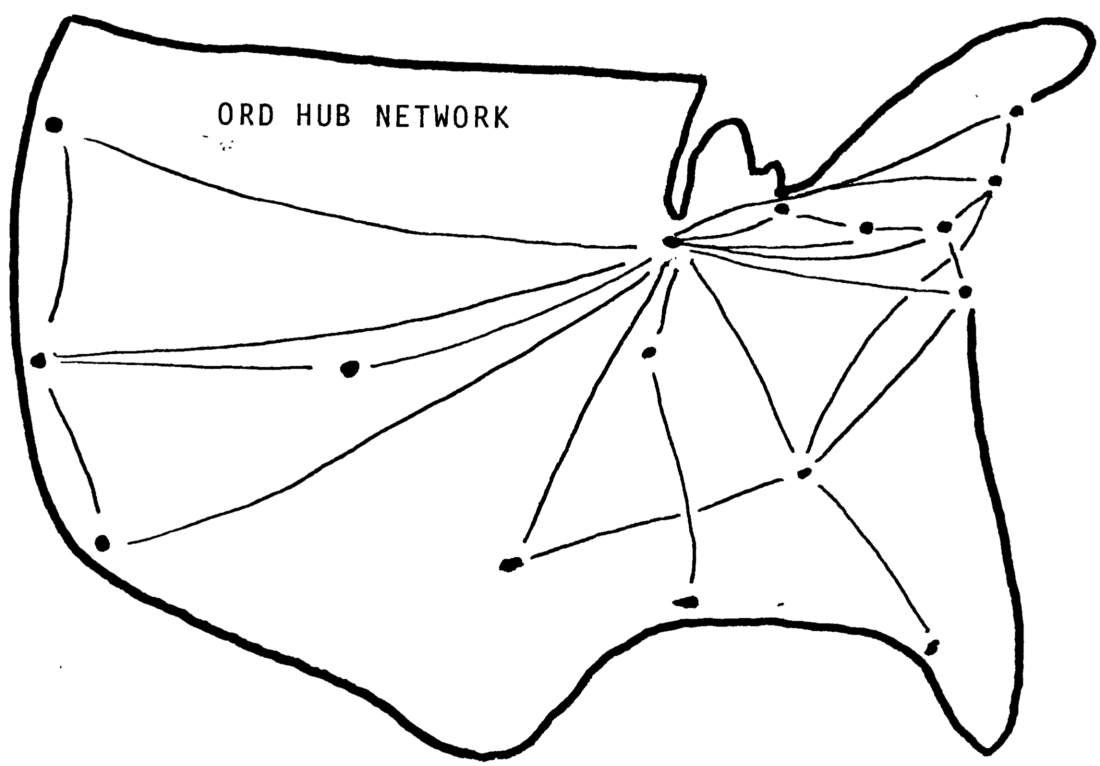
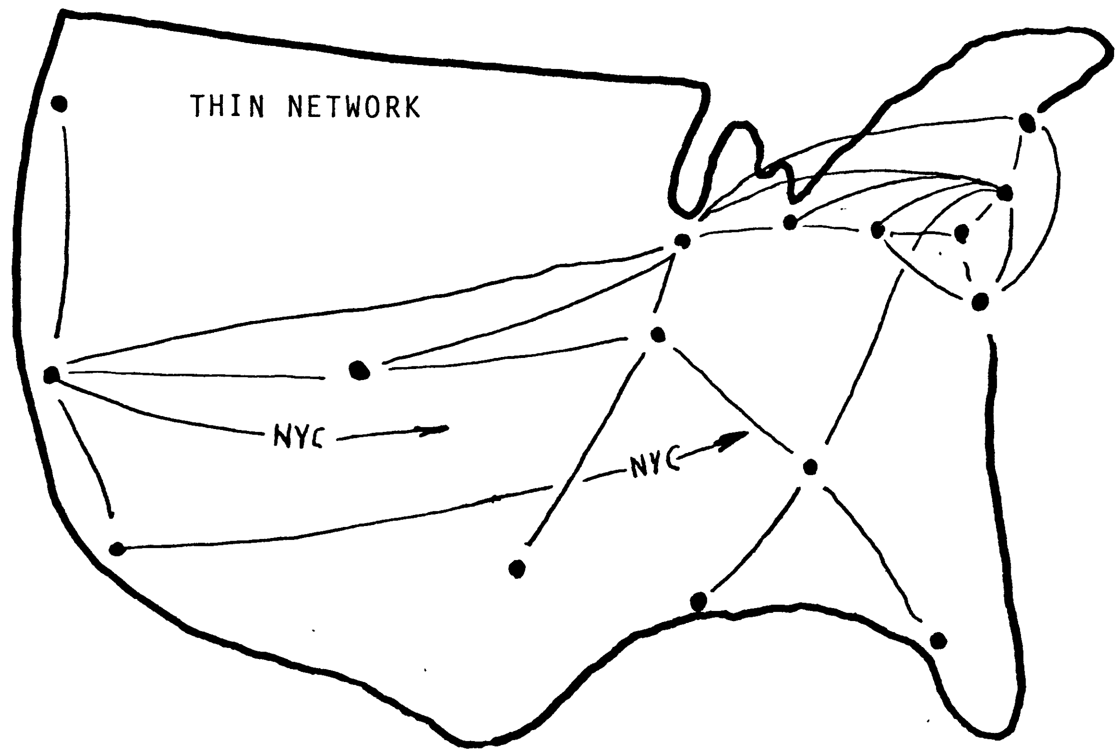


Figure 6.4.2c : Domestic Networks



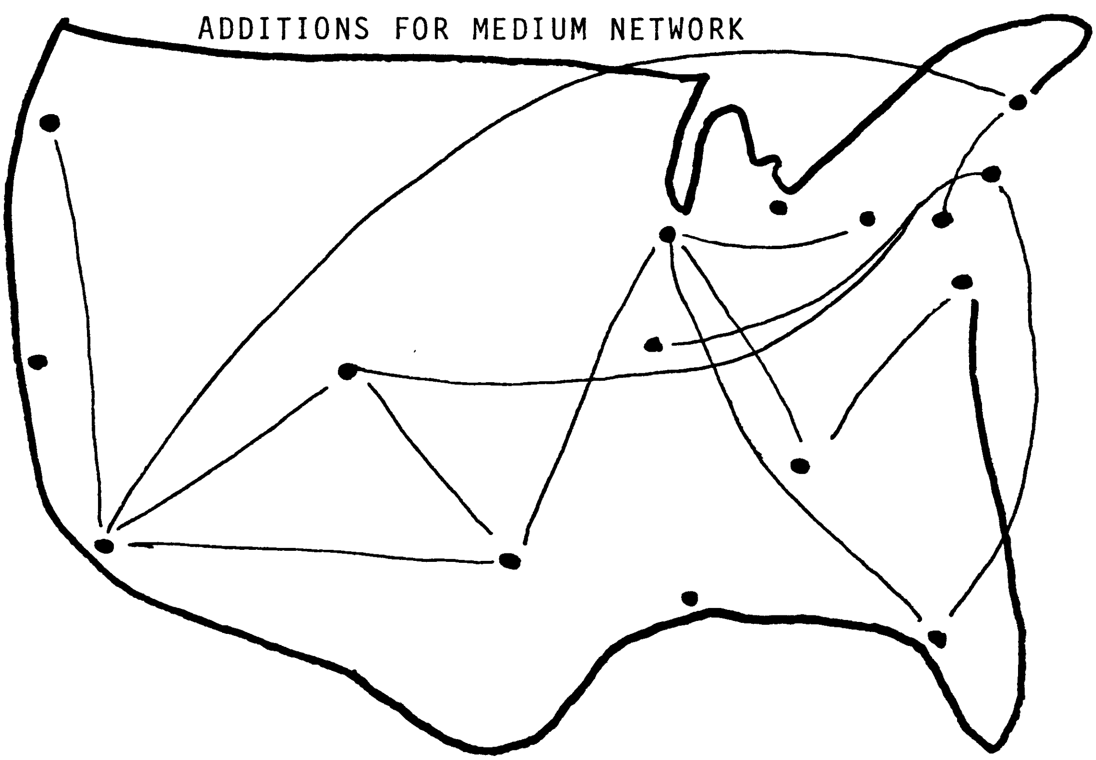
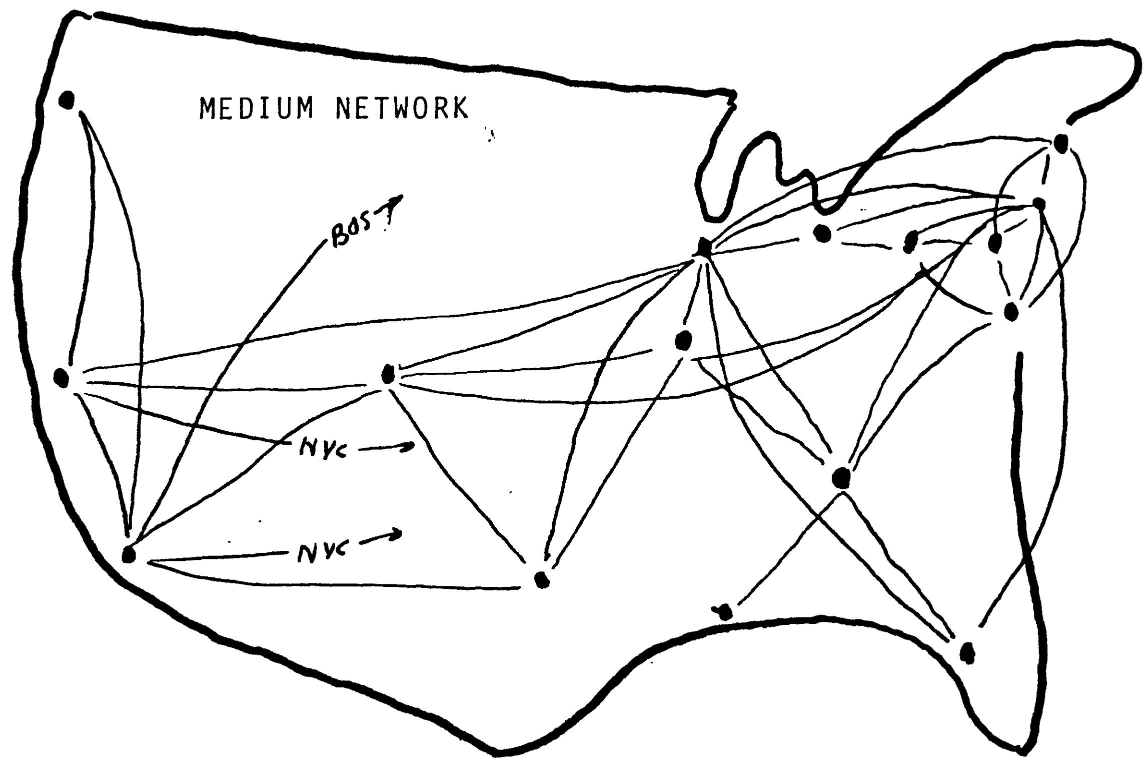
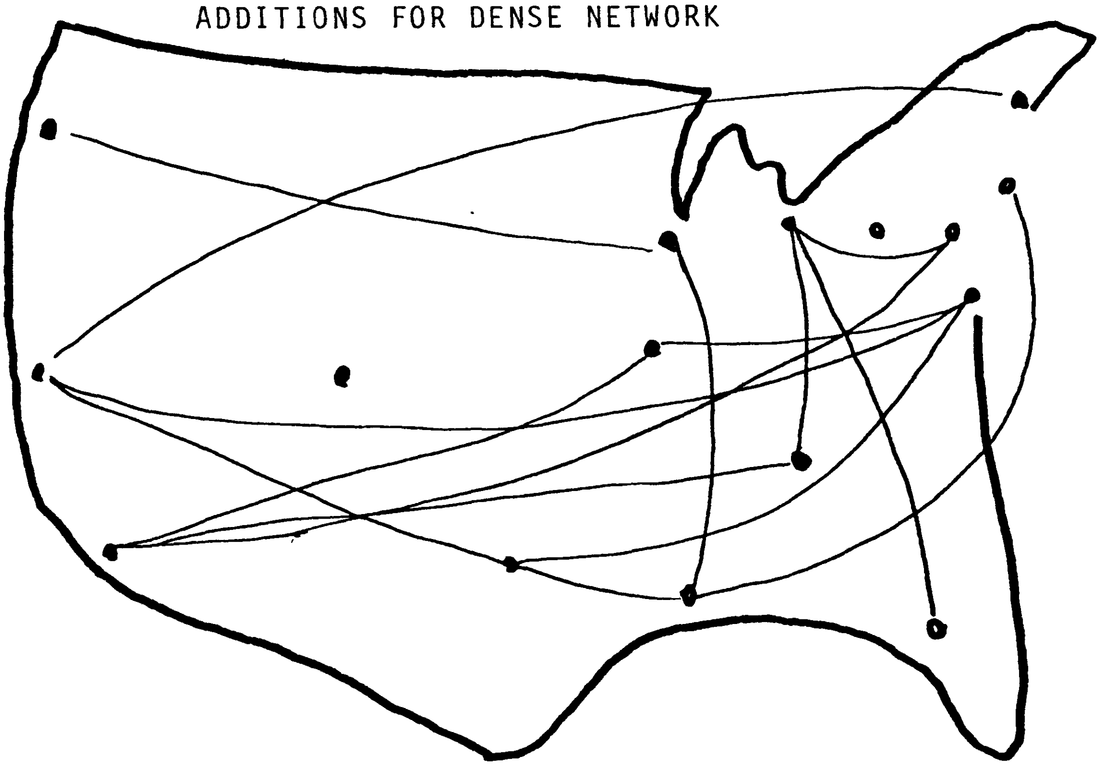


Figure 6.4.2d : Domestic Networks





of intermediate stops were small. The tree network was originally designed as part of an unpublished railroad network study where the link addition costs included the cost of building the rails.

