MODELING THE COMPETITIVE DYNAMIC AMONG AIR-TRAVEL ITINERARIES WITH GENERALIZED EXTREME VALUE MODELS

Gregory M. Coldren
Ph.D. Candidate, Department of Civil and Environmental Engineering
Northwestern University
737 Ridge Avenue #3H
Evanston, IL 60202
Phone: (847) 769-8080
Fax: (847) 491-4011
Email: g-coldren@northwestern.edu

1. INTRODUCTION

1.1 Preliminary Discussion

Passenger demand for air-carrier services is the foundation of the aviation industry. Understanding this demand is crucial for air-carriers, government policy makers, airport operators, air-carrier suppliers and researchers. Some studies of aviation demand forecast air-travel volumes for given levels of aggregation such as system, metropolitan region, airport or airport (city) pair. Other studies allocate air-travel volumes to air-carriers (or flights) at a given level of aggregation. These allocation studies estimate the relative importance of different carrier attributes (*e.g.* market presence, fare-levels, service quality) on demand share and quantify the tradeoffs that passengers make among these in their air-travel purchasing decisions. These studies have allocated system-wide airline demand to individual carriers (Nason 1981; Morash and Ozment 1996; Suzuki *et al.* 2001), airport-pair demand to individual carriers (Ghobrial and Soliman 1992; Nako 1992; Proussaloglou and Koppelman 1995) and airport-pair demand to flights (Yoo and Ashford 1996; Proussaloglou and Koppelman 1999; Bruning and Rueda 2000; Algers and Beser 2001).

All the above-mentioned allocation studies fall into at least one of the following categories: 1) studies based on data with a high level of geographic aggregation, 2) studies employing surveys with a very limited range of airport-pairs or 3) studies based on stated preference data which may be subject to bias (Ben-Akiva *et al.* 1989; Morrison 2000; Murphy *et al.* 2005). Additionally, a limitation of these studies is their failure to model air-travel demand at the itinerary level. An itinerary, as used here, is a leg (flight number) or sequence of legs connecting a given airport-pair. Itineraries are either nonstop, direct (a connecting itinerary not involving an airplane change), single-connect (a connecting itinerary with an airplane change) or double-connect (an itinerary with two connections)[1]. For a given day, an airport-pair may be served by hundreds of itineraries, each of which offers passengers a potential way to travel between the airports.

Itineraries are the products that are ultimately purchased by passengers and hence it is their characteristics that influence demand. In making their itinerary choices, travelers make tradeoffs among the characteristics that define each itinerary (*e.g.* departure time, equipment type(s), number of stops, route, carrier). Modeling these itinerary-level tradeoffs is essential to truly understanding air-travel demand.

Using comprehensive and official schedule and bookings data, this paper summarizes the framework established by Coldren *et al.* (2003), Coldren and Koppelman (2005a) and Coldren and Koppelman (2005b), and focuses on the problem of allocating airport-pair demands to individual itineraries linking the airport-pairs.

1.2 Air-travel Itinerary Share Models

Models that allocate airport-pair demand to the itinerary level are referred to as air-travel itinerary share models[2]. These models forecast the share of passengers expected to travel on each itinerary between any airport-pair. This is done by relating each itinerary's value (quality) to the total value of all itineraries in its respective airport-pair. The value of an itinerary is usually modeled as a function of its service characteristics (independent variables) and their corresponding preference weights (parameter estimates). Though itinerary share models vary with respect to the service characteristics that are included in their specifications, they all describe the impact that these service characteristics have on itinerary share. Additionally,

---

[1] These four classifications are referred to as an itinerary's "level-of-service" in this paper.

[2] These models are also referred to as network-planning or network-simulation models.

advanced itinerary share models allow the underlying competitive dynamic, if any, among air-travel itineraries to be measured.

Once an itinerary share model is estimated, these shares (probabilities) can be applied to airport-pair volume forecasts obtained from a separate model, allowing for itinerary-level passenger forecasts to be obtained. These itinerary-level forecasts can then be assigned to flight legs to obtain carrier market share at the flight-leg, airport-pair, region, system or any other level of aggregation.

When passenger forecasts from an itinerary share model are combined with fare and cost data, it becomes a powerful tool allowing an airline to evaluate its current schedule and network, as well as proposed network changes, alliance (merger) evaluations and “what-if” scheduling decisions (including modeling the potential actions of competitors) with respect to passengers, revenue and profitability. Thus, these models can be invaluable to a passenger airline’s tactical decision-making and strategic planning. Some common situations where these models aid carriers are: merger and acquisition scenarios, route schedule analysis, codeshare scenarios, equipment assignment scenarios, minimum connection time studies, price-elasticity studies, hub location studies, and equipment purchasing decisions.

Because itinerary share models guide critical decisions it is paramount that they be as accurate as possible since, for a given carrier, improving the forecasting ability of its itinerary share model will translate to improvements in its revenue management, schedule efficiency and profitability.

1.3 Quality of Service Index Itinerary Share Models

Many itinerary share models employ a demand allocation methodology referred to as quality of service index (QSI). QSI models[3] relate an itinerary's passenger share to its "quality" (and the quality of all other itineraries in its airport-pair), where quality is defined as a function of various itinerary service attributes and their corresponding preference weights. For a given QSI model, these preference weights are obtained using statistical techniques and/or analyst intuition. Once the preference weights are obtained and the QSI's are calculated for each itinerary, itinerary i's passenger share is given by:

$$S_i = \frac{QSI_i}{\sum_{j \in J} QSI_j} \tag{1.1}$$

where S_i is the passenger share assigned to itinerary i,
QSI_i is the quality of service index for itinerary i and
the summation is over all itineraries in the airport-pair.

QSI models are problematic on two fronts. First, a distinguishing characteristic of these models is that their preference weights (or sometimes subsets of these weights) are obtained independently from the other preference weights in the model. Thus, QSI models do not capture interactions existing among itinerary service characteristics (*e.g.* elapsed itinerary trip time and equipment, elapsed itinerary trip time and number of stops). Second, QSI models are not able to measure the underlying competitive dynamic that may exist among air-travel itineraries. This second inadequacy in QSI models can be seen by examining the cross-elasticity equation for the change in the share of itinerary j due to changes in the QSI of itinerary i:

$$\eta_{QSI_i}^{S_j} = \frac{\partial S_j}{\partial QSI_i} \frac{QSI_i}{S_j} = -S_i \tag{1.2}$$

---

[3] Information on QSI models obtained from TRB Transportation Research E-Circular E-C040 (2002) as well as the author's personal experience with the models.

The expression on the right side of Equation (1.2) is not a function of j. That is, changing the QSI (quality) of itinerary i will affect the passenger share of all other itineraries in its airport-pair in the same proportion. This is not realistic since, for example, if a given itinerary (linking a given airport-pair) that departs in the morning improves in quality, it is likely to attract more passengers away from the other morning itineraries than the afternoon or evening itineraries.

1.4 Motivation for Research

The research presented in this paper provides a discrete choice analysis framework for itinerary-level demand allocation modeling where the estimated models use aggregate logit-based share techniques[4]. In these models, a value function, the aggregate analog of utility, represents the relative desirability of each itinerary connecting an airport-pair for each day of the week. In specifying these value functions, itinerary-level and carrier airport-pair service-characteristic variables are used to relate each itinerary with its airport-pair demand share.

The motivation for the research contained in this paper is twofold given that the overarching goal is to model itinerary shares: 1) to understand the impact of different air-carrier service attributes on itinerary share and 2) to understand the underlying competitive dynamic between itineraries.

1.5 Contributions of Research

The research summarized in this paper contributes to the literature in three ways. First, the presented research systematically and comprehensively models aviation demand at the itinerary level. Second, unlike QSI-based itinerary share models, the logit-based models presented in this paper use techniques to simultaneously estimate parameters for the independent variables used in the model specifications. Explanatory variables used in this study such as level-of-service

---

[4] Aggregate logit shares are used (despite the availability of individual itinerary-level booking information) as the data do not contain information about traveler or trip characteristics.

indicators differing by market type, connection quality variables, departure time variables, and equipment size and type variables have previously not been reported in the literature. Additionally, implementation of the most basic itinerary share model presented in this paper by a major U.S. carrier yielded substantial improvements in the carrier's forecasting accuracy compared with its previous QSI-based model.

Third, generalized extreme value (GEV) (McFadden 1978) specifications are used to model the underlying competition among air-travel itineraries. These models provide flexibility in capturing inter-itinerary competition dynamics along a variety of dimensions. Most of these structures have never been used in the aviation demand literature and some are new to the logit share literature in general.

1.6 Outline

Section 2 outlines the conceptual and modeling framework of this paper. Section 3 details the impact of various itinerary service characteristics on itinerary share via the use of multinomial logit (MNL) models. Additionally, validation results from the implementation of these models by a major U.S. carrier are presented. Section 4 presents several variations of the nested logit (NL) model capturing the underlying competitive dynamic among air-travel itineraries along the dimensions of departure time, carrier and level-of-service. Section 5 advances these NL models by estimating ordered generalized extreme value (OGEV) and hybrid-OGEV models allowing for complicated inter-itinerary competition relationships to be measured. Finally, Section 6 provides some concluding remarks.

2. CONCEPTUAL AND MODELING FRAMEWORK

2.1 Introduction to Discrete Choice Analysis and Techniques

Discrete choice scenarios arise when decision-makers choose from a set of mutually exclusive and collectively exhaustive alternatives. Following the framework established by Domencich and McFadden (1975), discrete choice scenarios are described by four elements: a decision-maker, the alternatives available to the decision-maker, attributes of these alternatives and a decision rule.

The decision-maker is an individual person or group of persons (who make a common decision). The alternatives available to the decision-maker must be discrete, each of which has a vector of attributes. Finally, the decision rule commonly assumed is that decision-makers choose the alternative that maximizes their utility.

The analyst does not know the actual utility that decision-makers achieve from alternatives, however. Following convention, the utility of each alternative for a given decision maker is decomposed into two components: a deterministic component consisting of observable attributes of the alternative, and a random component representing unknown and/or unobservable components of the decision maker's utility. Further following convention, the utility of alternative i, U_i, is expressed as the sum of the deterministic component, V_i, and the random component, ε_i:

$$U_i = V_i + \varepsilon_i \tag{2.1}$$

Since ε_i is a random variable, the entire utility, U_i, is a random variable. V_i is generally assumed to be linear-in-parameters where:

$$V_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_n X_{ni} \tag{2.2}$$

That is, an alternative's deterministic value, V_i, is a linear function of explanatory variables (attributes, represented above by the $X_{ki}'s$) and their corresponding parameter estimates (represented above by the $\beta_k's$). This entire formulation of alternative utility is referred to as linear in parameters with additive disturbance (LPAD) (Manski 1973).

Since it is assumed that a decision-maker selects the alternative from his/her choice set (denoted by C) that maximizes his/her utility, and the analyst treats utilities as random, it is necessary to calculate the probability that a decision-maker will choose a given alternative. The probability that a decision-maker chooses alternative i from the choice set C with J alternatives, $\Pr ob\left(i:C\right)$, is given by:

$$\begin{aligned} &Prob\left(i:C\right) = Prob\left(U_i \geq U_j \quad \forall j\right) = \\ &\qquad Prob\left(V_i + \varepsilon_i \geq V_j + \varepsilon_j \quad \forall j\right) = Prob\left(\varepsilon_j - \varepsilon_i \leq V_i - V_j \quad \forall j\right) \end{aligned} \tag{2.3}$$

This is a multivariate cumulative distribution function for the differences between pairs of error terms. Specific assumptions on the distribution of the $\varepsilon's$ lead to different model structures. For example, the assumption that the error terms are independent and identically gumbel distributed across alternatives and decision-makers leads to the familiar multinomial logit model developed by McFadden (1974). However, to accurately model many choice situations (including the scenario in this research), it is necessary to relax the assumption of error independence across alternatives. The generalized extreme value family of models proposed by McFadden (1978) allow for this. These models are closed-form, are consistent with random utility maximization theory and assume that alternative error terms have a gumbel distribution with equal variance across alternatives (but not necessarily independently distributed across alternatives).

The parameters for a given GEV discrete choice model can be estimated using full-information maximum likelihood on a random or choice-based data sample. This yields parameter estimates that maximize the likelihood (posterior probability) of the observed sample choices conditional on the model.

2.2 Discrete Choice Scenario of Research Problem, Data Used and Data Aggregation Issues

The person or group who books an air-travel itinerary is the decision-maker in the modeling framework of this study. The alternative sets are modeled as all itineraries linking directional airport-pairs for each day of the week. For example, all Monday itineraries from Boston to Los Angeles constitute an alternative set, as do all Friday itineraries from Allentown to Tucson. Each itinerary in each alternative set is described by a set of attributes such as level-of-service, fare and departure time. Finally, it is assumed that the air-travel passengers being modeled choose itineraries with the intent of maximizing their utility.

Passenger bookings data for this research was obtained from a compilation of computer reservation systems (CRS). These are data sources containing detailed records of booked itineraries. CRS data is commercially available and compiled from several computer reservation systems including Sabre (Sabre Travel Network 2000, 2001), Galileo (Galileo International 2000, 2001) and Worldspan (Worldspan, L.P. 2000, 2001) as well as Internet travel sites such as Orbitz (Orbitz, L.L.C. 2000, 2001), Travelocity (Travelocity.com, L.P. 2000, 2001), Expedia (Expedia, Inc. 2000, 2001) and Priceline (Priceline.com, Inc. 2000, 2001). The CRS data is believed to include 90% of all domestic air-travel bookings during the study period. However, increasing use of direct-carrier bookings (via telephone and Internet) has since reduced the proportion of bookings reported by this source. The CRS data contains variables describing each booked itinerary such as trip origin and destination, leg(s) origin and destination, leg number, party size, flight number(s), leg(s) departure and arrival time(s), departure and arrival date, and airline.

Since the itinerary building rules employed by the different computer reservation systems differ to a limited degree, a major carrier’s itinerary building engine was used to generate the set of itineraries between all airport-pairs. This engine builds all feasible online itineraries (up to the double-connection level) using leg-based air-carrier schedule data obtained from the Official Airline Guide (OAG Worldwide Limited 2000, 2001). The OAG data contains the following variables: operating airline, codeshare airline (if a codeshare leg), origin, destination, flight number, departure and arrival time, equipment, days of operation, leg mileage and flight time. In building the itineraries, distance-based circuitry logic is used to eliminate unreasonable itineraries, and minimum and maximum connection times are incorporated to ensure that unrealistic connections are not allowed. Finally, itineraries are generated for each day of the week accounting for day-of-week differences in service offered.

