



A passenger demand model for air transportation in a hub-and-spoke network

Chieh-Yu Hsiao ^{a,b,*}, Mark Hansen ^a

^a Department of Civil and Environmental Engineering, Institute of Transportation Studies, National Center of Excellence of Aviation Operations Research (NEXTOR), University of California at Berkeley, Berkeley, CA 94720, USA

^b Department of Transportation Technology and Management, National Chiao-Tung University, 1001 University Road, Hsinchu 300, Taiwan, ROC

ARTICLE INFO

Article history:

Received 8 September 2010

Received in revised form 14 April 2011

Accepted 28 April 2011

Keywords:

Demand model

Route choice

Instrumental variables

Value of travel time

Panel data

ABSTRACT

This paper develops an air passenger model that deals with city-pair demand generation and demand assignment in a single framework. Using publicly available and regularly collected panel data, the model captures both time series and cross-sectional variation of air travel demand. The empirical analysis finds that pattern of correlations among alternatives can be described by a three-level nested logit model. Fare, frequency, flight time, direct routing, on-time performance, income, and market distance have significant effects on air demand. Correcting for the problem of endogenous air fares using instrumental variables yields more plausible estimates of price sensitivity and value of time.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

This paper presents a route-level air travel demand model for the US domestic airline network. Given supply characteristics on routes and regional demand-side variables, the model predicts passenger traffic for individual routes between specific airport pairs. It is based on random utility theory, in which route demand is generated from choices of whether or not to travel, what airports to fly from and to, what kind of route—direct or connecting—to use, and—if the route is connecting—what hub to fly through.

The model incorporates several advances from earlier work in this area. By incorporating the choice of whether or not to travel by air, it incorporates demand generation as well demand allocation, the traditional focus of random utility models. The model is based on publicly available and regularly updated data. Through use of an instrumental variable technique, it overcomes the problem of fare endogeneity that has bedeviled previous efforts to use these data for demand modeling. It includes on-time performance metrics as well as more traditional variables such as fare, frequency, and travel time.

Applications of the model are legion. It can be used with existing forecasts of future flight schedules, used for planning and investment analysis by FAA and NASA, to produce compatible forecasts of air passenger flows, which are currently lacking. Impacts of fare changes resulting from fuel price escalation or changes in aviation tax structures can be assessed. The model can be used to assess the effect of airport congestion on air traveler behavior and the resulting impact of traveler economic welfare. Finally, because the model is estimated on data streams that extend back many years, and are expected to continue into the future, it allows retrospective assessment of structural changes over time, and can be easily updated.

* Corresponding author at: Department of Transportation Technology and Management, National Chiao-Tung University, 1001 University Road, Hsinchu 300, Taiwan, ROC. Tel.: +886 3 513 1576; fax: +886 3 572 0844.

E-mail address: cyhsiao@nctu.edu.tw (C.-Y. Hsiao).

Our main purpose in this paper is to present the basic model, the data and methodology for estimating it, and key estimation results. More detailed treatment of the various applications appears in forthcoming articles. After a brief literature review in Section 2, Section 3 presents the theoretical model. Data and estimation procedures are discussed in Section 4, while Section 5 discusses estimation results and their implications. Section 6 offers conclusions and discusses prospects for future model enhancements.

2. Literature review

While air travel demand modeling have been the subject of considerable research, virtually all of the models are concerned with either the quantity of air traffic or the assignment of traffic, but not both. This limits model validity, applicability, and utility.

2.1. Demand generation model

We will refer to models of air travel quantity as demand generation models. The literature contains many such models. Units of observation include regions, airports, airlines, flight segments, city-pairs, airport-pairs, county-pairs, and country-pairs. Price, travel time, and flight frequency or schedule delay are typical supply-side variables in these models, while population, income, distance, and various measures of attraction (e.g. dummy variables for tourist destinations) are used to characterize the demand-side.

None of the demand generation models adequately deals with the availability of alternate routes for air travel between a given origin and destination. Models of segment traffic, such as [Abrahams \(1983\)](#), [Anderson and Kraus \(1981\)](#), and [Wei and Hansen \(2006\)](#), overlook network effects, such as availability of alternative routes and characteristics of complementary segments. In some cases ([Ippolito, 1981](#)), the data set is intentionally restricted to routes that are (in Ippolito's words), "more or less insulated." This may increase model validity, but at the expense of wide applicability.

The availability of multiple routes is also a problem for city-pair models, including those of [Kanafani and Fan \(1974\)](#), [De Vany and Garges \(1972\)](#), and [Bhadra \(2003\)](#). These models predict total traffic in city-pair markets. Although this traffic is normally divided among several routes, these models require a single set of representative supply-side variables. For example, the first two of the above works use for the travel time the lowest value among the alternatives, while Bhadra employs the average fare across all travelers in the city-pair as the price variable. It is easy to see circumstances in which these use of such variables can lead to misleading results—for example if the lowest travel time alternative also featured a very high fare or low frequency.

2.2. Demand assignment models

Demand assignment models explain the distribution of traffic—or the choice of individual travelers—among alternative modes, airports, routes, airlines, or other dimensions. Literature on such models has burgeoned in recent years, with development paralleling that of random utility models generally. Multinomial logit (MNL), nested logit (NL), mixed multinomial logit (MMNL) models, and specialized variants of these have all been applied.

Airport choice is one of the most widely studied topics in this literature. [Harvey \(1987\)](#), [Hansen \(1995\)](#), and [Windle and Dresner \(1995\)](#) all employ MNL models to analyze traveler choice of airport in multi-airport regions. NL models of airport-airline and airport-access mode choice have been developed by [Pels et al. \(2001, 2003\)](#). [Hess and Polak \(2005a,b\)](#) and [Pathomsiri and Haghani \(2005\)](#) have estimated MMNL models of airport choice. These models are based on airport passenger surveys, with access time, flight frequency, and fare the mainstay explanatory variables. Of these, fare has proven the most problematic, because of the multiplicity of fares available and the difficulty of determining the fares faced by individual choice makers. This has led to the omission of fares in some studies, and the use of average fares in others. [Pathomsiri and Haghani \(2005\)](#) note that the use of average fare often results in an insignificant or counterintuitive coefficient estimate. This may be the result of endogeneity bias. Since airline pricing and yield management systems often result in higher average fares on more popular routes, demand estimations that ignore simultaneity of supply and demand systems may give erroneous results. The estimated fare coefficient may also be affected by omitted service attributes that passengers value. Therefore, both simultaneity and omitted variables may lead the estimated coefficients that are biased upward (i.e., toward 0).

Route demand assignment models explain the market shares of routes serving the same O–D airport-pair or O–D city-pair. Route demand assignment model for city-pairs combine the airport demand assignment for multiple airport regions and the route demand assignment for airport-pairs. Earlier models of this type, such as [Kanafani and Fan \(1974\)](#) and [Kanafani et al. \(1977\)](#), considered settings in which non-stop routes are the dominant alternatives, such as San Francisco–Los Angeles. These are essentially airport-pair choice models.

Many airport-pair route assignment models have been developed, often as part of a supply–demand model that also predicts supply side behavior. Examples of MNL route assignment models include [Kanafani and Ghobrial \(1985\)](#), [Hansen \(1990, 1995\)](#), [Hansen and Kanafani \(1990\)](#), [Ghobrial and Kanafani \(1995\)](#) and [Adler \(2001, 2005\)](#). The NL model is also sometimes applied (e.g. [Weidner \(1996\)](#), and [Hsiao and Hansen \(2005\)](#)), with elemental alternatives nested according to routing type (direct/connecting). Other NL models, such as [Coldren and Koppelman \(2005\)](#), consider choice of both route and carrier, with

the latter used to define the nests. MMNL models of route/carrier choice, often using a combination of revealed and stated preference data, have also been developed (Coldren et al. (2003), Adler et al. (2005), Warburg et al. (2006)).