The 19 link network is a better representation of the minimum air network. We have labelled this the "discount" network because it relies heavily on three links (NYC-ORD, NYC-SFO, and SFO-ORD) as if these links were provided by some extremely cheap and convenient service. Otherwise the network is nearly identical to the tree network.

The 26 link network employing St. Louis as a central hub represents a quite dense series of link services from all cities to STL but requires only one intermediate stop for almost all trips. This network is more expensive than a similar hub complex focused just 250 miles north at Chicago. In the ORD hub network, the strong Chicago traffic no longer must detour through St. Louis, and substantial savings occur because of this. Both STL and ORD hub networks have a secondary hub at Atlanta.

Neither of these hub complexes has enough bypass flights or enough off hub through flights. The network labelled "thin" was designed by providing nonstop service to all markets of over 500 passengers and adding other links where necessary to avoid high circuity or excessive number of stops for individual markets. This thin network design is the first we have proposed which follows the style of domestic U.S. networks today. A substantial increase in stage length and hop length occurs over the two hub style networks, even though only two more links are used. Thus aircraft costs will be lower, averaged per mile for the fixed demand network.

However, this thin design still routes too many people through too many intermediate stops. There are still 1.71 stops per trip, and costs are not a minimum. Adding nonstop service to all previously neglected markets of 300 passengers or more increases the nonstop markets (links) by 50%, modestly increases the stage lengths, and brings the stops per trip to the range observed for U.S. domestic airline systems (≅1.36). Circuity is now quite small (5%), and cost is nearly a minimum. This solution is the "medium" network. It would be suitable for medium density U.S. domestic operations.

However we have a very dense network of demands. Yet fewer stops and more direct service is called for because of the large markets. Adding all previously neglected markets of over 200 passengers to the nonstop list increases the links to just over half the maximum. The result is the cheapest and most convenient network so far proposed. We label this the "dense" network.

The final network design is the all nonstop, 120 link network. (There is no picture of this in figure 6.4.2.) This design goes too far in providing direct service to markets better served multistop. Aircraft capacities are too small (67-83 seats on average).

Data concerning these 8 network designs are presented in tables 6.4.2-6.4.5. Table 6.4.2 shows the activity measures which allow us to calculate the costs. The passengers were routed around the networks using the minimum cost paths. Costs in this case included the $10 of passenger time for each intermediate stop. Passenger transfers refers to passengers passing through an intermediate city, i.e. intermediate stops. In the thinner half of the networks listed, most of these transfers will be connections. In the medium and dense networks, which are typical of U.S. trunk networks, the "transfers" would be through flight intermediate stops in almost every case.

A table of network descriptive indices is presented in 6.4.3, including the crucial index, total cost, including the cost of stops to passengers. We note that as links are added, stage length (averaged over aircraft) and hop length (averaged over passengers) increase. They are never equal because there are more passengers in short haul markets.

We can see by comparing the STL hub, the ORD hub, and the "thin" networks in table 6.4.3 how inconsistent any single index can be. The hub networks have nearly identical links and link miles, circuity, seats per mile, and stage lengths, but they have markedly different stops per trip, seats per departure, and hop length. The thin network, with only two more links, is different in all categories, except circuity.

We note that the average aircraft capacity (in either seats per departure or seats per mile) is outside of normal ranges for all but the medium and dense networks. The difficult design tradeoffs occur when aircraft capacities and stage length are being traded off against one another, which happens at capacities below 400.

As a final observation, we note that by dividing hop length into average trip length, which is hop length for the nonstop network, we get a crude estimate of the stops per trip. This is the relationship suggested in equation (6.1). This is also what was done in table A.2 in appendix A in creating such an index for the U.S. airlines. How well this index predicts actual stops per trip depends on the particular network, but within a set of designs for a specific network of demands, the index is more consistent.

Table 6.4.4 shows how the distribution of hop lengths changes as networks become denser. The nonstop distribution favors the longer hops as much as is possible. The distribution of hop lengths for thinner networks has a larger peak in the short lengths and a smaller tail at the long distances. The approach toward the nonstop distribution is not necessarily continuous or monotonic, as comparison of the two hub cases and the thin network shows.

The distribution of costs in table 6.4.5 brings out once again the basic network tradeoffs. Passenger costs ($10 per stop) and physical multistop costs (circuity, seat stop costs, and $6 of handling) are continuously reduced by adding links. At the same time, the cost of frequency rises. The stop costs are the departure portion of stage costs and the cost of frequency is the vehicle portion of aircraft costs. Thus in approximate terms aircraft capacity is being traded off against stage length.

The cost of frequency and of intermediate stops in the dense network is $807000 while the cost of frequency in the nonstop network is $1288000. There is a 13% increase in total costs

Table 6.4.2: U.S. Domestic Network, Stylized Services

load factors at 65%
each link served 8 times daily
activities for one-way services
total traffic is 51,600 passengers
costs as in section 6.2

network name	number of links	link miles	passenger transfers	passenger mi. (000)	seat departures	seat mi. (000)
tree	16	5999	148400	64200	308900	98800
discount	19	11914	80500	60600	204000	93200
STL hub	26	16098	71400	54600	190000	84000
ORD hub	27	16529	48400	55800	154500	85800
thin	29	20542	36400	55200	135900	84900
medium	45	33853	18800	52000	108700	80000
dense	67	61189	9100	49600	93800	76300
nonstop	120	141389	0	49300	79700	75800

Table 6.4.3: Domestic Network, Stylized Services; Indices

network name	passenger circuity	stops/ trip	seats/ departure	seats/ mile	stage (mi.)	hop (mi)	cost (000)
tree	1.30	3.88	2413	2059	375	320	$6377
discount	1.23	2.16	1342	977	627	457	4868
STL hub	1.11	2.36	913	652	619	442	4513
ORD hub	1.13	1.94	715	648	612	555	4077
thin	1.11	1.71	585	516	708	625	3836
medium	1.05	1.36	302	295	752	736	3490
dense	1.01	1.18	175	156	913	813	3448
nonstop	1.00	1.00	83	67	1178	951	3929

Table 6.4.4: DIstribution of hop lengths

network name	seats 0-500 mi	seats 501-1000 mi	seats 1001-200 mi	seats 2001+ mi
tree	68%	35%	-	-
discount	63	35	1%	1%
STL hub	70	21	9	-
ORD hub	53	38	9	-
thin	53	42	6	3
medium	43	40	11	4
dense	39	34	15	6
nonstop	33		21	12

Table 6.4.5: DIstribution of Costs

network name	passenger time costs (transfers)	physical multistop costs	cost of frequency	passenger nonstop costs
tree	23%	35%	1.4%	41%
discount	17	26	2.8	54
STL hub	16	21	4.1	59
ORD hub	12	18	4.7	65
thin	9.4	16	5.8	69
medium	5.4	8.6	10	76
dense	2.6	3.2	17	77
nonstop	0	0	33	67

going from network services to isolated nonstop services. Thus even a network of the densest markets in the world has significant cost savings from service integrated into a network.

We can illustrate this with greater relevance by another exercise on the same network of demands. This time we shall require that each market in the 120 be served at the load factor, fare, and frequency found optimal for service alone as in chapter 5. These requirements were listed in table 6.4.1. For this study the $10 charge for the passenger's time on an intermediate stop will not be levied. Instead we will insist that all stops be through flights and that for all multistop markets, the daily frequency of the multistop flights be decreased to the point where the decreased displacement time in the schedule offsets the time spent in the extra stop. This is a much more rigid requirement than the reduction of fare by $10.

Once again the network was designed to minimize costs for a network producing the required seats, frequencies, and passenger movements for each market. The actual minimum was found using an optimal network design linear program documented by this author in a previous work [34]. Some feeling for the market by market service requirements can be gained from table 6.4.6, which lists the activities summed at each airport and the average aircraft capacities and load factors required for the 15 markets involved at each city. (We note that the load factors are high and the aircraft capacities low by historic standards for these dense markets, but the figures are certainly within reason.)

Table 6.4.7 presents the results for the network design solution on a city by city basis. These numbers compare to the summed requirements from table 6.4.6. For the netovrk design, the departures at each airport are fewer than those indicated by totals of the single market services on table 6.4.6, and the aircraft size is bigger. Even in this dense network with severe restrictions on the tradeoff between service quality and fare, there is a noticeable amount of load building and network influence.

The system comparison at the bottom of table 6.4.7 shows that the network design in this constrained case involves a modest change in aircraft size and stage length from the all nonstop case and a very small reduction of total costs. This optimal network design is similar to the "dense" network design above in that the number of nonstop markets and the average seats per departure are nearly the same. However the design requirements in terms of frequency are lower for this complete specification than the stylized case, and the load factors and frequencies are distributed differently. This explains the cost and stage length differences.

The cost savings of network service over nonstop service are modest in both the stylized and the design specifications for the 120 market network of demands. This network of the densest markets in the U.S. provides the lower bound on the importance of network effects in achieving minimum costs. It would seem that a network is useful but not entirely necessary for the densest

Table 6.4.6 : Design Conditions for U.S. Domestic Network

city code	originating passengers	avg. stage length	design departures	avg. seats	load factors, /dep.	/mile
NYC	16480	912	136	171	.71	.74
BOS	7095	742	79	132	.68	.69
WAS	8735	670	96	134	.68	.69
PHL	3934	841	63	96	.65	.67
PIT	3858	579	62	96	.65	.66
DET	5643	662	74	114	.67	.67
ORD	12766	856	117	155	.70	.72
STL	4551	727	64	107	.67	.67
ATL	5777	819	74	116	.67	.67
TPA	2969	978	45	99	.67	.68
MSY	2586	932	44	90	.65	.67
DFW	5914	952	70	123	.69	.66
DEN	4988	1074	59	123	.69	.69
LAX	9270	1611	80	161	.72	.73
SFO	6511	1580	65	142	.71	.72
SEA	2503	1313	35	105	.68	.67
total or avg.	103,580	958	1163	129.4	.688	.703

Table 6.4.7: Results of Design Exercise on Dense Network

city summaries expresseed as a % of design values on table 6.4.6

city code	departures	average seats
NYC	97%	111%
BOS	74	136
WAS	91	119
PHL	79	132
PIT	86	145
DET	84	120
ORD	100	125
STL	89	170
ATL	84	150
TPA	62	163
MSY	61	183
DFW	78	159
DEN	85	141
LAX	78	129
SFO	79	127
SEA	76	134
total	84%	134%

System total statistics:

	network	nonstop	% network/nonstop
aircraft stage (mi)	793	838	95%
seat hop (mi)	834	938	89%
seats/departure	173	129	134%
stops per trip	1.12	1.00	112%
total cost (000)	$3124	$3206	97%
all nonstop markets	58	120	48%
non-nonstop pax cost (cost of freq & stops)	20%	22%	88%

markets.

6.5 Service in a Network

In chapter 5 we optimized the service in frequency, load factor, and fare for a market according to several objective functions. We concluded that services of equal value covered broad ranges of these parameters and that different designs had different distributions of benefits across consumers with different values of quality. We also suggested the optimum could be improved upon by combined price discrimination and quality differentials. (1) To this picture the network case has added the new degree of service quality which is called the number of intermediate stops and a new ability to price discriminate across city pair markets. We will not repeat the optimizations of chapter 5 for a network, since the problems of comparability of benefits across people with different values is complicated by the comparability across different city pairs in the network. Furthermore the only possible conclusion is that service is even more flexible in terms of its design parameters and the optima will be even more shallow.

