Collegiate Aviation Review International

Volume 36 | Issue 2 Peer Reviewed Article #6

10-3-2018

Understanding Determinants of Making Airline Route Entry and Exit Decisions: An Application of Logit Models

Canh Nguyen
Florida Institute of Technology

Cuong Nguyen
Ho Chi Minh City University of Technology

Understanding patterns of entry and exit decisions and determinants shaping the patterns are necessary for airline planners in drawing a robust route map and gaining their own competitive advantages. The study used logit models to exam the relationship between two separate binary dependent variables: entry versus no-entry, exit versus no-exit, and multiple independent variables. Dataset was extracted from the Bureau of Transportation Statistics DB1B for Quarter 1 of 2018, then was reconstructed based on original and destination (O&D) airport pairs to gain insights. The entry decision pattern model yielded seven significant factors: total passengers, average market fare, number of carriers, distance, low-cost carriers (LCC) existence, origin hub, and destination hub. In the meantime, the exit decision pattern model yielded all the seven aforementioned factors and two other significant factors: route type and the business model of the largest share airline. The findings made a practical implication to airline network planners in considering determinants affecting entry and exit decisions to build a more efficient and profitable network.

Recommended Citation:
Nguyen, C. & Nguyen, C. (2018). Understanding Determinants of Making Airline Route Entry and Exit Decisions: An Application of Logit Models. *Collegiate Aviation Review International*, 36(2), 92-116. Retrieved from http://ojs.library.okstate.edu/osu/index.php/CARI/article/view/7731

As a result of the Airline Deregulation Act in 1978, the U.S. airline industry has changed radically. Since then, airlines are able to freely make their own decisions as to where they should fly, what route market they should enter or increase frequencies, and what route market they should reduce or completely remove from the network. The newly deregulated industry witnessed an influx of new entrants with new business models such as low-cost carriers; Southwest Airlines is a typical and successful example thus far. Besides that, innovations in aircraft manufacturing has helped airlines operate flights more efficiently with lower costs, but higher capacity and longer range. Under all these conditions, the airline industry has been characterized as a free market economy, bringing out many flying opportunities for passengers; however, the competition among commercial airlines has become more intensive than ever. In order to survive in such a stiff competition, airlines are attempting to gain their own competitive advantages by building an efficient and profitable network. Understanding patterns of entry and exit decisions and determinants shaping these patterns are necessary for airline planners in drawing a robust route map.

The literature is replete with studies exploring key drivers and barriers to entering and exiting a given route. Baran (2018) examines the survival strategies of U.S. domestic airlines, which corresponded to route entry and exit decision and airfare competition. The dataset was a combination of Airline Origin and Destination (DB1B) and the U.S. Census Bureau in the period of 2011-2015. Baran restricted all entry and exit data associated with eight U.S. major airlines, and thus the dataset was $N = 2{,}111$ routes; however, there was no information about defining and measuring entry and exit decisions. Utilizing a logistic regression for a binary response, either entry or exit, the result showed a significant model with $\chi^2(10) = 120.59$, $p < .0001$, $R_L^2 = .0511$, and five significant factors: airline business model, distance, city population of origin airport, per capita income in original airport, and number of competitors. The limitation of the study was that Baran measured market concentration by calculating Herfindahl–Hirschman Index (HHI) based on the number of seats available per mile (ASM) airlines performed on each route without the consideration of the load factor. Our current study solved the problem by measuring the market concentration and market shares based on the number of passengers airlines transported on each route.

Abdelghany and Guzhva (2010) investigated entry and exit decisions by using a panel dataset of 38 quarters beginning in the first quarter of 1998 with the largest 10,000 city pairs in the U.S. domestic market. The dependent variable of the study was estimated by the differences in number of airlines between two consecutive quarters; positive changes indicated airline entries, while negative changes indicated airline exits. The independent variables in the analysis included market concentration measuring by Herfindahl–Hirschman Index (HHI), quarterly changes in market concentrations (ΔHHI_{t-1}, ΔHHI_{t-2}, and ΔHHI_{t-3}), market size measuring the number of passengers, quarterly change in market size (ΔPX_{t-1}), distance between city pairs, average one-way fare, and seasonality represented by three dummy variables for Quarter 2, 3, and 4 (comparing to the reference Quarter 1). The advantage of this study was the utilization of a panel data analysis that combined a cross-section and time-series analysis. the results of *F*-tests and Breusch-Pagan Lagrange Multiplier (LM) test failed to reject the null hypothesis and thus yielded a pooled ordinary least squares (OLS) estimation. All independent factors were significant at the preset alpha of 5% in the overall model, and the adjusted $R^2 = .09$. The disadvantage of the study was that the estimation of the dependent variable failed to capture changes in each unit of analysis (quarter-route) when, for example, each of three different airlines adds one flight into a given route, concurrently an airline removes one flight from the route. In such cases, although the response shows airline entry decisions (i.e., a positive response with two flights added into the route), there is still one exit decision in the route during the given quarter. This current study technically solved the problem by separately measuring entry and exit decisions in each route, which was fully discussed in definition and data construction sections. Previous research examined market characteristics: market density, distance, endpoint city populations and income, and hub effects together with competition-related factors (Boguslaski, Ito, & Lee, 2004; Ito & Lee, 2003; Oliveira, 2008).