In addition to model type, route assignment models differ with respect to the type of data used. Many employ aggregate data available from the US DOT 10% sample of air passenger itineraries. These data are comprehensive, readily available, and have been collected in a consistent manner for several decades. On the other hand, they contain no information about the traveler or non-airline portions of the trip. Models that consider such factors must be based on more specialized surveys, which are necessarily limited in temporal and spatial scope. While they afford a richer depiction of air traveler behavior within their domain, their generalizability is open to question. Limited sample sizes may also reduce the statistical reliability of models based on such data.

Most demand assignment models are concerned with individual choices or aggregate market shares of air travelers. Such models are completely distinct from demand generation models. A few demand assignment models incorporate demand generation as well. The MNL model of Adler (2001, 2005) incorporates the alternative of not traveling by air. The model is not estimated, and the likely correlation between the airline alternatives relative to the no-travel alternative is not addressed. There are also a number of intercity mode choice models in which air is one alternative. These generate air travel demand by attracting intercity travelers from other modes (Trani et al., 2003), and in some cases other destinations (Morrison and Winston, 1985). The mode choice models do not consider assignment beyond mode, and consider the overall volume of intercity travel to be fixed. They are, strictly speaking, neither demand generation nor demand assignment models as those terms are used here, but have some commonality with both.

3. The theoretical model

3.1. Conceptual framework

This research models city-pair air passenger demand at the route level.¹ In general, potential trips between two cities are derived from the socioeconomic activities in both cities. Potential travelers may have many choices regarding these potential trips. They may avoid air travel altogether by choosing different modes, such as auto and rail, or they may decide not to travel at all. Within the air mode, they may select different routes, of which airports and segments (non-stop links) are basic elements.

The general form of city-pair air passenger demand model is given by Eq. (1). The air traffic on a route is equal to the product of the market (city-pair) saturated demand and the market share of this route. The market saturated demand (or total potential demand) can be modeled as a function of socioeconomic and geographic characteristics of this market, such as populations of the origin and destination cities, or distance. The route market share is determined by a function of the vector of socioeconomic characteristics of this route, and supply characteristics for this route, its competing routes, and the “outside good.”

$$Q_{rt} = T_{m(r)t} \cdot MS_{rt} = T(D'_{m(r)t}) \cdot MS(D_{rt}, S_{rt}, S_{-rt}, S_{0t}) \quad (1)$$

where Q_{rt} is the air traffic on route r at time t ; $T_{m(r)t}$ is the saturated demand of the market (city-pair) m , served by route r , at time t ; MS_{rt} is a market share of route r at time t ; $T(\cdot)$ and $MS(\cdot)$ are a saturated demand function and a market share function, respectively; $D'_{m(r)t}$ is a market-specific (city-pair-specific) socioeconomic and geographic characteristic vector of market m , served by route r , at time t ; D_{rt} is a route-specific socioeconomic and geographic characteristic vector of route r at time t ; S_{rt} is a supply characteristic vector of route r at time t ; S_{-rt} contains the supply characteristic vectors of route r 's competitors at time t ; S_{0t} is a supply characteristic vector of the “outside good” 0 at time t .

In Eq. (1), $D'_{m(r)t}$ and D_{rt} include socioeconomic and geographic variables that respectively influence the size of the market and the share of traffic choosing a particular route. Typical variables used in the literature are population, income, employment of cities (metropolitan areas) defining the market, and distance. The market share variation of alternatives in a market is mainly explained by supply characteristics of these alternatives (S_{rt} , S_{-rt} , and S_{0t}). In other words, the market share of a route depends on attractiveness of its characteristics, compared to those of other routes and the “outside good” in the same market. Market characteristics can also affect the attractiveness of air routes compared to the non-air alternative. In long-haul markets, for example, the non-air alternative may be less attractive because there is less competition from other modes. Since airports and segments are basic elements of a route, supply characteristic vectors of a route should be composed of characteristics of the route, and of the airports and segments involved. Thus S_{rt} can be decomposed into three parts: $S_{rt} = \{S'_{rt}, S_{a(r)t}, S_{g(r)t}\}$, where $S_{a(r)t}$ and $S_{g(r)t}$ are characteristic vectors of the airports and the segment(s) served by route r at time t , respectively; S'_{rt} is a pure route characteristic vector of route r at time t . Typical supply characteristic variables include: air fare, travel time, and routing types (pure route variables), ground access time and airport delay (airport variables), and flight frequency (a segment variable).

¹ This conceptual model can be easily applied to the route-carrier level—simply differentiating routes by carriers. However, adding the carrier dimension may lead to a more complicated empirical model. In addition, while a route-carrier model might explain more details of air passenger demand than a route model, it is less likely to be applied to forecasting due to uncertainty of future carriers.

3.2. Saturated demand function

The saturated demand function defines the relationship between the total potential demand of markets and certain causal factors. Estimating this function is not straightforward because only the realized traffic, rather than the “potential” traffic, can be observed. Empirical demand studies in other sectors have employed two approaches. The first and more common approach is to base a saturation level on a socioeconomic variable. One might, for example, assume $T_{m(r)t} = \alpha * M_{m(r)t}$, where α is a proportionality factor and $M_{m(r)t}$ is the socioeconomic variable, such as population. The main advantage of this approach is its simplicity. However, in order to provide convincing results, justification and coefficient sensitivity tests for this assumption are needed. The second approach is to estimate a model for this function (e.g. estimate the parameter α). Because the saturated demand is a part of the whole demand model and the “potential” traffic cannot be observed, estimating the saturated demand model is more complicated. System equations and/or additional assumptions to simplify the estimation may be used in this approach. For example, Hansen (1996), and Wei and Hansen (2005) assumed that the total demand is much more than the total traffic in a market, and then separated the estimation of the saturated demand model from that of the whole demand model.

We employ the first approach here. While simple, it can be shown²—at least for the multinomial logit and nested logit model forms—that the proportionality factor setting may only affect the estimated intercept of the market share model if the proportionality factor is set large enough. Since the intercept is not the main coefficient of interest, this approach should work well. In addition, socioeconomic variables in the market share model (D_{rt}) can help to explain the market share difference between all routes and the outside good. Thus, the impacts of choosing an inappropriate parameter (e.g. α) and socioeconomic variables for $D'_{m(r)t}$ are reduced.

3.3. Market share function

Our market share function is based on random utility theory. The indirect utility of potential traveler i from route r at time t can be formulated as Eq. (2),

$$u_{irt} = \sum_{k=1}^K \beta_k x_{rtk} + \xi_{rt} + \mu_{irt} + \varepsilon_{irt}, \quad (2)$$

where x_{rtk} is an observable characteristic k of route r at time t , i.e., an observable supply characteristic vector S_{rt} ; β_k is a parameter to be estimated for characteristic k ; ξ_{rt} is a term to capture unobservable route characteristics at time t ; μ_{irt} is a term to capture individual taste deviations, which can be modeled as a function of individual characteristics and route characteristics; ε_{irt} is a stochastic term.

Assuming³ that every potential traveler chooses the alternative⁴ that gives the highest utility from all alternatives, and that ties occur with zero probability, the market share of route r at time t as a function of the characteristics of all alternatives competing in the market is given by integrating the population distribution functions of unobserved variables over the range of unobserved variables that induces the choice of route r at time t . An operational market share function needs to make assumptions on the population distribution functions, and then the integral can be calculated. Different assumptions on the population distribution functions lead to different discrete choice models. Three models—MNL, NL, and MMNL—are discussed below.