The generated itineraries are merged with the booked itineraries to assemble the estimation datasets for this research (the dependent variable in the models is the number of passengers who booked each itinerary). Finally, using the estimation data, aggregate GEV models are estimated with full-information maximum likelihood techniques using the GAUSS modeling software (Aptech Systems, Inc. 2005). The specific forms of the different models tested are described in Sections 3 – 5.

Even though the bookings data employed in this study is based on the choices of individual travelers, it does not include any information on the demographic characteristics of the individual that made the booking or any trip-related characteristics of the booking (such as income, business vs. leisure, number of days booked in advance of departure, duration of stay). Thus, since no individual data is available to identify differences among travelers, it may not be appropriate to count the full weight of the individual observations in calculating the statistics for the models. The most extreme adjustment can be accomplished by dividing the log-likelihood values for the models by the ratio of the number of booked passengers to the number of airport-pair, day-of-the-week combinations; and the parameter estimate t-statistics by the square root of this ratio. Intermediate adjustments that take account of the fact that individual choice behavior is observed can be justified as well. Statistics discussed in Sections 4 – 5 and presented in Tables 4 – 7 refer to the unadjusted values as well as the values obtained after adjusting (according to the ratio described above) due to the aggregate nature of the data.

3. MODELING THE IMPACT OF SERVICE CHARACTERISTICS ON AIR-TRAVEL ITINERARY SHARE USING THE MULTINOMIAL LOGIT MODEL

3.1 Introduction

An important motivation for developing itinerary share models is to gain an understanding of the impact of different service factors on itinerary share. In this section, aggregate MNL models are developed to model itinerary shares for each domestic[5] airport-pair in the United States. The itinerary attributes modeled are level-of-service, connection quality, carrier attributes, aircraft size and type, and departure time. Factors not directly modeled – though discussed – include elapsed itinerary travel time, carrier frequent flyer program effectiveness and service quality.

3.2 Modeling Framework

The United States was divided into five regions: each Continental time zone (East, Central, Mountain, West) and a region for Alaska and Hawaii. Using these regions, eighteen "entities" were defined: all sixteen combinations of the Continental time zones (East-East (E-E), East-Central (E-C), East-Mountain (E-M), East-West (E-W), . . . , West-West (W-W)), as well as an entity for Alaska and Hawaii to the Continental U.S. and an entity for the Continental U.S. to Alaska and Hawaii[6].

Using January 2000 data, aggregate multinomial logit models were estimated for each of these entities with a common specification using data from all airport-pairs in the respective entity. Under MNL model assumptions, it can be shown that the market share of passengers assigned to each itinerary between an airport-pair for a given day of the week is given by the following equation:

---

[5] Mexico-to-U.S. (and vice versa) and Canada-to-U.S. (and vice versa) airport-pairs are also included in the analysis.

[6] Airport-pairs within Hawaii or Alaska, and between Hawaii and Alaska were excluded from the analysis.

$$S_i = \frac{\exp(V_i)}{\sum_{j \in J} \exp(V_j)} \tag{3.1}$$

where S_i is the passenger share assigned to itinerary i,
$\exp(\ )$ is the exponential function,
V_i is the value (deterministic portion of utility) of itinerary i and
the summation is over all itineraries for the airport-pair-day-of-the week.

The independent variables used in these models are described in Table 1. Estimation results for five entities (the entities originating in the Eastern time zone and the West-East entity) are reported in this section[7]. The parameter estimates for the five models are reported in Table 2[8].

3.3 The Impact of Service Characteristics on Itinerary Share

3.3.1 The Impact of Level-of-Service on Itinerary Share

The estimation results presented here provide strong evidence of the importance of level-of-service on itinerary share. All parameter estimates for level-of-service, relative to the best level-of-service for the airport-pair[9], are large, negative and significant. The magnitude of preference among levels-of-service indicates that each reduction in level-of-service from the best available substantially reduces the value of the associated itinerary. These results provide evidence that passengers strongly prefer to avoid connections presumably due to their increased elapsed travel time, the inconvenience of switching planes, higher probability of delays, increased potential for lost baggage, etc.

---

[7] Similar results were obtained for the other thirteen entities.

[8] All parameter estimates in this table significant at the 0.05 level.

[9] Each airport-pair has a best level-of-service. For example, Boston – Los Angeles has nonstop service. Therefore, it is a “nonstop market”. However, the best level-of-service linking many airport-pairs is a single or even double connection.

The magnitude and relative differences among the level-of-service parameter estimates is very similar across these five entities. This result supports the idea that passengers are similarly sensitive to differences in level-of-service across the entire domestic system.

3.3.2 The Impact of Connection Quality on Itinerary Share

Several variables in the models measure the relative passenger preference for and quality of itineraries involving connections (in addition to the level-of-service variables). The results indicate that travelers making a connection strongly prefer the best connection (shortest ground time) among itineraries sharing a common leg at a transfer station. The models include two variables to measure the impact of second-best connection itineraries: the “second-best connection” and “second-best connection time difference” variables. Figures 1 and 2 show visual representations of two types of second-best connections. Both figures show Boston to Los Angeles single-connect itineraries with ORD (Chicago) as the connecting station and time represented along the vertical axis. The #123 – #789 connection is a “second-best connection” in both figures.

In addition, passengers compare all connections linking an airport-pair (possibly involving different connecting airports) with respect to elapsed time (flight plus ground time) differences. The “best connection time difference” variable captures passenger aversion to both excess ground and flight time. The magnitude of this parameter is approximately 1/2 to 2/3 of the “second-best connection time difference” parameter estimate suggesting that the comparison between connecting itineraries sharing a common leg carries more importance than the comparison across different routings.

A related comparison of itinerary connection quality is the ratio of each itinerary’s distance to that of the shortest itinerary distance for the airport-pair (“distance ratio” variable).

The parameter estimates for this variable were negative in all of the models indicating that circuitous itineraries have lower value than more direct itineraries.

3.3.3 Impact of Carrier Presence, Fares and Carrier on Itinerary Share

Carrier market presence can be represented as the amount of service provided in the airport-pair (nonstop frequency for example) or as a function of the carrier's overall presence in the origin and/or destination airports (percentage of departures for example). Several functional forms of carrier presence were experimented with in the models. The specification yielding the most reasonable results was the "point of sale weighted airport presence" variable included in the present models. The parameter estimates for this variable are positive matching *a priori* expectations since increased presence at an airport allows a carrier to offer passenger's greater frequencies and more destinations. Additionally, increased carrier presence provides marketing opportunities and enables carriers to attract more frequent flyer customers.

Itineraries with higher fares have a lower value than lower-cost itineraries (other things being equal). Fare data for models estimated in this paper was obtained from the "Superset" data source (Data Base Products, Inc. 2000, 2001). The parameter estimates for the "fare ratio" variable in the models are negative as expected. Though the fare data employed is superior to other revealed preference aviation demand studies, it is important to recognize that the data is based on averages for each carrier across all itineraries for each airport-pair. Detailed fare data for specific itineraries and ticket classes was not available for this research.

Carrier (represented by dummy variables in the models) is an important component of itinerary choice. These variables capture the many influences that a carrier's image has on itinerary choice such as marketing effectiveness, frequent flyer program quality, food quality, safety record, on-time performance, propensity to overbook, baggage handling record, flight crew friendliness, etc. These parameter estimates are all highly significant and (for a given carrier) varied greatly across the entities reflecting differences in a carrier's quality of service and image in different entities. The carrier-specific constants have been suppressed from Table 2 for proprietary reasons. However, exclusion of these parameter estimates does not impact the behavioral interpretation of the models.

3.3.4 Impact of Aircraft Size and Type on Itinerary Share

Clearly, the type(s) of aircraft used on an itinerary and the number of seats offered on these aircraft types affects the value of the itinerary. The estimation results confirm the belief that passengers prefer large (mainline) jets to regional jets to propeller aircraft due to their increased speed, more comfortable cabins and higher perceived levels of safety. Also, passengers prefer larger aircraft to smaller aircraft (within an aircraft type). Figure 3 shows the contribution of aircraft type and number of seats to itinerary value for the East-East entity displaying typical aircraft sizes of 15-40 for propeller aircraft, 50-80 for regional jet aircraft and 100+ for mainline aircraft. Value is determined by applying the appropriate equipment variable and the corresponding seats variable. Passenger sensitivity to an increase in aircraft size is particularly strong for smaller aircraft.

3.3.5 Impact of Departure Time on Itinerary Share

The five entities displayed in Table 2 show distinctly different preference patterns for departure time. However, in all cases early morning (before 7 A.M.) and late evening (after 8 P.M.) itineraries are generally not preferred. Complete interpretation of these parameter values depends on the flight schedules supplied in each entity. A negative parameter indicates that the time period is not preferred after taking account of less frequent flight offerings during the period. These departure time-of-day preferences reflect different origin and destination constraints across entities. For example, itineraries departing after 8 P.M. have a higher preference in the West-East entity than the westbound entities. This is a result of these travelers taking advantage of “red eye” itineraries allowing them to arrive on the East Coast early the following morning.

3.3.6 Impact of Elapsed Time, Frequent Flyer Program Effectiveness and Service Quality on Itinerary Share

Elapsed itinerary trip time was not included in the models presented in this paper because it did not yield significant parameter estimates. This is due, primarily, to the other variables in the model specification that indirectly measure elapsed time. These include the level-of-service indicators, the connection quality variables (second-best connection variables, best connection time difference, distance ratio) and the aircraft size and type variables.

Due to data unavailability, the impact of frequent flyer programs and service quality (*e.g.* on-time performance, baggage handling record, overbooking rates) on itinerary share were not explicitly modeled in this research. However, it is believed that the carrier constants included in the presented models capture many of the service quality effects as well as the attractiveness of a given carriers frequent flyer program.

3.4 Validation

The full set of eighteen logit-based models was implemented as a component of a major U.S. carrier’s itinerary share model. For a given month, validation was undertaken by allocating demand to itineraries between each airport-pair and assigning the itinerary-level demand to the flight-segments that define each itinerary. For each of the carrier’s segments the total number of forecasted passengers was compared to onboard passenger count data. Absolute errors were averaged across segments and compared to predictions using the carrier’s previous QSI-based itinerary share model for the twelve months of 1999, August 2001, March 2002 and May 2002.

The results, reported in Table 3, are consistently superior for the logit-based models for every month based on model estimation from a single month (January 2000). This suggests very good model stability from month to month. Note that similar results were obtained when validation was applied to the selected months in 2002; that is, even including months after the September 11th, 2001 terrorist attacks.

3.5 Summary and Conclusions

This section summarizes the influence of various service attributes on itinerary share. Eighteen aggregate multinomial logit models are estimated, each covering a major region-to-region (or intra-region) of the United States.

Detailed analysis is performed on the relative importance of itinerary level-of-service, connection quality, aircraft type and size, departure time, carrier presence, fares and carrier on itinerary share. The parameter estimates for all of these itinerary characteristics have an intuitive interpretation and are consistent over all eighteen entities. Validation results using a major U.S. carrier's onboard segment-level data consistently favor the logit-based models relative to previously used QSI-based models.

Additionally, itinerary service attributes controllable by a carrier to improve market share are identified. Among these, the most important are the provision of higher levels of service (more nonstop and direct itineraries) suggesting the potential market value of moving away from the dominant hub-and-spoke network structure of the major legacy carriers, flying mainline jets, matching service offerings with customer departure time preferences, optimizing connection times and maintaining a substantial presence in primary markets.

4. MODELING THE UNDERLYING COMPETITIVE DYNAMIC AMONG AIR-TRAVEL ITINERARIES WITH NESTED LOGIT MODELS

4.1 Introduction

Section 3 modeled itinerary shares using aggregate multinomial logit functions of the itineraries' attributes. MNL models are adequate for describing the impact of service attributes on airport-pair itinerary share. However, the across-itinerary independence of the itinerary error terms (inherent in the derivation of the MNL model structure) implies that all itineraries "compete" equally with each other for a given airport-pair-day-of-the-week. That is, the underlying competition among air-travel itineraries is assumed to be "uniform" when modeled with a MNL function. This property of the MNL model is apparent by examining the cross-elasticity equation for the change in the probability of itinerary j due to changes in an attribute of itinerary i :

$$\eta_{X_{ik}}^{P_j} = \frac{\partial P_j}{\partial X_{ik}} \frac{X_{ik}}{P_j} = -P_i X_{ik} \beta_k \tag{4.1}$$

where X_{ik} is the value of itinerary $i's$ k^{th} attribute and β_k is attribute $k's$ parameter estimate. Note that the expression on the right side is not a function of j . That is, changing an attribute of itinerary i affects all other itineraries in the same proportion.

The central hypothesis of this section is that the underlying competition among air-travel itineraries for a given airport-pair and day-of-the-week is not uniform. Rather, it is asserted that groups of itineraries sharing one or more common attributes will exhibit more competition (as measured by cross-elasticities) amongst themselves than with itineraries not sharing these attributes.

In this section, it is hypothesized that the competition among air-travel itineraries is differentiated by proximity in departure time, carrier, level-of-service or a combination of these dimensions. Clearly, itineraries with similar departure times share characteristics that air travelers consider in their itinerary selection process. Thus, itineraries within a given time period are likely to compete with each other more than with itineraries in a different time period. Similarly, it is likely that itineraries of a given carrier compete more with each other than with itineraries of different carriers due to loyalty factors including frequent flyer program affiliations. Finally, it is believed that itineraries with the same level-of-service share many characteristics that differentiate them from itineraries with different levels-of-service.

Subsection 4.2 outlines the modeling framework for this section and Section 5. Subsection 4.3 estimates a “base” MNL model. Due to the belief that the competition among itineraries is differentiated by proximity in departure time, carrier or level-of-service, the constraints of the MNL model are then relaxed in Subsection 4.4 by estimating two-level nested logit models. These models permit itineraries to be grouped (nested) by departure time, carrier or level-of-service. In addition to the value function parameter estimates, these models also contain an inverse logsum parameter representing the level of itinerary competition within nests. For a given two-level NL model, the estimated value of the inverse logsum parameter (along with the overall model fit) indicates whether it is valid to group the itineraries according to the selected nesting structure.

It is likely that the competition among air-travel itineraries is differentiated by proximity along more than one of the above-mentioned dimensions. To address this, a two-level weighted nested logit (WNL) model is estimated in Subsection 4.5 combining the competitive results of different two-level NL models. This model allows the simultaneous consideration of parallel two-level nesting structures (each structure yields an inverse logsum parameter estimate indicating the amount of itinerary competition within the nests of that structure) with a weight parameter indicating the relative importance of each structure. The formulation of this model is similar to the principles of differentiation models developed by Bresnahan *et al.* (1997).