However we do want to point out that the network effect can be a powerful one in reducing costs. Our network designs focused on cost reduction and the unique network quality index, intermediate stops. A complete optimization would take into account the reduction in the costs of frequency and adjust service levels upward as well as fares down. In this light the single link service levels with frequency costs halved, which were presented in table 5.1.12 of chapter 5, suggest the direction that medium density services might take in an optimal network.

Because of reduced frequency cost, network services should be more frequent and cheaper than services to markets of the same density and distance served in isolation. Marginal seat costs are not cheaper and may often be higher due to intermediate stops. Therefore load factor in a network should be at or above the isolated market optimum.

This reduction of frequency costs occurs particularly in the case of thin markets served onestop by the concatenation of two denser nonstop links. This is just where we would like to see savings from the viewpoint of making competition viable in medium density markets. For instance, in the design network of section 6.4, Boston to San Francisco is served onestop 5 times: 2 times via Chicago, 2 times via St. Louis, and once via Denver. In this case denser markets are being joined end to end to create the Boston to San Francisco frequencies at almost no extra cost. Thus Boston to San Francisco schedules can be duplicated without duplicating the entire nonstop frequency costs. This makes

---

(1) That is different qualities offered at prices differing more than costs.

competition more viable.

6.6 Other Network Issues

Our discussion has focused on the one issue of network design which is always an issue, the tradeoff between aircraft capacity and stage length. Other issues are relevant in specific cases.

The question of backhaul, deadhead links, and weak links is a dominant cost issue when it arises as a problem. Fortunately these problems do not come up in reasonably dense and well connected networks of demand such as U.S. trunk airlines enjoy. Still the problems can come up in locally thin parts of a network, and they do arise for the regional and commuter carriers. All we have to say about such issues is that where demand network is limited by regulation of access to interrelated markets, substantial cost penalties can result.

The second major issue we have neglected is whether passengers experience connnecting or through service at their intermediate stops. In reality a transportation network is a network not only in space but also in time through the day. In this sense the network design problem can be divided into three steps: (i) establishing the links, their frequency and capacity, (ii) establishing the through movement of passengers and aircraft at each node and (iii) establishing the timing of nonstop and through services. Ideally we would like to solve all three problems simultaneously. Practically, computers cannot solve such problems for more than a few nodes and links. So we have collapsed steps (ii) and (iii) into step (i) by assuming that reasonable through services and timing can occur and building those assumptions into our cost and service measurements.

Neglect of the through vs. connecting issue is lamentable since the tradeoffs in cost and value are undoubtedly relevant at today's level of network intensity. It may be that the predominance of through service in the U.S. networks represents a higher quality of service and cost than many travelers would consider ideal. In this light it is interesting to note the discount connecting traffic between the North and Florida at Atlanta. Unfortunately, neither our demand model nor our network models are sufficiently sharp tools to allow us to make any informed comment on these issues.

Finally, network competition may take the less than benevolent form of one firm dominating a complex of markets or specific hubs in such a way that all other firms face only a sparse and thus high cost network of demands. Dominance of networks by dominance of specific hubs is an issue we have not explored. But we illustrated in section 6.4 both Chicago- and St. Louis- based hubs and then we saw in later discussion of the Boston to San Francisco market that service could be through either hub. Inter-hub competition is very likely viable.

6.7 Summary and Conclusions

Network design influences the single market optima by reducing frequency costs and raising capacity costs. Thus the gap between average cost and marginal cost for a market served in a network is narrower than for the market served alone. Also the definitions of average and marginal are even more elusive and less relevant.

Networks make service possible in thin markets and competition easier in medium markets by reducing frequency costs and increasing average aircraft size. U.S. domestic networks seem to be operating in a reasonably well connected way, allowing only modest numbers of intermediate stops and using reasonably efficient aircraft capacities.

Pricing in a network allows for discrimination among markets as well as among customers. But since average costs for passengers in different markets are almost impossible to define with any relevance, discrimination is hard to prove or to criticize. It is perhaps an issue of social concern how costs and benefits should be distributed among markets of different densities and distances. For instance, should thin markets pay the average costs of service alone, or should they pay the extra costs of their inclusion in the network? The prices will differ by factors of 2 or more depending on the allocation.

Finally, the problem of optimizing a network of services for a network of demands involves so many tradeoffs of cost and quality and so much flexibility of design and performance that specific designs are hard to fault and general rules are impossible to establish.

7 Summary and Conclusions

This work purports to be a systems analysis of scheduled air transportation. True to the concept of a systems analysis, the discussion has developed a series of sub-models for supply, demand, for load factor, for single link service, and for network design. Also true to the concept of a systems analysis, the models have broad applicability to a number of issues of design, operation, regulation, and competition in airline services. Many of the concepts carry over usefully to other common carrier modes. But there has been one driving force behind all these discussions: almost every matter discussed was a consequence of the economies of scale in aircraft capacity as illustrated in chapter 2 and reproduced here in figure 7.1. Also conceptually important was the definition of the market for transport services as the city pair and of the quality of service in terms of frequency and load factor. In this light we may review the ground we have covered.

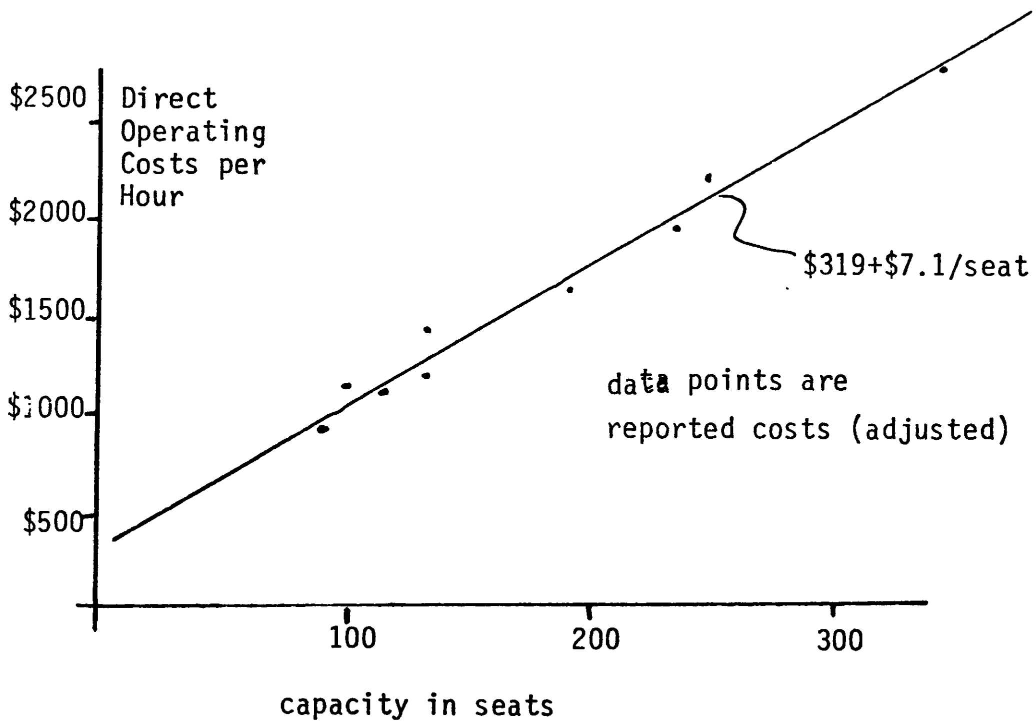
Chapter 2 established the relationship of aircraft cost per stage to capacity. The linear rise of costs from some positive value at zero capacity was derived both from design considerations and reported cost data. Because airlines can choose aircraft covering a broad range of capacities along this line, they experience separable costs of frequency and capacity. Our work is nearly unique in developing the consequences of this cost structure.

Chapter 2 suggested that these economies of aircraft capacity tended to make scheduled service a natural monopoly, but full development of this argument awaited further definition of the market and the service.

Chapter 3 defined the market for scheduled air service as the city pair. Considerations of substitutability of use borrowed from the legal definition of monopoly meant that exclusive service of a city pair by a single carrier established a local monopoly. With the assumption that one or only a few scheduled services would be available in each market, it was shown that there would be a distribution of benefits among consumers due to their different values of the particular service quality offered. It was also shown that competiers would tend to duplicate service levels, failing to braoden the distribution of benefits. Although distributional considerations have political and social implications for air service, their existence has been ignored in past work. It seems possible that recent political pressure on airlines comes not from their excess profits (earnings have been low enough) but from the history of neglect of the low value of time segment of the market.

Chapter 4 developed the concept of load factor in transportation services. Load factor was shown to be a part of the design quality of service. Load factor can be adjusted independent of frequency by changing aircraft capacity. Load factor affects service availability; load factor is not a casual consequence of supply and demand mismatching. A 55% average load factor may be a perfect accommodation of a steady supply to a fluctuating demand.

Figure 7.1: Aircraft Cost vs. Capacity



Source: See figure 2.2.3

Chapter 4 also developed the point that an oligopolistic approach to competition requires service matching among competitors not only in overall quality as shown in chapter 3, but specifically in the load factor dimension. Discussion assuming conventional reservations services showed that competition might be unstable in the direction of higher load factors.

Chapter 5 integrated the models of chapters 2, 3, and 4 to produce single market technical performance abilities for scheduled air service. Several definitions of optimal performance were explored. Maximum profits was shown to be far from a fair operating condition for a single firm. Maximum surplus turned out to (1) require subsidy and (2) be a very small improvement over zero loss operations. Maximum traffic at zero loss was adopted as a reasonable optimality criterion, although maximum revenues at zero loss came closer to reproducing existing conditions.

Optimal services were shown to vary in frequency, fare, and load factor among markets of different densities and distances. This is the first time all three degrees of freedom have been allowed to vary continuously, so the nature of the optima is of particular interest. The most often neglected (1) freedom, adjustments in aircraft capacity, was shown to be quite important. Load factor was by far the least variable among the optima.

It was also demonstrated that if any one of the aspects of perceived price - fare, frequency, or load factor - were artificially constrained, a new optimum nearly as good could be found by adjusting the remaining degrees of freedom. Optima were found to be quite broad and shallow. This flexibility of design makes regulation of service ineffective at altering basic equilibria.

The final chapter added to the picture considerations associated with the design of networks of service. Networks were shown to be a way of sharing and reducing the costs of frequency among markets. Larger aircraft capacities are achieved at the expense of shorter stage lengths and intermediate stops. Cost minimum configurations were quite shallow. Explicit design and performance depended on details of the list of demands served.

Network efficiencies reduce frequency costs and therefore allow more competition and more service to low density markets. The issue is raised that economies of local network intensity may lend a monopoly tendency to network operations. Further discussion of network design and competition was not undertaken, since the germane point was the necessity of networks of service in achieving cost minimums for service in a market.

From the considerations of chapters 2 through 6 we expect to see at most two or three airline competitors in large airline markets. Where aircraft capacities in use are in the medium to large range (20C to 400 seats), the natural monopoly is weak. However, offerings of distinctly different service levels may be

---

(1) neqlected in analysis, not in practice.

rare.

Unfortunately chapter 5 suggests that regulation of service is extremely difficult since the optimum is so hard to define and the dimensions of service are so flexible.

Chapter 6 suggests that where interlaced networks of services compete, even low density markets may receive competitive service. However, even when large aircraft are employed it is incorrect to assume that all the economies of scale have been exploited. Larger aircraft capacities in a network are usually achieved at the expense of extra stops. Because of the expense of those stops, average costs in a network will be above the single link average costs for the aircraft capacities in use. The natural cost advantage of an airline network using 200 seat aircraft can be as strong or stronger than the natural monopoly that exists on a single link served by 100 seat aircraft.

Network considerations of vehicle frequency costs suggest that as demand among cities grows, the optimal network design will change from a sparse spanning tree pattern to a highly interconected nonstop pattern. The index of this growth is the number of intermediate stops per passenger trip. At the same time, single link discussions suggested that service should take the form of a single quality for thin markets and multiple price/quality combinations for denser markets. Thus overall system growth should move from sparse networks and a single service option to multiply connected networks and multiple service levels. U.S. domestic operations for trunk carriers appear to be in an intermediate stage of this process.

Overall conclusions from this work are that competition is possible among networks and that regulation is difficult. However, there are dangers. Competition may not lead to a distribution of choices of service quality and price. Further, network design may isolate particular medium density markets or particular cities from competition and local exploitation can occur.