The most common and simple model assumes that (1) potential travelers are homogeneous so that $\mu_{irt} = 0$; and (2) the ε_{irt} 's are independent and identically distributed (i.i.d.) across travelers, routes, and time with a type I extreme value distribution. This leads to the MNL. If we normalize the utility from the outside good alternative to zero ($\sum_{k=1}^K \beta_k x_{0tk} + \xi_{0t} = 0$), the market share of route r at time t is

$$MS_{rt} = \frac{\exp\left(\sum_{k=1}^K \beta_k x_{rtk} + \xi_{rt}\right)}{1 + \sum_{j \in R(m(r)t)} \exp\left(\sum_{k=1}^K \beta_k x_{jtk} + \xi_{jt}\right)}, \quad (3)$$

where $R(m(r)t)$ represents all routes available in the market served by route r at time t .

The NL model gives more flexible substitution patterns than the MNL model and still keeps the computational simplicity and tractability of the MNL model. The correlations of the stochastic terms in the NL model are specified by a variance component structure, instead of assuming that the stochastic terms are i.i.d. As a result, an alternative is more likely to substitute for an alternative in the same nest, than for an alternative in different nest. The nesting structure has to be determined. In our route choice model, since different routes of a market may share the same airports and/or segments, routes can be grouped

² Refer to Hsiao (2008) for more details.

³ Further discussions and formulas about these assumptions can be found in the discrete choice literature (e.g. McFadden (1981)) and its applications, such as Berry et al. (1995) and Nevo (2001).

⁴ While this assumption may be unrealistic for analyzing some products, it is easier to justify in a route choice model, since for each realized trip a traveler always uses only one route.

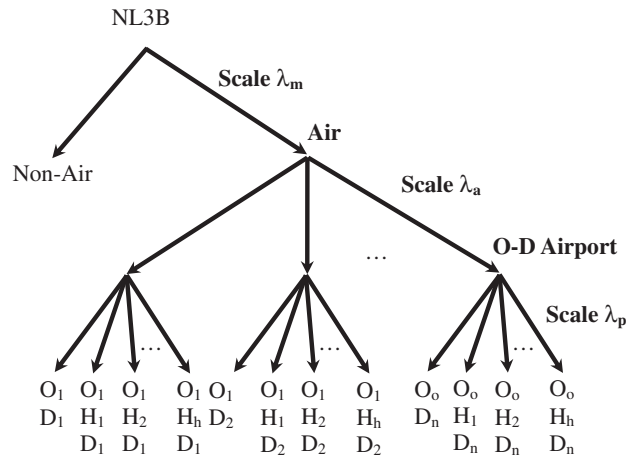


Fig. 1. Nesting structure—three-level nested logit (NL3B).

by their common characteristics. Although this provides a priori information on the possible nest structure, the final nesting structure needs to be determined empirically as discussed in Section 4.

The MMNL model provides the most flexible substitution patterns among these three models, but also has the greatest computational complexity because the integral defining market shares of the MMNL model cannot be computed analytically. The MMNL model allows individual heterogeneity ($\mu_{irt} \neq 0$), i.e., potential travelers may have different preferences for route characteristics. The individual deviations (μ_{irt}) can be modeled as a function of individual characteristics and route characteristics. Detailed information about the aggregate MMNL model can be found in Berry et al. (1995) and Nevo (2001), among others.

4. The operational model

4.1. Model forms and nesting structures

This research employs the aggregate NL form for the market share function, and also estimates the aggregate MNL model for purposes of comparison. The empirical objective of this research focuses on the coefficients and ratios of coefficients, and the NL model can serve this purpose well.⁵ Moreover, the NL model provides a good balance between flexibility and computational complexity. For the nesting structures of the models, routes are grouped in a nest by assuming that the routes with more common characteristics are more likely to be closer competitors. The common characteristics used in the empirical analysis include (1) air routes or the non-air alternative, (2) origin–destination (O–D) airport pair, and (3) routing type (direct or connecting route). Based on different combinations of these characteristics, five nesting structures are examined—including one MNL, one two-level NL (NL2), two three-level NL (NL3A and NL3B), and one four-level NL (NL4) model. The NL4 model differentiates alternatives by characteristics (1), (2), and (3), from top to bottom, The NL3B model distinguishes alternatives by characteristics (1) and (2), as shown in Fig. 1, while the NL3A model uses (1) and (3). The NL2 model only considers characteristics (1). The estimated ratio(s) of scale parameters of an NL model can be used to determine whether the nested logit model is consistent with utility-maximizing behavior,⁶ and whether the higher-level NL model collapses to a lower-level NL (or MNL) model.

An air route in Fig. 1 is presented by its origin airport (O), destination airport (D), and connecting (hub) airport (H), if applicable. For example, in the city-pair market O–D, the route O_1D_1 is the direct route from the origin airport 1 to the destination airport 2. We only consider routes with at most one connecting airport as alternatives. Thus, there is only one H for each connecting route. Removing routes with more than one connection makes the models more tractable with little loss of generality, since the vast majority of US domestic trips involve less than two connections.

4.2. Specification

According to the proposed demand model, Eq. (1), route demand is determined by functions of socioeconomic and supply characteristic vectors. Referring to (1), we use origin and destination populations for $D'_{m(r)}$, i.e., the basis for estimating mar-

⁵ For instance, Brownstone and Train (1999) mentioned that “If indeed the ratios of coefficients are adequately captured by a standard logit model, as our results and those of Bhat (1996) and Train (1998) indicate, then the extra difficulty of estimating a mixed logit or a probit need not be incurred when the goal is simply estimation of willingness to pay, without using the model for forecasting.”

⁶ Refer to Train (2003) and Ben-Akiva and Lerman (1985) for details.

ker saturated demand. We use origin and destination incomes for D_{rt} (so $D_{rt} = D_{m(r)t}$). Finally, for supply variables S_{rt} (and S_{-rt}) we use air fare, scheduled flight time, flight frequency, on-time performance, market distance, and routing type (direct or connecting).

4.2.1. Market saturated demand

This research assumes a maximum number of potential trips in a market based on population. Simply put, we assume that the more people that could travel in a city-pair market, the more people that will travel. The potential number of trips for a city-pair m at time t is specified as a function of the city-pair population, Eq. (4).

$$T_{m(r)t} = \alpha * M_{m(r)t} = \alpha * \text{Population}_{m(r)t} \quad (4)$$

where α is the proportionality factor; $M_{m(r)t}$ is the observable socioeconomic variable chosen for reflecting the potential total traffic; $\text{Population}_{m(r)t}$ is the geometric mean of populations of the city-pair m served by route r at time t ; for each city, the population of the metropolitan area served by an airport or an airport system is used.

The proportionality factor α is set to be 10 per quarter; i.e. we assume that every unit of population may make as many as 10 trips per quarter. Ten is a large number of potential trips for intercity travel. The real number of air trips is, of course, much smaller than this potential. Sensitivity tests for this setting are performed to check the robustness of the model parameters to the assumed α value.

4.2.2. Income

Income is used to capture the economic activities that generate air travel demand and potential travelers' purchasing power. Both economic activity and purchasing power are expected to have positive impacts on air travel demand. Thus, higher income level is expected to generate more air trips. The geometric mean⁷ of incomes of two cities is used as an explanatory variable for the city-pair demand. For each city, the income variable is measured by the per capita personal income (in constant dollars, based on the 4th quarter of 2004) of the metropolitan area served by an airport or an airport system.

4.2.3. Price

For our price variable we use the average fare paid by passengers traveling on a route in 2004 (the 4th quarter) constant dollars. As mentioned in Section 2.2, the air fare variable may be endogenous, because of supply and demand simultaneity and/or omitted variables. As a result, the fare coefficient estimated by ordinary least squares (OLS) method is likely biased upward—toward zero. The inferred fare elasticities and the value-of-time may therefore be underestimated and overestimated, respectively. This research applies the instrumental variables (IV) estimation⁸ to solve the endogeneity problem.

