Journal of Aircraft
Vol. 45, No. 1, January–February 2008

Choice of Route Networks: A Qualitative Model for Overland and Overwater Routes

Dipasis Bhadra* and Brendan Hogan†
The MITRE Corporation, McLean, Virginia 22102

DOI: 10.2514/1.30137

In this paper, airlines and other users' route choice behaviors in the national airspace system of the United States have been examined. A binary logit regression model was used to determine the choice of routes by the users employing instrument flight rules. By categorizing routes into two types—over land and over water—and using a sample of almost 98,500 observations from the Federal Aviation Administration's Enhanced Traffic Management System, it was found that operational characteristics, such as distance, commercial flights, and altitudes of the flights are critical in determining route choice. Inflationary environment, caused by sectoral jet fuel prices and overall economy-wide prices, leads to the choice of overwater routes. Busier months as well as peak times of the day lead to the choice of overwater routes. Finally, airspace redesign tends to make overwater routes more attractive vis à vis than overland routes.

I. Introduction

THE National Airspace System (NAS) in the United States is structured primarily around a web of air transportation markets linked through a network of 519 commercial airports located in and around 369 metropolitan statistical areas (MSAs). The total number of origin–destination (O&D) markets in the NAS ranges somewhere between 36,000 and 40,000 pairs depending upon the season and economic cycles. Understanding airline network flows [see Fig. 1; enhanced traffic management system (ETMS)] [1] serving this complex web of markets is a key to understanding the tactical issues involving air traffic flows in the NAS. Furthermore, strategic planning involving future NAS infrastructure and personnel at the service delivery points (SDPs) must take into account evolution of the air traffic network over time. [At present, 519 tower facilities (118 of them are towers with radar coverage), 167 terminal radar approach control (TRACON) facilities and 21 en route centers within the United States (20 in the contiguous United States, or CONUS, plus Anchorage, Alaska) are considered to be SDPs. However, a better understanding of market networks and resultant routes will also allow one to understand the properties of the air traffic flows at the sector level, that is, the lowest form of SDPs.] Structural forces such as demography and economics, combined with market forces that include competition and airport position in the overall network, ultimately determine service to market pairs [2]. However, airspace redesigns accompanied with ever-changing forces in the marketplace (e.g., jet fuel price [3,4]) and status of aircraft equipage (e.g., particular avionics) may also have significant influence over how airlines decide to fly a particular segment where such choices are available.

U.S. airlines have been restructuring quite rigorously [5], especially over the last five years. Had it not been for the steadily increasing fuel prices since 2002 accompanied with a falling yield [6], the U.S. airlines might have turned a profit for the full year ending in 2005. Expenditures on jet fuel (23.4%) almost parallel labor costs (24.5%) in overall costs for the airlines [6]. Average domestic yield, that is, price per seat mile that a revenue passenger pays for travel is a measure of revenue strength, has been declining domestically during 2000–2005.

In this paper, the relationship between airlines' choice of routes and the factors influencing them has been explored. As a case study, the choice of overland vs overwater routes to serve O&D segments for a subset of similar NAS routes has been empirically examined. The determinants of network flows and their evolution have tactical and strategic values for the efficient allocation of NAS resources. This is especially true when these resources are severely constrained by, for example, congestion at terminals and airspace. It is the goal of this present work to shed light on these issues through an empirical analysis of routing behavior based on flight plan data. The remainder of the paper is organized as follows. Section II provides a brief background and reviews the literature. Section III lays out the data and the methodology. Section IV provides the empirical model and findings. Section V concludes the paper.

II. Background and Literature

The choice of market network influences how airlines choose their routes to serve O&D demands and connections [7]. Airlines assign their aircraft fleet by taking into account the O&D demands and fares; network characteristics of their own, codeshare partners and competitors; fleet size and composition; and fleet operating characteristics and associated costs [8,9]. These assignments are constrained by operational considerations, including crew restrictions, maintenance requirements, availability of fleet, and other requirements. A solution to the airlines' optimization problem (e.g., revenue maximization or cost minimization) routinely assigns daily flight itineraries to their optimal path [10]. The optimized aircraft fleet structures, results of fleet assignment modeling (or FAM), thus represents dynamic value chains which are positioned and sequenced through an itinerary. Combined with projected characteristics of the day (wind flow, for example), the FAM produces requested flight paths or routes that serve the optimized itinerary for an airline.

For the purposes of modeling and studying flight activity in the NAS, it is useful to understand the routing activity of NAS users. Brunetta et al. [11] have shown that the routing choices by users of European airspace vary according to many factors including time of day, day of the week, and type of operator. Specifically, they examined the 45 most congested airport pairs in Europe for a test week of data. They observed that for each of these markets there is one route that is preferred by the vast majority of users. However,

Presented as Paper 7725 at the 6th ATIO/AIAA Conference, September, 2006, Wichita, KS; received 6 March 2007; revision received 9 July 2007; accepted for publication 10 July 2007. Copyright © 2007 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. Copies of this paper may be made for personal or internal use, on condition that the copier pay the $10.00 per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923; include the code 0021-8669/08 $10.00 in correspondence with the CCC.

*Principal Economist, Center for Advanced Aviation System Development (CAASD), 7515 Colshire Avenue; dbhadra@mitre.org. Chair, Economics Technical Committee AIAA. Senior Member AIAA. (Corresponding Author).

†Senior Simulation and Modeling Engineer, Center for Advanced Aviation System Development (CAASD), 7515 Colshire Avenue. Member AIAA.

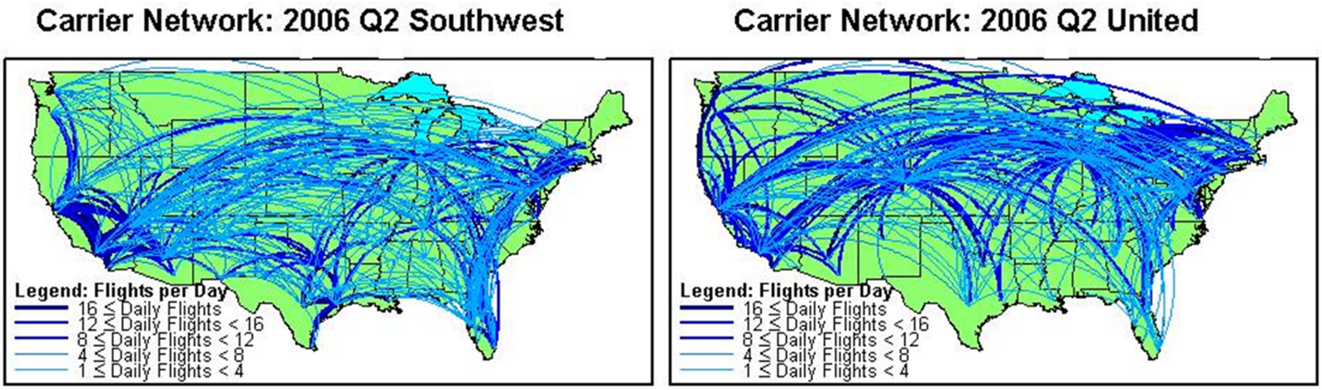

Fig. 1 Differentiated airline network.