Appendix A: Indirect Operating Costs for Airlines

Introduction

Indirect operating costs (IOCs) constitute about 40% of airline costs. These costs cover the sales, administrative, passenger traffic, and aircraft handling expenses of an airline. Aircraft direct operating costs (DOC) receive considerable study not only by the manufacturers and airlines but also by consultants, government agencies, academics, and the pilots' union This is not the case with respect to IOCs. Although IOCs may not have been the subject of as much detailed work, but they have been estimated time and again for different studies of airline systems It has always been assumed that these costs depend on activity measures such as revenue passengers, revenue passenger miles, aircraft departures, aircraft miles, seat departures, and seat miles.

Little is known of the relationship of IOC to these activities Our models assume that IOC are proportional to these activities, and that the costs (with the exception of landing fees) are the same throughout the network. Thus it is assumed that the labor involved in passenger boardings, or in sales or aircraft handling, is the same price and is used with the same efficiency in Boston as in Boise. In a sense these inputs are treated in the same manner as aircraft block hours.

The arguments for constant proportionality of IOC to activities at all levels and in all geographic locations are at least partially convincing: First, we do not have the data to prove the contrary. Second, many of the measures, such as aircraft miles, seat miles, and passenger miles, are not station related. Third, managers in the industry are not uncomfortable with assuming no economies of scale with station size and little variation with region. Thus we may for the moment accept the assumption of proportionality of IOC to total activity levels. The big problem lies in determining what activity measures the IOCs are marginally proportional to.

In this appendix estimates are made by referring to two studies of airline data, one by Simpson and Taneja [47] and the other by Douglas Aircraft [48]. The results of these studies are used to guide our own statistical analyses of cost figures reported by the domestic trunk airlines to the CAB.

As in chapter 2 we are not hypothesis testing using statistical methods. Nor are we calibrating a production function for the intermediate products of revenue passenger miles handled, boardings handled, aircraft dispatched, etc. We are trying to calibrate cost relationships we deduce to exist from engineering considerations (i.e. from examining the underlying activities). Unfortunately, the statistical evidence is so poor it is difficult to calibrate these relationships.

Difficulty of Statistical Calibrations

The airlines report their expenditures on IOCs to the CAB. They also report activity statistics sufficient to determine airline passenger miles, passenger boardings, aircraft miles,

Table A.1: Correlation Matrix for Activities and IOC categories

U.S. Domestic Trunk Airlines, airline totals.
1970-1974 Annual Figures
Constant Dollars
Strike affected data deleted

column no.		(2)	(3)	(4)	(5)	(6)
IOC's	(1) Passenger Service Cost	.964	.977	.970	.837	.892
	(2) Aircraft & Traffic Servicing	1.00	.978	.959	.899	.942
	(3) Promotion & Sales	.978	1.00	.968	.892	.928
	(4) Revenue Passenger Miles	.959	.968	1.00	.844	.905
	(5) Aircraft Departures	.899	.892	.844	1.00	.969
	(6) Passenger Boardings	.942	.928	.905	.969	1.00

seat miles, etc. It would seem straightforward to perform a multi-variable regression and accept the results for the cost coefficient calibration. The difficulty is this: all activities and costs are highly correlated (see table A.1). The first activity measue introduced in the regression, often revenue passenger miles, explains all costs. (R^2 is near 0.9 for any activity variable and any cost category.) All this implies is a nearly complete dependence of IOC on activity in general. The next variable introduced in the regression usually increases the standard error of the estimate, with very small improvement in R^2. This occurs because of the high degree of collinearity between the first and second activity measure. The result is any traffic measure alone seems to explain all IOC. There is little to choose among the accuracy of one over the other.

This means the choice of IOC dependencies must be guided almost exclusively by judgments of the cause and effect relationships for these costs.

This is most unfortunate because the ratio of per aircraft to per passenger costs and the ratio of per departure to per mile costs are important parameters in the network design problem. Yet it is these very ratios that are impossible to discover from regression of IOC data.. For instance, we believe the cost of servicing aircraft departures and passenger boardings) at an airport is proportional to the number of those movements. We also know that longer flights require a little more to be done, so we expect a small fraction of the costs to depend on either passenger or aircraft miles. But if we put miles in the calibration, the regression will assign the majority of the costs against them, slighting the more fundamental activities, departures and boardings.

The lack of clear indication of cost dependence on different activities occurs for several reasons. First, there is relatively little variation in the ratios of passengers to departures or aircraft miles to departures among the airlines (see table A.2). This is equivalent to saying the activities ratioed are collinear. But the similarity among the airlines is of real world intent. The CAB sought in the 1960's and 1970's to allow all airlines a fair profit while enforcing a system-wide fare structure which is uniform across market densities. Less profitable airlines were given new markets generally held to be profitable. By this process, any airline with costs significantly off average due to a lack of long haul or high density routes is brought back to average costs by normalizing the important ratios of activity variables. This produces roughly equal profits at equal fares, but it also prevents the variations necessary to establish statistically the functional relationships we seek.

This objection could be overcome by comparisons between trunk and local service carriers, whose activity ratios are quite different. Unfortunately, there is an appreciable difference in the quality of passenger service associated with local service operations, so IOC comparisons between local service airlines and the trunks are not comparing the same outputs. This phenomenon may extend itself to the trunks themselves. That is, the

Table A.2: 1973 Activity Ratios for U.S. Domestic Trunk Airline Networks

	PAX/DEP	RPM/ACM	SEATS/AC	LF	RPM/PAX	ACM/DEP	Trip/Stage
American	53.8	72.1	143	.530	976	728	1.34
Eastern	47.4	59.9	112	.553	644	512	1.26
United	54.2	70.2	137	.547	894	690	1.30
Braniff	43.0	61.8	126	.502	727	505	1.44
Continental	42.5	68.6	142	.484	879	545	1.61
Delta	45.7	64.9	127	.512	611	430	1.42
National	46.0	81.5	164	.496	860	485	1.77
Northwest	43.7	73.5	190	.409	1009	601	1.68
Western	49.5	67.9	118	.576	804	587	1.37
average	47.3	68.9	140	.512	823	565	1.46

PAX=passengers; DEP=departures; RPM=revenue passenger miles; ACM=aircraft miles; SEATS=aircraft capacity; AC=aircraft; LF=load factor; Trip=RPM/PAX; Stage=ACM/DEP; Trip/Stage is a surrogate for stops per passenger trip consistent ticket price among markets with intrinsically different costs may produce variations in the quality of service rather than in the expense of that service.

Finally, we can lay some blame on the aggregation of our data. Each airline has somewhat different amounts of discount, charter, cargo, and first class traffic (see table A.3). These customers receive different treatment at different cost, and these variations may mask small changes in the more basic activities which we might otherwise observe.

The objections to regression studies are many, and an examination of past attempts can only serve to illustrate the point. This has not prevented several authors from addressing the task. This section compares the published work of Taneja and Simpson [47] at MIT, the published efforts of the Douglas Aircraft Corporation [48], and some small unpublished work of this author at MIT.

Airline IOCs are subdivided by the CAB into four categories. Aggregate costs in each of the four categories is reported by each airline on CAB form 41. We shall address each category in turn, as was done by Taneja and Douglas Aircraft.

Passenger Servicing

The cost of stewardesses and food is reported under this heading. The industry's average expenditure on this account per revenue passenger miles was $0.0098/RPM. (1) In this light the figure by Taneja of $0.0106/RPM seems reasonable. However, it would seem that an account made up largely of stewardess salaries is probably proportional to passenger travel time, not distance, i.e there should be an amount per boarding which is over 200 times the cost per mile. This amount covers the time spent on the ground, taxiing, or in terminal area maneuvers. In this light, the figure by Douglas Aircraft for local service carriers is more reasonable. Douglas Aircraft suggests $0.0067/RPM +$0.47/BOARDING.

The cost per boarding seems low, but a regression analysis by the author of domestic trunk carriers' annual operations 1970-1974 produced similar results. The statistical correlation between passenger service cost and the activities passenger boardings and passenger miles does not stand out as singularly more accurate that any of the other correlations available. However, the dependence makes the greatest sense. The regression results were $0.0093/RPM +$0.65/BOARDING.

A second regression was made using the airlines' expansion paths' slopes. Each airline's costs and activities (1970-1974) were averaged and the annual changes from average were used as data for the regression. This eliminates the influence of whatever small fixed costs exist in data for this category. The results were different, and did establish a substantial zero distance cost. The result was $0.00122/RPM +$4.052/BOARDING.

---

(1) Throughout this discussion 1976 dollars will be employed exclusively. All figures were converted using the consumer price index.

Table A.3: Sources of income for U.S. trunk airlines (%'s)

	coach	1st class	charter	freight & other
American	71	17	1	11
Eastern	78	12	1	9
United	73	13	0	14
Braniff	74	16	1	19
Continental	76	12	1	11
Delta	77	14	0	9
National	78	12	0	10
Northwest	75	12	2	11
Western	86	7	1	6

Table A.4: Comparison of Preditions for Promotion And Sales Costs

1976 dollars throughout

Taneja regressions:

trunk airlines: \$2.24/pax + \$.0122/rpm

local airline: \$.018/rpm

Douglas Aircraft regressions:

local airlines: \$1.61/pax + \$.0063/rpm

Swan regressions (trunks):

long run regression: \$2.47/pax + \$.0071/rpm

shortrun regression: \$3.08/pax + \$.0016/rpm +\$126/departure

There is a significant difference between the results of the two regressions performed on the same data. The different results would be important in any system design. Yet there is essentially no methodological way of choosing between the two. The straightforward regression among the airlines' total costs should have found the long run expansion path of an airline. In fact it did compare with a regression of 1972 data alone. The second regression addressing the year to year differences may be the short run expansion path. (1) On the other hand, the year to year differences were not obscured by variation in fixed (zero passenger) costs among the airlines, so it may be the true long run relationship.

The author feels that an average of the two results, or even a sum slightly favoring the first regression, is reasonable. The educated estimate might be $0.0067/RPM + $2.68/BOARDING. This is reasonable at both transcontinental and short haul distances.

Aircraft and Traffic Servicing

This account includes the costs associated with handling an aircraft on the airport surface plus the costs of boarding passengers and luggage. This category seems to disprove the value of regression entirely; the best correlation of the cost of these departure-based activities is with revenue passenger miles. The logical dependence of this cost category on aircraft departures and passenger boardings cannot be demonstrated by statistical analysis.

This situation persuaded both Taneja at MIT and Douglas Aircraft to avoid the use of passenger boardings as an index. Only Douglas Aircraft even included aircraft departures, but the indicated cost dependence on the activity was slight. This author obtained similar results from both the short run and the long run regression analyses. However, with mileage-based activities not allowed, the causal variables could be brought into the calibration with the following results for the short run regression:

Cost=$7.94/BOARDING+$183.2/DEP

(2) But the question remains, can we rely entirely on logic to disassociate traffic handling costs from revenue passenger miles in the face of such persistent statistical evidence?

Promotion and Sales

Promotion and sales costs largely pay for the travel agent's commission on ticket sales, the costs of reservations, and the costs of advertising. The largest of these expenditures is

---

(1) We shall call this second set of analyses the short run regression.

(2) As discussed above, we emphatically refuse to judge these calibrations on statistical grounds. R-squared and F statistics have been suppressed.

agents' commissions, which are a flat percentage of ticket price. Ticket price is approximately $17/BOARDING +$0.074/RPM, (1) Thus promotion and sales costs are proportional and passenger miles. All of the regressions, Taneja's, Douglas Aircraft's, and this author's, confirmed this relationship. The regressions were in rough agreement, and the ratios of per boarding costs to mileage costs were roughly that of the ticket price. For the sake of completeness, the result of the several regressions are presented in table A.4.

In a situation in which fares are not predicted by such a formula, it seems more reasonable to express these costs as a surcharge on other costs rather than a function of boardings and miles This assumes that in the future fares depend on costs and promotion and sales costs are a fraction of fares. The author recommends the use of 12% of fares or 13.6% of other costs.

General and Administrative

These costs are the costs of management, and are traditionally called overheads. As such it is normal to treat them as a constant percent of all other expenses. (2) Douglas Aircraft takes the view that depreciation and rental expense should be exempted from overhead charges. With this logic they achieve an estimate of 5.4% of all non-depreciation expenses, including direct operating costs (DOC). This estimate agrees nicely with the 1973 domestic trunk average of 5.6%. The author recommends the use of one or the other of these figures.

Summary

Regression studies of airline IOC do not clearly indicate any cause and effect relationships between measures of different aspects of airline activities and the appropriate cost categories Nor are reasonable hypotheses even favored by the statistical evidence. As a consquence, it is justifiable to use judgmental cost assignments rather than to follow blindly results of regression studies in estimating cost coefficients. Because of the high degree of collinearity among variables, it is usually possible to select a convincing regression output to support any reasonable assignment. (3)

The best judgmental assignments for the IOC of a U.S. domestic airline appear to come to the following total:

Overhead and sales costs of 19% of all other costs (except depreciation)

Indirect costs for passenger and aircraft of
$10.62/BOARDING

---

(1) In the 1960's and 1970's fares were nearly linear with distance and independent of density. These figures are representative of 1976 coach fares with tax.