Although the access costs may also affect travelers' decisions on routes, particularly for the airport choice in multiple airport systems, this research does not explicitly specify the access cost variables in the model mainly due to the data availability. The effects of access costs are partially captured by the airport dummy variables. In addition, applying the IV to air fare should eliminate the impact on the fare coefficient from omitting access cost variables.

4.2.4. Scheduled flight time

Among air alternatives in a market, a route with longer scheduled flight time is expected to be less attractive, other factors being equal. The scheduled flight time variable for an air route is defined as the sum of gate-to-gate scheduled time of flight segments of the route. The gate-to-gate scheduled time of a segment is determined by averaging over scheduled flights on the segment, using individual flight data from FAA's Airline Service Quality Performance (ASQP) database (FAA, 2007).

4.2.5. Flight frequency

The greater the number of flights, the more convenient traveling between two cities is. From the viewpoint of travel time, higher flight frequency reduces the time difference between desired and actual schedule arrival/departure time, also known as schedule delay. In addition, higher frequency is more likely to keep a traveler close to his or her original schedule when unexpected events, such as flight cancellations and delays, happen.

As suggested by Hansen (1990), flight frequency is taken in logarithmic form for two reasons. First, marginal effects of flight frequency on route utilities are expected to be diminishing with increasing number of flights. Second, a route alternative can be considered as an aggregation of individual flights, and frequency is therefore a measure of the size of the route alternative. The logarithmic form is the most suitable⁹ for a characteristic that captures the size of an aggregated alternative.

Since flight frequency is a segment characteristic, a route utility function may include several frequency variables. This research specifies three frequency variables—one for direct routes and two¹⁰ for connecting routes. We differentiate the frequency effects for connecting routes by taking maximal and minimal numbers of flights on two segments. The hypothesis is that the minimum frequency is more critical to the connecting service, and thus a given fractional flight frequency increase on the

⁷ Employing the geometric mean implies that market demand of a city-pair is not affected by income of one city if the other city has zero income.

⁸ Refer to Section 4.4 for details.

⁹ Refer to Ben-Akiva and Lerman (1985) for more details.

¹⁰ This research discards routes with three or more segments, which carry about 5% of passengers, to simplify the analysis. Thus, every connecting route has two segments.

segment with lower frequency should increase service attractiveness more than an equivalent change on the segment with higher frequency.

4.2.6. On-time performance

Whereas travelers accept most characteristics of the service (e.g. fare and scheduled travel time) before their trips, on-time performance is realized during the trip, and thus becomes an important determinant of travelers' ultimate satisfaction¹¹. Better on-time performance may thereby attract more traffic to the route in the future.

We use airport "average delay per flight"¹² to capture on-time performance effects. Although average delay by segment better reflects on-time performance of a route, potential travelers are more likely to have on-time performance information for airports than segments.¹³ Moreover, there are "negative delay" cases, in which flights arrive or depart early, whose effects we also investigate. Average positive and negative delays are calculated by separating early and late flights. The hypothesis is that negative delay should have a smaller effect on demand than positive delay.

Since potential travelers do not know their flight delays when they choose their routes, they may consider expected flight delay as one of the service characteristics. This research uses flight delays of previous period(s) to capture this phenomenon. More precisely, the hypothesis is that potential travelers make decisions based on recent information—defined as one and four quarters (subscripted as $t - 1$ and $t - 4$, respectively) before the quarter of the observation (subscripted as t). While the $t - 1$ term provides the most recent on-time performance information, the $t - 4$ term captures the most recent results for the same season.

4.2.7. Routing type

The more connections required by a route, the lower its convenience. Thus, potential travelers usually prefer direct routes over connecting routes, all else equal. This research, since it considers only direct and one-connection routes, employs a dummy variable to capture connection disutility.

4.2.8. Market distance

Market distance may affect potential travelers in two ways: mode choice and propensity to travel. Since the model includes a non-air alternative, mode competition should be taken into account in order to estimate the total market share of air routes in a market. Travelers are more likely to choose air service over slower surface modes in long-haul markets than in short-haul markets. The marginal effect of distance on mode choice is expected to attenuate at longer distances, once air travel has become the dominant mode. Distance also effects the overall proclivity toward travel in a city-pair market. As suggested by the literature on transportation geography, interactions between cities are likely to diminish with distance, and with that the propensity to travel. This offsets the mode choice effects. The net effect is indeterminate at shorter distances and is likely to be concave because of the mode choice component. We therefore employ a specification that allows for this.

4.2.9. Other factors

In addition to the above causal factors, this research specifies several sets of dummy variables to capture unobserved fixed effects. The first set of dummy variables is for connecting (hub) airports. Twenty-nine dummy variables are used—one for each of the 30 benchmark airports,¹⁴ except for Tampa International Airport (TPA) which is used as the benchmark airport. Another set of dummy variables captures fixed effects of origin and destination airports. They are only specified for airports in multiple airport systems because potential travelers do not have a chance to choose among terminal airports in single airport systems. This set of dummy variables may capture, for instance, differences in airport accessibility. The third set of dummy variables captures seasonal and yearly fixed effects. People may be more (or less) likely to travel in certain seasons or years, for reasons not captured by socioeconomic variables in our model. For example, after 9/11, people curtailed air travel because of security concerns as well as the increased hassle of more stringent screening.

4.3. Data

To estimate the model, this research compiles a panel data set that includes variables for major US domestic routes over 40 quarters—all quarters between year 1995 and 2004. The raw data is from five sources. Fare and route information is extracted from Hub, which contains US data from DOT's Airline Origin and Destination Survey (DB1B), a 10% sample of airline

¹¹ For example, Ross and Swain (2007) argued that "industry surveys consistently identify departure punctuality as a key determinant of consumer satisfaction, especially on shorter flights."

¹² One may calculate the total delay time by summing the time differences between actual and schedule time for all flights, or for all delayed flights, defined by a delay threshold. This research chooses the first approach.

¹³ Although delay statistics by airline, by airport, and even by flight number are all available in the United States, no delay statistics by segment are directly available for potential travelers. Even though potential travelers may find the percentage of on-time of a flight (not a segment) on the Internet when they book, the percentage cannot reflect the delay level of the flight because the same percentages of on-time flights may represent significantly different delay levels.

¹⁴ To simplify the empirical work, this research only includes direct routes and routes connecting at one of the 30 benchmark airports in the sample. Refer to Federal Aviation Administration (2001) for these benchmark airports. The Federal Aviation Administration developed capacity benchmarks for 31 of the busiest airports in 2001. Since this research only considers the trips in the continental United States, the Honolulu International airport (HNL) is removed from the connecting airport list.

tickets, “cleaned” through cross-checking with other data sources (Data Base Products, 2004a, 2005a). Flight frequency data is extracted from Onboard Domestic (Data Base Products, 2004b, 2005b), a commercial product based on US DOT’s T-100 database. Scheduled flight time and on-time performance variables are calculated from FAA’s Airline Service Quality Performance (ASQP) database (FAA, 2007). Income and population by metropolitan are downloaded from the Regional Economic Information System, US Bureau of Economic Analysis (2006). Unit jet fuel cost of US domestic operations, used in our instrument for average fare, is calculated from Air Transport Association’s Fuel Cost and Consumption Report (2005).