there are variations from this trend on certain markets with respect to the time of day and day of week. For example, it was observed that one carrier serving the Dublin to Paris Charles de Gaulle market chose a specific, shorter route for each of its weekend flights, while using longer routes for each of its weekday flights. Further, a low-cost carrier in the same market chose the shorter route for just one of its three daily flights during the week, while it used the shorter option for all six of its daily flights on weekends. These choices likely have to do with the congestion users have come to expect on certain routes during specific days and times of day. However, it is interesting to note that the users' individual business models respond to this situation in slightly different ways. One of the conclusions with regard to this topic [11] was that charter operators, in general, are more creative and adaptive in their choice of routes than other user categories. In this present paper, NAS routing practices have been examined to determine if there is a similar level of variation in user behavior in the United States. In addition, the trends in the routing behavior are explored over time to gain insight into how patterns have adapted in response to rising fuel prices, airspace modifications, and other external factors [12].

To help illuminate the routing decisions made by NAS users, the current research focuses on a subset of markets that involve overwater options. In addition to the usual complexities associated with selecting a route, including time, congestion, and fuel considerations, these markets have the added financial component for airlines of extra aircraft equipage requirements necessary to fly offshore. For example, flights that travel more than 100 n miles from the nearest point of land must have enough life jackets and life rafts for every occupant as well as emergency signaling devices for the rafts and radio communications equipment capable of transmitting back to a land facility [13]. The users' incentives for providing this extra level of equipage for their aircraft are the possible time and fuel savings from flying these offshore routes.

For an illustration of the essential tradeoffs between choice of routes and cost savings, consider Fig. 2a which demonstrates three common routing options from the New York City (NYC) area to South Florida, along with an estimated distance for each. The shortest routing is to fly over land until the Carolinas and then cut directly over the ocean toward South Florida using the Atlantic Routes (ARs). Flying the ARs can usually save users about 70 n mile compared to the all-land route, which is usually a combination of the J121 and J79 airways over the southern portion of the trip. A third option shown in the figure is the water route A761, which is significantly longer than the others but can provide a faster overall trip to the destination in the event of severe congestion or weather delays on the preferred routes. For example, a 60-min departure delay out of the NYC area for the preferred route can be adequately made up by avoiding the departure delay taking A761 even though it adds 30 min of extra en route time [14]. Clearly, there are significant incentives for users in which the combined earlier arrival time and added flight distance becomes an attractive option in the face of heavy delays. The tradeoffs for the Houston–South Florida route present a similar story (Fig. 2b): a savings of 90 n mile evaluated against the cost of equipage to fly the offshore Q routes.

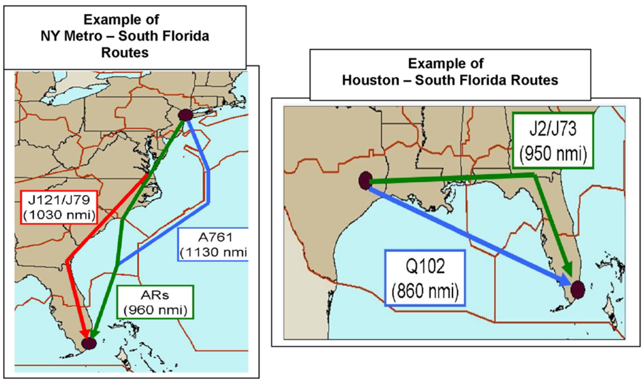

Fig. 2 Overland and overwater routes.

Congestion within the ZMA (Miami) center causes significant arrival delays at the South Florida airports, including Miami, Fort Lauderdale, and Palm Beach. The delays at Miami ripple through the entire NAS and often cause reroutes in flights to and from the Caribbean. Similarly, delays at Fort Lauderdale ripple through the NAS eastern corridors. The heavy flow of traffic between New York's John F. Kennedy and Fort Lauderdale that has been stimulated by the success of JetBlue is often impacted by delays in ZMA airspace. The Federal Aviation Administration (FAA) has traditionally dealt with these problems using the standard traffic flow management (TFM) [TFM is a process where the FAA attempts to balance the supply of scarce resources (i.e., airports and airspace capacity) to that of demand from users applying combinations of air traffic control (ATC) tools and procedures within the collaborative decision making (CDM) framework (see [15]).] initiatives, such as ground delay programs (GDPs) and ground stop (GS) programs. (Ground stops, ground delay, and airborne holding are the usual procedures of congestion management at busy airports. These procedures are exercised frequently when congestion mounts up due to weather patterns and/or some other bottlenecks.) Clearly, these are ad hoc measures and do not result in efficient allocation of terminal and airspace capacity. Observing these, the FAA implemented the Florida Airspace Optimization (FAO) in October 2005. The FAO, generally speaking, is a series of airspace modifications consisting of 1) new sectorization in the Washington en route center (ZDC) and ZMA; 2) introduction of new destination-specific overwater routes to increase north–south capacity; and, 3) new area navigation (RNAV) and standard terminal arrival routes (STARs) to eliminate complex crosses and merges at Miami, Fort Lauderdale, and Palm Beach [16].