(2) Taneja regresses against RPM and aircraft miles.

(3) It is also possible to select equally convincing regressions to support unreasonable assignments. That is the point.

+$183.2/DEP
+$0.0067/RPM.
All these figures are in 1976 dollars.

Appendix B: Calibration of Air Travel Demand Model

We have deduced that air travel demand should follow the formula:

$$D = k_1 \cdot (F + v \cdot T)^{\alpha} \tag{B.1}$$

Where F is the fare and T is the total travel time. The question is, what are reasonable values for v and α? The only attempt to calibrate a model with such a formulation involved several complicating factors. This author [3] calibrated a non-linear model split model for intercity passenger travel in the 100 to 400 mile range. The model included predictions for total travel in the city pairs based on an intervening opportunities formulation as well as air modal split effects. With this formulation and the data available, no value of v could be shown to produce a well-defined minimum for summed squared error. Therefore v was arbitrarily set to $7 per hour, that being near the hourly income for the year in question (1976). With v at $7/hr., α was best in the neighborhood of -0.7.

The entire work is very speculative, relying as it does on a formulation involving a number of separate effects and a very limited amount of short haul data. A far better calibration for air travel demand was performed by Eriksen [11]. His model had the form:

$$D = k_1 \cdot F^{\gamma} T^{\beta} \tag{B.2}$$

Eriksen's calibration involved medium and long haul markets further disaggregated into low, medium, and high density classes. His values for γ and β are given in table B.1. These values are fare and time elasticities. That is to say that the relevant slopes are:

$$\frac{\partial D}{\partial F}\frac{F}{D} = \gamma \simeq -1.0 \qquad (B.3)$$

$$\frac{\partial D}{\partial T}\frac{T}{D} = \beta \simeq -0.5 \qquad (B.4)$$

Equation B.1 states that the slopes are:

$$\frac{\partial D}{\partial F}\frac{F}{D} = \frac{F}{F+vT}\alpha \qquad (B.5)$$

$$\frac{\partial D}{\partial T}\frac{T}{D} = \frac{vT}{F+vT}\alpha \qquad (B.6)$$

Setting (B.3) equal to (B.5) and (B.4) equal to (B.6) requires the relevant slopes of the two formulations to be equal. This process produces the following simple results:

$$v = \frac{F}{T}\frac{\beta}{\gamma} \qquad (B.7)$$

$$\alpha = \gamma + \beta \qquad (B.8)$$

Equation (B.7) tells us that the value of time is the same as the relative tradeoff of time and money in (B.2). Equation (B.8) tells us that the formula summing time and fare effects (B.1) requires an elasticity equal to the sum of time and fare elasticities.

Table B.1 Eriksen Demand Calibration

Values for fare elasticity γ:

	medium haul	long haul	
low density	-.60	-0.5	
medium density	-.89	-2.1	-1.0
high density	-.58	-1.3	

Values for time elasticity β:

	medium haul	long haul	
low density	-.57	-.14	
medium density	-.99	-.45	-0.5
high density	-.53	-.43	

Data from Eriksen [11]

Table B.2 Calibration for $\beta/\gamma=2$, $\alpha = -1.5$.

market distance	assumed frequency	1976 fare	calculated time	imputed v
400	3	$54	3.00 hrs	$ 8.9/hr
400	6	$54	2.07	12.9
400	9	$54	1.76	15.2
800	2	$82	4.72	8.6
800	4	$82	3.33	12.2
800	6	$82	2.86	14.2
1500	1	$130	8.91	7.3
1500	3	$130	5.18	12.6
1500	5	$130	4.43	14.7
2000	2	$167	7.09	11.7
2000	4	$167	5.69	14.5
			average	$12.1/hr

calculated time from T = distance/507 + .37 + 5.6/frequency

imputed v from v = fare/2·time

Equation (B.8) is satisfied in our case by α=-1.5. Equation (B.7) cannot be satisfied in all markets. Table B.2 provides a reasonable average value for participating travellers in 1976. A somewhat lower figure would be appropriate for markets including low fare/low service trips not take in 1976 due to an absence of such service offering in the market.

Appendix C: Frequency, Displacement Time, and Competition

Frequency and Displacement Time

A large number of flights in a city pair market increases the quality of the service by reducing the amount of time passengers spend waiting for a departure. The usual measure of the time inconvenience for a published schedule is the displacement time either forward or backward from desired departure times to available flights. Even with peaking in the time of day distribution of desired departure times, well designed schedules can reduce the average displacement time to one quarter of the average headway [22]. Thus for a 24 hour day, one would expect an average displacement time (Td) of

$$Td = h/\text{frequency}; \; h = 6 \qquad (C.1)$$

Steve Eriksen in his Ph.D. thesis [11] examined a level of service index which reflects only this displacement time and the aircraft travel time. Eriksen calculated this index using published airline schedules [27] and time of day distributions of demand derived from considerations of convenient departure and arrival times and matters of flight time and time zone changes. From data for 1974 we (1) culled 31 markets with (a) all nonstop or all one stop services (i.e. constant aircraft times), (b) between 3 and 21 departures per day, and (c) one carrier with over 80% of the market. These markets and their displacement times are presented in table C.1. There is a good distribution of distances and densities.

Using these times and frequencies, equation (C.1) was calibrated by least squares regression. The results predicted h as:

$$h = 6.19 \pm 0.26 \text{ (hrs)} \qquad (C.2)$$

The relationship was quite significant ($R^2=.95$, $t=24.$) A regression for (C.1) allowing a constant term produced nearly identical results.

However, one question must be asked. Why is h greater than 6 hours? With 4 to 8 hours of almost no demand at night, h should be in the neighborhood of 5 hours, not 6. We must conclude that either airline schedules imperfectly match time of day peaking of demands or Eriksen's time of day distributions are off. Probably both effects contribute.

We can make one adjustment which reduces the estimate of h. There was a certain amount of head to head scheduling even in these "monopoly" markets. Removing frequencies within one half an hour of earlier frequencies and recalibrating produced

$$h = 5.73 \text{ hrs} \pm 0.25 \qquad (C.3)$$

---

(1) The data gathering and manipulation for this work was done by D.F.X. Mathaisel. Any conceptual errors are this author's.

(R^2=.95; t=23.) This we take as the best estimate we can get for h.

Competition by Frequency

The logic of displacement times extends itself to competitive scheduling practices. For airlines with otherwise identical service, departure schedules are competitive tools. With only 2 departures in a market, airlines do best to take different parts of the day for their departure. This avoidance of each other's times stimulates a bigger net market. But the fifth departure time only reduces expected travel time by 17 minutes, which is unlikely to stimulate much demand. For fixed or nearly fixed total demand, competitive scheduling between two airlines becomes a sort of a game to see who can capture the greatest fraction of the day as closest to their departures. This scheduling game has been explored by de Neufville [38]. The only stable points are (i) two competitors scheduled head to head or (ii) one airline alone. (1) Any dominance of frequency by one carrier allows him to capture more traffic per departure than the smaller competitor by matching the small schedule and then strategically locating the remaining flights. Even a 50-50 split by interleaved departures is unstable due to end of day effects. The 50-50 head to head solution is the only stable one for two carriers.

Unfortunately such schedule duplication is a waste of resources. The same service convenience can be provided by using larger aircraft (one for each head to head pair of departures) with a reduction in cost.

Fortunately this process of head to head scheduling occurs less often than theory would suggest. Often one carrier dominates. In the 48 markets found in Eriksen's data with between 7 and 21 departures (2) and maximum market shares below 80%, the average maximum market share was 57%. The standard deviation was 12 percentage points, indicating a broad distribution, as is shown in table C.2.

It must be stated at the outset that the head to head theory was suspect. Network scheduling and routing effects influence airline schedules in a single market, perhaps as much as considerations of intra-market competition. We tested in a crude way this theory by developing an index of competition.

Index of Competition Defined

We define an index of competition I according to the formula

$$Td = (h/frequency) \times I \qquad (C.4)$$

Where Td is our familiar expected displacement time. The dimensions of I are departures per departure time. An I of 2.0 means there are two departures for every departure time of worth, or a head to head schedule.

---

(1) The theory even extends to three or more carriers matching schedules.

(2) All nonstop or all one-stop departures.

There are three possible predictors for I in a market. Each has different implications for the type of competition that is going on.

Head to Head Competition: Ih

One predictor for I is Ih=n, where n is the number of competitors. With distributions of market share as in table C.2, this is a hopelessly bad predictor for I. Airlines are not matching frequency shares exactly evenly through the day. However they may be matching in some parts of the day. If we take the schedule of the dominant carrier as entirely useful, and all smaller schedules as matched head to head by this schedule, the predictor Ih becomes:

$$Ih = 1/MSmax \qquad (C.5)$$

Where MSmax is the maximum market share. (1)

For our 48 competitive markets (table C.2) we calculated actual indices of competition (Ia) and compared them with Ih. The actual index of competition was calculated by the "half hour" rule. Any departure within 30 minutes after a previously counted departure was not counted. (2) Eleven out of the 48 markets appeared to be more or less head to head. But all the rest of the markets had far smaller Ia's. (3) The mean Ia was 1.34; the mean predicted by the head to head hypothesis was 1.85.

Our suspicion that head to head scheduling is not the norm seems supported.

Collusion

Another possibility for scheduling behavior is a careful interleaving of different airlines' departures to avoid each other's departure times. This is the norm under collusive operations in European markets. It maximizes total schedule convenience for the market while allowing service by more than one airline. Under these arrangements, all departures would count. The index of competition under the collusive hypothesis would be Ic=1.00. (4)

For the 48 competitive markets, only 7 points seemed to fit the collusive hypothesis even approximately. In fact an I of

---

(1) Market share is used as a surrogate for frequency share for reasons associated with the original use of this research. This situation is a lamentable weakness in the present analysis.

(2) No other useful rule could be implemented given the form of Eriksen's data.

(3) An Ia greater than Ih would mean a carrier was scheduling head to head with itself. This occurs once in a while due to the limitations of our definitions.

(4) In our 31 near monopoly markets, the average Ia was 1.08, indicating a little head to head scheduling of the existing competitive departures and some inefficiencies in our definitions and processes.

1.00 was about a standard deviation below the mean observed for Ia, Ia=1.34.

Collusive scheduling definitely occurs in some markets, but it is not the dominant or representative behavior.

Blind Competition

A third hypothesis was created before the data was selected and it turned out to fit the average behavior pretty well. This hypothesis was that airlines by and large ignored each others time of day schedules in scheduling their own flights. Thus some flights (perhaps half by our "half hour" definition) would be head to head and some interleaved. The whole process would have lots of random variation in the results. However a predictor for the index of competition for blind competition Ib assuming half matching flights would be:

$$Ib = 1/(MSmax+.5(1-MSmax)) \quad (C.6)$$

The term in the denominator is the first pair of terms of the Taylor series expansion of $\sqrt{MSmax}$ about 1.0. Thus Ib becomes approximately

$$Ib \cong 1/\sqrt{MSmax} \quad (C.7)$$

It is this form (C.7) we used as the predictor for the index of competition.

Figure C.1 shows that 30 out of 48 points fell closer to this prediction than either the head to head or the collusive predictions. Furthermore, the average value was right and the root mean square error was by far the lowest of all the predictors, as shown in table C.3. (1)

Apparently the blind competitive hypothesis does describe the trend and predict the average wasted departure times reasonably well for 2 and 3 competitor markets.

Further Discussion

Frankly, the evidence is less than overwhelming. An attempt was made to eliminate the weakness in defining the actual index of competition by testing directly for displacement time vs. schedule in competitive situations. That is, given (C.3), what predictors do we get for h using these three predictors for index of competition? We calibrated Eriksen's displacement times against that predicted by a choice of I divided by the actual frequency for our 48 markets plus 19 more city pairs with frequencies between 3 and 6. (Table C.4 presents the added data points.) (2)

The results of these regressions are presented in table C.5. They need some interpretation. First, the Ia's by the half hour

---

(1) The comparison is not fair since the other two hypotheses can have errors in only one direction.

(2) Low frequency markets were not used before because the definition of head to head by the "half hour" rule did not fairly apply, but these new points greatly add to the data base for displacement times.

