predicting sector flows that, although it is relatively easy to apply, can readily indicate potential areas of excess traffic loading. Based on empirical data, the method preserves general patterns of network flow without specifying the actual geometry of each aircraft's flight. For ATC network planners, such a method for predicting traffic distribution could provide a useful tool for ensuring safe and efficient movement of air traffic.

ACKNOWLEDGMENT

I wish to acknowledge the assistance and support of the Federal Aviation Administration's Technical Center, Atlantic City, New Jersey, in pursuing this work. Portions of the study were supported by the U.S. Department of Transportation.

REFERENCES

1. N.W. Polhemus. Modeling Aircraft Flow in Air Traffic Control Systems. Proc., Transportation Research Forum, 1974, Vol. 15, pp. 526-539.
2. J.R. Ford and D.R. Fulkerson. Flows in Networks. Princeton Univ. Press, Princeton, NJ, 1962.
3. R.B. Potts and R.M. Oliver. Flows in Transportation Networks. Academic Press, New York City, 1972.
4. T.W. Anderson and L.A. Goodman. Statistical Inference About Markov Chains. Annals of Mathematical Statistics, Vol. 28, No. 1, 1957, pp. 89-110.
5. C. Chatfield. Statistical Inference Regarding Markov Chain Models. Applied Statistics, Vol. 22, No. 1, 1973, pp. 7-20.
6. G.T. Duncan and L.G. Lin. Inference for Markov Chains Having Stochastic Entry and Exit. Journal of the American Statistical Association, Vol. 67, No. 340, 1972, pp. 761-767.

Publication of this paper sponsored by Committee on Airfield and Airspace Capacity and Delay.

Analyzing Ticket-Choice Decisions of Air Travelers

SCOTT D. NASON

This paper examines the nature of the problem that faces air travelers confronted with choosing from among a variety of air fares, each associated with different service characteristics, and the problem of forecasting these decisions. A theoretical framework is developed that views the problem at the level of the individual traveler; the ticket-type choice is expressed in terms of the individual's socioeconomic characteristics, the characteristics of the trip in question, and the level of service associated with each available alternative. Logit models are suggested as the preferable functional form on the basis of theoretical and computational grounds, and the properties of logit models are briefly described. A pilot application of the method is presented for a two-alternative situation (full fare versus standby) by using a small sample of interview data collected from departing passengers at Boston's Logan Airport. A calibrated model is presented that demonstrates a statistically significant relationship between the ticket-type choice and the fare, fare differential, trip purpose, automobile ownership (as a proxy for income), and the passenger's perception of the delays that may be expected if flying standby. This application merely demonstrates a method and could easily be improved by using the airlines' onboard surveys for estimation.

Events during the last few years have substantially altered the air-travel-demand forecasting requirements of the individual airlines. Until recently, the number of different fares available was quite limited, and differences among the fare packages available from individual airlines were almost nonexistent. In this environment, the crucial requirements were for an aggregate estimate of the size of an individual city-pair market, which may or may not have been based on the level of service available in that market, and a carrier's share of the total, based on a measure of that carrier's frequency share (or a more-sophisticated model that took into account the timing of those flights).

With the advent of deregulation, pricing freedom has emerged as a major factor that influences air-travel-demand decisions. Discount fares have stimulated new travel. Just as important to airline marketing departments is the impact on the yield per passenger or per passenger mile, which is affected by the passenger's choice of ticket type, as well as the impact of discount fares on the passenger's carrier-choice decisions. Passengers have always made minor distinctions between carriers on the basis of food, cabin attendants, or advertisements, but more and more there is a tangible economic incentive to choose one carrier over another. Examples include the unlimited-mileage tickets available on Eastern and Allegheny; the straight price reductions offered in some markets by National, Braniff, Texas International, World, and Transamerica (among others); and half-price coupon offers from United and American.