Clearly, the two-level WNL model has advantages over the more restrictive two-level NL model structure. However, this model cannot capture the inter-itinerary competition dynamic that may exist among itineraries sharing a common attribute within another attribute. For example, (as mentioned above) it is believed that itineraries sharing a common time period compete more closely with each other than with itineraries of different time periods. However, within these time periods it is likely that itineraries of the same carrier exhibit even more competition amongst themselves. As a result, Subsection 4.6 presents the estimations of three-level nested logit models. These models group (nest) itineraries at an upper-level for a given dimension, and within each upper-level nest group itineraries according to a second dimension (lower-level nest). In addition to the value function parameter estimates, these models yield upper and lower-level inverse logsum parameter estimates indicating the differential amount of itinerary competition within the upper and lower-level nests. The estimated values of these inverse logsum parameters (along with the overall model fit) indicate whether a three-level nesting specification is supported.

All of the above mentioned models are members of the generalized extreme value family of models (McFadden 1978) and are shown to outperform the base MNL model with respect to statistical tests and behavioral interpretations, leading to a clearer understanding of the air-travel inter-itinerary competition dynamic.

4.2 Modeling Framework

As mentioned above, the motivation in developing the itinerary share models of this section (and Section 5) is to understand the underlying competitive structure of air-travel itineraries. Using

May 2001 data, models are estimated using all airport-pairs from the East to the West (as determined by time zone) regions of the United States and Canada. All models estimated in this section and Section 5 have a common specification with respect to the independent variables representing itinerary service characteristics[10].

As discussed in Subsection 2.2, due to the aggregate nature of the data, the log-likelihood values for the models estimated in this section and Section 5 are adjusted by dividing by the ratio of the number of booked passengers to the number of airport-pair, day-of-the-week combinations; and the parameter estimate t-statistics by the square root of this ratio. For the estimation data, the log-likelihood adjustment factor is $469{,}078/14{,}893 = 31.50$; and the t-statistic adjustment factor is $Sqrt\left(31.50\right) = 5.61$.

4.3 Multinomial Logit Model

The base or reference model for this section is the MNL model reported in Table 4[11]. The parameter estimates are reported in groups corresponding to level-of-service, connection quality, carrier attributes, aircraft type, and departure time variables. The parameter estimates for this model are very similar to the estimates from Section 3, where the interpretation of these estimates is detailed. For this section and Section 5 it is sufficient to state that the value function parameter estimates have the correct sign, are of reasonable magnitude, and all are significant at the 0.05 level after adjustment.

---

[10] Some value function variables (e.g. point of sale weighted airport presence, aircraft size) had insignificant or incorrect (sign) parameter estimates in the advanced models. Therefore, the value function specification for models contained in this section and Section 5 is simpler than that found in Section 3.

[11] In Tables 4 and 5, all parameter estimates significant at the 0.05 level after the adjustment procedure. In both tables, the significance of value function parameter estimates is with respect to their hypothesized value of zero; the significance of inverse logsum parameter estimates is with respect to their hypothesized value of one.

4.4 Two-Level Nested Logit Model

Initial two-level nested logit estimations assumed nesting based on each of the three dimensions described above. In these models, itineraries are grouped into nests according to departure time (morning, 5:00 – 9:59 A.M.; midday, 10:00 A.M. – 3:59 P.M; evening, 4:00 P.M. – Midnight), carrier (six major U.S. carriers and a group of “other” carriers) or level-of-service (nonstop, direct, single-connect, double-connect). Visual representations of the two-level time NL model and the two-level carrier NL model are shown in Figures 4 and 5, respectively.

With a two-level nested logit specification, the share of passengers assigned to each itinerary between an airport-pair for a given day of the week is given by:

$$S_i = S_{n'} \times S_{i|n'} = \frac{\exp\left(\frac{1}{\mu}\Gamma_n\right)}{\sum_{n'\in N}\exp\left(\frac{1}{\mu}\Gamma_{n'}\right)} \times \frac{\exp(\mu V_i)}{\sum_{i'\in n}\exp(\mu V_{i'})} \tag{4.2}$$

where S_i is the passenger share assigned to itinerary i,
$S_{n'}$ is the passenger share assigned to nest n',
$S_{i|n'}$ is the passenger share assigned to itinerary i given nest n',
μ is the inverse logsum parameter associated with the nests,
$\Gamma_n = \ln\left(\sum_{i'\in N_n}\exp(\mu V_{i'})\right)$ and
V_i is the value of itinerary i.

The inverse logsum parameter[12] must be greater than one to ensure consistency with utility maximization theory. The inverse logsum parameter estimates can be interpreted as indicating the amount of "competition" among itineraries sharing a common nest with larger values indicating a higher level of substitution within nests.

The estimation results for two-level NL models with itineraries nested by departure time and carrier are reported in Table 4 showing that itineraries within a common time period and itineraries flown by the same carrier have common attributes that passengers consider in their itinerary selection process. These models reject the hypothesis that the MNL model is the true model at the 0.001 level after adjustment. On the other hand, grouping itineraries by level-of-service did not yield theoretically acceptable results (the inverse logsum parameter was estimated to be less than one, which is inconsistent with utility theory). This was surprising since it seems likely that an itinerary within a given level-of-service nest would share many characteristics with the other itineraries within the same nest and thus should have higher cross-elasticities among themselves than with itineraries of different levels-of-service.

4.5 Two-Level Weighted Nested Logit Model

The two-level weighted nested logit model simultaneously estimates parallel two-level nesting structures (each structure is equivalent to a two-level NL model) with a weight parameter indicating the relative importance of each structure. Each itinerary in each alternative set appears twice in the model, once in each of the parallel structures. The WNL model can be shown to be a special case of the generalized nested logit model (Wen and Koppelman 2001).

---

[12] In many other studies, the logsum parameter is represented by θ and must be less than one. In the current derivation, μ, the inverse logsum, is equal to $\frac{1}{\theta}$ leading to identical results.

Due to the strong empirical results from the two-level time and carrier NL models, a two-level WNL model was estimated with a time structure and a carrier structure (see Figure 6). The share of passengers assigned to each itinerary between an airport-pair for a given day of the week is:

$$\begin{aligned} S_i &= w_t \times S_t \times S_{i|t} + w_c \times S_c \times S_{i|c} \\ &= w_t \times \frac{\exp\left(\frac{1}{\mu_t}\Gamma_t\right)}{\sum_{t' \in N} \exp\left(\frac{1}{\mu_t}\Gamma_{t'}\right)} \times \frac{\exp(\mu_t V_i)}{\sum_{i' \in t} \exp(\mu_t V_{i'})} \\ &\quad + w_c \times \frac{\exp\left(\frac{1}{\mu_c}\Gamma_c\right)}{\sum_{c' \in N} \exp\left(\frac{1}{\mu_c}\Gamma_{c'}\right)} \times \frac{\exp(\mu_c V_i)}{\sum_{i' \in c} \exp(\mu_c V_{i'})} \end{aligned} \qquad (4.3)$$

where c represents the carrier nests,
t represents the time nests,
w_c is the weight given to the carrier structure and
$w_t = 1 - w_c$ is the weight given to the time structure.

Estimation results for this model are reported in Table 4. The inverse logsum parameters for both the time and carrier nests are significantly greater than one (at all levels of significance after adjustment) indicating increased itinerary competition among itineraries sharing a common time period or carrier. The weight parameter is close to ½ and significantly different than zero or one (at all levels of significance after adjustment) indicating that each portion of the structure is important. Finally, this model outperforms both the two-level time and carrier NL models at the 0.001 level after adjustment.

4.6 Three-Level Nested Logit Model

Six three-level nested logit model specifications were estimated representing all possible three-level combinations for the three itinerary dimensions under study (upper-level time and lower-level carrier (time, carrier); carrier, time; time, level-of-service; level-of-service, time; carrier, level-of-service; level-of-service, carrier).

For the three-level NL models, the share of passengers assigned to each itinerary between an airport-pair for a given day of the week is given by:

$$S_i = S_{m'} \times S_{n'|m'} \times S_{i|n'}$$

$$= \frac{\exp\left(\frac{1}{\mu_m}\Gamma_m\right)}{\sum_{m' \in M} \exp\left(\frac{1}{\mu_m}\Gamma_{m'}\right)} \times \frac{\exp\left(\frac{\mu_m}{\mu_n}\Gamma_n\right)}{\sum_{n' \in N} \exp\left(\frac{\mu_m}{\mu_n}\Gamma_{n'}\right)} \times \frac{\exp\left(\mu_n V_i\right)}{\sum_{i' \in n} \exp\left(\mu_n V_{i'}\right)} \tag{4.4}$$

where $S_{m'}$ is the passenger share assigned to upper-level nest m',

$S_{n'|m'}$ is the passenger share assigned to lower-level nest n' given upper-level nest m',

μ_m is the inverse logsum parameter associated with the upper-level nests,

μ_n is the inverse logsum parameter associated with the lower-level nests,

$\Gamma_n = \ln\left(\sum_{i' \in N_j} \exp\left(\mu_n V_{i'}\right)\right)$ and

$\Gamma_m = \ln\left(\sum_{n' \in N_m} \exp\left(\frac{\mu_m}{\mu_n}\Gamma_{n'}\right)\right)$.

In order to be consistent with utility maximization theory, both upper and lower-level inverse logsum parameters must be greater than one and the lower-level inverse logsum parameter must be greater than the upper-level inverse logsum parameter. The requirement that the lower-level inverse logsum parameter be greater than the upper-level inverse logsum parameter implies that itineraries within the same lower-level nest (and hence within the same upper-level nest) share the most attributes and compete more closely with each other than with other itineraries. Itineraries sharing a common upper-level nest (but not a lower-level nest) have less competition among themselves than with itineraries that share the same lower-level nest, but a greater level of competition than with itineraries in a different upper-level nest.

Of the six three-level NL models estimated, only two satisfied the inverse logsum conditions described above. These models, reported in Table 5, are for time, level-of-service and time, carrier. A visual representation of the three-level time, level-of-service NL model is shown in Figure 7 and the three-level time, carrier NL model is shown in Figure 8. The time, carrier model rejects both the time and carrier two-level NL models at the 0.001 level after adjustment.

The time, level-of-service model does not reject the two-level time NL model after adjustment. However, it does improve upon the two-level NL time model before adjustment and both its inverse logsum parameters are significant after adjustment (they are significantly different from each other after adjustment as well). Regardless, the marginal significance of this model implies that this three-level nesting specification may not be valid.

These three-level NL results indicate that there is moderate itinerary competition among itineraries sharing a common time period and greater competition among itineraries sharing both time period and carrier or (to a lesser extent) time period and level-of-service. This demonstrates the importance of conditioning the within carrier (level-of-service) competition dynamic by time period. However, note that the three-level time, carrier NL model implies that itineraries of the same carrier, but of different time periods, do not exhibit much competition amongst themselves.

4.7 Summary and Conclusions

This section shows that the competition among air-travel itineraries is not “uniform”. Thus, itinerary share models employing multinomial logit methodology are not adequate. Two-level nested logit models are estimated showing that itineraries sharing a common time period or carrier (but not level-of-service) exhibit a strong amount of competition amongst themselves. Using these results, a two-level weighted nested logit model with parallel time and carrier nesting structures is estimated. It significantly rejects the standard two-level NL models and has advantages over the more restrictive NL model structure.

Three-level nested logit models are estimated. The results of these models show that itineraries sharing a common time period have a moderate amount of competition amongst themselves, while itineraries sharing both time period and carrier (and to a lesser extent time period and level-of-service) exhibit a strong amount of competition amongst themselves.

Finally, while the estimations for models with itineraries nested by level-of-service were generally not significant, it is still reasonable to expect that increased competition exists within level-of-service nests (especially when the level-of-service nests are within upper-level time period nests as demonstrated by the marginal significance of the time, level-of-service three-level NL model). Regardless, it appears that the underlying competition among air-travel itineraries can almost fully be described by nesting itineraries by the departure time and carrier dimensions.

5. MODELING THE PROXIMATE COVARIANCE PROPERTY OF AIR-TRAVEL ITINERARIES ALONG THE TIME OF DAY DIMENSION

5.1 Introduction

The nested logit models estimated in Section 4 demonstrate the importance of considering the differential competition among air-travel itineraries connecting airport-pairs. In particular, structures were estimated showing that inter-itinerary competition is differentiated by departure time, carrier and (to a lesser extent) level-of-service.

These nested logit model structures (those that nest itineraries by departure time) group itineraries by arbitrary discrete time periods, however. This imposes unrealistic constraints on the departure time-of-day competition dynamic; for example, it implies that an itinerary within a given nest will compete more closely with an itinerary sharing the nest than with an itinerary in an adjacent nest that is closer in departure time.

The models presented in this section capture a more complicated and realistic itinerary competition structure (for the time of day dimension) than the variations of the nested logit model. It is hypothesized that – within an airport-pair – the amount of competition between itineraries is differentiated by the proximity in their departure times. This property, named “proximate covariance” by Small (1987), implies that itineraries that are “closer” to each other (by departure time) exhibit a higher amount of substitution/competition with each other than with itineraries that are more separated in time. The level of substitution/competition between itineraries increases the closer they are to each other. Models estimated in this section capture this property by grouping (nesting) itineraries (according to their departure times) into narrow time periods and ordering these time periods from early morning to late evening. These models are consistent with the hypothesis that an itinerary will compete most closely with itineraries in the same narrow time period and less closely as the difference in time periods increases.

This section begins by estimating several ordered generalized extreme value models (Small 1987). Small's development of the OGEV model was for the case of distinctly ordered alternatives (*e.g.* a household decision scenario of how many automobiles to own). In the current application, the OGEV structure is used to model the underlying competition among air-travel itineraries (for a given airport-pair-day-of-the-week) along the time of day dimension. As will be shown in the next subsection, the nesting structure of these models consists of overlapping time periods where each itinerary is allocated to contiguous nests according to allocation parameters. The values and significance of these allocation parameters indicate whether the assumption underlying the OGEV model (i.e. the proximate covariance property) is valid.

As demonstrated in Subsection 4.6, it is desirable to model the intra-carrier (and potentially intra-level-of-service) competition dynamic within an upper-level time of day structure. The OGEV models described in the preceding paragraph cannot accomplish this (since they only model the inter-itinerary competition dynamic along the time of day dimension). To rectify this, "hybrid" OGEV models are estimated in this section. These models incorporate the traditional OGEV model structure (described above) at the upper level with a GEV component such as the NL model at the lower level.