In order to simplify the empirical work and/or ensure reliable data, the raw data is filtered by several rules.¹⁵ After the data were filtered, 1660,569 route-quarter observations—including 96 thousand direct route-quarters and 1.56 million connecting route-quarters—remained to estimate the model. The 1660,569 route-quarter observations corresponded to 213,917 market-quarters, 76,629 routes, and 6133 markets. In addition, it is necessary to associate airports with metropolitan regions since the model predicts travel between regions rather than specific airports. This research follows the definition of a MAS proposed by Hansen and Weidner (1995).¹⁶

4.4. Model estimation

The basic strategy for estimating aggregate logit models is to transform market share functions and then estimate parameters by linear regression. For MNL models, the market share of route r at time t is given by Eq. (3). The difference between natural logarithms of market shares of two alternatives (r and r') is described as Eq. (5). Regressing the left hand side of the equation on differences of explanatory variables gives estimates of the parameters of interest (β_k 's).

$$\ln(MS_{rt}) - \ln(MS_{r't}) = \sum_{k=1}^K \beta_k (X_{rtk} - X_{r'tk}) + (\zeta_{rt} - \zeta_{r't}) \quad (5)$$

Alternative-pairs need to be determined before running the regression. One simple way is to use the outside good (non-air) alternative of which utility is normalized to zero as the base alternative (r') for every route. As a result, Eq. (5) can be simplified to Eq. (6), in which there is no need to differentiate explanatory variables. Another way is to pick an alternative randomly as the base alternative (r') for other alternatives.

$$\ln(MS_{rt}) - \ln(MS_{0t}) = \sum_{k=1}^K \beta_k X_{rtk} + \zeta_{rt} \quad (6)$$

For NL models, estimations become more complicated. One possible solution is to derive an equation, which is similar to Eq. (6) but adding conditional market share term(s) and its (their) coefficient(s), for each nesting structure.¹⁷ However, while this approach seems to provide a convenient way to estimate parameters, additional exogenous variables are required since the conditional market shares are endogenous. This research does not choose this approach because finding valid instrumental variables (IVs) becomes harder as the number of endogenous variables requiring them increases.

This research sequentially estimates NL models by decomposing NL models into MNL models. More precisely, a nested logit model is estimated by nest and from bottom level to top level. Within a nest, an MNL model is estimated by applying Eq. (5),¹⁸ in which the base alternative is randomly picked. Each level (except for the level involving the fare variable, in which the method of two stage least squares is used), is estimated by OLS and then the inclusive value(s) of nest(s) at this level are calculated. Inclusive values of nests of a lower level are added into a higher level as an explanatory variable, the coefficient of which is the ratio of scale parameters. When estimating the NL models the utility of the non-air alternative is normalized to zero, and the scale parameters of the bottom nests are set to one.

This research applies the instrumental variables method to solve the endogeneity problem of air fare. The instrumental variable for air fare is defined as the product of the route distance and unit jet fuel cost (in 2004 dollars per gallon). This variable captures the cost of offering the service, and thus affects the price of the service, but is expected to have no direct impact on market shares. Since this research applies Eq. (5), in which all variables are differences in attribute levels between two alternatives, the difference in distance-fuel cost product between two routes is used as the IV for their fare difference.

¹⁵ For simplicity, only domestic itineraries that (1) are with one or two coupons, (2) are between top 100 origin and destination airports, and (3) either direct routes or connecting at 30 benchmark airports are included in the sample. Some routes are discarded because of their unreasonable average yields or low frequency. In addition, while calculating on-time performance from ASQP database, some flights are not included because their records are considered to be outliers. Refer to Hsiao (2008) for more details.

¹⁶ They defined a MAS using two criteria: airports operating in a metropolitan area and existing competition for local passengers. However, some airports are not in the sample due to their low traffic. This affects the definition of MASs used in this research: a MAS may involve fewer airports or become a single airport system. Refer to Hsiao (2008) for the complete list of MASs.

¹⁷ Refer to Hsiao (2008) for more information about this approach.

¹⁸ Since the models are estimated by applying Eq. (5), all fixed effects with the same values for all alternatives in a nest are differentiated out. Thus, the estimates implicitly take these effects into account.

Table 1
Estimation results.

Variable	(1) MNL-OLS	(2) MNL-IV	(3) NL3B-OLS	(4) NL3B-IV
Fare (hundreds of 2004 dollars)	−0.178*** [0.005]	−1.410*** [0.411]	−0.160*** [0.005]	−1.546*** [0.206]
ln(Frequency)—Direct (flights per quarter)	1.282*** [0.015]	1.212*** [0.030]	1.337*** [0.016]	1.240*** [0.029]
ln(Max frequency of two segments)—Connecting (flights per quarter)	0.408*** [0.010]	0.501*** [0.034]	0.440*** [0.009]	0.627*** [0.030]
ln(Min frequency of two segments)—Connecting (flights per quarter)	0.793*** [0.007]	0.883*** [0.035]	0.822*** [0.007]	0.957*** [0.023]
Scheduled flight time—Direct (minutes)	−0.018*** [0.000]	−0.005 [0.004]	−0.019*** [0.000]	−0.004 [0.002]
Scheduled flight time—Connecting (minutes)	−0.018*** [0.000]	−0.008* [0.004]	−0.019*** [0.000]	−0.006** [0.002]
Dummy for direct routes (=1, if direct route)	3.353*** [0.145]	4.477*** [0.421]	3.874*** [0.141]	6.066*** [0.397]
Positive hub arrival delay _{t-1} (minutes per flight)	−0.004*** [0.001]	−0.008*** [0.002]	−0.002** [0.001]	−0.006*** [0.002]
Positive hub arrival delay _{t-4} (minutes per flight)	−0.004*** [0.001]	−0.012*** [0.003]	−0.002** [0.001]	−0.007*** [0.002]
Inclusive value of level 3 (parameter= λ_p/λ_a)			0.937*** [0.011]	0.664*** [0.014]
Inclusive value of level 2 (parameter= λ_a/λ_m)			0.711*** [0.009]	0.795*** [0.010]
Inclusive value of level 2 *market distance			−0.008*** [0.001]	−0.012*** [0.001]
Market distance (hundreds of miles)	0.148*** [0.005]	0.150*** [0.010]	0.018*** [0.005]	−0.024*** [0.005]
ln(market distance)	1.261*** [0.042]	0.844*** [0.145]	1.888*** [0.048]	1.575*** [0.046]
Per capita personal income of market (thousands of 2004 dollars)	−0.665*** [0.006]	−0.637*** [0.012]	0.015*** [0.003]	0.038*** [0.003]
Constant (level 1)	0.003 [0.003]	−0.001 [0.006]	−17.316*** [0.116]	−16.229*** [0.102]

Notes: (1) Standard errors in brackets are robust to heteroskedasticity, serial correlation and market cluster effects; (2) All regressions include hub dummy variables for connecting routes, origin and destination airport dummy variables for MASs, and year and quarter dummy variables for time fixed effects; (3) MNL models are estimated by Eq. (5), in which the base alternative is randomly picked.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

4.5. Estimation results

Because lower-level nested logit and multinomial logit models are special cases of higher-level nested logit models, this research estimates proposed nesting structures from higher-level to lower-level nested logit models, including multinomial logit models, until a model that is consistent with utility maximization is found. The NL4 and NL3A estimation results are not consistent with utility maximization. The NL3B model, shown in Fig. 1, is found to be the highest-level NL model that is consistent with utility maximization.¹⁹ Thus, the NL3B models, estimated by OLS and IV methods, are summarized in Table 1. The MNL models with the same explanatory variables are also presented for comparison.

As shown in Table 1, most coefficients of explanatory variables are statistically significant and have expected signs. Although all estimated fare coefficients illustrate negative fare impacts on demand, the fare coefficients from IV estimates are more reasonable. This can be seen from their inferred values of travel time (VOTs). Recall that when air fare is endogenous,²⁰ its coefficient estimated by OLS is more likely biased toward zero and thus the inferred VOTs are overestimated. As shown in Table 2, estimates from OLS method—column (1) and (3)—give unreasonable high VOTs, especially for values of scheduled flight time: all the inferred values of scheduled flight time are greater than \$614 per hour (39 times larger than the median wage rate of 2004). While literature on transportation economics suggests a wide range of VOTs,²¹ inferred VOTs from OLS estimates are still out of these ranges. In contrast, fare coefficients from IV estimations are larger (in absolute values) than those

¹⁹ The consistency is determined by the estimated ratio(s) of scale parameters of these models. Although the estimated ratios of scale parameters are different for different specifications, the conclusions of the consistency are the same under different experiments of specifications.