As noted in [16], the benefits of the FAO come from delay reduction, improvements in the Caribbean routes, reduction in flight times, and the associated financial savings. A significant amount of delay reduction is due to the introduction of additional ARs. Efficacy of the ARs is best expressed in a sharp increase in usage by commercial air carriers (AC) and low-cost commercial carriers (AC:

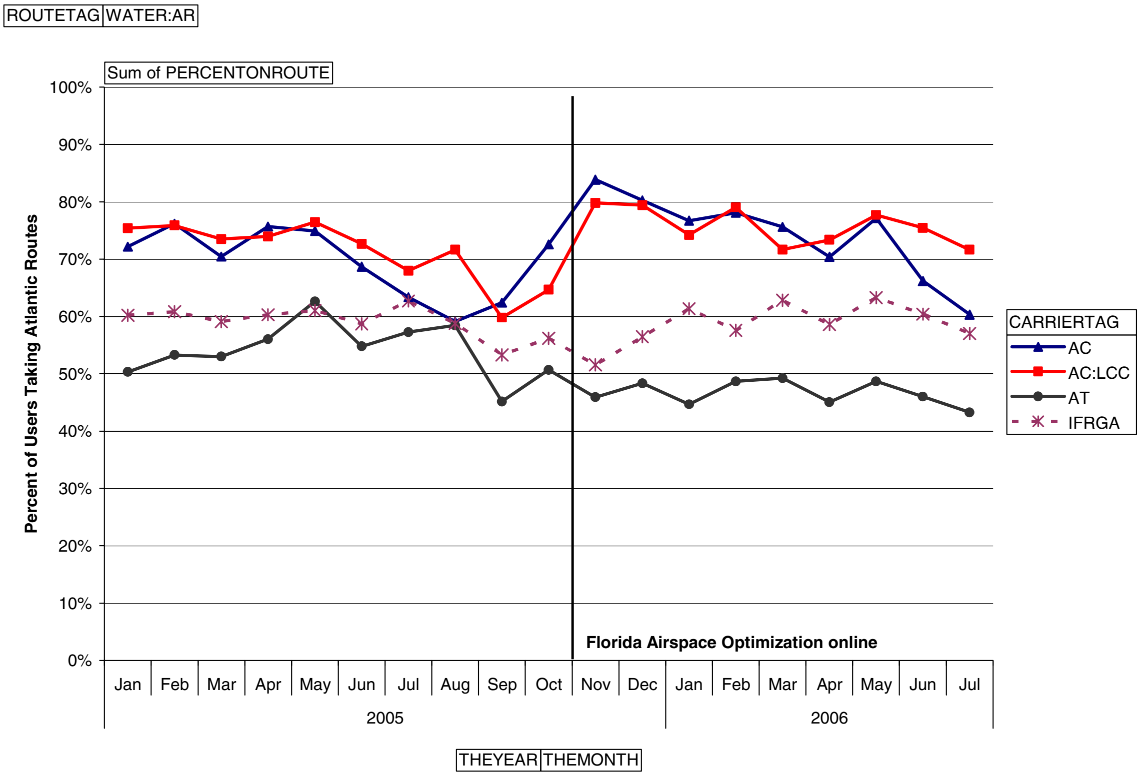

Fig. 3 Northeast to South Florida water route usage.

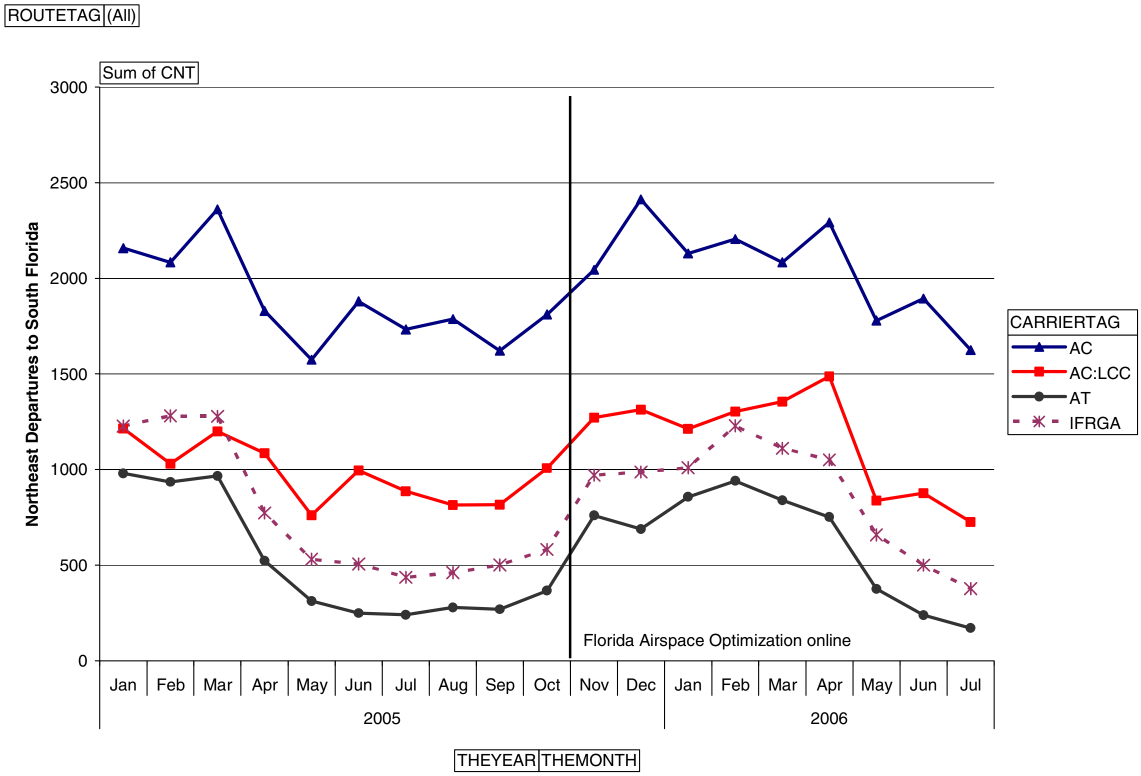

Fig. 4 Air traffic volume from Northeast states to South Florida.

LCC) following the FAO coming online at the end of October 2005 (Fig. 3). Notice, however, air taxi (AT) and general aviation (IFRGA) operations did not demonstrate the same jump in the use of offshore routes. This suggests that these user groups are less likely to have the necessary equipment on their aircraft to use the offshore routes when compared to their air carrier counterparts.

Although the overall usage of the ARs increased following the introduction of FAO, it is interesting to note the drop in AR usage by commercial carriers during the summer months. As shown in Fig. 4, the volume of flights in these markets drops during this same time period each year, but it is not immediately clear why the proportion of flights using the shorter ARs decreases as well.