5.2 Ordered Generalized Extreme Value Model

A visual representation (for a generic airport-pair-day-of-the-week) of an OGEV model with six time periods in which each itinerary is allocated to two nests is presented in Figure 9. A visual representation of an OGEV model with eight time periods in which each itinerary is allocated to three nests is presented in Figure 10. The share of passengers assigned to each itinerary between an airport-pair for a given day of the week is as follows:

$$S_{i\subset k}=\sum_{j=k}^{k+M}P\left(i\subset k\mid N_j\right)P\left(N_j\right) \tag{5.1}$$

where $i\subset k$ indicates that itinerary i departs during time period k,

$M+1$ is the number of nests to which each itinerary is allocated,

N_j is nest j that includes alternative i (where $j=1,2,\ldots,K+M$),

K is the total number of time periods,

$P\left(i\subset k\mid N_j\right)$ is the probability of choosing alternative i from nest j and

$P\left(N_j\right)$ is the (unobserved) probability of choosing nest j.

The components of equation (5.1) can be expanded in terms of the probability of choosing a specific itinerary, i, from nest j to which it is allocated as follows:

$$P\left(i\subset k\mid N_j\right)=\frac{\alpha_{j-k}\exp\left(\mu V_i\right)}{\sum_{i'\subset k'\in N_j}\alpha_{j-k'}\exp\left(\mu V_{i'}\right)} \tag{5.1a}$$

where $\sum_{i'\subset k'\in N_j}$ is the summation over all itineraries, i', belonging to nest j,

α_{j-k} is the allocation parameter for an itinerary belonging to time period k assigned, in part, to nests $j=k,k+1,\ldots,k+M$ subject to $\alpha_i\geq 0$ and $\sum_{i=0}^{M}\alpha_i=1$,

μ is the inverse logsum parameter associated with the nests and

$V_{i'}$ is the deterministic portion of the utility for alternative i'.

and the probability of choosing nest j is as follows:

$$P\left(N_j\right)=\frac{\exp\left(\frac{1}{\mu}\Gamma_{N_j}\right)}{\sum_{\forall N_m}\exp\left(\frac{1}{\mu}\Gamma_{N_m}\right)} \tag{5.1b}$$

where $\Gamma_{N_j}=\ln\left(\sum_{i'\subset k'\in N_j}\alpha_{j-k'}\exp\left(\mu V_{i'}\right)\right).$

The estimation results for the models represented in Figures 9 and 10 are reported in Table 6[13]. The significance of the inverse logsum parameters (relative to one) and the allocation parameters (relative to both zero and one) indicate increased inter-itinerary competition for both within and proximate time periods. Both models reject the two-level NL time model (Table 4) at the 0.001 level after adjustment. Additionally, these OGEV models are behaviorally superior to the nested logit model since they allow for differential itinerary competition across time period boundaries. The statistical and behavioral superiority of these OGEV models confirms the belief that itinerary competition is differentiated by proximity in departure time (higher competition with close proximity).

The three-allocation OGEV model significantly rejects the two-allocation OGEV model at all levels of significance after adjustment. Additionally, it is behaviorally superior since (for a given itinerary) it yields four differential "levels" of inter-itinerary competition: itineraries sharing the same time period, itineraries in adjacent time periods, itineraries that are separated by two time periods and itineraries that are separated by three or more time periods (the two-allocation OGEV model allows for three differential levels of inter-itinerary competition).

---

[13] In Table 6, all parameter estimates significant at the 0.05 level after the adjustment procedure. The significance of value function parameter estimates is with respect to their hypothesized value of zero; the significance of inverse

Examining the cross-elasticity equations of this three-allocation OGEV model for the change in the probability of itinerary j due to changes in an attribute of itinerary i illustrates these relationships. If itinerary i is three or more time periods away from itinerary j, the elasticity is given by:

$$\eta_{X_{im}}^{P_j} = \frac{\partial P_j}{\partial X_{im}} \frac{X_{im}}{P_j} = -X_{im} \beta_m P_i \tag{5.2}$$

where X_{im} is the value of itinerary $i's$ m^{th} attribute and β_m is the parameter corresponding to attribute m. This is the same elasticity formula as that obtained for the MNL model. However, if itinerary j belongs to time period k and itinerary i belongs to time period $(k-2)$[14], the elasticity is given by:

$$\eta_{X_{im}}^{P_j} = -X_{im} \beta_m \left[P_i + \frac{(\mu - 1) P(i \mid N_k) P(j \mid N_k) P(N_k)}{P_j} \right] \tag{5.3}$$

This elasticity is larger in magnitude than the elasticity in equation (5.2) since μ must be larger than one. Next, if itinerary j belongs to time period k and itinerary i belongs to time period $(k-1)$[15], the elasticity is given by:

$$\eta_{X_{im}}^{P_j} = -X_{im} \beta_m \times \left[P_i + \frac{(\mu - 1) \left[P(j \mid N_k) P(i \mid N_k) P(N_k) + P(j \mid N_{k+1}) P(i \mid N_{k+1}) P(N_{k+1}) \right]}{P_j} \right] \tag{5.4}$$

---

logsum parameter estimates is with respect to their hypothesized value of one; the significance of allocation parameter estimates is with respect to their hypothesized values of zero and one.

[14] An analogous formula applies if itinerary i belongs to time period $(k+2)$.

[15] An analogous formula applies if itinerary i belongs to time period $(k+1)$.

This elasticity is larger in magnitude than the elasticities in equations (5.2) and (5.3). Finally, if itineraries j and i both belong to time period k, the elasticity is given by:

$$\eta_{X_{im}}^{P_j} = -X_{im}\beta_m \times \left[P_i + \frac{(\mu-1)\left[\begin{array}{l} P(j \mid N_k)P(i \mid N_k)P(N_k) + P(j \mid N_{k+1})P(i \mid N_{k+1})P(N_{k+1}) \\ +P(j \mid N_{k+2})P(i \mid N_{k+2})P(N_{k+2}) \end{array}\right]}{P_j} \right] \tag{5.5}$$

This elasticity is larger in magnitude than the elasticities in equations (5.2 – 5.4).

Hybrid OGEV specifications incorporating inter-itinerary competition along the carrier or level-of-service dimensions, under the time dimension, are estimated in the following subsection.

5.3 Three-Level Nested Logit Ordered Generalized Extreme Value Model

Following the results obtained from the three-level NL models in Subsection 4.6, three-level nested logit ordered generalized extreme value (NL-OGEV) models are estimated where the OGEV model structure is incorporated in the upper level of the three-level NL model structure. These models have itineraries allocated to nests at the upper level according to an OGEV structure and nested at the lower level by carrier or level-of-service. Visual representations of a three-level time, carrier NL-OGEV model (with itineraries allocated to two OGEV nests) and a three-level time, carrier NL-OGEV model (with itineraries allocated to three OGEV nests) are shown in Figures 11 and 12, respectively. Similar representations would show the corresponding three-level time, level-of-service NL-OGEV models.

For the three-level time, carrier NL-OGEV models, the share of passengers assigned to each itinerary between an airport-pair for a given day of the week is given by:

$$\begin{aligned} S_{i\subset k,c} &= \sum_{j=k}^{k+M} P\left(OGEV_j\right) P\left(c \mid OGEV_j\right) P\left(i \mid c, OGEV_j\right) \\ &= \sum_{j=k}^{k+M} \frac{\exp\left(\frac{1}{\mu_{OGEV}}\Gamma_j\right)}{\sum_{j'\in J}\exp\left(\frac{1}{\mu_{OGEV}}\Gamma_{j'}\right)} \times \frac{\exp\left(\frac{\mu_{OGEV}}{\mu_{NL}}\Gamma_c\right)}{\sum_{c'\in C}\exp\left(\frac{\mu_{OGEV}}{\mu_{NL}}\Gamma_{c'}\right)} \times \frac{\alpha_{j-k}\exp\left(\mu_{NL}V_i\right)}{\sum_{i'\subset k',c}\alpha_{j-k'}\exp\left(\mu_{NL}V_{i'}\right)} \end{aligned} \tag{5.6}$$

where $P\left(OGEV_j\right)$ is the passenger share assigned to the j^{th} upper-level OGEV nest,

$P\left(c \mid OGEV_j\right)$ is the passenger share assigned to carrier $c's$ lower-level NL nest given the j^{th} upper-level OGEV nest,

$P\left(i \mid c, OGEV_j\right)$ is the passenger share assigned to itinerary i given lower-level carrier NL nest c and upper-level OGEV nest j,

μ_{OGEV} is the inverse logsum parameter associated with the upper-level OGEV nests,

μ_{NL} is the inverse logsum parameter associated with the lower-level carrier NL nests,

$\Gamma_c = \ln\left(\sum_{i'\subset k',c}\alpha_{j-k'}\exp\left(\mu_{NL}V_{i'}\right)\right)$ and

$\Gamma_j = \ln\left(\sum_{c'\in C}\exp\left(\frac{\mu_{OGEV}}{\mu_{NL}}\Gamma_{c'}\right)\right)$.

and similarly for three-level time, level-of-service NL-OGEV models. Consistent with the three-level NL model, the OGEV and NL inverse logsum parameter estimates must be greater than one and the lower-level NL inverse logsum parameter must be larger than the upper-level OGEV inverse logsum parameter.

The estimation results for the three-level NL-OGEV models are reported in Table 7[16]. The three-level time, level-of-service (two-allocation) NL-OGEV model does not reject the two-allocation OGEV model (Table 6) after adjustment. However, it does improve upon the two-allocation OGEV model before adjustment, both its inverse logsum parameters are significant after adjustment (they are significantly different from each other after adjustment as well), and the allocation parameter is significant after adjustment.

The three-level time, carrier (two-allocation) NL-OGEV model rejects the three-level time, carrier NL model (Table 5) and the two-allocation OGEV model (Table 6) at the 0.001 level (before and after adjustment).

The time, carrier (three-allocation) NL-OGEV model rejects the three-level time, carrier NL model (Table 5), the three-allocation OGEV model (Table 6) and the time, carrier (two-allocation) NL-OGEV model (Table 7) at the 0.001 level (before and after adjustment). Of the models presented up to this point in the paper, the three-level time, carrier (three-allocation) NL-OGEV model has the best overall model statistics (by far). Additionally, its inverse logsum and allocation parameter estimates are all highly significant after adjustment. This indicates a high level of competition among itineraries flown by the same carrier within the same, adjacent or plus/minus two time periods.

These strong three-level NL-OGEV results indicate that imposing an upper-level OGEV structure and a lower-level NL structure on the itinerary competition dynamic dramatically improves upon the more rigid NL model structure. In addition to reinforcing the finding from Section 4 that the within carrier and (to a lesser extent) within level-of-service competition dynamic should be conditioned by time period, the significance of the OGEV allocation parameters in these models indicate that itineraries do indeed have several differential levels of competition (with respect to departure time) with other itineraries. That is, the closer itineraries are to each other (with respect to departure time) the more they will compete with each other.

---

[16] The three-level time, level-of-service (three-allocation) NL-OGEV model did not yield reasonable results. Additionally, in Table 7, parameter estimates in bold not significant at the 0.05 level after the adjustment procedure.

5.4 Summary and Conclusions

This section presents models capturing complicated and realistic inter-itinerary competition dynamics along the time of day dimension. Each of the models tests the hypothesis that air-travel itineraries (for a given airport-pair-day-of-the-week) exhibit proximate covariance; that is, the amount of competition (substitution) between itineraries increases as the difference in their departure times decreases. Variations of the nested logit model are not capable of capturing this phenomenon.

Two and three-allocation OGEV models are estimated. Both of these models show that air-travel itineraries do indeed exhibit the proximate covariance property. The three-allocation OGEV model captures, for each itinerary, four differential levels of competition with respect to other itineraries in its airport-pair along the time of day dimension (depending on the proximity of the itineraries' departure times). These OGEV models are the first in the aviation demand literature to capture the proximate covariance property.

Advanced hybrid OGEV models are estimated incorporating an OGEV structure at the upper level with a GEV (in particular, NL) structure on the lower level. In addition to capturing the proximate covariance property of air-travel itineraries, these models also measure differential inter-itinerary competition dynamics along the carrier or level-of-service dimensions. Of these models, the three-level time, carrier (three-allocation) NL-OGEV model yielded superior model statistics and behavioral interpretations. This model is the preferred specification of the paper.

Finally, the models estimated in this section are shown to have advantages over the more restrictive model structures presented in Section 4, leading to a clearer understanding of the air-travel itinerary competition dynamic.

6. FINAL DISCUSSION

Though the purpose of this research is to present the benefits of logit-based air-travel itinerary share models, it should be noted that there are certain advantages to QSI-based models. The advantages of QSI models lie in their simplicity, intuitive interpretation of model parameters and adequate forecasting ability. Logit model estimations on the other hand (in particular, the models presented in this paper) require large amounts of data, specialized software, and a trained analyst to estimate the model and interpret the model results. Additionally, by definition, when using a logit-based itinerary share model the analyst cannot “turn off” variables as is possible with QSI-based models. However, due to their simultaneous and optimal estimation of parameters (allowing for variable interactions to be captured), forecasting superiority and (most importantly) ability to model complex itinerary substitution patterns (allowing for rapidly changing market conditions to be accurately modeled), the advantages of logit-based models far outweigh the disadvantages.

The first goal of this research was to understand the impact of different air-carrier service attributes on itinerary share. The presented models contain independent variable specifications linking itinerary and airline characteristics to itinerary share. Explanatory variables used in this study such as level-of-service indicators differing by market type, connection quality variables, departure time variables, and equipment size and type variables have previously not been reported in the literature.

The second goal of this research was to model the underlying competitive dynamic among air-travel itineraries. To accomplish this, increasingly complex generalized extreme value specifications (in particular, variations of the nested logit and ordered generalized extreme value models) are used. Some of these structures have never been used in the aviation demand literature and some are new to the logit share literature in general. The advanced models estimated in this paper have advantages over the more restrictive model structures and outperform these models with respect to forecasting accuracy, statistical tests and behavioral interpretations, leading to a clearer understanding of the air-travel itinerary competition dynamic.

ACKNOWLEDGMENT

A Dissertation Year Fellowship awarded to the author from Northwestern University's Transportation Center supported this research. The author is very thankful for this support.

REFERENCES

Algers, S., M. Beser, 2001. Modelling Choice of Flight and Booking Class – A Study using Stated Preference and Reveled Preference Data. International Journal of Services Technology and Management 2 (1/2), 28-45.

Aptech Systems, Inc. 2005. GAUSS. Maple Valley, WA, USA.

Ben-Akiva, M., T. Morikawa and F. Shiroishi, 1989. Analysis of the Reliability of Stated Preference Data in Estimating Mode Choice Models. Selected Proceedings of the Fifth World Conference on Transport Research 4, 263-277.

Bresnahan, T. F., S. Stern and M. Trajtenberg, 1997. Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s. RAND Journal of Economics 28 (0, Special Issue), S17-S44.

Bruning, E., V. Rueda, 2000. The Importance of Frequent Flyer Programs as a Barrier to Entry into Air Travel Markets. Journal of Transportation Law, Logistics and Policy 67 (4), 367-380.