²⁰ Tests for endogeneity of air fare based on the proposed instrumental variable appear that air fare is endogenous for different specifications.

²¹ For example, Small and Winston (1999) summarized estimates of value of time by transportation mode. The range, for different modes and trip types, is from 6 to 273% of wage rate. They also described that air travelers have a very high VOT—the VOT for air travelers for vacation trips is 149% of wage rate, estimated by Morrison and Winston (1985).

Table 2
Inferred values of travel time.

Time type	(1) MNL-OLS	(2) MNL-IV	(3) NL3B-OLS	(4) NL3B-IV
Scheduled flight time—Direct	614.4 (3852%)	21.3 (134%)	721.7 (4525%)	16.8 (105%)
Scheduled flight time—Connecting	623.8 (3911%)	32.9 (206%)	705.5 (4423%)	24.1 (151%)
Positive hub arrival delay _{t-1}	124.4 (780%)	33.6 (210%)	63.7 (399%)	22.5 (141%)
Positive hub arrival delay _{t-4}	138.2 (867%)	49.3 (309%)	68.8 (431%)	27.6 (173%)

Notes: (1) Units of VOTs: dollars per hour in 2004 dollars; (2) VOTs as percentages of wage rate are shown in parentheses. The US median wage rate of 2004—\$15.96 per hour (Bureau of Labor Statistics, 2008)—is used to calculate these percentages.

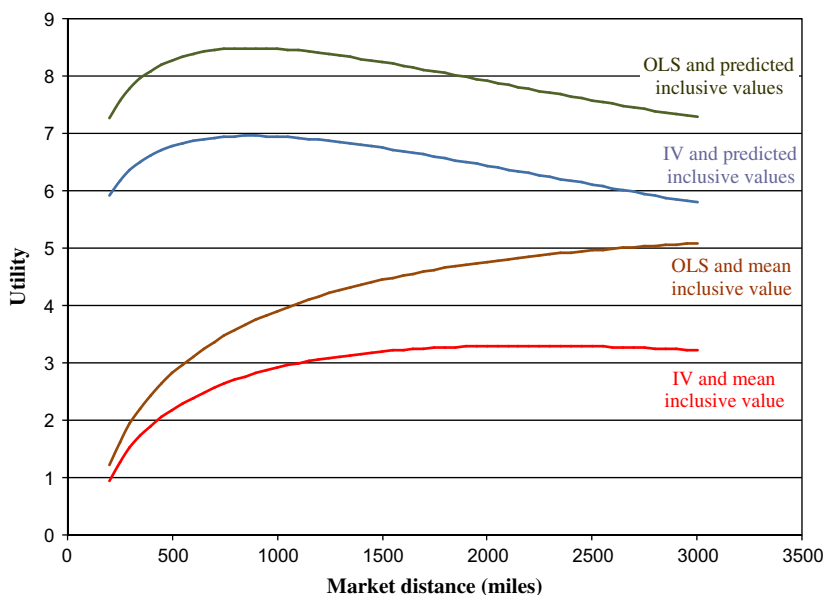


Fig. 2. Market distance effects.

from OLS estimations, and provide sensible VOTs—at least in the same order as those reported in the literature. For example, the value of scheduled flight time of direct routes, given by the preferred model—column (4)—is \$16.8 per hour (105% of wage rate).

The estimated frequency coefficients confirm the hypothesis that the minimum frequency is more critical to the connecting service, and thus a proportional flight frequency increase on the segment with lower frequency increases service attractiveness more than an equivalent change on higher frequency segment. Although all coefficients of scheduled flight time indicate that travelers prefer routes with shorter scheduled flight time, only the IV estimates suggest significantly different marginal effects for different routing types. The NL3B-IV estimates show that a 1-min increase of scheduled flight time on connecting routes have a larger²² (about 1.4 times) impact of utility than that on direct routes, while the NL3B-OLS estimates give almost equal marginal effects for both routing types. As a result, the IV estimates imply larger VOTs for connecting routes than for direct routes, given that the fare coefficients are—by construction—identical for both routing types. This result has two possible explanations. First, travelers may feel more comfortable spending their time on direct flights than on connecting ones. On the former, for example, they do not have to worry about missing their subsequent flights due to flight delay and/or finding gates. Second, there may be nonlinear effects of flight time that translate into the observed differences in coefficient estimates. Given a city-pair market, scheduled flight time of a connecting route is normally greater than that of a direct route. The nonlinear effects would make travelers less likely to choose a connecting route with flight time much longer than that of a direct route.

²² The hypothesis that the scheduled flight time coefficient of connecting routes is less than or equal to that of direct routes is rejected at the 5% significance level.

Positive hub arrival delay of one and four quarters before the decision quarter are the only significant delay variables in our NL3B-IV estimation, although many on-time performance metrics²³ were tried. This suggests that potential travelers make decisions based on recent available information—including most recent impressions and seasonal effects—on positive hub arrival delay. When choosing among connecting routes, travelers avoid connecting at airports with high expected delay in certain seasons. For NL3B models, the coefficient differences between the two hub delay variables are not statistically significant,²⁴ implying that potential travelers weigh on-time performance of the two periods (one and four quarters before the decision quarter) equally. In addition, we expect a 1-min hub delay increase has a larger impact on demand than an equivalent change in scheduled flight time of a connecting route, because (1) delay disturbs travelers' original schedules and plans, and (2) travelers dislike travel time uncertainty. The NL3B-IV estimates confirm this hypothesis: the sum of two hub delay coefficients is more negative²⁵ than the coefficient of scheduled flight time. The sum is the appropriate basis for comparison because, if delay levels shift upward or downward to a new steady state, the change will be affect both of the lagged variables.

After controlling for the other factors the coefficients of the direct route dummy variable still indicate that potential travelers strongly prefer direct routes than connecting routes, regardless of specifications and estimation methods. This reflects the layover time required for a connecting flight, as well as the physical effort and psychological stress associated with making—and sometimes missing—a connection.

While the ratio of scale parameters (λ_p/λ_a) based on the OLS estimation implies that the correlation of the total utilities for two air routes sharing the same O-D airport pair is very low, the ratio based on the IV estimation implies that the correlation is moderate. The large difference between estimated ratios of scale parameters from the two estimation methods further demonstrate the importance of correcting for the endogenous air fare problem.

The estimated ratios of scale parameters (λ_a/λ_m) from both OLS and IV estimates are consistent with utility-maximization for a reasonable range of market distance. Longer-haul markets have lower ratios of scale parameters (λ_a/λ_m), indicating that the correlations of the total utilities among O-D airport pairs (and thus among routes) in longer-haul markets are higher. Thus, in longer-haul markets route attribute changes are more likely to shift traffic between air routes as opposed to affecting total air market traffic. In shorter-haul markets, air routes are more likely to compete with other modes (non-air alternative), such as auto and rail.

As shown in Fig. 2, estimates of the NL3B models yield concave effects of market distance on air route demand—the marginal effects are decreasing as distance increases, given a reasonable range of inclusive values of level 2. Considering the cases where inclusive value depends on market distance,²⁶ both the OLS and IV estimates imply that air routes have the highest demand potential in markets of distance 850–900 miles, all else equal. For markets of distance shorter than that range, the distance effects reflect declining competition from competing modes, which causes air demand to increase with distance; in long-haul markets, the effect is reversed, presumably due to negligible mode competition and decreasing propensity to travel. These findings are somewhat supported by the National Household Travel Survey (US Department of Transportation, 2001), which shows that mode share for air increases with distance and air becomes the dominant mode starting from the markets of distance 750–999 miles.