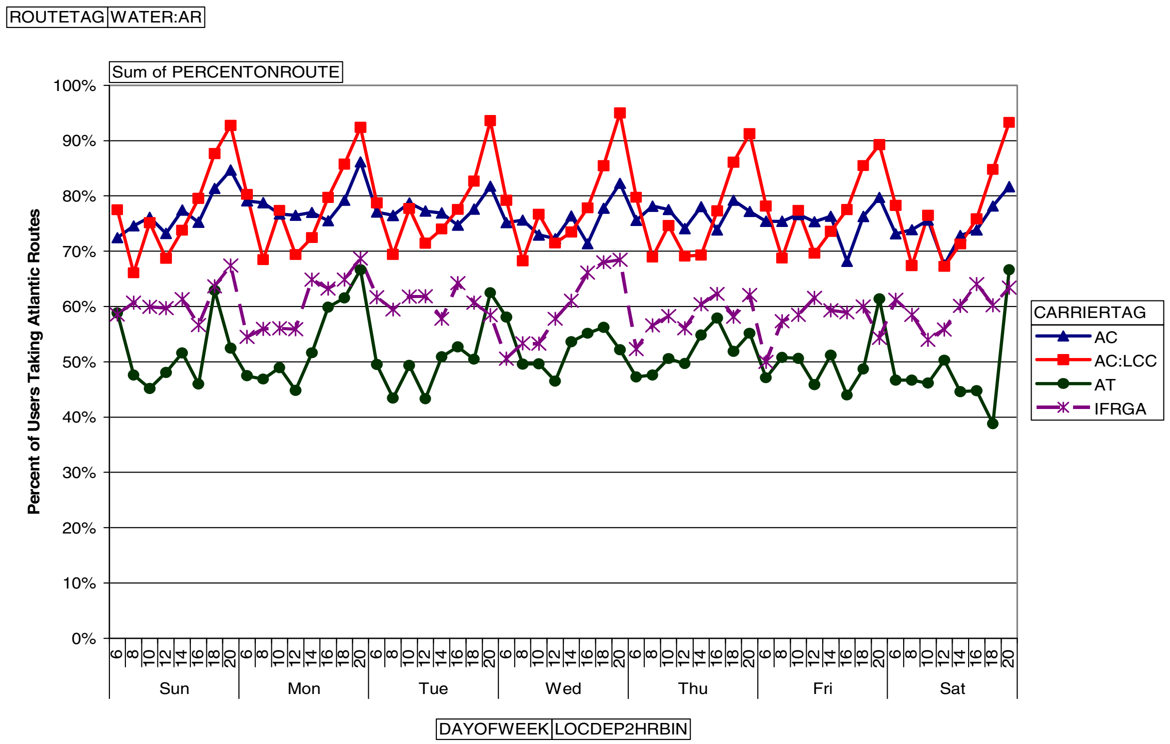

Fig. 5 Weekly pattern in Northeast to South Florida route usage.

Another level of complexity in route decision can be revealed by examining the weekly pattern in users' behavior (see Fig. 5). During the busiest late-morning to midafternoon periods (local departure hours 0800–1600 in the horizontal axis), approximately 75% of commercial carrier flights take the ARs. Interestingly, this jumps to between 80 and 90% for the late-evening flights (i.e., local departure hours 1800–2000) when the overall volume is lower. This perhaps indicates that delay and congestion are still driving factors in route choice despite the airspace improvements that have been made via FAO. However, it may be too early to draw conclusions due to the fact that there may be further adoption by carriers of new routes resulting with perhaps lower congestion and delays.

Considering both the north and the south flows along the ARs, the benefits of FAO are significant. Total annual benefit of the program has been estimated to be $18 million [16].

The previous discussion demonstrates that airlines' routing depends on many factors, including economy-wide factors influencing the choice of routes and the procedures that are being implemented. In Sec. IV, an empirical framework has been proposed to both demonstrate and quantify these relationships. Before that, Sec. III briefly describes the data and empirical methodology.

III. Data

The FAA's ETMS data [1] have been used as the basis for the quantitative analysis for this research. The ETMS data contain records of flights operated under instrument flight rules (IFR) that were processed by the host computer system and worked by FAA controllers. The ETMS flights are based on the radio detection and ranging (RADAR) data thus eliminating some potential bias in the reporting procedures followed by different air route traffic control centers (ARTCCs) that is perhaps present in air traffic activities data system (or, ATADS) data [1].

ETMS is the driver for the FAA's TFM system, and therefore, is updated with data from all IFR flights. Flight plans (FZ) in ETMS represent the sequences of planned routes as host computers receive them from airlines and carriers who submit IFR flight plans. In addition to FZ messages, arrivals (AZ), departures (DZ), radar hits (TZ), handoffs between en route centers (UZ), and route of flight (RT) messages are frequently used. Other messages, somewhat

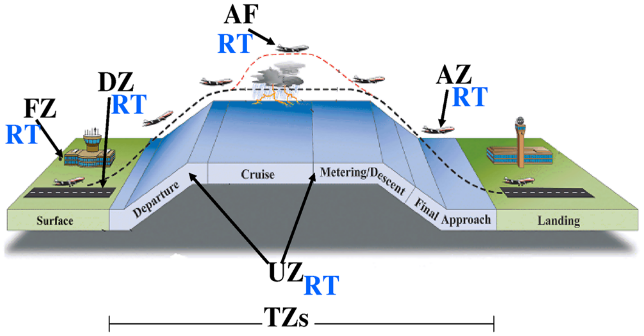

Fig. 6 Phases of flight in the NAS.

infrequently used, include flight plan amendments (AF), flight plan cancellations (RZ), and transoceanic (TO). It is evident that these messages have tremendous value, both in terms of analyzing strategic flow decisions to predict congestion on the ground and airspace, as well as unearthing tactical positions of a flight (see Fig. 6 [17]). Subsequently, RTs are often used in conjunction with DZs in deriving TFM negotiating tools whereas FZ, DZ, TZ, RT, and UZ are used to drive the traffic situation display.

As mentioned earlier in this paper, the scope of ETMS flights considered in this research are those flights that have reasonable options to either fly over water or over land to reach their destination. Specifically, flights from Texas and the Northeast states (Connecticut, Massachusetts, Maine, New Hampshire, New Jersey, New York, Rhode Island, Vermont) to South Florida are included. The airports that are considered to be South Florida for this work are as follows: Boca Raton, Palm Beach International, Ft. Lauderdale International, Ft. Lauderdale Executive, Ft. Myers Page Field, Ft. Myers International, Naples, Miami International, and Stuart.