Coldren, G. M., F. S. Koppelman, 2005. Modeling the Competition among Air-travel Itinerary Shares: GEV Model Development. Transportation Research – Part A 39 (4), 345-365.

Coldren, G. M., F. S. Koppelman, 2005. Modeling the Proximate Covariance Property of Air-travel Itineraries along the Time of Day Dimension. Transportation Research Record, In Press.

Coldren, G. M., F. S. Koppelman, K. Kasturirangan and A. Mukherjee, 2003. Modeling Aggregate Air-travel Itinerary Shares: Logit Model Development at a Major U.S. Airline. Journal of Air Transport Management 9 (6), 361-369.

Data Base Products, Inc. 2000, 2001. Superset. Dallas, TX, USA.

Domencich, T. A., D. McFadden, 1975. Urban Travel Demand – A Behavioral Analysis. American Elsevier Publishing Company, Inc., New York.

Expedia, Inc. 2000, 2001. Expedia.com. Bellevue, WA, USA.

Galileo International 2000, 2001. Galileo CRS. Parsippany, NJ, USA.

Ghobrial, A., S. Y. Soliman, 1992. An Assessment of Some Factors Influencing the Competitive Strategies of Airlines in Domestic Markets. International Journal of Transport Economics 19 (3), 247-258.

Manski, C. F., 1973. The Analysis of Qualitative Choice. Ph.D. Dissertation, The Massachusetts Institute of Technology.

McFadden, D., 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In: Frontiers in Econometrics. P. Zarembka (Ed.). Academic Press, New York, 105-142.

McFadden, D., 1978. Modeling the Choice of Residential Location. Transportation Research Record 673, 72-77.

Morash, E. A., J. Ozment, 1996. The Strategic Use of Transportation Time and Reliability for Competitive Advantage. Transportation Journal 36 (2), 35-46.

Morrison, M., 2000. Aggregation Biases in Stated Preference Studies. Australian Economic Papers 39 (2), 215-230.

Murphy, J. J., P. G. Allen, T. H. Stevens and D. Weatherhead, 2005. A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation. Environmental and Resource Economics, In Press.

Nako, S. M., 1992. Frequent Flyer Programs and Business Travellers: An Empirical Investigation. The Logistics and Transportation Review 28 (4), 395-414.

Nason, S. D., 1981. The Airline Preference Problem: An Application of Disaggregate Logit. Presented at the AGIFORS Symposium, Santa Barbara, California.

OAG Worldwide Limited 2000, 2001. Official Airline Guide. Bedfordshire, LU5 4HB, United Kingdom.

Orbitz, L.L.C. 2000, 2001. Orbitz.com. Chicago, IL, USA.

Priceline.com, Inc. 2000, 2001. Priceline.com. Norwalk, CT, USA.

Proussaloglou, K., F. Koppelman, 1995. Air Carrier Demand: An Analysis of Market Share Determinants. Transportation 22 (4), 371-388.

Proussaloglou, K., F. S. Koppelman, 1999. The Choice of Air Carrier, Flight, and Fare Class. Journal of Air Transport Management 5 (4), 193-201.

Sabre Travel Network 2000, 2001. Sabre GDS. Southlake, TX, USA.

Small, K. A., 1987. A Discrete Choice Model for Ordered Alternatives. Econometrica 55 (2), 409-424.

Suzuki, Y., J. E. Tyworth and R. A. Novack, 2001. Airline Market Share and Customer Service Quality: A Reference-dependent Model. Transportation Research – Part A 35 (9), 773-788.

Transportation Research E-Circular E-C040, 2002. Aviation Demand Forecasting: A Survey of Methodologies. Transportation Research Board, Washington, D.C.

Travelocity.com, L.P. 2000, 2001. Travelocity.com. Dallas, TX, USA.

Wen, C-H, F. S. Koppelman, 2001. The Generalized Nested Logit Model. Transportation Research – Part B 35 (7), 627-641.

Worldspan, L.P. 2000, 2001. Worldspan CRS. Atlanta, GA, USA.

Yoo, K. E., N. Ashford, 1996. Carrier Choices of Air Passengers in Pacific Rim: Using Comparative Analysis and Complementary Interpretation of Revealed Preference and Stated Preference Data. Transportation Research Record 1562, 1-7.

TABLE 1: Description of Explanatory Variables

Variable	Description
Level-of-Service	Dummy variable representing the level-of-service of the itinerary (nonstop, direct, single-connect or double-connect) with respect to the best level-of-service available in the airport-pair.
Second-Best Connection	For connection itineraries sharing a common leg, a dummy variable indicating that the itinerary is not the best connection (with respect to ground time) for the given incoming or outgoing leg at a transfer airport.
Second-Best Connection Time Difference	If the second-best connection indicator equals one, this variable measures the ground time difference between the itinerary and the best connection itinerary.
Best Connection Time Difference	Elapsed time difference between an itinerary involving a stop or connection and the fastest itinerary involving a stop or connection for each airport-pair independent of transfer airport.
Distance Ratio	Itinerary distance divided by the shortest itinerary distance for the airport-pair multiplied by 100.
Point of Sale Weighted Airport Presence	Carrier origin and destination presence (determined by percentage of departures) weighted by industry airport-pair point-of-sale percentages divided by 100 to give units of percent from 0 to 100.
Fare Ratio	Carrier average fare divided by the industry average fare for the airport-pair multiplied by 100.
Carrier	Dummy variable representing carriers having more than 0.5% of itineraries in the entity. All other carriers are combined together in a single category.
Code share	Dummy variable indicating whether any leg of the itinerary was booked as a code share.
Regional Jet	Dummy variable indicating whether the smallest aircraft on any part of the itinerary is a regional jet.
Propeller Aircraft	Dummy variable indicating whether the smallest aircraft on any part of the itinerary is a propeller aircraft.
Mainline Jet Seats	If an itinerary involves neither a regional jet nor a propeller aircraft leg, this variable measures the number of seats on the smallest aircraft for the itinerary.
Regional Jet Seats	If an itinerary includes a regional jet leg (but no propeller aircraft leg), this variable measures the number of seats on the smallest regional jet aircraft for the itinerary.
Propeller Aircraft Seats	If an itinerary includes a propeller aircraft leg, this variable measures the number of seats on the smallest propeller aircraft for the itinerary.
Departure Time	Dummy variable for each hour of the day (based on the local departure time of the first leg of the itinerary).

TABLE 2: MNL Itinerary Share Models

Explanatory Variables			**Entity**		
	E-E	**E-C**	**E-M**	**E-W**	**W-E**
Level-of-Service					
Nonstop Itinerary in Nonstop Market[17]	0.0000	0.0000	0.0000	0.0000	0.0000
Direct Itinerary in Nonstop Market	-1.6819	-1.5253	-1.7168	-1.6615	-1.5349
Single-Connect Itinerary in Nonstop Market	-3.1213	-2.8760	-2.7087	-2.9729	-2.8858
Double-Connect Itinerary in Nonstop Market	-7.7557	-7.2970	-7.5112	-7.2130	-6.5395
Direct Itinerary in Direct Market	0.0000	0.0000	0.0000	0.0000	0.0000
Single-Connect Itinerary in Direct Market	-0.7901	-0.7768	-1.0989	-1.0615	-0.9847
Double-Connect Itinerary in Direct Market	-4.9442	-4.5565	-4.3308	-4.4475	-4.5869
Single-Connect Itinerary in Single-Connect Market	0.0000	0.0000	0.0000	0.0000	0.0000
Double-Connect Itinerary in Single-Connect Market	-3.0953	-3.1205	-2.5431	-2.8209	-2.7105
Connection Quality					
Second-Best Connection	-0.5560	-0.5058	-0.5322	-0.7269	-0.4377
Second-Best Connection Time Difference	-0.0157	-0.0190	-0.0178	-0.0162	-0.0180
Best Connection Time Difference	-0.0108	-0.0094	-0.0093	-0.0104	-0.0088
Distance Ratio	-0.0125	-0.0116	-0.0210	-0.0173	-0.0207
Carrier Attributes					
Point of Sale Weighted Airport Presence	0.0024	0.0100	0.0078	0.0071	0.0077
Fare Ratio	-0.0018	-0.0038	-0.0045	-0.0035	-0.0040
Carrier Constants	--------	--------	--------	--------	--------
Code share	-1.5911	-2.1255	-1.9220	-2.1658	-2.2082
Aircraft Size and Type					
Mainline Jet	0.0000	0.0000	0.0000	0.0000	0.0000
Regional Jet	-0.7046	-0.8317	-0.8844	-0.7079	-0.4110
Propeller Aircraft	-1.2420	-1.0599	-0.8081	-0.9988	-0.9269
Mainline Jet Seats	0.0041	0.0037	0.0047	0.0032	0.0036
Regional Jet Seats	0.0117	0.0101	0.0144	0.0080	0.0052
Propeller Aircraft Seats	0.0246	0.0146	0.0157	0.0184	0.0212
Departure Time					
Midnight - 5 A.M.	-1.1653	-1.1577	-1.2341	-1.2037	-0.7228
5 – 6 A.M.	-0.4653	-0.4577	-0.5341	-0.5037	0.1482
6 – 7 A.M.	0.0000	0.0000	0.0000	0.0000	0.0000
7 – 8 A.M.	0.2865	0.2693	0.2179	0.2297	0.0570
8 – 9 A.M.	0.2836	0.4130	0.3983	0.3168	0.2354
9 –10 A.M.	0.2046	0.3204	0.4156	0.3413	0.0311
10 - 11 A.M.	0.1219	0.2571	0.4539	0.3425	0.0592
11 - 12 noon	0.1022	0.2539	0.3842	0.2670	0.0196
12 – 1 P.M.	0.1452	0.3137	0.4252	0.2977	0.0570
1 – 2 P.M.	0.1732	0.3676	0.2298	0.2594	0.0219
2 – 3 P.M.	0.2622	0.4742	0.1964	0.2462	-0.0988
3 – 4 P.M.	0.3438	0.5254	0.2136	0.2557	0.1175
4 – 5 P.M.	0.4230	0.5853	0.2476	0.2097	0.0873
5 – 6 P.M.	0.4667	0.5959	0.2419	0.2324	0.1779
6 – 7 P.M.	0.4054	0.5516	0.2491	0.1889	-0.2437
7 – 8 P.M.	0.2047	0.4627	-0.0099	0.0472	0.2914
8 – 9 P.M.	-0.0362	0.0866	-0.0984	-0.2243	0.1417
9 – 10 P.M.	-0.3565	-0.2917	-0.6878	-0.3924	0.0541
10 – Midnight	-0.6468	-0.7454	-0.7637	-0.3170	-0.0573
Rho-square w.r.t. Zero	0.377	0.369	0.372	0.396	0.385

---

[17] "Nonstop Market" means the "best" level-of-service available in the airport-pair is a nonstop itinerary, "Direct Market" means the best level-of-service available in the airport-pair is a direct itinerary, etc.

TABLE 3: Onboard Segment-Level Model Validation Analysis (Mean Absolute Percentage Deviation)

Month	QSI-Based Model	Logit-Based Model	Difference
January 1999	19.52	17.91	1.61
February 1999	18.98	17.21	1.77
March 1999	18.40	16.48	1.92
April 1999	17.31	15.94	1.37
May 1999	17.22	16.23	0.99
June 1999	18.33	16.87	1.46
July 1999	16.70	14.91	1.79
August 1999	16.41	14.57	1.84
September 1999	17.77	15.93	1.84
October 1999	18.13	15.66	2.47
November 1999	16.14	14.42	1.72
December 1999	19.28	17.73	1.55
August 2001	17.47	15.79	1.68
March 2002	17.91	16.14	1.77
May 2002	19.08	17.73	1.35
Average	**17.91**	**16.23**	**1.68**

TABLE 4: Itinerary Share Models: MNL, Two-Level NL's and Two-Level WNL

	Model			
Explanatory Variables	**MNL**	**2-Level NL Time**	**2-Level NL Carrier**	**2-Level WNL: Time | Carrier**
Level-of-Service				
Nonstop Itinerary in Nonstop Market	0.0000	0.0000	0.0000	0.0000
Direct Itinerary in Nonstop Market	-1.9595	-1.6271	-1.7798	-1.4253
Single-Connect Itinerary in Nonstop Market	-2.8371	-2.3540	-2.5695	-2.0624
Double-Connect Itinerary in Nonstop Market	-6.6264	-5.4663	-5.7872	-4.6364
Direct Itinerary in Direct Market	0.0000	0.0000	0.0000	0.0000
Single-Connect Itinerary in Direct Market	-0.7370	-0.6207	-0.6579	-0.5323
Double-Connect Itinerary in Direct Market	-3.9250	-3.2331	-3.4217	-2.7375
Single-Connect Itinerary in Single-Connect Market	0.0000	0.0000	0.0000	0.0000
Double-Connect Itinerary in Single-Connect Market	-2.6015	-2.1915	-2.3043	-1.8997
Connection Quality				
Second-Best Connection	-0.4208	-0.3331	-0.3216	-0.2396
Second-Best Connection Time Difference	-0.0087	-0.0071	-0.0074	-0.0058
Distance Ratio	-0.0135	-0.0109	-0.0131	-0.0103
Best Connection Time Difference	-0.0056	-0.0047	-0.0051	-0.0041
Carrier Attributes				
Fare Ratio	-0.0060	-0.0052	-0.0039	-0.0033
Carrier Constants (Proprietary)	--------	--------	--------	--------
Code share	-1.8601	-1.5241	-1.6861	-1.3383
Aircraft Type				
Mainline Jet	0.0000	0.0000	0.0000	0.0000
Regional Jet	-0.4560	-0.3856	-0.4225	-0.3464
Propeller Aircraft	-0.4201	-0.3496	-0.3658	-0.2919
Departure Time				
5 – 6 A.M.	-0.2184	-0.1931	-0.2084	-0.1814
6 – 7 A.M.	0.0000	0.0000	0.0000	0.0000
7 – 8 A.M.	0.1385	0.1118	0.1235	0.0964
8 – 9 A.M.	0.2381	0.1907	0.2150	0.1663
9 –10 A.M.	0.2646	0.2135	0.2365	0.1848
10 - 11 A.M.	0.2672	0.1873	0.2412	0.1619
11 - 12 noon	0.2290	0.1643	0.2168	0.1507
12 – 1 P.M.	0.2476	0.1761	0.2293	0.1593
1 – 2 P.M.	0.1614	0.1043	0.1507	0.0956
2 – 3 P.M.	0.1686	0.1058	0.1599	0.0982
3 – 4 P.M.	0.1856	0.1219	0.1709	0.1100
4 – 5 P.M.	0.0960	0.0486	0.0934	0.0523
5 – 6 P.M.	0.0972	0.0490	0.0840	0.0429
6 – 7 P.M.	0.1760	0.1179	0.1535	0.1007
7 – 8 P.M.	0.0833	0.0443	0.0857	0.0502
8 – 9 P.M.	-0.0803	-0.0807	-0.0563	-0.0541
9 – 10 P.M.	-0.2587	-0.2243	-0.2131	-0.1778
10 – Midnight	-0.3407	-0.3179	-0.2847	-0.2546
Inverse Logsum Parameter (Time)	--------	1.2244	--------	1.5435
Inverse Logsum Parameter (Carrier)	--------	--------	1.1768	1.4519
WNL Weight Parameter (Time Structure)	--------	--------	--------	0.5364
Log Likelihood at Zero	-2,173,197	-2,173,197	-2,173,197	-2,173,197
Log Likelihood at Convergence	-1,558,186	-1,557,443	-1,556,663	-1,555,632
Adjusted Log Likelihood at Convergence	-49,466	-49,443	-49,418	-49,385
Rho-square w.r.t. Zero	0.2830	0.2833	0.2837	0.2842