4.5.1. Summary and discussion

Regarding nesting structures, the NL3B models outperform the MNL models. First, the NL3B models confirm the non-homogeneous correlations among alternatives, implying that the MNL models incorrectly portray substitution patterns among routes. Second, while the MNL models give similar patterns of coefficients for route level variables, their income (a market level variable) coefficients are, implausibly, negative.

The NL3B-IV model is the preferred model, since its estimates and implications are more sensible, especially for the results of level 2 and 3, which imply reasonable VOTs and correlations of total utilities for air routes. Correcting for the endogeneity problem of air fare also helps to determine the appropriate nesting structure, since the ratios of the scale parameters in NL models are affected by the endogeneity problem. Note that it is possible that observed flight frequency is endogenous, because of supply and demand simultaneity. However, flight frequency is a segment characteristic and each segment may serve many routes and markets; that is, flight frequency is not solely determined by specific route traffic. Therefore, the endogeneity bias caused by frequency may not be severe since the proposed model is a route demand model. This research, hence, only focuses on the remedy for bias caused by the air fare variable.

Recall that to implement the proposed model this research assumes saturated demand levels, depending on city-pair population. The results of lower levels are not affected by the assumption because of the estimation method (Eq. (5)). Only estimates of level 1 may be affected by the assumption. We take the NL3B-IV model as the base case, which assumes every unit of population may make 10 trips per quarter, and change the assumption to different numbers of trips (1, 5, 10, and 50 trips). The results show that the estimates change very little, except for the intercept. Thus the key estimation results are insensitive to our admittedly arbitrary choice of saturation traffic function.

²³ As discussed in Section 4.2, this research investigates (1) departure delays of origin and hub airports, and arrival delays of hub and destination airports; (2) positive and negative delays; and (3) delays of one and four quarters before the decision quarter. The total number of delay variables is 16 ($4 * 2 * 2$).

²⁴ All *p*-values are greater than 0.52.

²⁵ The null hypothesis that the sum of hub delay coefficients is less than or equal to the coefficient of scheduled flight time is rejected at the 5% significance level.

²⁶ The distance effects are partially determined by inclusive values. In the figure, the inclusive values are either set to their mean values or to the predicted values that are determined by functions of distances. We regress inclusive value on market distance (and a constant term) to get each function.

Corrections for standard errors of higher level coefficients may be needed. Because the sequential estimation does not carry variances of inclusive values into higher levels, the standard errors of higher level coefficients in these levels are usually underestimated. The standard errors presented in Table 1 are not corrected since most of the coefficients are very significantly different from zero.

4.6. Case study

In this section, we illustrate an application of the model by assessing the potential impact of a capacity enhancement at a hub airport.

A tremendous amount of money has been and will be spent on improvements meant to reduce flight delays. Applying the proposed model, benefits of delay reductions, which are important for justifying the investments, can be quantified. In this section, we conduct a policy experiment, based on 2004 data, on on-time performance to demonstrate one application of the model. The experiment focuses on delay changes at a specific hub airport—using Chicago O'Hare International Airport (ORD) as an example. The planned airport capacity enhancement at ORD may result in this.

Changes in air passenger traffic volumes²⁷ and their components can be used to assess the impacts. The effects on both ORD and the rest of the system are of interest. If a proposed project is expected to reduce the current delay of ORD by 25%, the NL3B-IV model predicts an increase of 422 thousand connecting passengers (about 4.5% of the original connecting volume) annually at ORD. The increased volume of traffic is from three sources: (1) 68 thousand passengers change from direct routes to routes connecting through ORD; (2) about 155 thousand passengers are attracted from the other 29 hubs; and (3) 200 thousand passengers are from the potential travelers who chose other modes or did not travel. From the viewpoint of the whole air system, the net effect is an increase of 200 thousand passengers. The total benefit of the project is approximately 31.4 million (in 2004 US dollars), estimated by applying the rule of one-half.

5. Conclusions

This research develops a city-pair air demand model and applies it to the air transportation system of the United States. The contributions of the paper to the literature are in both methodology and empirical findings. In terms of methodology, the proposed model incorporates several advances from earlier work in this area. By incorporating the choice of whether or not to travel by air, it incorporates demand generation as well demand allocation. It is based on publicly available and regularly updated data, but, through use of an instrumental variable technique, overcomes the problem of fare endogeneity that has bedeviled previous efforts to use these data for demand modeling.

Our empirical analysis includes on-time performance metrics as well as more traditional variables such as fare, frequency, and travel time. The analysis finds that the pattern of correlations among alternatives can be captured by the three-level nested logit (NL3B) model, which implies that a route is more likely to compete with another route of the same O-D airport pair than the routes of the other O-D airport pairs, and is least likely to be substituted by the non-air alternative. The NL3B model estimated by instrumental variable method (NL3B-IV) is the preferred model since it provides more sensible values-of-time and correlations of total utilities for alternatives than those of NL3B-OLS.

The empirical analysis also suggests that (1) air fare is endogenous and correcting the endogeneity problem by the IV method significantly improves the fare coefficient and its implications; (2) the minimum frequency is more critical to the connecting service; (3) the inferred values of scheduled flight time are \$16.6/h for direct routes and \$24.1/h for connecting routes, both in 2004 dollars; (4) when choosing among connecting routes, travelers avoid connecting at airports with high expected delay; (5) under steady state a 1-min hub delay increase has a larger impact on demand than an equivalent change in scheduled flight time of a connecting route; (6) there is a concave relationship between market distance and air route demand; (7) in a longer-haul market route attribute changes are more likely to shift traffic between routes as opposed to affecting total air market traffic.

Model forms and choice sets can be modified to suit different applications. Studies with different purposes may need other approaches to calculate saturated demand. As discussed in Section 3.2, one solution is to estimate a model for saturated demand. Moreover, in order to recognize heterogeneity among potential travelers and allow more flexible substitution patterns, the mixed logit model may be considered, if the price of computational complexity is affordable.

The proposed model does not consider the case that potential travelers may choose other destination cities. This might be a problem in markets involving many leisure trips. Adding characteristics of other cities as explanatory variables and/or including destination alternatives in the choice set can be solutions for this problem. Another issue related to choice sets is that for simplicity, the proposed model does not differentiate routes by carriers. The model can be extended to the route-carrier level when needed, although new nesting structures have to be examined.

In our empirical study we only consider routes with at most one connecting airport as alternatives. Although this makes the model more tractable with little loss of generality, it produces smaller networks than the true situations, and makes the substitution patterns of alternatives somewhat restrictive. Further development is necessary before we can use the model in very thin markets where multi-stop routings are fairly common.

²⁷ The number of passengers in the data set is a 10% sample from US DOT's Airline Origin and Destination Survey (DB1B). All the traffic levels presented in the experiment are converted into 100% levels by multiplying a factor of 10.

Since our main purpose in this paper is to present the basic model and key estimation results, we leave more detailed treatment of the various applications for future work. For example, we use a market distance variable to capture the distance effect on air demand. Another alternative is to estimate different models for different distance ranges (e.g. short-haul and long-haul). While our approach keeps the model simple, the alternative approach allows different estimates for all factors, which may be more flexible. In addition, much useful information, for example demand elasticities, values of time, and coefficient changes over time, can be derived from this model. These will be the subject of subsequent work.