This paper examines the nature of these new decisions that face air travelers and proposes a technique that should prove useful in analyzing the passenger's ticket-choice decisions. The ticket-type choice is viewed within the context of the entire trip-planning process. Each individual's decision is based on that person's characteristics and the characteristics of each available alternative--travel time, price, reservation, length-of-stay restrictions, etc. This type of problem has exact parallels in other decision-making processes, and the modeling of personal preferences, which is well developed elsewhere, is adapted to the problem at hand.

TRIP-PLANNING PROCESS

There are several decisions involved in planning a trip by air; these include (a) a decision to travel somewhere, (b) a choice of destination or destinations and departure and return times, (c) a decision to fly in preference to other modes of travel, and (d) a selection of the least-expensive and most-convenient flight and ticket combination. For many trips, some of these decisions may be trivial or made simultaneously with other decisions. The first three (or even all four) are likely to be made simultaneously and without much

hesitation in a number of cases, such as a necessary business trip from New York to Los Angeles. On the other hand, each decision may be distinct and nontrivial, as in the planning of a summer vacation.

The primary concern here is in analyzing the fourth decision, the flight and ticket-type choice. To that end, it is assumed merely that the first three decisions precede this one. That is, passengers determine the nature of their trip and then obtain the most-favorable flight and fare available. Problems that arise due to the less-common case in which flights or fares determine the destination choice or the duration of the stay must also be considered.

DETERMINANTS OF FLIGHT AND TICKET-TYPE CHOICE

Unlike standard air-travel-demand models, in which average population statistics and average fare and travel time characteristics are used to estimate aggregate demand, the ticketing problem should be viewed at an individual (or disaggregate) level. Each traveler's choice must be expressed in terms of the individual's characteristics and the characteristics of the choices available for the given trip.

Personal characteristics are used to approximate the different tastes of potential travelers. For example, it is generally conceded that individuals who have a high income place a higher value on travel-time savings and thus that income is a valuable indicator of how a traveler will trade off time or convenience versus money. A number of other personal characteristics may conceivably affect an individual's travel behavior; these may include age, sex, travel experience, or other factors.

A second class of personal characteristics relates to the particular trip that is being made. The purpose of the trip may serve as an important indicator of other underlying choice determinants, which include who is paying for the trip, the length of stay (which may affect eligibility for discount fares), and the flexibility of departure and arrival dates and times. The trip characteristics are likely to determine which fares, flights, or both are available to the traveler and may again affect the value of travel time or the certainty with which advance plans may be relied on.

Finally, the level-of-service (LOS) characteristics of the available alternatives are instrumental in the traveler's choice from among them. Alternatives will generally be differentiated by price and may be characterized by different amenities (aircraft type, food, drinks, movies, etc.); by different booking requirements (minimum and maximum stay; advance reservations, payment, or both; and cancellation fees); and by different trip-time characteristics (night flights, change of planes, lack of guaranteed space).

Naming these factors and acknowledging their importance is, of course, far easier than expressing the magnitude of their impact. The latter problem may best be handled, however, by obtaining information about the above-named factors from many travelers, which would include the flight and fare-type decisions they made, and by attempting to infer what other people would do as a function of their personal and travel characteristics.

LOGIT MODEL

The modeling of individual travel choices has progressed substantially in the last six to eight years. During this period a number of applications have helped to develop and refine the technical problems involved in modeling individual behavior. For a combination of theoretical, empirical, and computational reasons, the logit model has emerged as the most-common and practical model for problems of this type and appears to be well suited to the problem at hand. A more-complete discussion of logit and other disaggregate modeling techniques as well as the advantages and disadvantages of each has been given by McFadden (1) and by Richards and Ben-Akiva (2).

In order to apply a logit model to a choice problem, it is necessary to determine the set of alternatives that faces each individual and to define the attractiveness (utility) of each. The utility function for each alternative is expressed simply as a linear function of the traveler's socioeconomic characteristics, the trip characteristics, and the LOS variables for the given alternative in the relevant city-pair market. The sign and magnitude of the function coefficients are determined by means of an estimation process described below.