The specific elements of ETMS data that have been studied for this work are FZs and the system generated RT messages before departure. For each flight data for departure airport, arrival airport, carrier, date, departure time of day, day of week, route, equipment, and altitude have been kept. This database was developed for all of calendar year 2005 as well as January through July 2006. For compactness and ease of analysis, individual flight records are not retained, but rather the data are aggregated as much as possible by the fields mentioned above. One complicating factor in doing this is that the route string in ETMS flight plan records often contains subtle differences that are irrelevant to the goals of this study, but yet get recorded as distinct routes due to the minor discrepancies in the text string. Examples of this are the alternating use/absence of the prefix "K" for U.S. airports (e.g., KATL vs ATL), the numeric digit following standard instrument departure (SIDS)/STARS, the duplicate string when an airport code is followed by its very high frequency omnidirectional radio range (VOR) of the same name, and a four character coordination fix time appended to the end of a route string. Differences in the characteristics mentioned here are not to be considered distinct routes and a "soundex" algorithm has been used to ignore these minor variations. (Soundex is a phonetic algorithm designed to return a like result for homophones. In this case we use a custom algorithm to return a like text string for routes that would be considered identical if read by a pilot or controller but contain subtle character differences such as those mentioned in the text.)

In addition to the flight plan fields described previously, a couple of additional fields have been derived to add value to the database. These include tags for the type of route and type of flight operator, as well as the distance along the route (as calculated iteratively from waypoint to waypoint). The route tag that has been used indicates whether it was a land or water route, and if it is a water route, an identifier was included to distinguish Atlantic Route (AR), the severe weather playbook A761, or a Gulf of Mexico Q route. This determination was made based on searching the route string for key airways corresponding to each of the categories. The flight operator tag was developed to give an indication of the high level categories of AC, AC:LCC, AT, or general aviation (GA). This determination was made based on the Federal Aviation Regulations (FAR) operating part number that each specific carrier code is registered under. For this purpose the air carriers are considered to be parts 121, 121/135, and 129, the air taxis part 135, and general aviation parts 91, 133, 137, and 125. The designation of low-cost carrier is reserved for the six operators—Southwest, Air Tran, JetBlue, Spirit, ATA, and Frontier.

These operational data were then combined with relevant economic variables, such as price (in cents/gallon) of jet fuel (or kerosene oil), an overall measure of inflationary environment as represented by the index of consumer prices (1982–84 = 100), and commercial airlines' demand (millions of barrels/day) for jet fuel. These data come from the U.S. Government's Energy Information Agency [4].

IV. Analytical Framework

It is evident that route choice reflects current users' behavior and affects future user behavior via its impact on congestion and delays. To capture the mechanics of this tradeoff, let us assume for simplicity that a representative or a *generic* user chooses only two classes of routes ($k = i, j$). [In the present context, this can be used to represent any carrier choosing routes, either commercial scheduled and/or other IFR users. For our demonstration here (see [18] for use in the context of aircraft choice by airlines), we do not make any distinction in the objective function of these users.] Then, the individual user's profit maximization (π_k) can be stated as

$$\pi_k = p_k \mathrm{pax} - c_k - R \tag{1}$$

where p_k is the average price per passenger on route k, pax is the number of passengers on board, c_k is the variable cost of the employing route, and R captures a host of other factors considered fixed, for example, stage length, types of network, origin and destination airports, day of the week, time of day, and season.

The first order condition for maximization (with respect to route choice) yields

$$p(\mathrm{pax})'_k - c'_k = 0 \tag{2}$$

In general, a particular route should be employed to the point where marginal revenue [first term on the left-hand side of Eq. (2)] equals marginal cost (constant for both choices). This condition can then be used to derive the optimal demands for routes i and j. For route i to be chosen, the indirect profit from i (Π_i) must be greater than the indirect profit from j (Π_j), or

$$\Pi_i(p_i c'_k; R) > \Pi_j(p_j c'_k; R) \tag{3}$$

and vice versa. The choice will find an equilibrium for the user at the margin when $\Pi_i = \Pi_j$.

The problem of choosing one route over another [Eq. (3)] can be transformed into a marginal framework by specifying the discrete choice probabilities as the dependent variable [18,19]. Assume that the profits in the ith and jth route can be best approximated by the following function linearly linking (via coefficient a) to exogenous variables, X:

$$y_i = \sum_{k=1}^{K} a_{ki} X_k + v_i \tag{4}$$

and

$$y_j = \sum_{k=1}^{K} a_{kj} X_k + v_j \tag{5}$$

$i = i,\ j$ choices of routes, $k = 1 \cdots K$ exogenous variables, and v represents error terms.

Let us define a difference in profits between two routes (Y^*) as the following:

$$Y^*_{ij} = y_i - y_j \tag{6}$$

or

$$= \sum_{k=1}^{K} (a_{ki} - a_{kj}) X_k + (v_i - v_j)$$

which can be written as

$$Y^*_{ij} = \sum_{k=1}^{K} b_{ik} X_k - \mu_i$$

when$Y^*_{ij} > 0$ (implying $\mu_i < \sum b_{ik} X_k$) if route i is chosen and $Y^*_{ij} < 0$ (implying $\mu_i < \sum b_{ik} X_k$) if route j is chosen. This allows us to write the probabilistic statement as the following:

$$P(i)(Y_{ij} = i) = P_i(Y^*_{ij} > 0) = P\left(\mu_i \sum_{k=1}^{K} b_{ik} X_k\right) \tag{7}$$

If μ_i is a continuous variable and has logistic distribution and Y_i is binary, then the above choice can be written as a binary logit model; if Y_i has multiple categories (i.e., polytomous), then the above equation can be written as a multinomial logit model. Although water routes can be further subdivided (see Fig. 2a), the route choices are restricted to over land and over water in this paper.

A. Binary Logit

A linear probability model does not guarantee the predicted values of the choice to lie between (0, 1). Thus, it requires a process of translating the values of the attribute X (i.e., vector containing explanatory variables explaining the choice) to a probability which ranges in value from 0 to 1. Furthermore, the property that such a transformation would allow increases (decreases) in X to be associated with an increase (decrease) in the dependent choice variable for all values of X should also be maintained. Together, these requirements suggest the use of the cumulative probability function (F). F is defined as having its value equal to the probability that an observed value of a variable (for every X) will be less than or equal to a particular X. The range of F is then (0, 1) because all probabilities lie between 0 and 1. The resulting probability distribution $P(i)$ may be expressed as follows:

$$P(i) = F(\alpha + \beta X_i) = F(Z_i) \tag{8}$$

where α and β are the parameters of the model and F represents the distribution. Common models in this category include probit (standard normal), logit (logistic), and gompit (extreme value) specifications for the F function. The two cumulative probability functions, the normal (probit) and the logistic, have been used widely in the literature and among practitioners [20,21].