TABLE 5: Itinerary Share Models: Three-Level NL's

Explanatory Variables	Model 3-Level NL: Time, LOS	 3-Level NL: Time, Carrier
Level-of-Service		
Nonstop Itinerary in Nonstop Market	0.0000	0.0000
Direct Itinerary in Nonstop Market	-1.6479	-1.6570
Single-Connect Itinerary in Nonstop Market	-2.3401	-2.3802
Double-Connect Itinerary in Nonstop Market	-5.5099	-5.2215
Direct Itinerary in Direct Market	0.0000	0.0000
Single-Connect Itinerary in Direct Market	-0.5935	-0.6362
Double-Connect Itinerary in Direct Market	-3.2467	-3.1347
Single-Connect Itinerary in Single-Connect Market	0.0000	0.0000
Double-Connect Itinerary in Single-Connect Market	-2.2118	-2.1679
Connection Quality		
Second-Best Connection	-0.3290	-0.2504
Second-Best Connection Time Difference	-0.0071	-0.0063
Distance Ratio	-0.0108	-0.0112
Best Connection Time Difference	-0.0047	-0.0049
Carrier Attributes		
Fare Ratio	-0.0051	-0.0036
Carrier Constants (Proprietary)	--------	--------
Code share	-1.5082	-1.5408
Aircraft Type		
Mainline Jet	0.0000	0.0000
Regional Jet	-0.3827	-0.4019
Propeller Aircraft	-0.3459	-0.3294
Departure Time		
5 - 6 A.M.	-0.1925	-0.2152
6 - 7 A.M.	0.0000	0.0000
7 - 8 A.M.	0.1099	0.1163
8 - 9 A.M.	0.1880	0.1944
9 -10 A.M.	0.2108	0.2148
10 - 11 A.M.	0.1850	0.1971
11 - 12 noon	0.1625	0.1808
12 - 1 P.M.	0.1740	0.1845
1 - 2 P.M.	0.1029	0.1152
2 - 3 P.M.	0.1045	0.1175
3 - 4 P.M.	0.1199	0.1328
4 - 5 P.M.	0.0455	0.0577
5 - 6 P.M.	0.0457	0.0383
6 - 7 P.M.	0.1139	0.1068
7 - 8 P.M.	0.0416	0.0546
8 - 9 P.M.	-0.0818	-0.0586
9 - 10 P.M.	-0.2234	-0.1982
10 – Midnight	-0.3161	-0.2821
Upper-Level Inverse Logsum Parameter (Time)	1.2124	1.0667
Lower-Level Inverse Logsum Parameter (Carrier)	--------	1.3568
Lower-Level Inverse Logsum Parameter (LOS)	1.2376	--------
Log Likelihood at Zero	-2,173,197	-2,173,197
Log Likelihood at Convergence	-1,557,435	-1,554,227
Adjusted Log Likelihood at Convergence	-49,442	-49,341
Rho-square w.r.t. Zero	0.2833	0.2848

TABLE 6: Itinerary Share Models: Two and Three-Allocation OGEV's

Explanatory Variables	**Model** **2-Allocation OGEV**	 **3-Allocation OGEV**
Level-of-Service		
Nonstop Itinerary in Nonstop Market	0.0000	0.0000
Direct Itinerary in Nonstop Market	-1.6049	-1.5549
Single-Connect Itinerary in Nonstop Market	-2.3157	-2.2380
Double-Connect Itinerary in Nonstop Market	-5.3363	-5.1295
Direct Itinerary in Direct Market	0.0000	0.0000
Single-Connect Itinerary in Direct Market	-0.6106	-0.5905
Double-Connect Itinerary in Direct Market	-3.1604	-3.0386
Single-Connect Itinerary in Single-Connect Market	0.0000	0.0000
Double-Connect Itinerary in Single-Connect Market	-2.1618	-2.0988
Connection Quality		
Second-Best Connection	-0.3161	-0.2966
Second-Best Connection Time Difference	-0.0070	-0.0067
Distance Ratio	-0.0107	-0.0102
Best Connection Time Difference	-0.0046	-0.0044
Carrier Attributes		
Fare Ratio	-0.0051	-0.0050
Carrier Constants (Proprietary)	--------	--------
Code share	-1.4842	-1.4229
Aircraft Type		
Mainline Jet	0.0000	0.0000
Regional Jet	-0.3764	-0.3631
Propeller Aircraft	-0.3435	-0.3329
Departure Time		
5 - 6 A.M.	-0.1825	-0.1749
6 - 7 A.M.	0.0000	0.0000
7 - 8 A.M.	0.2335	0.2050
8 - 9 A.M.	0.3132	0.2808
9 -10 A.M.	0.3360	0.2993
10 - 11 A.M.	0.3069	0.2991
11 - 12 noon	0.2812	0.2532
12 - 1 P.M.	0.2978	0.2622
1 - 2 P.M.	0.2321	0.2052
2 - 3 P.M.	0.2285	0.1957
3 - 4 P.M.	0.2425	0.2348
4 - 5 P.M.	0.1883	0.1677
5 - 6 P.M.	0.1888	0.1686
6 - 7 P.M.	0.2507	0.2293
7 - 8 P.M.	0.0638	0.0756
8 - 9 P.M.	-0.0521	-0.0371
9 - 10 P.M.	-0.1814	-0.1612
10 – Midnight	-0.2607	-0.2412
Inverse Logsum Parameter	1.2607	1.3182
Alpha 1 (Allocation Parameter)	0.2215	0.0728
Alpha 2 (Allocation Parameter)	--------	0.2520
Log Likelihood at Zero	-2,173,197	-2,173,197
Log Likelihood at Convergence	-1,557,214	-1,556,869
Adjusted Log Likelihood at Convergence	-49,435	-49,424
Rho-square w.r.t. Zero	0.2834	0.2836

TABLE 7: Itinerary Share Models: Three-Level NL-OGEV's

Explanatory Variables		**Model**	
	3-Level NL-OGEV (2-Allocation): Time|LOS	**3-Level NL-OGEV (2-Allocation): Time|Carrier**	**3-Level NL-OGEV (3-Allocation): Time|Carrier**
Level-of-Service			
Nonstop Itinerary in Nonstop Market	0.0000	0.0000	0.0000
Direct Itinerary in Nonstop Market	-1.6459	-1.6060	-1.5840
Single-Connect Itinerary in Nonstop Market	-2.3144	-2.3175	-2.2837
Double-Connect Itinerary in Nonstop Market	-5.4438	-4.9888	-4.8732
Direct Itinerary in Direct Market	0.0000	0.0000	0.0000
Single-Connect Itinerary in Direct Market	-0.5766	-0.6203	-0.6148
Double-Connect Itinerary in Direct Market	-3.2063	-3.0155	-2.9596
Single-Connect Itinerary in Single-Connect Market	0.0000	0.0000	0.0000
Double-Connect Itinerary in Single-Connect Market	-2.2079	-2.1151	-2.0968
Connection Quality			
Second-Best Connection	-0.3123	-0.2144	-0.1936
Second-Best Connection Time Difference	-0.0069	-0.0059	-0.0057
Distance Ratio	-0.0107	-0.0107	-0.0101
Best Connection Time Difference	-0.0046	-0.0048	-0.0048
Carrier Attributes			
Fare Ratio	-0.0051	-0.0034	-0.0034
Carrier Constants (Proprietary)	--------	--------	--------
Code share	-1.4720	-1.4865	-1.4594
Aircraft Type			
Mainline Jet	0.0000	0.0000	0.0000
Regional Jet	-0.3749	-0.3947	-0.3894
Propeller Aircraft	-0.3409	-0.3176	-0.3138
Departure Time			
5 – 6 A.M.	-0.1829	-0.1911	-0.1922
6 – 7 A.M.	0.0000	0.0000	0.0000
7 – 8 A.M.	0.2356	0.2638	0.2308
8 – 9 A.M.	0.3151	0.3407	0.3092
9 –10 A.M.	0.3380	0.3622	0.3193
10 - 11 A.M.	0.3047	0.3424	0.3363
11 - 12 noon	0.2790	0.3205	0.2747
12 – 1 P.M.	0.2958	0.3370	0.2833
1 – 2 P.M.	0.2327	0.2645	0.2196
2 – 3 P.M.	0.2288	0.2572	0.2067
3 – 4 P.M.	0.2415	0.2718	0.2534
4 – 5 P.M.	0.1911	0.2201	0.1878
5 – 6 P.M.	0.1912	0.2060	0.1840
6 – 7 P.M.	0.2523	0.2678	0.2471
7 – 8 P.M.	0.0600	0.0948	0.1058
8 – 9 P.M.	-0.0531	**-0.0014**	**0.0122**
9 – 10 P.M.	-0.1794	-0.1141	-0.0944
10 – Midnight	-0.2585	-0.2018	-0.1837
Upper-Level OGEV Inverse Logsum Parameter (Time)	1.2325	1.0896	1.1087
Lower-Level NL Inverse Logsum Parameter (Carrier)	--------	1.4539	1.5196
Lower-Level NL Inverse Logsum Parameter (LOS)	1.2718	--------	--------
Alpha 1 (Allocation Parameter)	0.1903	0.1787	0.0205
Alpha 2 (Allocation Parameter)	--------	--------	0.2425
Log Likelihood at Zero	-2,173,197	-2,173,197	-2,173,197
Log Likelihood at Convergence	-1,557,199	-1,553,430	-1,552,661
Adjusted Log Likelihood at Convergence	-49,435	-49,315	-49,291
Rho-square w.r.t. Zero	0.2835	0.2852	0.2855

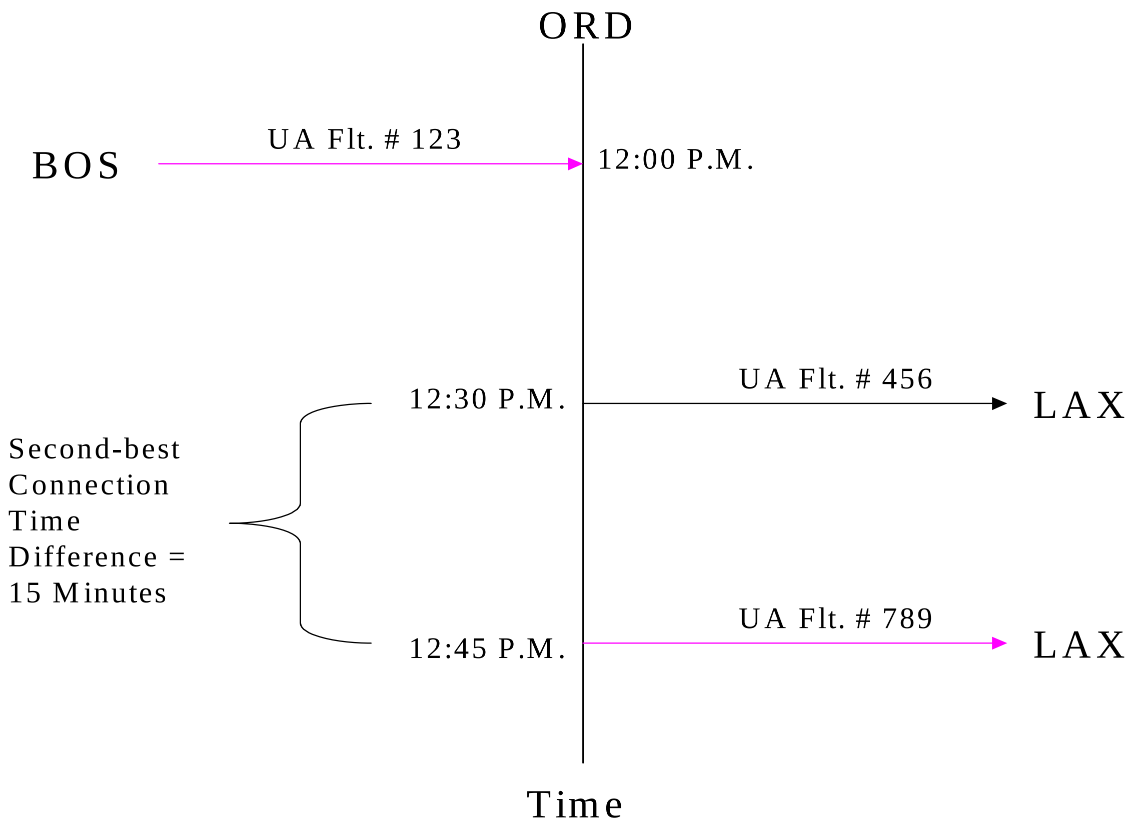

FIGURE 1: Visual Representation of Second-best Connection and Second-best Connection Time Difference