The utility functions (which express the relative attractiveness of each alternative) may be used to estimate the probability that any given alternative will be selected. This probability should, of course, increase with the utility of the alternative and should approach a maximum value of 1 for very high utilities or a minimum of 0 for very low ones. Also, the sum of the probabilities for the set of available alternatives must be 1 (i.e., only one alternative must be selected).

A number of possible functional forms meet the above criteria, but the logit model has generally been chosen for a variety of statistical reasons. In the logit model, the probability of selecting alternative i is expressed as follows:

$$P_i = \exp(U_i)/\Sigma\exp(U_j) \quad \text{for all j's} \tag{1}$$

This function yields an S-shaped curve that approaches 0 and 1 at extreme values of U_i. In addition, it should be noted that the curve is steepest (that is, changes in utility have the greatest effect) for values near i's proportional share of the market. Small changes in utility are relatively unimportant for very popular or very unpopular alternatives.

Computational features of the logit model make it relatively easy to set the probability of one or more alternatives to zero. Thus, for example, it would be necessary to make supersaver fares unavailable to travelers in markets in which such a fare is not offered. It would be prudent to make it similarly unavailable to business travelers who had little advance knowledge of the timing and destination of their trips. Such travelers do not assign a low utility to supersaver fares; they are simply unable to take advantage of them. Similarly, it may be argued that some alternatives are unavailable to many travelers because of a lack of knowledge of their existence.

DATA REQUIREMENTS

In order to estimate a logit model of passengers' ticket-type choice, it is necessary to have data on the decisions made by air travelers when faced with a known set of alternatives as well as data on passengers' socioeconomic characteristics. The requirements for information about the passenger and for a determination of which alternatives are actually available to a given traveler tend to argue against the use of waybill data for model estimation. All the necessary data could be obtained, however, by means of the airlines' on-board surveys.

These surveys should inquire about a few socioeconomic characteristics that may influence the

passenger's decision, such as age, income, and travel experience. The major portion of the survey must elicit information about the trip being made; this would include its origin, destination, flight, fare class, purpose, duration, advance notice, flexibility, ticket purchaser, size of party, and other similar information.

The total number of data points obtained is not so important as the distribution of those observations. Specifically, it should be recognized that no inferences regarding an alternative may be made unless it was the choice of at least some of the respondents. That is, in order to identify the characteristics of trips made under the budget fares, for example, it is necessary to obtain observations from some who chose budget fares over other alternatives and some who chose other fares in preference to budget fares. It is the comparison of these two groups of people and their trips that yields the ability to predict what others may choose. Similarly, it should be recognized that carrier-choice models would require data from more than one carrier.

As a general rule, a random sample of a few thousand passengers would be more than adequate to support any such modeling effort that did not have specific goals regarding some lightly used alternative. Another option would be to use a choice-based sampling method, in which users of some alternative are specifically surveyed in order to ensure a minimum number of such observations. Known statistical methods can correct whatever biases may have resulted from nonrandom sampling.

Although coverage of every interesting alternative is necessary, it is not essential that every city-pair market be surveyed. In fact, the models should be totally independent of the markets for which they are calibrated. Market characteristics (such as fares, travel times, or distance) are factors in the ticket-type choice, but the effects of these variables should be uniform wherever basic behavior patterns are consistent. Thus an important feature of these models is that they may be estimated on one set of markets and applied to a different set of markets by substituting the appropriate personal and market data as input.

This geographic transferability is matched to a lesser extent by the temporal stability of the coefficients. So long as people do not alter their personal preferences with regard to the various LOS measures, the models will continue to hold their validity. This assumption will probably be valid for a few years. Shifts in use from one alternative to another will result (in the short run) from changes in the personal characteristics of travelers (such as changes in income or travel experience) or from other modeled variables such as the purposes of trips, the fares charged, or travel times.

ANALYZING NEW ALTERNATIVES

It follows from the discussion above that an alternative that does not exist and therefore that could not have been chosen in the sample requires special treatment. These options are, needless to say, perhaps the most interesting from a forecasting standpoint. Such alternatives fall into two classes and present differing magnitudes of concern.