Figure C.1: Indices of Competition

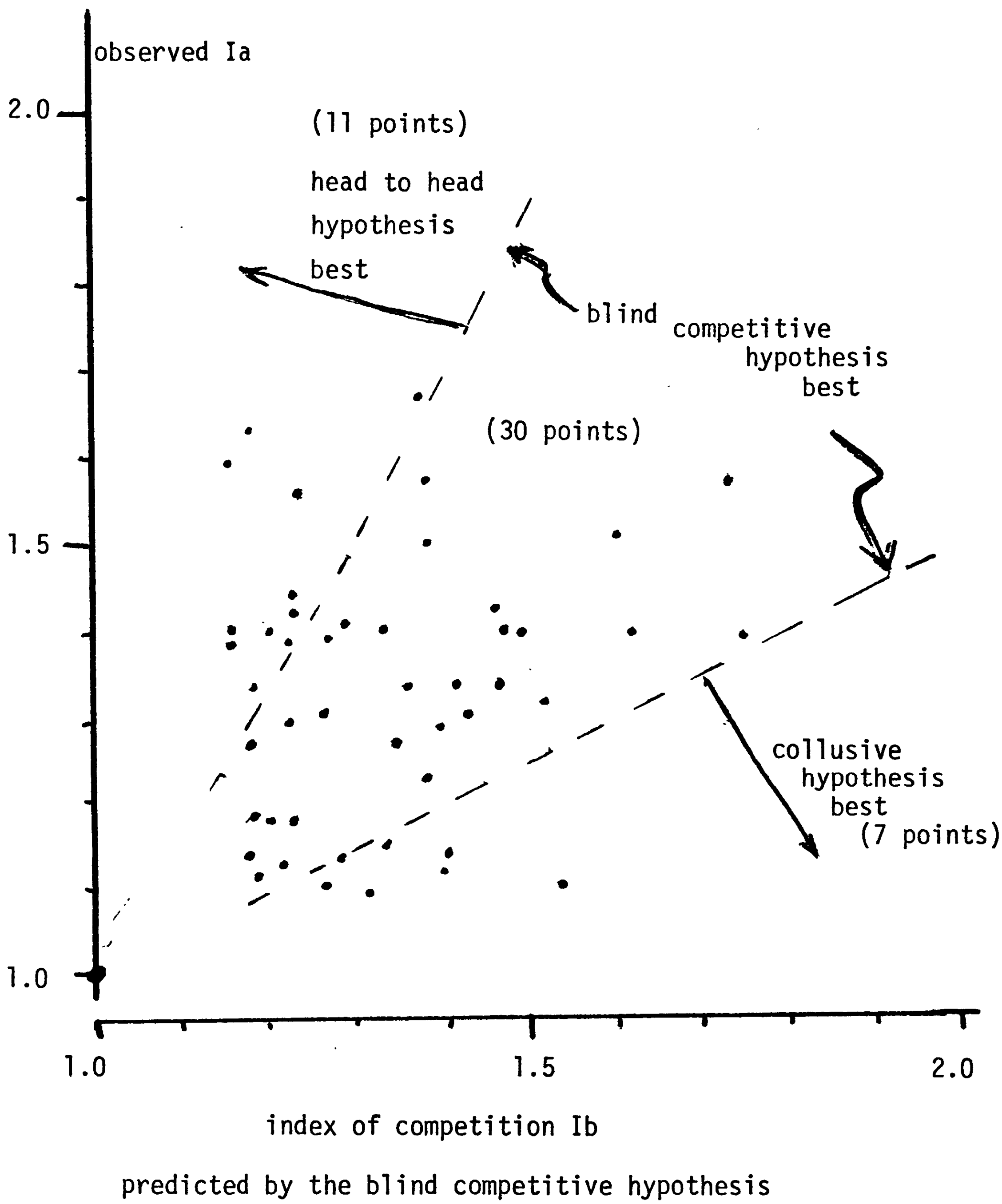


Table C.1: Data for 31 "Monopoly" Markets

City Pair Airport Codes from ref 27	largest market share	service frequency	unduplicated frequency	displacement time
SLC LAX	1.00	8	6	.90
MKC STL	.81	16	12	.44
STL MKC	.82	15	13	.42
ALB NYC	1.00	11	9	.56
PHL ORF	.90	9	8	.74
ORF PHL	.91	8	7	.67
NYC ALB	1.00	7	7	.82
ORD OMA	1.00	7	7	.90
OMA ORD	1.00	7	7	.83
ORF RIC	1.00	7	6	1.18
PIT ALB	1.00	6	6	1.01
TUC DEN	1.00	6	5	.96
LAX SLC	1.00	6	5	.91
RIC ORF	1.00	6	6	1.36
ICT OKC	.89	5	5	1.00
ATL CVG	1.00	5	5	1.30
DAL TUC	1.00	5	5	1.76
DAY PIT	1.00	5	5	.98
PIT DAY	1.00	5	5	1.05
DEN TUS	1.00	5	4	1.31
RDU WAS	1.00	5	5	.90
WAS RDU	1.00	5	5	1.22
BOS BTT	1.00	4	4	1.04
DTT BOS	1.00	4	4	1.23
LAS RNO	1.00	4	4	1.45
RNO LAS	1.00	4	4	1.40
ALB BOS	1.00	3	3	1.78
BOS ALB	1.00	3	3	1.56
TUS DAL	1.00	3	3	2.11
BIS MOT	.98	3	3	3.13
MOT BIS	.87	3	2	2.45

Table C.2: Data for 48 Competitive Markets

City Pair code	MSmax	total frequency	unduplicated frequency	displacement time
MSY HOU	.45	18	13	.36
ATL MSY	.72	14	11	.86
HOU MSY	.46	17	12	.42
SFO SEA	.67	14	9	.58
PWB BOS	.43	17	13	.37
ORD LAX	.28	21	11	.61
MSY ATL	.72	13	8	.58
BOS PWM	.47	16	12	.50
DFW OKC	.70	13	11	.45
OKC DFW	.60	14	10	.49
SEA SFO	.52	15	10	.53
DFW TLS	.66	13	9	.51
PIT WAS	.55	14	11	.45
MEM ATL	.74	11	8	.57
DEN SFO	.74	11	8	.81
SFO DEN	.74	11	7	.77
WAS PIT	.46	14	10	.49
MKE MSP	.61	12	10	.61
MEM STL	.57	12	11	.46
DEN SLC	.48	13	10	.62
MIA WAS	.67	11	8	.81
PHL ORD	.62	11	10	.69
STL MEM	.62	11	8	.81
MKC OMA	.50	12	9	.60
ATL DFW	.70	10	9	1.11
WAS MIA	.65	10	7	.97
SLC DEN	.33	14	9	.53
MSP MKE	.52	11	7	1.13
OMA MKC	.52	11	9	.71
STL ATL	.68	9	7	.68
LAX PDX	.67	9	8	.77
ORD DAL	.52	10	6	.69
EAS SFO	.42	11	10	.51
ATL MEM	.60	9	8	.69
RDU RIC	.72	8	7	.74
ATL STL	.72	8	6	.90
PDX LAX	.51	9	7	.93
DTT PIT	.50	9	8	.60
SFO WAS	.33	11	8	.63
DFW LUB	.55	8	7	.76
PIT DTT	.54	8	6	.90
ROC ORD	.70	7	5	1.33
ORD TUS	.70	7	6	1.07
LUB DFW	.50	8	7	.70
LNK OMA	.65	7	6	.89
SFO NYC	.39	9	6	.83
NYC SFO	.38	9	8	.96
TUS ORD	.56	7	5	2.08

Table C.3 Average Indices of Competition for 48 markets

	average	deviation from observed
observed Ia	1.34	0.0
head to head Ih	1.85	0.69
collusive Ic	1.00	0.38
blind compet. Ib	1.35	0.20

Table C.5: Predictions for h

method	h	σ	t	R^2
Ia=actual	6.25	.22	28	.92
Ib=1/ MSmax	5.43	.21	26	.91
Ih=1/MSmax	3.90	.18	22	.88
Ic=1.00	7.31	.27	27	.92

Table C.4: 19 Competitive Markets

City Pair code	MSmax	total frequency	unduplicated frequency	displacement time
DAY STL	.75	6	6	.90
ORD ROC	.73	6	5	.89
DEN SEA	.68	6	6	1.26
DEN NYC	.68	6	3	1.81
MEM MSY	.62	6	5	1.32
SEA DEN	.56	6	4	1.53
NYC DEN	.56	6	5	1.47
PIT CVG	.52	6	5	1.19
MSY MEM	.52	6	6	1.20
OKC ICT	.72	5	5	1.27
STL DAY	.67	5	5	.99
CVG PIT	.58	5	5	1.42
LEX CVG	.59	4	4	1.64
CVG LEX	.57	4	3	1.62
OMA LNK	.54	4	4	2.69
MSU JAN	.72	3	3	1.53
JAN MSY	.54	3	3	3.58
PHL LAX	.43	3	3	1.71
LAX PHL	.42	3	3	2.03

rule produced a value for h very near that for monopoly markets. So our half hour rule was not unforgivably bad. Second, all four methods produced similar R2 and standard errors. This merely suggests that displacement time vs. frequency relationship is strong.

Third, there is a very weak indication that the blind competitive hypothesis is best, based on a combination of low standard error and a value of h closest to our best estimate from equation (C.3).

Summary and Conclusions

Displacement time depends on frequency. In competitive markets the usefulness of total market frequency in reducing average displacement times is reduced. But the reduction is only half that expected from head to head scheduling practices. The waste is about that expected if airlines schedule blindly against each other.

Better work should be done along these lines.

Appendix D: Analytical Solution to the Buffer Size Problem

The probability for a potential passenger being denied space on a flight according to the formulation discussed in chapter 4 is

$$Pd=\frac{1}{\mu}\int_{c}^{\infty}\phi(d)\cdot(d-c)\ dd \tag{D.1}$$

Where the following definitions apply:

Pd=the probability of being denied space

c= the flight's capacity

$\phi(d)$ = the distribution of demand d for a flight, which is assumed to be Gaussian with mean μ and standard devieation σ.

This formula can be evaluated analytically.* We define the non-dimensional version:

$x = (d-\mu)/\sigma$

$X = (c-\mu)/\sigma$

ϕ_0= the standard unit normal distribution

Φ = the integral of ϕ_0, the cumulative normal.

With these variables (D.1) becomes:

$$Pd = \frac{\sigma}{\mu}\int_{X}^{\infty}\phi_0(x)\cdot(x-X)\ dx$$

$$= \frac{-X\cdot\sigma}{\mu}\int_{X}^{\infty}\phi_0(x)dx + \frac{\sigma}{\mu}\int_{X}^{\infty}\phi_0(x)\ x\ dx$$

We note that

$$x\ dx = 1/2\ d(x^2)$$

---

*The derivation was suggested by equation (2) in [52].

and also

$$\phi_0(x) = 1/\sqrt{2\pi}\cdot e^{-x^2/2}$$

So we may state

$$Pd = \frac{-X\cdot\sigma}{\mu}\cdot\Phi(x)\Big|_X^\infty + \frac{\sigma}{\mu}\int_{X^2}^\infty \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}\cdot\frac{1}{2}\cdot d(x^2)$$

The integral may be evaluated to become

$$Pd = \frac{-X\sigma}{\mu}\Phi(x)\Big|_X^\infty - \frac{\sigma}{\mu}\,\phi_0(x)\Big|_X^\infty$$

$$Pd = \frac{\sigma}{\mu}\{X\cdot(\Phi(X)-1)+\phi_0(X)\} \qquad \text{(D.2)}$$

We recall that X may be interpreted as the buffer in seats between the mean demand and the capacity, non-dimensionalized by the standard deviation of the demand.

We see from (D.2) that for constant Pd, as the ratio σ/μ rises the term in brackets may become smaller. This means the buffer non-dimensionalized by the standard deviation may fall.

The expression (D.2) must still be evaluated numerically since there is no closed form expression for Φ. We have performed these evaluations for a representative set of values for μ/σ in table D.1. The middle two columns of D.1 illustrate the difference between load factors designed to meet the 2% criterion for Pd as we have developed it and load factors designed using the simpler assumption that constant buffer in terms of σ produces constant service quality. The range of actual turnaway rates for constant buffer (in σ's) and different

Table D.1: Buffers Needed for Constant Pd, Lost Turnaway Model

Pd = 2%

μ/σ	for Pd=2%, buffer in σ's	nominal load factor	load factor at buffer=1.36·σ	Pd for buffer=1.36σ
1.0	1.67	37%	42%	4.0%
1.5	1.49	50%	52%	2.7%
2.0	1.36	60%	60%	2.0%
2.5	1.26	66%	65%	1.6%
3.0	1.17	72%	69%	1.3%
10.0	0.50	95%	88%	0.2%

Table D.2: Ranges of Load Factors for Constant Pd

assumed load size μ	35% cyclic σ_c	random $\sigma_r = \sqrt{2\mu}$	total σ	μ/σ	load factor for 2% Pd
20	7	6.32	9.43	2.12	61%
35	12.25	8.37	14.83	2.36	64%
50	17.50	10.00	20.16	2.48	66%
75	26.25	12.25	28.97	2.59	67%
100	35	14.14	37.75	2.65	67%
150	52.5	17.32	55.28	2.71	68%
200	70	20.00	72.80	2.75	69%
400	140	28.28	142.82	2.80	70%
∞	-	-	-	2.86	70%

random variance estimated assuming travel
in groups of 2 (as in [52])

distributions (in μ/σ) is broad. Equation (D.2) says that for constant X, the turnaway rate is proportional to σ/μ. This is illustrated in the last column of D.1.