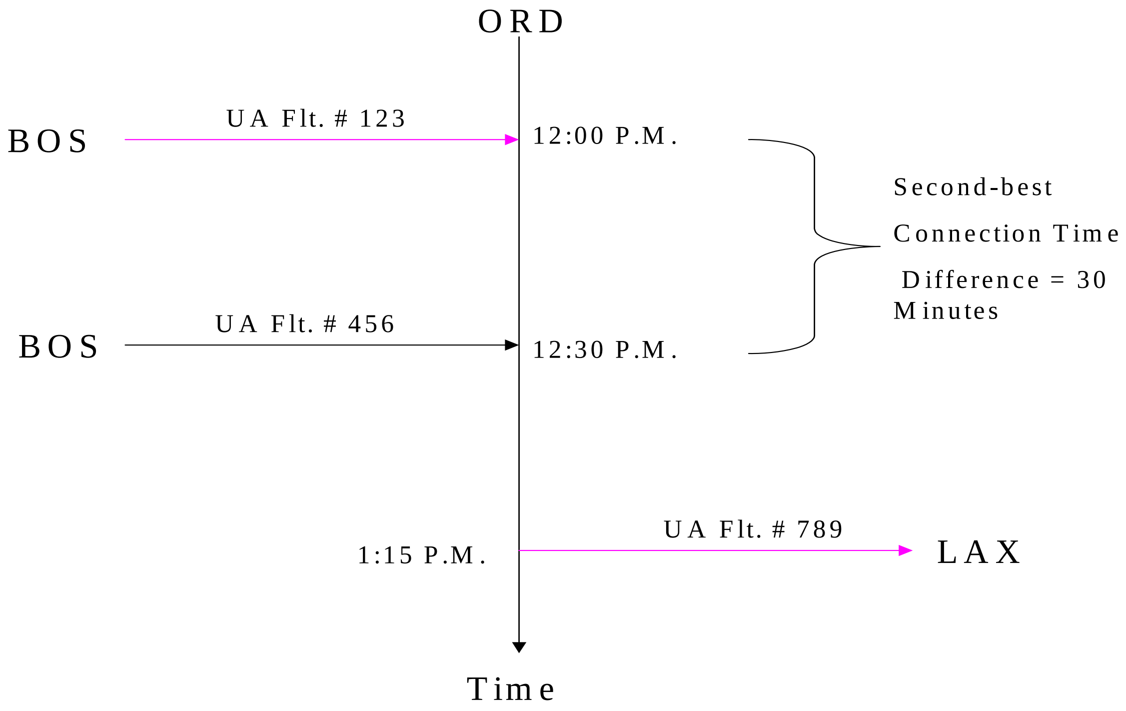

FIGURE 2: Visual Representation of Second-best Connection and Second-best Connection Time Difference

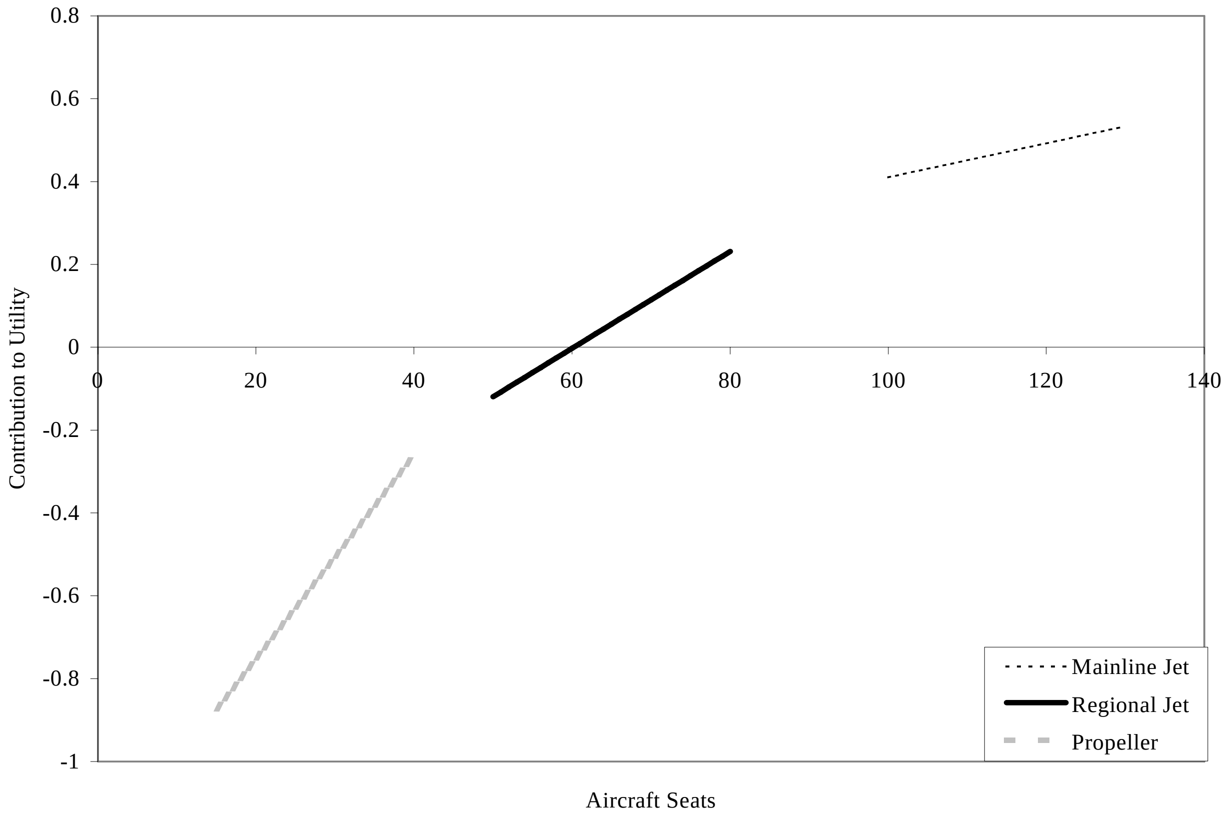

FIGURE 3: Contribution of Aircraft Size and Type to Itinerary Value for the East-East Entity