References

- Abrahams, M., 1983. A service quality model of air travel demand: an empirical study. *Transportation Research Part A* 17 (5), 385–393.
- Adler, N., 2001. Competition in a deregulated air transportation market. *European Journal of Operational Research* 129 (2), 337–345.
- Adler, N., 2005. Hub-spoke network choice under competition with an application to western Europe. *Transportation Science* 39 (1), 58–72.
- Adler, T., Falzarano, S., Adler, G.S., 2005. Modeling service trade-offs in air itinerary choices. *Transportation Research Record: Journal of the Transportation Research Board* 1915, 20–26.
- Air Transport Association, 2005. Fuel cost and consumption report. <<http://www.airlines.org/economics/energy/MonthlyJetFuel.htm>> (accessed 27.04.05).
- Anderson, J.E., Kraus, M., 1981. Quality of service and the demand for air travel. *Review of Economics and Statistics* 63 (4), 533–540.
- Ben-Akiva, M., Lerman, S., 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, MA.
- Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. *Econometrica* 63 (4), 841–890.
- Bhat, C., 1996. Accommodating variations in responsiveness to level-of-service measures in travel mode choice modeling. Working paper, Department of Civil Engineering, University of Massachusetts at Amherst.
- Bhadra, D., 2003. Demand for air travel in the United States: bottom-up econometric estimation and implications for forecasts by origin and destination pairs. *Journal of Air Transportation* 8 (2), 19–56.
- Brownstone, D., Train, K., 1999. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics* 89 (1–2), 109–129.
- Bureau of Economic Analysis, 2006. Local area personal income. <<http://www.bea.gov/bea/regional/reis/default.cfm?catable=CA1-3§ion=2>> (accessed 11.07.06).
- Bureau of Labor Statistics, 2008. Weekly and hourly earnings data from the current population survey. <<http://data.bls.gov/cgi-bin/surveymost>> (accessed 20.03.08).
- Coldren, G.M., Koppelman, F.S., Kasturirangan, K., Mukherjee, A., 2003. Modeling aggregate air travel itinerary shares: logit model development at a major US airline. *Journal of Air Transport Management* 9 (6), 361–369.
- Coldren, G.M., Koppelman, F.S., 2005. Modeling the competition among air-travel itinerary shares: GEV model development. *Transportation Research Part A* 39 (4), 345–365.
- Data Base Products, 2004a. Hub and gateway training manual. Dallas, Texas.
- Data Base Products, 2004b. Onboard domestic/international operations manual—windows version. Dallas, Texas.
- Data Base Products, 2005a. Hub 1995–2004 (CD-ROMs). Dallas, Texas.
- Data Base Products, 2005b. Onboard domestic 1995–2004 (CD-ROMs). Dallas, Texas.
- De Vany, A.S., Garges, E.H., 1972. A forecast of air travel and airport and airway use in 1980. *Transportation Research* 6 (1), 1–18.
- Federal Aviation Administration, 2001. Airport capacity benchmark report 2001. Washington, DC, USA.
- Federal Aviation Administration, 2007. About airline service quality performance (ASQP). <<http://aspm.faa.gov/information.asp>> (accessed 01.07.07).
- Ghobrial, A., Kanafani, A., 1995. Future of airline hubbed networks: some policy implications. *ASCE Journal of Transportation Engineering* 121 (2), 124–134.
- Hansen, M., 1990. Airline competition in a hub-dominated environment: application of non-cooperative game theory. *Transportation Research Part B* 24 (1), 27–43.
- Hansen, M., 1995. Positive feedback model of multiple-airport system. *ASCE Journal of Transportation Engineering* 121 (6), 453–460.
- Hansen, M., 1996. Airline frequency and fare competition in a hub-dominated environment. Unpublished paper.
- Hansen, M., Kanafani, A., 1990. Airline hubbing and airport economics in the pacific market. *Transportation Research Part A* 24 (3), 217–230.
- Hansen, M., Weidner, T., 1995. Multiple airport systems in the United States: current status and future prospects. *Transportation Research Record: Journal of the Transportation Research Board* 1506, 8–17.
- Harvey, G., 1987. Airport choice in a multiple airport region. *Transportation Research Part A* 21 (6), 439–449.
- Hess, S., Polak, J.W., 2005a. Mixed logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management* 11 (2), 59–68.
- Hess, S., Polak, J.W., 2005b. Accounting for random taste heterogeneity in airport choice modeling. *Transportation Research Record: Journal of the Transportation Research Board* 1915, 36–43.
- Hsiao, C.-Y., 2008. Passenger Demand for Air Transportation in a Hub-and-Spoke Network. Ph.D. Dissertation, University of California at Berkeley.
- Hsiao, C.-Y., Hansen, M., 2005. Air transportation network flows: equilibrium model. *Transportation Research Record: Journal of the Transportation Research Board* 1915, 12–19.
- Ippolito, R.A., 1981. Estimating airline demand with quality of service variables. *Journal of Transport, Economics and Policy* 15 (1), 7–15.
- Kanafani, A., Fan, S.L., 1974. Estimating the demand for short-haul air transport system. *Transportation Research Record* 526, 1–15.
- Kanafani, A., Ghobrial, A., 1985. Airline hubbing: some implications for airport economics. *Transportation Research Part A* 19 (1), 15–27.
- Kanafani, A., Gosling, G., Taghavi, S., 1977. Studies in the demand for short haul air transportation. Special report. Institute of Transportation Studies, Berkeley.
- McFadden, D., 1981. Econometric models of probabilistic choice. In: Manski, C., McFadden, D. (Eds.), *Structural Analysis of Discrete Data*. MIT Press, Cambridge, MA, pp. 198–272.
- Morrison, S., Winston, C., 1985. An econometric analysis of the demand for intercity passenger transportation. In: Keeler, T. (Ed.), *Research in Transportation Economics*, vol. 2. JAI Press Inc, pp. 213–237.
- Nevo, A., 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69 (2), 307–342.
- Pathomsiri, S., Haghani, A., 2005. Taste variations in airport choice models. *Transportation Research Record: Journal of the Transportation Research Board* 1915, 27–35.
- Pels, E., Nijkamp, P., Rietveld, P., 2001. Airport and airline choice in a multiple airport region: an empirical analysis for the San Francisco Bay Area. *Regional Studies* 35 (1), 1–9.
- Pels, E., Nijkamp, P., Rietveld, P., 2003. Access to and competition between airports: a case study for the San Francisco Bay Area. *Transportation Research Part A* 37 (1), 71–83.
- Ross, A., Swain, A., 2007. Fighting flight delay. *Operations Research/Management Science Today*, April 2007.
- Small, K., Winston, C., 1999. *The Demand for Transportation: Models and Applications*. Essays in Transportation Economics and Policy: A Handbook in Honor of John R. Meyer. Brookings Institution Press, Washington, DC, USA.
- Train, K., 1998. Recreation demand models with taste differences over people. *Land Economics* 74 (2), 230–239.
- Train, K., 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Trani, A.A., Baik, H., Swingle, H., Ashiabor, S., 2003. Integrated model studying small aircraft transportation system. *Transportation Research Record: Journal of the Transportation Research Board* 1850, 1–10.

- US Department of Transportation, 2001. Long distance transportation patterns: mode by trip distance. National Household Travel Survey, Washington, DC, USA.
- Warburg, V., Bhat, C., Adler, T., 2006. Modeling demographic and unobserved heterogeneity in air passengers' sensitivity to service attributes in itinerary choice. *Transportation Research Record: Journal of the Transportation Research Board* 1951, 7–16.
- Wei, W., Hansen, M., 2005. Impact of aircraft size and seat availability on airlines demand and market share in duopoly markets. *Transportation Research Part E* 41 (4), 315–327.
- Wei, W., Hansen, M., 2006. An aggregate demand model for air passenger traffic in the hub-and-spoke network. *Transportation Research Part A* 40 (10), 841–851.
- Weidner, T., 1996. Hubbing in US air transportation system: economic approach. *Transportation Research Record: Journal of the Transportation Research Board* 1562, 28–37.
- Windle, R., Dresner, M., 1995. Airport choice in multiple-airport regions. *ASCE Journal of Transportation Engineering* 121 (4), 332–337.