In a practical sense μ/σ changes with μ because of the random component of σ. Such changes produce a spread of 8 points in load factor required for constant denial rate across load sizes in reasonable ranges. This is shown in the final column of table D.2. Again, the changes in load factor are on the smae order as changes due to designs optimized for different densities and distances as in chapter 5.

Appendix E: Estimate of Cyclic Variation in Demand

This appendix estimates the variability of demand caused by the combination of monthly, daily, and hourly traffic cycles. The random variable d is defined as the demand for a randomly selected flight in a schedule. $\phi(d)$ is the distribution of that random variable. $\phi(d)$ is not directly observable, but an analogous distribution of load factor about its mean is. We call the distribution of either loads or load factor $\phi(l)$. We understand that the distribution will be nondimensionalized by its mean, so the distinction between load and load factor is not relevant. We will attempt to recreate the distribution $\phi(d)$ about its mean by examining the traffic on large sections of airline networks. Available data present the traffic totaled in several different ways. These totals reveal several independent demand cycles. By combining these separate influences, we estimate $\phi(l)$ in general.

Unfortunately, the distribution we get will be for all travel markets in general. Also, it will be per hour rather than per departure. This is an important point because departure frequencies do match to some degree demand cycles. Furthermore night coach fares also adjust a highly cyclic demand to a less cyclic supply. The net result is that our estimate of traffic distributions is an imperfect estimate of the demand per departure distributions.

Nevertheless, there is value in this sustained numerical effort. First, the components of the variability of load factors are observed separately and their relative importance measured. Second, detailed reconstruction of ϕ will develop a fuller and less theoretical understanding of the problems of demand variability and the difficulty of its measurement. Third, as an end result we will have a rough support for our estimate of $\mu/\sigma = 2$, which we use in later sections. And finally, economies of scale in aircraft capacity are reinforced by efficiencies associated with random part of this variability.

Cycles of Demand

The most fundamental cycle of demand is weekly. Most airline schedules are identical on a day-to-day basis while demand volume is not. Figure E.1 displays the cycles of traffic through the week. A schedule with capacity sufficient to exactly accommodate Friday's mean demand would obtain an average load factor of 87% through the week.

The most interesting statistic from Figure E.1 is the estimate of the standard deviation of weekly demand. The standard deviation is best estimated as 9.7% of the mean. (1) If the demand in each hour of the day went through the same day of the week cycle, this would be the contribution of the day of the week to the standard deviation of demand. Since some hours, such as noon, are less cyclic than others, say 5:00 p.m., the data in figure E.1 represent a lower bound.

---

(1) "Best" estimate is route mean square deviation from mean, adjusted by 7/6 to allow for the loss of a degree of freedom from calculating the mean.

A similar examination of total travel in the months of the year for all airlines and distances reveals a monthly standard deviation of 9.6% of the mean. Monthly variations for different markets are undoubtedly different, so this is again a lower bound. On the other hand, when we relate demand to capacity in order to obtain expected standard deviation of load factor, we must bear in mind that capacity is also adjusted seasonally among markets so the <u>unmatched</u> variability in monthly demand may be near 9.6% of the market or even less.

Convolution of weekly and monthly cycles is possible, since the way the data were gathered guarantees their independence. Representing the weekly cycle as in Figure E.2 and the monthly as in Figure E.3 and multiplying out the probabilities produces the distribution in E.4. This is an estimate of the day-to-day variability over the year. The standard deviation is approximately the root mean square of the sum of the two previous standard deviations, or 13.6%. The distribution is far from Gaussian, but is moving in that direction.

There are also hour by hour cycles of demand which must be satisfied by available aircraft capacity. Unlike daily or monthly cycles, hourly variations in demand occasionally fall so low that it is economic to idle equipment rather than operate with low loads. Aircraft are seldom involved in flying or being loaded for more than 14 hours in a day, (1) and less than 10 hours are needed for maintenance. This means some time is idle due to low demand. This often occurs at night.

Figure E.5 shows demand for the 14 hours of the day between 7 a.m. and 9 p.m. This is for all markets, so peaks and valleys of individual markets cancel out. The estimated standard deviation is 33% of the mean. Thus if aircraft are used equally through the 14 hour day, hourly variations in traffic far outweigh weekly or annual cycles in influencing load factor. Hour by hour cycles dominate the standard deviation of air passenger demand. The convolution of all three effects produces a standard deviation of approximately 35.5% of the mean. Hourly variability alone accounts for 32.8%.

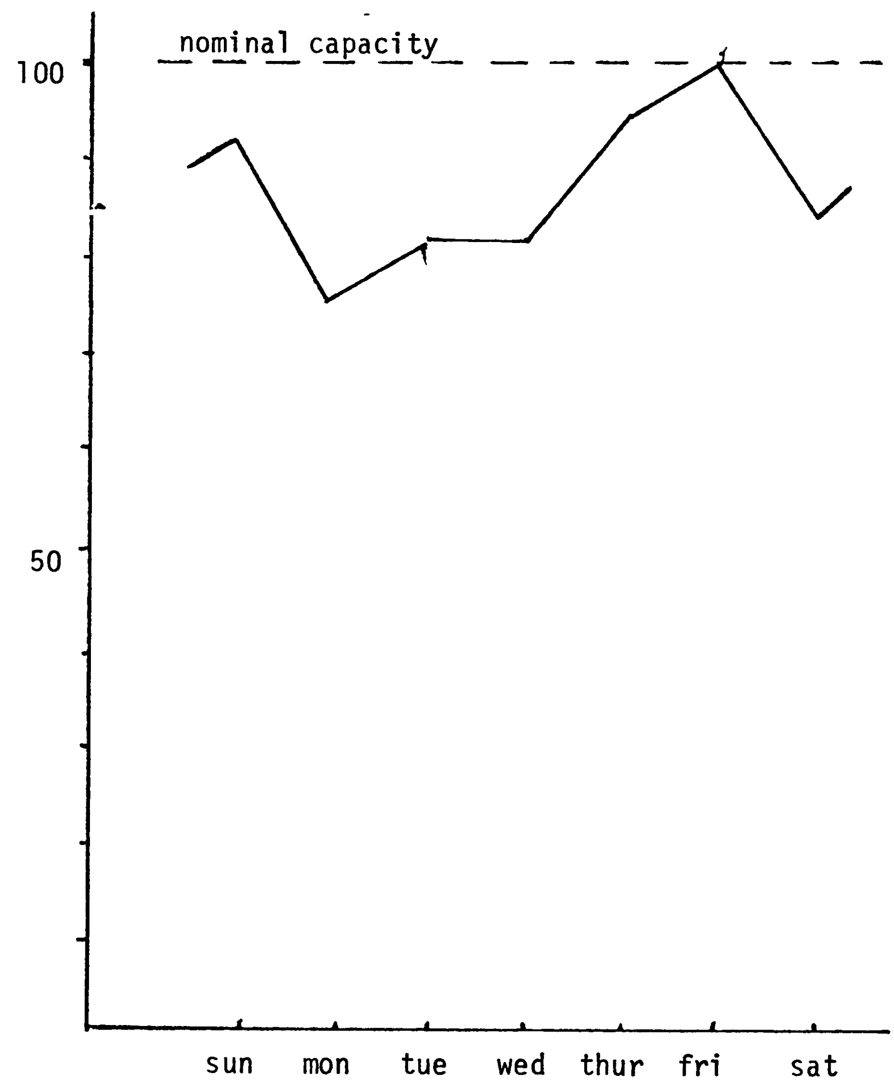
These three cycles, monthly, daily, and hourly, combine to make that part of the variability in the load on a departure which we have considered to be independent of the load size. We named that part of the variability of demand the cyclic variability and we have now described the cycles which influence it. Our estimate of the cyclic variability is just under 40% of the mean demand. (2)

---

(1) See Figure 2.3.1 in Chapter 2 for an illustration of this.

(2) Because of the use of large totals of traffic to estiamte the variations in O(1), random components of the variability have averaged out of all observations in this section.

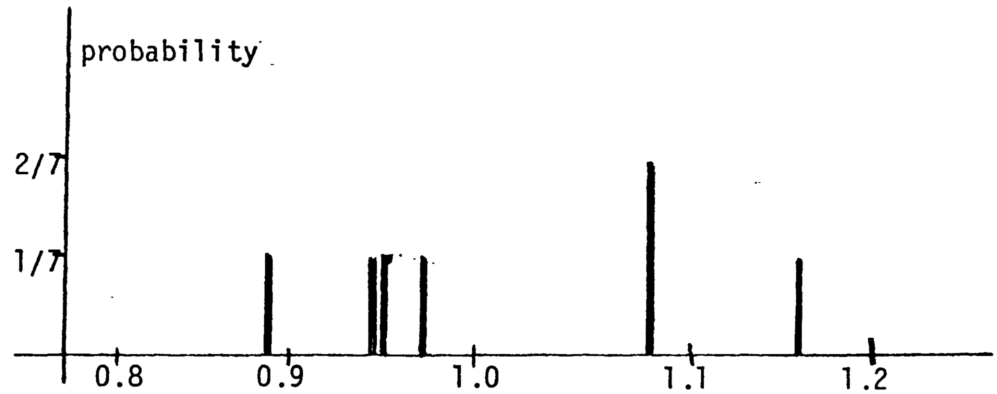
Figure E.1: Air passenger traffic cycle by day of week



Source: CAB Discount Fare Policy, Federal Register, Vol 42, no 100, p 18 (ref 40).

Data is summed passenger boardings for four trunk airlines for week ended 14 Aug 76. Trip lengths of 501 to 1000 miles are included. Precise definition of data measure is somewhat speculative due to incomplete explanation in original reference.

Figure E.2 Distribution of traffic through the week by day

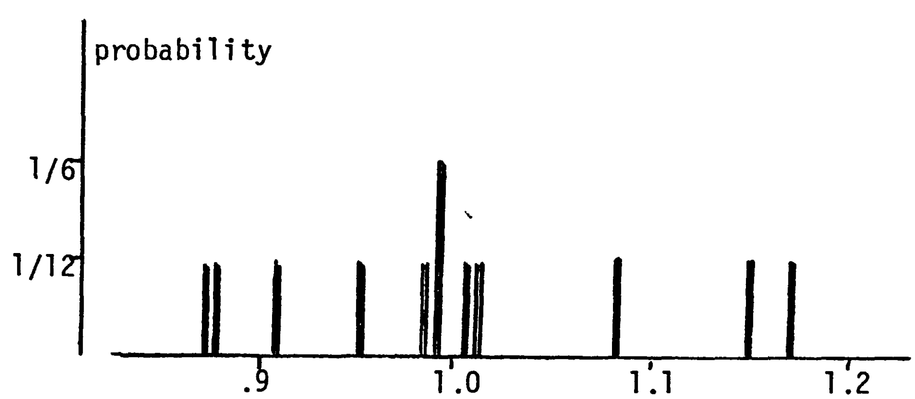


For markets 501-1000 miles; week ended 14 Aug 76 for four trunk airlines

Average is 87% of maximum; $\sigma/\mu = 9.7\%$

Source: CAB Discount Fare Policy; Federal Register, Vol 42, no 100, p 18. (see comment, figure E.1)