First, many new alternatives are in fact only extensions of existing alternatives. A longer minimum stay or advance-purchase requirement, for example, may be analyzed within the present framework if the variables that reflect these requirements are cardinal rather than ordinal variables. Thus, an explanatory variable equal to the number of days before departure that the ticket must be purchased may be calibrated on observations of 7- and 14-day requirements and extrapolated (cautiously) to 30 days or some other length of time.

New fares or flights that do not have comparable existing examples require special treatment. In some cases it will be possible to infer reasonable values for the coefficients in the new utility function. More often, it is necessary to survey passengers regarding their hypothetical use of the new type of fare. There are, of course, dangers involved in inferring behavior patterns from individuals' stated actions in hypothetical situations, but such problems may be overcome if special care is taken.

SAMPLE APPLICATION

The modeling technique outlined above has been applied in one pilot study by using sample data collected for a hypothetical standby ticket. [A complete description of this study may be obtained from the author.] The limitations of such a data set should be recognized, and the following example should be viewed as a demonstration of a technique rather than a presentation of a usable model. In addition to this study, however, the modeling technique has been successfully applied to the mode-choice decisions of nonbusiness travelers in France by using a home-interview survey of 2000 households.

Passengers at Logan Airport were surveyed to determine the price differential required to induce them to forgo the benefits of the guaranteed seat. This value of a reservation varies with income and the passenger's perceived level of inconvenience if denied boarding. The two alternatives that face each passenger, therefore, are to fly under a full-fare ticket at a known fare and travel time or to fly standby at a lower fare and an estimated delay due to the possibility of not obtaining a seat. Each passenger survey must then be used to generate a set of ticket-type choices over a range of reservation prices. The passenger is assumed to choose the standby ticket for all reservation prices greater than the indicated value and a full-fare ticket for all smaller price differentials.

Utility functions for the two alternatives were defined by the following variables:

$$U = f(\text{FARE, EXPECTED DELAY, BUSINESS, CARS, FLORIDA, PROFESSION}) \quad (2)$$

where

EXPECTED DELAY = function of probability that a seat may not be obtained if flying standby (PNOSEAT) and length of delay if a seat is not obtained (DELAY);
BUSINESS = dummy variable: 1 for business trips, 0 otherwise;
CARS = number of cars in household;
FLORIDA = dummy variable: 1 if destination was Florida, Bermuda, or Caribbean (survey was conducted in February when these markets are well traveled); and
PROFESSION = dummy variables: PROFEM = 1 if employed, 0 otherwise; PROFR = 1 if retired, 0 otherwise (i.e., PROFEM = 0, PROFR = 0 for unemployed, students, etc.).

The dummy for business is an attempt to distinguish

Table 1. Some sample results.

Destination	Dummy Variables						Probability of Using Full Fare
	RESPR ($)	BUSINESS	PNOSEAT	DELAY	PROFEM	CARS	
St. Louis	40	No	0.75	5	1	1	0.35
	40	No	0.25	5	1	1	0.23
	20	No	0.75	5	1	1	0.68
	0	No	0.75	5	1	1	0.86
Los Angeles	40	No	0.75	5	1	1	0.94
	40	No	0.10	5	1	1	0.65
	40	No	0.10	5	1	2	0.74
	100	No	0.50	5	1	2	0.08
New York	5	No	0.50	1	1	2	0.58
	20	No	0.50	1	1	2	0.33

Note: Dummy variables are defined in Equations 2 and 5.

passengers who have not paid for their ticket. CARS, PROFEM, and PROFR are proxies for income, which were reluctantly used because true income figures were unobtainable.

There is no theoretically sound basis for inclusion of a dummy variable for Florida. In fact, one would expect that a sound data base would not require such a market segmentation for model estimation. The fact that it is necessary here reflects some difference between winter vacation spots and other destinations that is not captured by other variables. For the purposes of this application, it is helpful to view the set of all airline trips as being divided into business, winter vacation, and all other.