To specify the binary logit, let Y represent a dummy variable that equals 1 when the route choice 1 is chosen and 0 when the other category is chosen. [Instead of strictly defining one category of route, one can also put all others in one category. For example, the choice of one category (J routes in Fig. 2) and all others (i.e., all non-J routes from Fig. 2) can be defined under this binary choice.] Then assume that each individual choice Z_i^* represents the critical cutoff value which translates the underlying index into a choice decision, such as

$$\text{individual choice for} \quad \begin{array}{l} \text{route} = 1 = \text{overwater if } Z_i > Z_i^* \\ \\ \text{route} = 0 = \text{overland if } Z_i \leq Z_i^* \end{array} \tag{9}$$

In this case, the threshold is set to zero, but the choice of a threshold value is irrelevant as long as a constant term is included in X_i. The logit model assumes that Z_i^* is a cumulative distribution function for the logistic distribution, so that the probability that Z_i^* is less than (or equal to) Z_i can be computed from the probability distribution function. The standardized cumulative probability distribution function for the logistic distribution is written as

$$\begin{aligned} P(i)(y_i = 1 \mid x_i, \beta) &= e^{\mu\beta' x_{in}} / (e^{\mu\beta' x_{in}} + e^{\mu\beta' x_{jn}}) \\ &= 1/(1 + e^{\mu\beta'(x_{il} - x_{jn})}) \end{aligned} \tag{10}$$

where x_{in} and x_{jn} are vectors describing the attributes (n dimensions) of alternatives i (route = 1) and j (route = 0); μ is a scale parameter that is positive in value. When the parameters of the Z_i are linear, the parameter μ cannot be distinguished from the overall scale of the βs [22]. Oftentimes, μ is assumed to be equal to 1. By construction, the variable P_i will lie between (0,1). P_i is the probability that an event occurs, that is, probability of the choice of category 1 (or, J routes).

The empirical framework can be specified as follows:

$$\begin{aligned} P_i(y_i = j \mid x_i, \beta) \quad &= \alpha_{ij} + \beta_1(\text{distroute}) + \beta_2(\text{carrierID}) \\ (j = 1, 2) \quad &+ \beta_3(\text{alt}) + \beta_4(\text{JKTCUUS}) + \beta_5(\text{JFFAAUS}) \\ &+ \beta_6(\text{CICPIUS}) + \beta_7(\text{gaspintcpi}) \\ &+ \beta_8(\text{weekdayeffect}) \\ &+ \beta_9(\text{monthdummy}) + \beta_{10}(\text{FAO}) + \varepsilon_i \end{aligned} \tag{11}$$

where distroute = distance in nautical miles between flight segments, carrierID is a dummy variable assuming 0 (when IFR flights are not commercial scheduled airlines) and 1 (for commercial airlines, both low cost and others); alt is cruise altitude in feet, JKTCUUS is the price of jet fuel (or kerosene oil) in cents/gallon; JFFAAUS is a commercial airlines' demand for jet fuel, CICPIUS is the index of consumer prices (1982–1984 = 100) used by the Energy Information Agency [4] for a proxy of overall inflation in the economy, gaspintcpi is a variable that represents the interactions between jet fuel price and overall consumers' price index (i.e., JKTCUUS × CICPIUS); weekdayeffect is a dummy representing weekdays (=1 when days of the week are Monday–Friday) and weekend (=0 if Saturday and Sunday); and monthdummy is a dummy representing months of the year (=1 during May–September; and =0 in all other months). Finally, FAO representing a dummy signifying when FAO went into effect (=1 if months are after October 2005 and =0 if not). The ε_i term represents the residual errors.

B. Empirical Results

The entire data set is split between overland (one-third) and overwater routes (two-thirds). The total numbers of these routes are 98,489 (Table 1). These data were combined with the jet fuel and other price and usage information to estimate and quantify the binary logit model using the Statistical Analysis System, SAS© [21].

Table 1 Overland and overwater route choices

Ordered value	RouteTagNum	Frequency
Overland routes	0	31,684
Overwater routes	1	66,805
Total observations, N		98,489

Table 2 Estimated model of route choices

Logistic regression of route choice The LOGISTIC procedure *Testing global null hypothesis: $\beta = 0$*			
Test	Chi square	DOF	Pr > ChiSq
Likelihood ratio	16,993.57	10	<0.0001
Score	15,661.81	10	<0.0001
Wald	13,223.50	10	<0.0001
Type 3 analysis of effects			
Effect	DOF	Wald chi square	Pr > ChiSq
distroute	1	10,162.13	<0.0001
carrierID	1	4,600.71	<0.0001
alt	1	1,546.05	<0.0001
JKTCUUS	1	9.38	0.0022
JFFAAUS	1	13.79	0.0002
CICPIUS	1	3.63	0.0567
gasintcpi	1	10.50	0.0012
weekdayeffect	1	6.74	0.0094
monthdummy	1	17.81	<0.0001
FAO	1	42.12	<0.0001

The essential results of the binary choice model of overland routes (choice = 0) and overwater routes (choice = 1) are presented in Table 2. As evident, the selected model that is expressed in terms of exogenous variables presented in Eq. (11) and Table 2 is highly significant. The overall statistics of the logit model as captured by the likelihood ratio and Wald statistics are highly significant implying that the choice of the model is statistically significant.

Furthermore, the Wald chi-square test of statistical relevance indicates that all exogenous variables are statistically significant. (In effect, this is a likelihood ratio test for the null hypothesis that all coefficients are zero; the test for individual coefficients is called Wald statistics. This is easily computed by dividing estimated coefficients by their standard error and squaring the result [21].) As expected, the distance of the routes (distroute), carrierID dummy, and altitude all play significant roles in deciding between overland and overwater routes. The FAO plays a significant role as well, and so do the weekdays and month dummies.

Interestingly, however, the choice of routes depends on economy-wide factors, such as jet fuel and broad prices. For example, jet fuel prices (JKTCUUS), consumers' price index (CICPIUS) (somewhat less statistically significant), and an interaction between the two (gaspintcpi) play important roles in determining the choice between overland and overwater routes. Although some other efforts by airlines to increase fuel efficiencies have been noted (see [3] for a chronicle), the effect of jet fuel prices on route choices, to our knowledge, has not been researched or documented in the past. These choices, controlled for other crucial variables such as routes, distance, etc., can contribute significantly to airlines' fuel efficiency. Finally, overall sectoral activities, as captured by JFFAAUS, are also important factors contributing to the choice of route.