FIGURE 4: Two-Level NL Time Model Structure

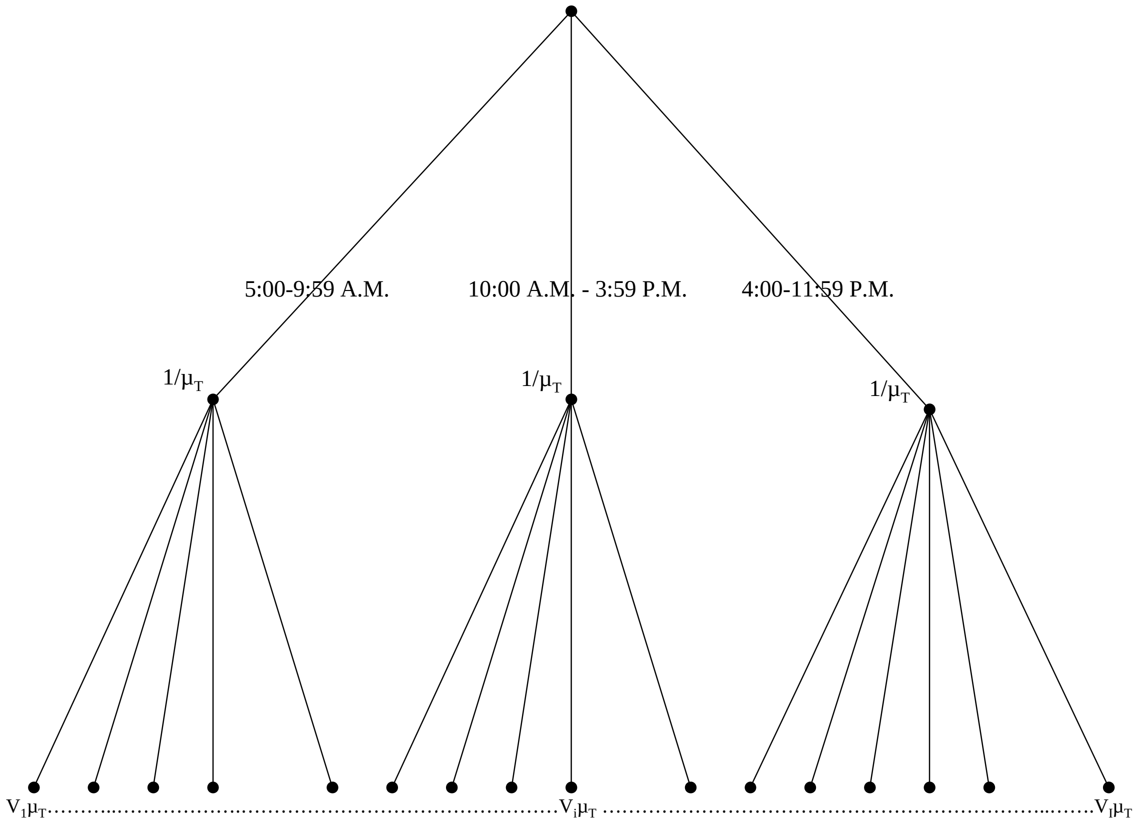

FIGURE 5: Two-Level NL Carrier Model Structure

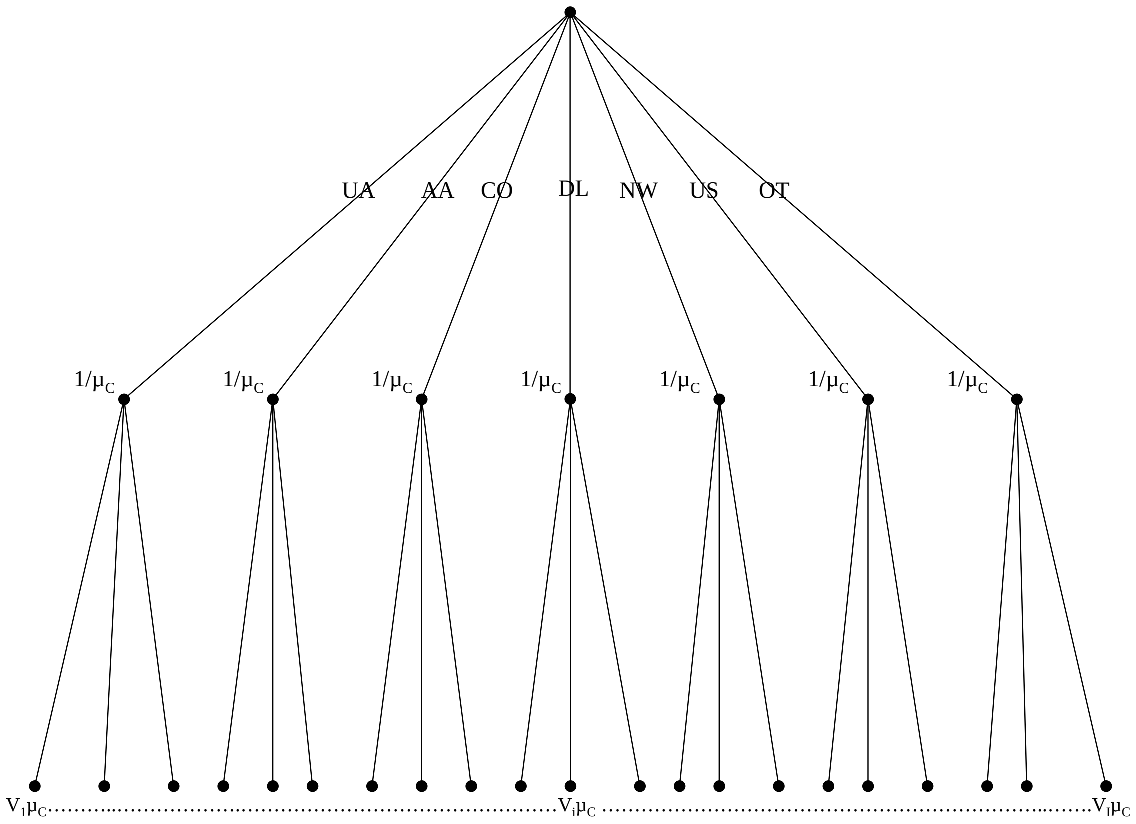

FIGURE 6: Two-Level WNL Time | Carrier Model Structure

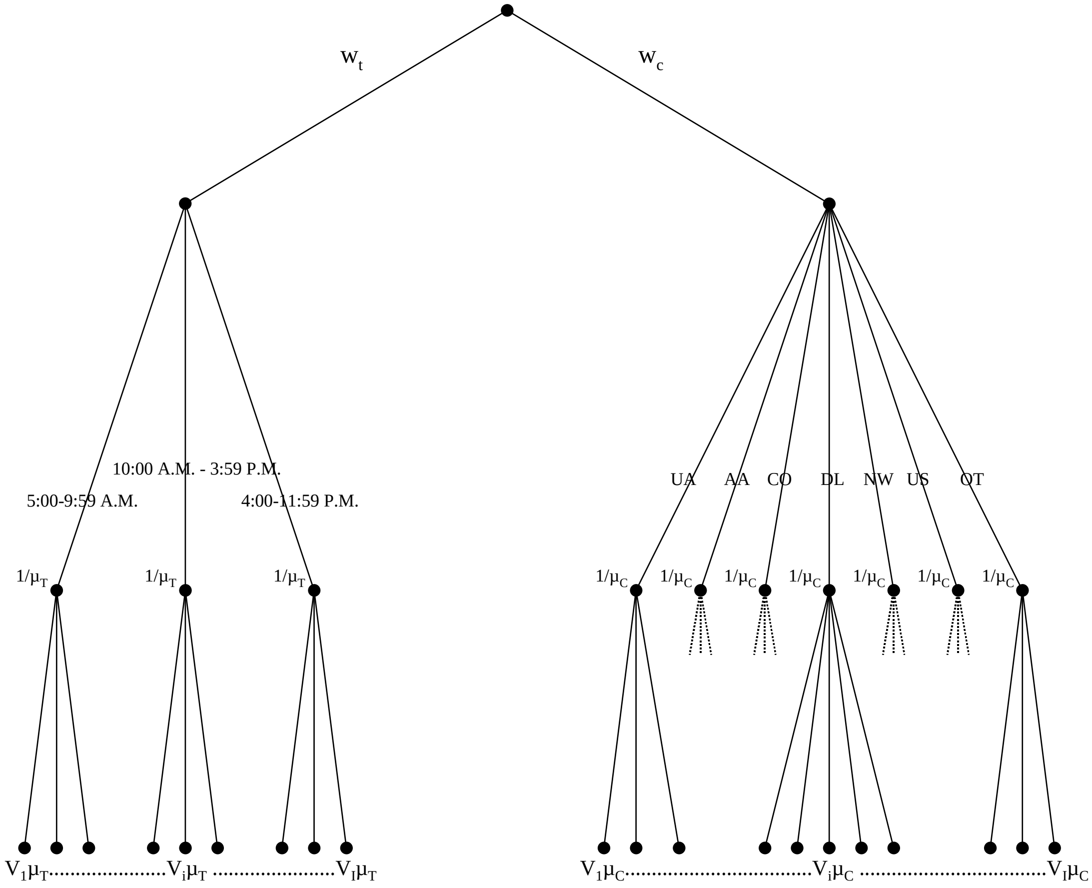

FIGURE 7: Three-Level NL Time, Level-of-Service Model Structure

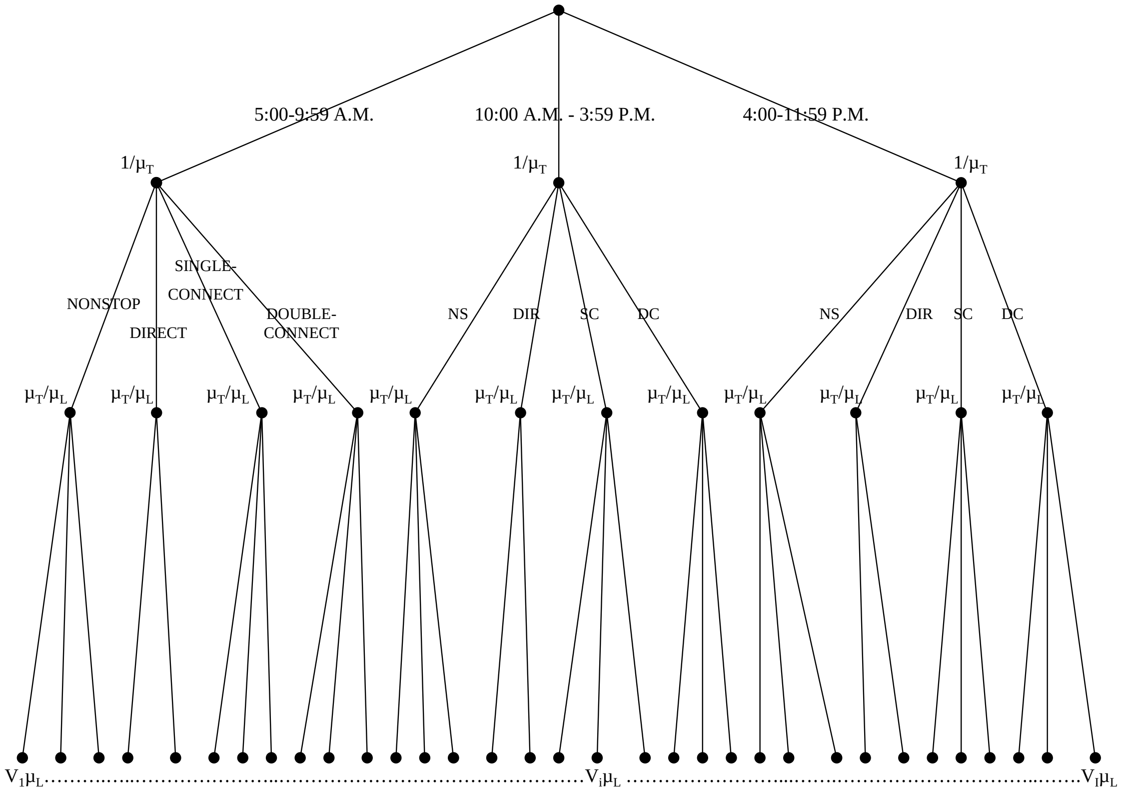

FIGURE 8: Three-Level NL Time, Carrier Model Structure

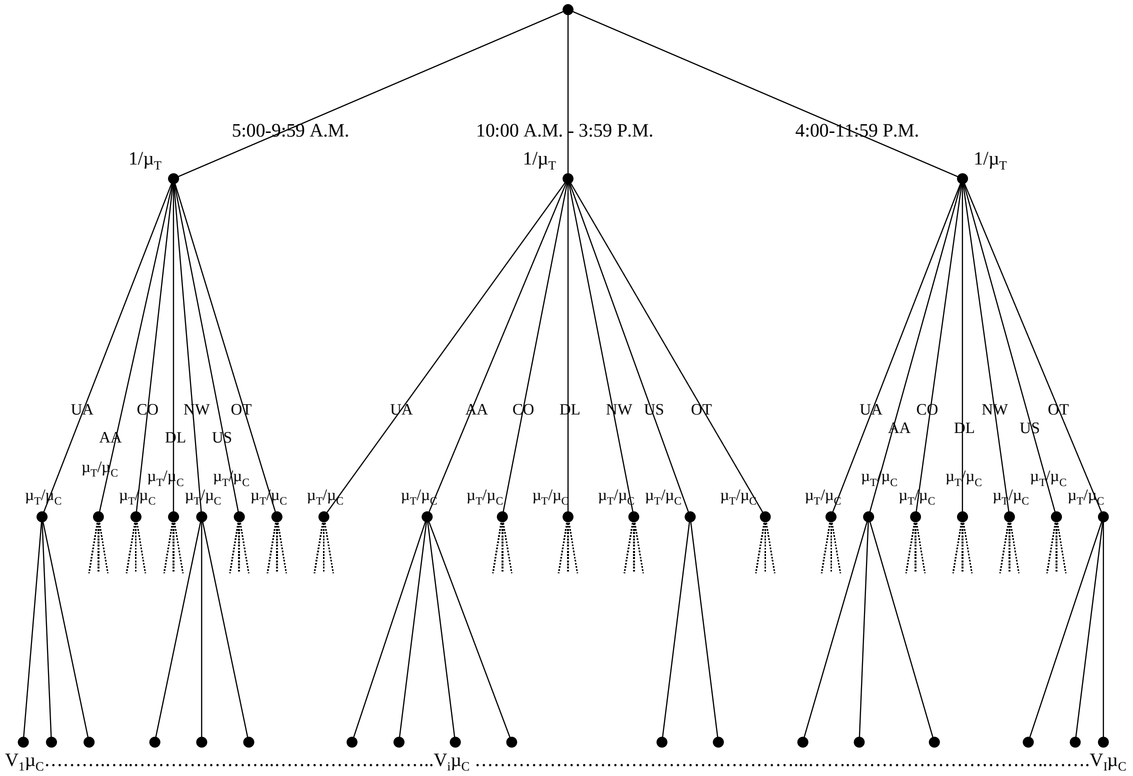

FIGURE 9: Two-Allocation OGEV Model Structure

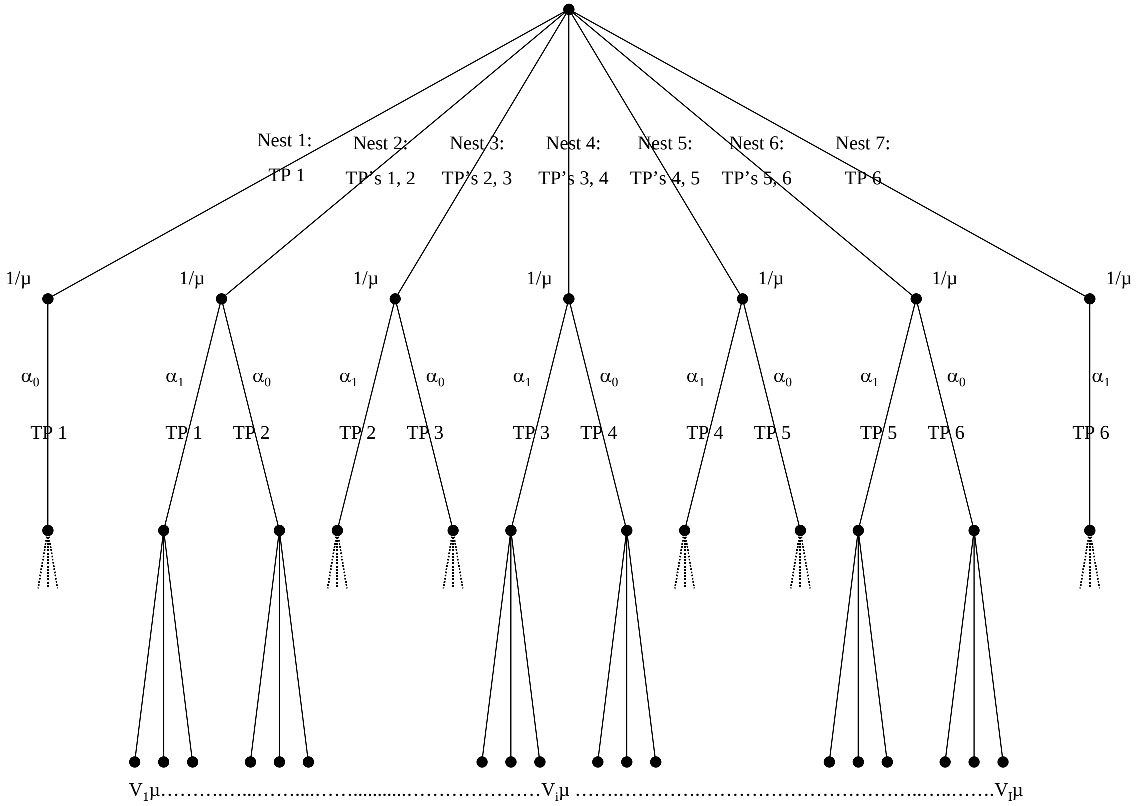

FIGURE 10: Three-Allocation OGEV Model Structure

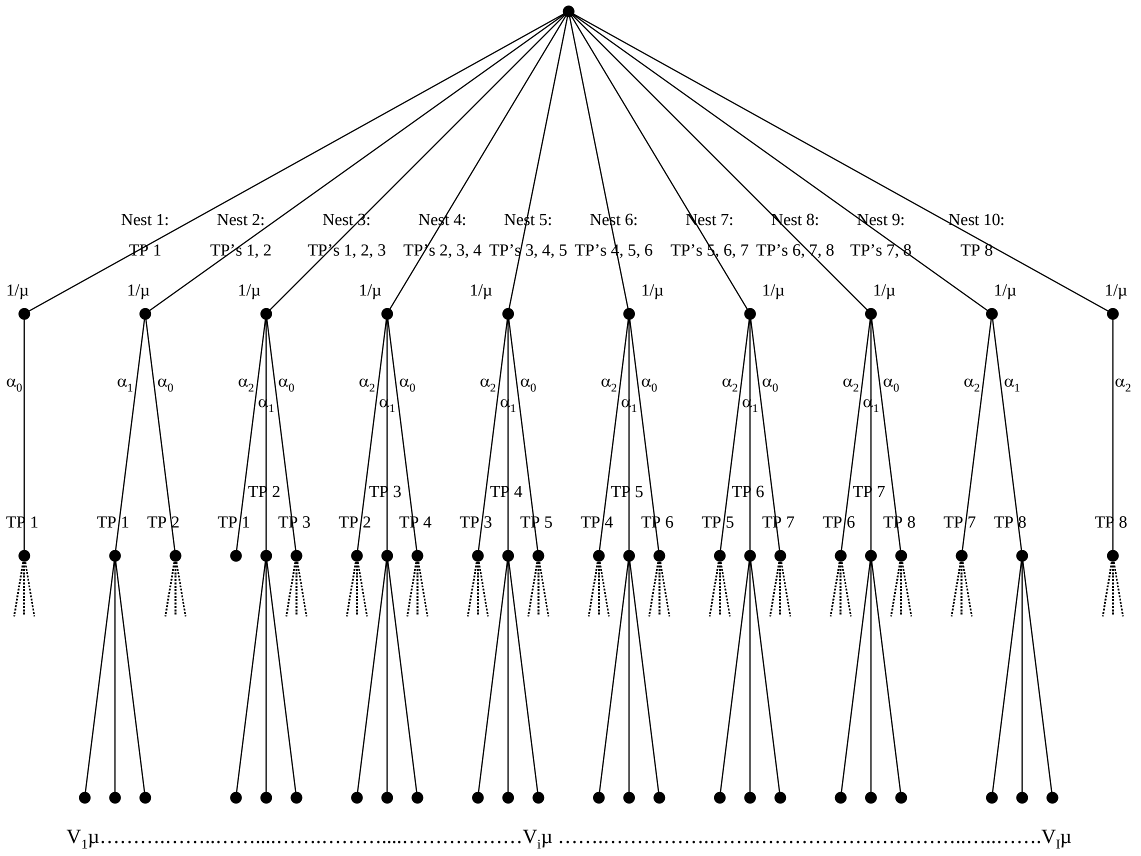

FIGURE 11: Three-Level Time, Carrier NL-OGEV (Two-Allocation) Model Structure

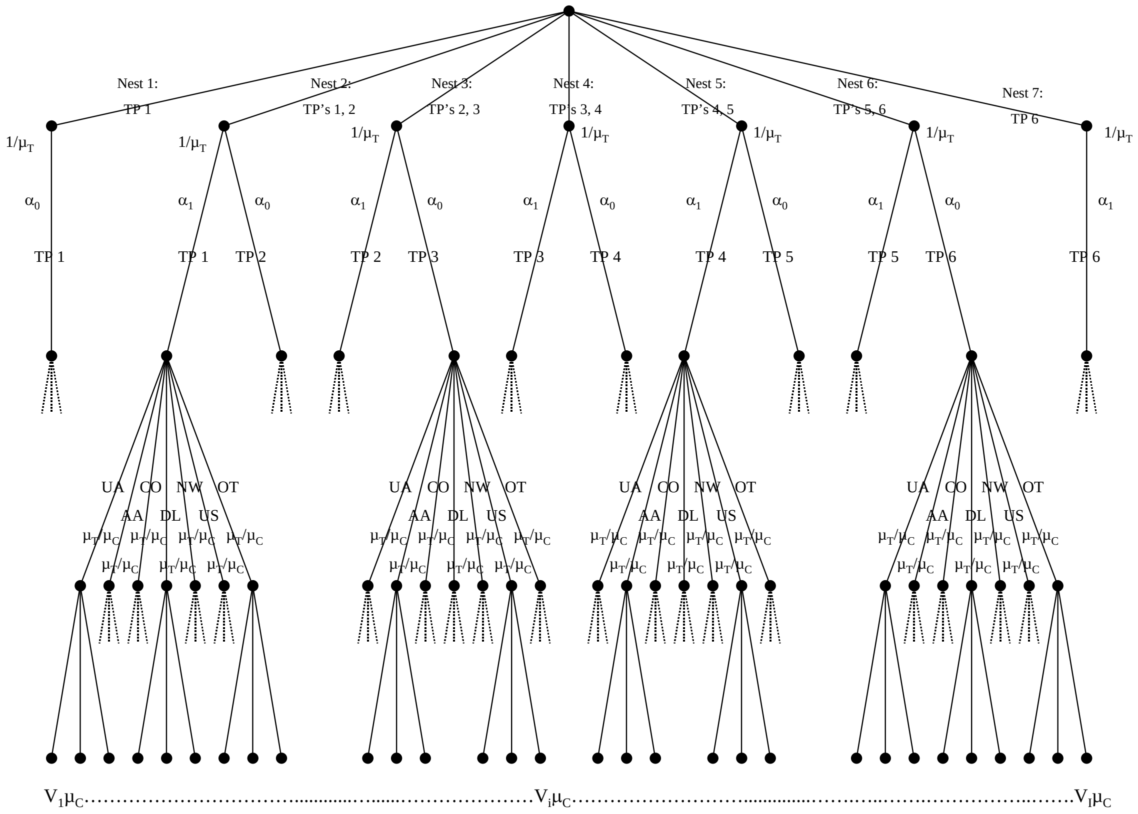

FIGURE 12: Three-Level Time, Carrier NL-OGEV (Three-Allocation) Model Structure