By using the formula in Equation 1, it is possible to express the probability of choosing full fare as follows:

$$P_{FULL\ FARE} = \exp(U_{FULL\ FARE})/[\exp(U_{FULL\ FARE}) + \exp(U_{STANDBY})] \quad (3)$$

By dividing the numerator and the denominator by $\exp(U_{FULL\ FARE})$, it can be shown that

$$P_{FULL\ FARE} = 1/\{1 + \exp[-(U_{FULL\ FARE} - U_{STANDBY})]\} \quad (4)$$

Thus the probability of choosing full fare depends only on the difference between the two utilities. This property of logit models is used to simplify the estimation process.

The equation to be estimated is given as follows:

$$\begin{aligned} U_{FULL\ FARE} - U_{STANDBY} = {} & A_0 + A_1(\text{RESPR}) + A_2(\text{FULL FARE}) \\ & + A_3(\text{EXPECTED DELAY}) \\ & + A_4(\text{BUSINESS}) + A_5(\text{CARS}) \\ & + A_6(\text{FLORIDA}) + A_7(\text{PROFEM}) \\ & + A_8(\text{PROFR}) \end{aligned} \quad (5)$$

where RESPR is the difference between full fare and standby (or reservation price).

This model was estimated and the following model was obtained:

$U_{FULL\ FARE} - U_{STANDBY}$ = -2.247[-6.0] - 0.069 [-16.5]RESPR + 0.0211[8.79]FULL FARE + 1.178[4.35] EXPECTED DELAY + 0.9867[4.63]FLORIDA - 0.7511 [-2.81]BUSINESS + 0.5585[1.91]PROFEM + 0.0746[0.19] PROFR + 0.4332[4.10]CARS.

The numbers in brackets are t-statistics and indicate that all coefficients except the professional classes are significant at the 95 percent confidence level. In addition, all coefficients except the dummy for business have the anticipated signs. The statistics for the logit model's goodness-of-fit measures are (a) log of likelihood function = -356, (b) percent right = 88, and (c) -2 log-likelihood ratio = -1132 (with 9 df). These are comparable in some respects to the R^2- and standard-error measures from regression results.

By substituting this calculation into Equation 5, it is possible to express the probability of using full fare for any observation or prospective traveler. The estimated probabilities of choosing the full-fare option are shown in Table 1 for a range of values. The observed probabilities seem reasonable for values of the independent variables that lie near the center. This logit model breaks down, however, as RESPR goes to zero or as PNOSEAT goes to zero. Clearly, the probability of reserving should go to 1 and 0, respectively, for these two cases. That these two results cannot be achieved is due to the near-linear utility function, which does not capture the asymptotic effects as these limits are approached.

MODEL APPLICATIONS

In order to use such a model in a marketing framework, it is necessary that the model provide data for the population as a whole so that price elasticities and revenue estimates may be determined.

Direct elasticities can be calculated from the logit results by the following formula:

$$\epsilon_{ikt} = (\partial P_{it}/\partial X_{ikt})(X_{ikt}/P_{it}) = (1 - P_{it})(X_{ikt})(\beta_k) \quad (6)$$

where

ϵ_{ikt} = elasticity for person i on mode t with respect to variable k,
P_{it} = probability that i will choose t,
X_{ikt} = value of kth variable for i on mode t, and
β_k = coefficient of kth variable.

For the first example in Table 1, this means that we can predict that this person will choose full fare 35 percent of the time and that a 1 percent increase in the price differential will cause a 1.8 percent decrease in that figure [ϵ = (1 - 0.35) (40) (-0.069) = -1.8]. In this range, reservations are price elastic; that is, passengers are highly sensitive to price changes. To an airline, which presumably would be concerned about encouraging full-fare passengers if revenues less the increased costs of carrying additional passengers will not decline, it is important not to operate in the highly price-elastic range. It can be shown that RESPR = $30 implies that the probability of traveling full fare = 0.52 and therefore that ϵ = (1 - 0.52) (30) (-0.069) = -1.00. The airline can thus expect to increase revenue (from the group of passengers represented by the first row of Table 1), since it decreases the price differential toward $30.