Table 3 reports the model significance of the exogenous variables. The variable distroute is a significant variable that negatively impacts the choice from overland (base or reference choice) to overwater routes. In other words, as distance increases between market segments, it is likely that users will choose fewer overwater routes and more overland routes. An explanation for this is that the distance advantage of overwater routes, as a fraction of the total trip distance,

Table 3 Estimated parameters of route choices in the NAS: analysis of maximum likelihood estimates

Parameter	Routetagnum	DOF	Estimate	Standard error	Wald chi square	$Pr >$ ChiSq
Intercept	1	1	−0.66	3.50	0.04	0.85
distroute	1	1	−0.01	0.00	10162.13	<0.0001
carrierID	1	1	1.17	0.02	4600.71	<0.0001
alt	1	1	0.01	0.00	1546.05	<0.0001
JKTCUUS	1	1	0.06	0.02	9.38	0.00
JFFAAUS	1	1	0.57	0.15	13.79	0.00
CICPIUS	1	1	3.44	1.81	3.63	0.06
gasintcpi	1	1	−0.03	0.01	10.50	0.00
weekdayeffect	1	1	0.04	0.02	6.74	0.01
monthdummy	1	1	0.13	0.03	17.81	<0.0001
FAO	1	1	0.40	0.06	42.12	<0.0001

Table 4 Route specific performance of the estimated model

	Perfect matching	Not matching	Total predicted observations
Overland routes	10,535	7,323	17,858
Overwater routes	59,482	21,149	80,631

is less noticeable for longer markets which may tilt the preference toward land routes in this sample.

Second, increased presence of commercial flights (i.e., carrierID = 1 if commercial scheduled flights; 0 if not) makes it likely that overwater routes will be chosen. The water routes can only be used when aircraft are equipped with proper avionics, safety jackets, and life rafts on board. Passenger aircraft flying commercial scheduled flights are more likely to have these on board, and hence, likely to fly overwater as opposed to overland routes. This confirms our initial observations (see Fig. 2) that air taxis and general aviation do not take advantage of the FAO new overwater routes at as high a rate as scheduled air carriers. Third, altitude (alt) matters in the choice; the higher the altitude, the more likely overwater routes will be chosen. This seems intuitive in that the offshore airspace should have less complexity than the overland routes and users would be more likely to be cleared for optimal higher cruise altitudes.

Fourth, higher jet fuel prices (JKTCUUS) make it more likely that overwater routes will be chosen. So does the overall consumers' price index (CICPIUS). In other words, a general inflationary environment makes overwater routes more common as a choice than over land. Faced with increasing jet fuel prices, for example, airlines have to balance the tradeoffs between additional costs of flying over water (i.e., complying with safety rules and proper avionics) evaluated against the potential benefit (e.g., reduced delays, and in many cases, reduced distances) of choosing overwater routes vs equivalent cost and benefit of flying over land. Results indicate that higher jet fuel prices indeed result in more commonly choosing overwater routes than overland routes by airlines and other users. Fifth, the larger the sectoral consumption for jet fuel (JFFAAUS), the more likely it is that overwater routes will be chosen. This result is perhaps indicative of the fact that overwater routes are a small percentage of the total traffic flow, and hence, factors driving overland routes are better explained by commercial airlines' consumption of jet fuel. Sixth, interactions between jet fuel prices and consumers' price index tend to influence overland route choice.

Seventh, both weekday (=0 if weekend, and =1 when weekdays) and months dummy (May–September = 1) tend to a higher likelihood for overwater route choices. This finding suggests that summer months tend to be important for overwater route choice even though these are not the months when traffic peaks for the routes in the data set (see Figs. 4 and 5). In other words, something else other than traffic, likely convective weather (that corresponds with these months), perhaps tends to explain overwater choice.

Finally, airspace redesigning influences route choice significantly. As represented by the FAO procedure that went into effect at the end of October 2005, it was found that the dummy variable representing that cutoff time played a key role in the choice of overwater versus overland routes. As the economy-wide pressure increases from rising jet fuel prices, lower fares, and relatively few opportunities for restructuring left, it is anticipated that airlines and other users will increasingly turn their attention to airspace redesign and request improvements optimizing routes as the way to optimize on their fuel use, and hence, further cost cutting [23].

As indicated elsewhere [24], properties of the qualitative choice models are somewhat different and not so straightforward compared to linear regression models. Despite these difficulties, the estimated parameters from Table 3 can be used to compute the probability of two choices. Overall predicted choices can then be compared to the observed choices. Figure 7 reports these results. Notice, by construction, if one choice is predicted below the observed responses (e.g., predicted overland choice captures 56% of the actual observations), then the other choice will be overestimated (i.e., 120% in the case of overwater). In total, they add up to the total number of observations.

However, a decomposition of the choices (Table 4) reveals that the model captures over one-third of the total observed overland choices (i.e., 10,535 of the total of 31,684 overland routes) while predicting almost 90% (59,482 of the total 66,805) of overwater choices

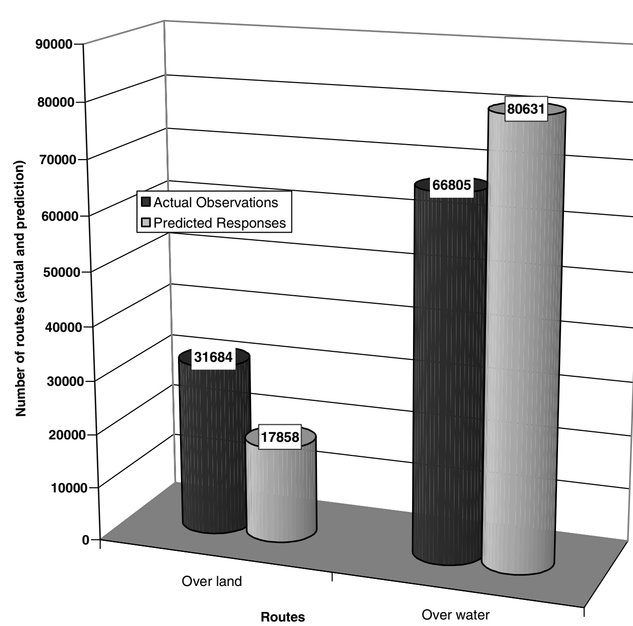

Fig. 7 Predicted choices versus actual observations.

correctly. From the predicted routes (Table 4), however, the model matches perfectly almost 60% (i.e., 10,535 of the total 17,858) for overland routes while overwater routes matched approximately 74%. Looked at this way, the estimated model performs fairly well.