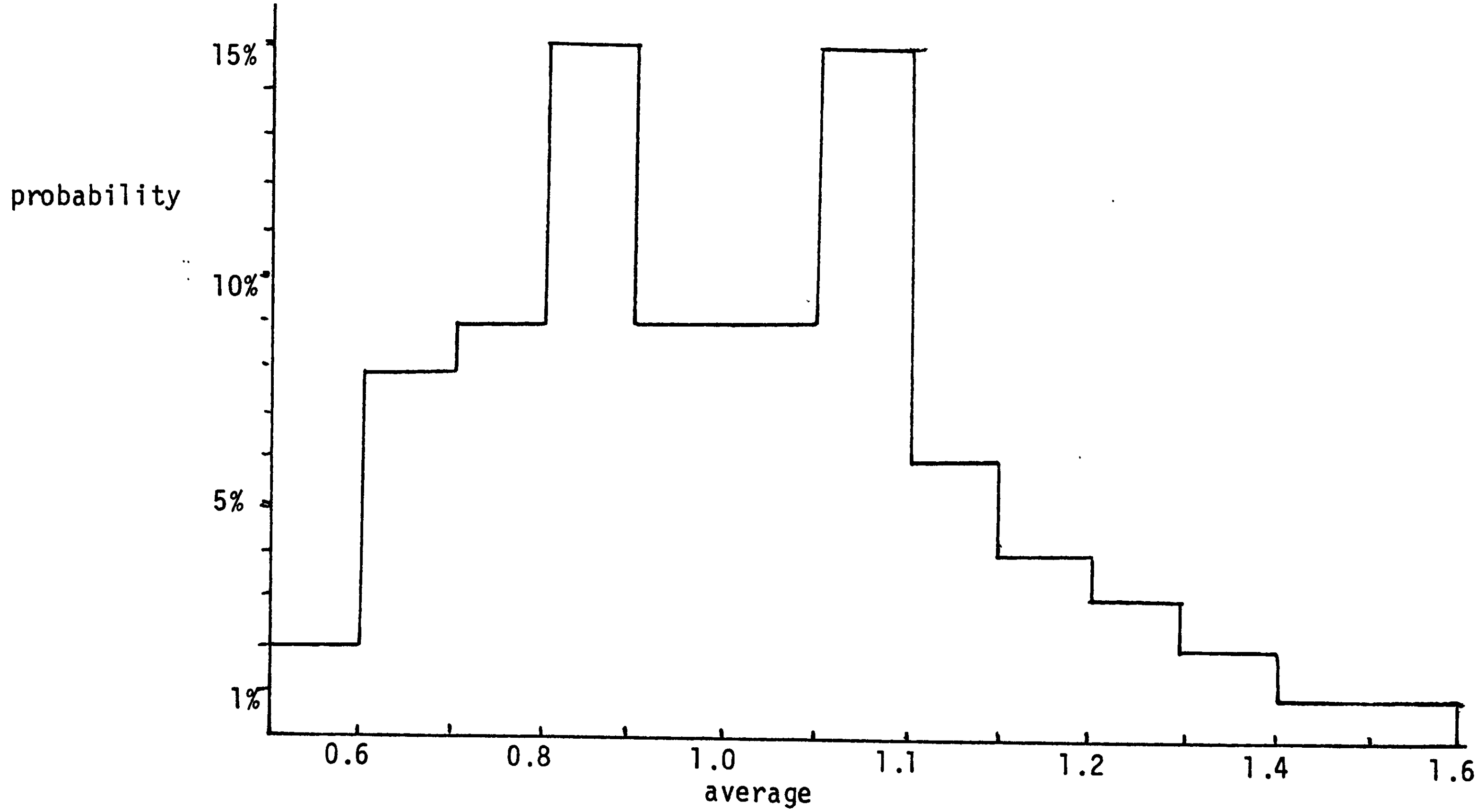
Figure E.3: Distribution of traffic through year by month



For year ending sept 76. Data adjusted to remove growth trend of 1% per month. Data is systemwide passenger miles for four trunk airlines.

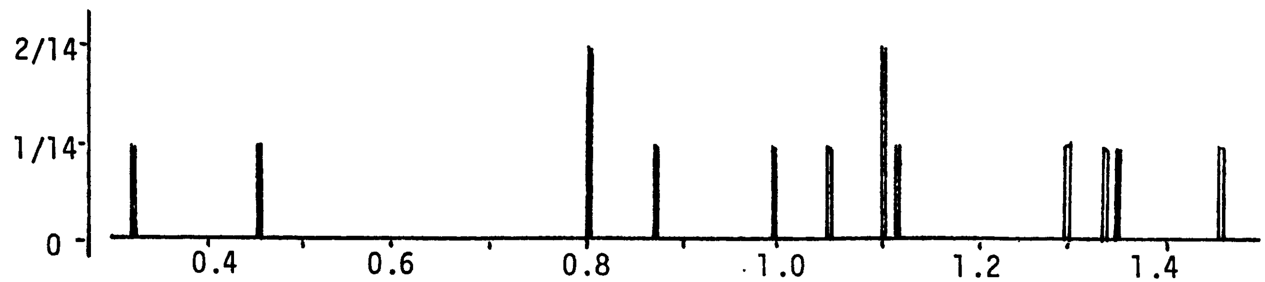
Source: Ibid., p 14.
Average load is 85% of maximum; $\sigma/\mu = 9.6\%$

Figure E.4: Convolution of day of week and month of year distributions of traffic



Average is 74% of maximum (data from figures E.3 and E.2)

Figure E.5: Distribution of traffic Through the Day by Hour



Average load is 69% of maximum

Source: CAB Discount Fare Policy, Federal Register, Vol 42, no 100, p7. (ref 40) Data for the month of February, 1976; 7am to 9 pm. Four trunk airlines. See comment on figure E.1.

References -- number to author conversion list

1 Air Transport Assn.
2 Caves
3 Blumer & Swan
4 Chamberlin
5 CAB
6 CAB
7 CAB
8 CAB
9 Dorman
10 Douglas & Miller
11 Eriksen
12 Fruhan
13 Mohring
14 Gordon & de Neufville
15 Wycoff
16 Grenau
17 Hotelling
18 Greig
19 Lankford
20 Cherrington
21 Nason
22 Neuve-Eglise & Simpson
23 Roberts
24 Sercer
25 Poole
26 Simpson
27 Official Airline Guide
28 Jordan
29 Wheatcroft
30 Wong
31 Eads
32 Gallant, Scully, & Lange
33 DeVany
34 Swan
35 DeVany
36 Anderson
37 Roberts
38 de Neufville & Gelerman
39 Janes
40 Federal Register
41 Matheisal & Swan
42 Swan
43 Simpson
44 Kahn
45 Taneja
46 Simpson
47 Simpson & Taneja
48 Douglas Aircraft
49 Bard
50 de Neufville & Mira
51 Rosenberg
52 Shlifer & Vardi

References

[1] Air Transport Association, Standard Method of Estimating Comparative Direct Operating Costs of Turbine Powered Transport Airplanes, December 1967.

[36] Anderson, James E., "Efficiency in the Market for Airline Services," Discussion paper no. 73, Boston College Economics Department, 1976.

[49] Bard, Y., Nonlinear Parameter Estimation, Academic Press, 1974.

[3] Blumer, T., and Swan, W. An Analysis of Short Haul Air Passenger Demand, NASA CR 152157, 1978.

[2] Caves, Richard E., Air Transport and its Regulators, Harvard University Press, 1962.

[4] Chamberlin, E.H., Theory of Monopolistic Competition, Harvard Press, 1958.

[20] Cherrington, Paul W., Airline Price Policy: A Study of Domestic Airline Passenger Fares, Harvard Press, 1958.

[8] Civil Aeronautics Board (CAB), Aircraft Operating Cost and Performance Report, 1976.

[7] CAB, Origin and Destination Survey, table 8, 1976.

[5] CAB, Domestic Passenger Fare Investigation, U.S. GPO 003-006-00071-9,

1976. [40] CAB, "Discount Fare Policy," Federal Register, 24 May 1977, part VIII.

[50] de Neufville, Richard and Mira, L.J., "Optimal Pricing Policies for Air Transport Networks," Transportation Research, vol. 8, pp 181-192, 1974.

[38] de Neufville, Richard and Gelerman, W., "Planning for Satellite Airports," Transportation Engineering Journal, August 1973.

[33] DeVany, A.S., "The Fffect of Price and Entry Reguation on Airline Output, Capacity, and Efficiency," Bell Journal of Economics, Spring 1975.

[35] DeVany, A.S., "Revealed Value of Time in Air Travel," Review of Economics and Statistics, vol 16, Feb., 1974.

References

[9] Dorman, Gary J., Airline Competition: A Theoretical and Empirical Analysis, Berkeley Department of Economics Working Paper, SC-7603, Sept. 1976.

[48] Douglas Aircraft, Study of Short Haul Aircraft Economics, NASA CR-137686, Sept. 1975.

[10] Douglas, G.W. and Miller, J.C.,, Economic Regulation of Domestic Air Transport, Brookings, 1974.

[31] Eads, George, The Local Service Airline Experiment, Brookings, 1972.

[11] Eriksen, Steven, Policy Oriented Multi-Equation Models of U.S. Domestic Air Passenger Markets, MIT Ph. D., 1977.

[12] Fruhan, William E., Flight for Competitive Advantage: A Study of the U.S. Domestic Trunk Air Carriers, Boston, 1972.

[32] Gallant, R., Scully, M., and Lange, W., Analysis of VSTOL Aircraft Configurations for Short Haul Air Transportation Systems, MIT FTL report FT-66-1, 1966.

[14] Gordon, S. and de Neufville, R., "Design of Air Transportation Networks," Transportation Research, vol. 7, pp207-222, 1973.

[18] Greig, John A., Rationalization of Airline Networks within Western Europe with Special Reference to Swissair, Ph. D., University of London, 1975.

[16] Gronau, R., "The Value of Time in Passenger Transportation: the Demand for Air Travel," National Bureau of Economics Research Occasional Paper 109. Columbia University Press, 1970.

[17] Hotelling, Harold, "Stability in Competition," Economic Journal, vol. 39, March 1929.

[39] Jane's All the Worlds Aircraft, Jane's Yearbooks, London, 1976.

[28] Jordan, William, Airline Regulation in America, Johns Hopkins, 1970.

[44] Kahn, Alfred E., The Economics of Reguation, Wiley & Sons, 1970.

[19] Lankford, R.L. and Pina, E.I., Methods for Analysis,

References

Management, and Marketing of Scheduled Airline Capacity, Boeing Aircraft, 1977.

[41] Mathaisel, Dennis, and Swan, W.M., "Airline Competitive Scheduling Behaviour," never published. For this reason, the data and analysis is presented in Appendix C of this thesis.

[13] Mohring, Herbert, Transportaion Economics, Ballenger Co., 1976.

[21] Nason, Scott D., Competition and Efficiency in Regulated Airline City-Pair Markets, MIT CTS report 77-16, MIT 1977.

[22] Neuve-Eglise, M.J. and Simpson, R.W., A Method for Determining Optimum Vehicle Size and Frequency of Service for a Short Haul V/STOL Air Transportation System, MIT FTL Report R-68-1, 1968.

[27] The Official Airline Guide, North American Edition, Reuben Donnelly, 1974.

[25] Poole, E.C., Costs - A Tool for Railroad Management, Simmons-Boardman Publishing Corp, N.Y. 1962.

[37] Roberts, P.O., A Set of Simplified Multimodal Cost Models for Use in Freight Studies, MIT CTS report 77-27, 1877.

[23] Roberts, P.O., Forecasting Freight Flows Using a Disaggregate Freight Demand Model, MIT CTS Report 76-1, 1976.

[51] Rosenberg, Arne, Air Travel in Europe, National Swedish Consumer Counsel, 1970.

[24] Sercer, R.W., Block Times for Passenger Aircraft, MIT FTL memo M73-11, 1973.

[47] Simpson, R.W. and Tanaja, N.K., A Multi-Regression Analysis of Airline Indirect Operating Costs, MIT FTL report R67-2, Dec. 1967.

[46] Simpson, R.W., Airline Market Share Analysis, class notes fo 16.74 at MIT.

[43] Simpson, R.W., Class Notes for Airline Economics. See also [22]

References

[26] Simpson, R.W., "Technology for Design of Transport Aircraft," MIT-NASA Workshop; Airline Systems Analysis, MIT FTL Report R72-7, 1972.

[42] Swan, W.M., "Network Study of Subsidized Air Service," Journal of Aircraft, Vol. 13, no 4, April 1976.

[34] Swan, W.M., Economics of Airline Networks: A Mathematical Model, MIT CTS report 77-10, 1977.

[52] Shlifer, E. and Vardi, Y., "An Airline Overbooking Policy," Transportation Science, vol. 9, no 2, 1975.

[45] Taneja, Nawal K., The Commercial Airline Industry, Lexington Books, 1976.

[29] Wheatcroft, Steven, Economics of European Air Transport, Harvard Press, 1956.

[30] Wong, R.T., Accelerating Bender's Decomposition for Network Design, MIT Ph. D., Feb. 1978.

[15] Wycoff, D.D. and Maister, D.H., THe Domestic Airline Industry, Chapter 7, Lexington Books, 1977.