This method should prove particularly helpful in analyzing not only fare changes, but also such LOS characteristics as minimum-stay requirements, advance-purchase requirements, and preferences for certain types of aircraft. Once these variables have entered into the utility functions and coefficients have been estimated, it is possible to measure the impact of changes in these characteristics. Even more helpful will be the ability of the analyst to identify the reactions of various market segments, which will permit the type of price discrimination necessary to induce new business without diverting revenues from full-fare passengers.

In order to project aggregate demand for a given alternative, it is necessary to have information about the potential travelers and their potential trips. The total number of trips in a market and the characteristics of trip makers must be forecast externally from these ticket-type models by using the carrier's standard methods. This process may involve sophisticated models of trip generation or may be based on something as simple as a projected market-growth trend and an assumption that the socioeconomic distribution of passengers remains unchanged. In either case, aggregation methods require that the carrier's passengers be grouped into some number of relatively homogeneous cells. The model must then be applied to each cell separately; this will forecast the ticket-type choice of its members and accumulate the aggregate shares for each alternative.

CONCLUSIONS

The above discussion has set forth a proposed method for analyzing the many new factors that affect the airline-passenger flight and ticket-selection process. The model relies on a statistical technique that is well tested in other behavioral modeling disciplines and particularly in modeling transportation mode-choice decisions.

A pilot application of the model was performed on a set of survey observations by using a two-alternative choice set--full fare or standby--but is easily extended to any number of alternatives and is adaptable to many types of distinguishing characteristics, such as booking requirements, length-of-stay requirements, and time-of-day restrictions. A complete data set for estimation could easily be obtained by using the on-board surveys made by the carriers. With the richer data base, many of the simplifying assumptions in the pilot application could be relaxed, which would provide a sound model to aid carriers in their complex marketing decisions.

REFERENCES

1. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics (P. Zarembka, ed.), Academic Press, New York, 1973.
2. M. Richards and M. Ben-Akiva. A Disaggregate Travel Demand Model. Heath, Lexington, MA, 1975.

Publication of this paper sponsored by Committee on Aviation Demand Forecasting.

Assessing the Safety and Risk of Air Traffic Control Systems: Risk Estimation from Rare Events

ALLEN C. BUSCH, BRIAN COLAMOSCA, J. STUART HUNTER, AND NEIL W. POLHEMUS

To assess the safety and risk of current and proposed air traffic control route-separation standards, it is necessary to estimate the frequency of occurrence of extremely rare events. Since direct estimates of collision risk from historical data require sample periods that are unacceptably long, alternative methods are necessary. This article describes a probabilistic model for collision risk and its use in the North Atlantic airspace; it includes a discussion of sequential testing designed to determine whether current navigational performance meets a specified target level of safety.

A problem frequently encountered by analysts who wish to determine the level of risk associated with a particular transportation system is that of estimating the frequency of rare events. Most catastrophic transportation accidents, such as the midair collision of two commercial airliners, occur so infrequently that estimates of the accident rate are difficult to obtain directly. Consequently, probabilistic models are often constructed to describe the various factors that must occur to cause an accident. Estimates of the rate of occurrence of these factors are then obtained separately and combined later in an overall-risk computation.

An example of such an indirect approach is that of collision-risk methodology, first proposed by Reich (1) to estimate the risk of midair collisions between aircraft strategically separated in the lateral, longitudinal, and vertical dimensions and subsequently applied to determine route spacing in the North Atlantic and Central East Pacific regions (2-4). Essentially, the model factors the occurrence of a collision into three events (lateral overlap, longitudinal overlap, and vertical overlap), all of which must occur simultaneously to create a collision. Since the frequency of each event is several orders of magnitude higher than the frequency of a collision, it can be estimated in a sufficiently short period of time. If we assume that the three events are independent, their probabilities can then be multiplied to estimate the probability of a collision.

This paper examines some of the important estimation problems raised in applying collision-risk