There are quite a few areas of future research. Most importantly, an extension of this research can be made to include other overland and overwater routes in the NAS in order to reach generalized conclusions. In addition, the choices can be understood further, and perhaps better, if aircraft information is used. Aircraft identifications from the data can be used to estimate the magnitude of the economic benefit of these choices. This research can be used by calculating the benefit of overwater route choices against overland routes. Finally, geospatial locations of airports play critical roles in determining the route choices. Any future research should also give special attention to this aspect of route choice. These are some of the remaining tasks for any future research.

V. Conclusions

In this paper, airlines and other users' route choice behaviors for a subset of routes in the NAS have been examined. A binary logit regression model was used to determine the choice of routes by the IFR users for traffic between northeastern states and Texas to South Florida. By categorizing routes into two broad types, over land and over water, it was found that distance, commercial flights, and altitudes of the flights are important variables behind route decisions. Furthermore, inflationary environment, due to sectoral jet fuel prices or due to overall economy-wide price pressure, makes the choice of overwater routes more likely. Busier months as well as peak times of the day lead to the choice of overwater routes. Finally, airspace redesign as captured by the representative Florida airspace optimization tends to make overwater routes more likely vis à vis overland routes.

Acknowledgments

D. Bhadra and B. Hogan would like to thank Gerald Dorfman, Debra Pool, and George Solomos of the MITRE/CAASD (Center for Advanced Aviation System Development), and three anonymous referees for their suggestions and insightful comments leading to substantial improvement in the paper.

References

[1] Federal Aviation Administration. "FAA Operations and Performance Data," U.S. Department of Transportation, Washington, D.C., 2007, http://www.apo.data.faa.gov/ [retrieved 5 July 2007].

[2] Bhadra, D., and Hechtman, D., "Determinants of Airport Hubbing in the U.S.: An Econometric Framework," *Public Works Management and Policy*, Vol. 9, No. 1, July 2004, pp. 26–50. doi:10.1177/1087724X04265082

[3] Air Transport Association. "Fuel-Related Challenges Facing U.S. Airlines," http://www.airlines.org/economics/energy [retrieved 5 July 2007].

[4] Energy Information Agency, U.S. Department of Energy, http://www.eia.doe.gov/ [retrieved 5 July 2007].

[5] Gittel, J. H., von Nordenflycht, A., and Kochan, T. A., "Mutual Gains or Zero Sum? Labor Relations and Firm Performance in the Airline Industry," *Industrial and Labor Relations Review*, Vol. 57, No. 2, 2004, pp. 163–180.

[6] Air Transport Association. "Fare and Finance," http://www.airlines.org/economics/finance [retrieved 5 July 2006].

[7] Bhadra, D., "Demand for Air Travel in the United States: Bottom-Up Econometric Estimation and Implications for O&D Forecasts by O&D Pairs," *Journal of Air Transportation*, Vol. 8, No. 2, Sept. 2003, pp. 19–56.

[8] Lederer, P. J., and Nambimadom, R. S., "Airline Network Design," *Operations Research*, Vol. 46, No. 6, Nov.–Dec. 1998, pp. 785–804.

[9] Bhadra, D., Gentry, J., Hogan B., and Wells M. T., "Future Air Traffic Timetable Estimator," *Journal of Aircraft*, Vol. 2, No. 42, April 2005, pp. 320–328.

[10] Lohatepanont, M., and Barnhart C., "Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment," *Transportation Science*, Vol. 38, No. 1, Feb. 2004, pp. 19–32. doi:10.1287/trsc.1030.0026

[11] Brunetta, L., Righi, L., Andreatta, G., and Odoni, A., "Analysis of Alternative Routes for Selected City Pairs," Tech. Rept. EUROCONTROL Innovative Route Charging Schemes Work Package 6 Pt. II, Dec. 2003.

[12] Heimlich, J., "Fuel Challenges Facing U.S. Airlines," *Global Aviation Fuel*, June 2006, pp. 28–31.

[13] U.S. Government Printing Office, Electronic Code of Federal Regulations, "Title 14: Aeronautics and Space; Subchapter F—Air Traffic and General Operating Rules," http://www.gpoaccess.gov/ecfr [retrieved 10 Aug. 2006].

[14] Damato, J., "Reduce East Coast Departure Delays: Fly the Offshore Radar Routes," NBAA presentation, 29 Sept. 2005, http://web.nbaa.org/public/ops/airspace/fao/OffshoreRadarRoutesNEW.ppt [retrieved 9 Aug. 2006].

[15] Wanke, C., Callaham, M., Greenbaum, D., and Masalonis, A., "Measuring Uncertainty in Airspace Demand Predictions for Traffic Flow Management Applications," AIAA Paper 2003-5708, 11–14 Aug. 2003.

[16] Federal Aviation Administration. "New FAA Airspace Plan Speeds Florida Flights," Southern region press release, http://www.faa.gov/news/press_releases/news_story.cfm?newsId=6406 [retrieved 5 July 2007].

[17] Bhadra, D., "Air Traffic Performance by Market Segments," *Journal of the Transportation Research Forum*, Vol. 46, No. 2, 2007, pp. 47–62.

[18] Bhadra, D., "Choice of Aircraft Fleets in the U.S. Domestic Scheduled Air Transportation System: Findings from a Multinomial Logit Analysis," *Journal of the Transportation Research Forum*, Vol. 44, No. 3, 2005, pp. 143–162.

[19] McFadden, D., "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, edited by P. Zarembka, Academic Press, New York, 1974, pp. 105–142.

[20] Hosner, D. W., and Lemeshow, S., *Applied Logistic Regression*, Wiley, New York, 2nd ed., 2000.

[21] Allison, P. D., *Logistic Regression Using the SAS System: Theory and Application*, SAS Press, Cary, NC, 2001.

[22] Ben-Akiva, M., and Lerman, S. R., *Discrete Choice Analysis: Theory and Applications to Travel Demand*, MIT Press, Cambridge, MA, 1984.

[23] Bisignani, G., "Speech at the Regional ANSP Conference," AirServices Australia, Aug. 18, http://www.iata.org/pressroom/speeches/2006-08-18-01.htm [retrieved 13 Sept. 2006].

[24] Pindyck, R. S., and Rubinfeld, D. L., *Econometric Models and Economic Forecasts*, 3rd ed., McGraw–Hill, New York, 1991.