Based on the review of the literature, the motivation to conduct the current study was to partially replicate with the latest dataset as well as mostly using more robust estimation in the dependent variables, entry and exit decisions. The current study also followed and examined factors suggested by previous research relative to distinguishing patterns of airline entry and exit in the U.S. domestic market. The different points were that we decided to omit several suggested variables and added new variables into the analyses because we reconstructed the raw dataset and gained more insights. Specifically, we used the variable of total passengers in each route to represent the demand as opposed to using the cities' population and income, and we obtained some new variables such as LCC existence and the business model of the airline with the largest share.

Purpose

The purpose of the study was to identify factors that would shape route entry and exit decision patterns of commercial airlines in the U.S. domestic market. The two variables of interest, *entry patterns* and *exit patterns* on the city-pair market, were independent of each other. The targeted research factors consisted of 11 independent variables (IVs): total passengers, number of departures, average market fare, number of carriers, market concentration, distance between origin and destination airports, route type, existence of low cost carriers, business model of airline with the largest share, origin hub airport, and destination hub airport. The study was restricted to the U.S. domestic routes that have both origin and destination airports located within the United States. In addition, the study only considered route markets that had at least one operation of a commercial airline. The study was cross-sectional in nature and used the 2018 dataset of Quarter 1 that was archived in the Bureau of Transportation Statistics (BTS).

Research Questions

The research questions that guided the current study were as follows:

Research question 1: When examined from the entry pattern model, what is the relationship between the targeted variables and the dichotomous response variable that distinguished between airline route entry and non-entry decisions?

Research question 2: When examined from the exit pattern model, what is the relationship between the targeted variables and the dichotomous response variable that distinguished between airline route exit and non-exit decisions?

Hypotheses

The corresponding research hypotheses were as follows:

Research hypothesis 1: When examined from the entry pattern model, at least one of the targeted variables will have predictive value relative to distinguishing between airline route entry and non-entry decisions.

Research hypothesis 2: When examined from the exit pattern model, at least one of the targeted variables will have predictive value relative to distinguishing between airline route exit and non-exit decisions.

Definitions of Variables

Dependent Variables

The study consisted of two separate dependent variables of an airlines, the entry patterns and exit patterns, which are both critical strategies that an airline takes into consideration either to increase its market share or to exit from routes. Given a route, entry and exit patterns were initially constructed by the differences in the number of departures between Quarter 1 of 2018 and Quarter 4 of 2017. The positive differences across the given route were counted as entry decisions, which means that airlines either entered for the first time or increased their frequency. In the meantime, the negative differences across the given route were counted as exit decisions, which means that airlines either stopped their air service or reduced their flight frequency. If the difference returned 0 in that route, there were no entry and exit decisions. For example, the route from ABE to ATW in Quarter 1 of 2018 compared to Quarter 4 of 2017. There were two entry decisions: Delta Air Lines (DL) with a 10-flight increase and United Airlines (UA) with a 1-flight increase; at the same time, one exit decision was made by American Airlines (AA) with a 1-flight decrease, compared to Quarter 4 of 2017. Using dummy coding strategies, *entry patterns* were then coded in favor of nonzero values (i.e., 1 was coded as entry decisions made versus 0 coded as no change in flight departures across city-pair routes). Similarly, *exit patterns* were coded in favor of nonzero values (i.e., 1 was coded as exit decisions made versus 0 coded as no change in flight departures across city-pair routes). It was noted that the two entry and exit decisions were independent of each other; therefore, we gained two separate variables, entry patterns and exit patterns.

Independent Variables

Total passengers were an aggregated number of passengers carried by all airlines on a given city-pair route.

Number of departures was used interchangeably with the aggregated number of flights operated by all airlines on a given route.

Average market fare was defined as the average price of passenger transportation service that all airlines offered in a city-pair market.

Number of carriers counted all airlines operating on a given route.

Market concentration was scored by the Herfindahl–Hirschman Index (HHI) for each route, measuring the market structure and competition level of the market. Generally, the score is computed by a summation of squared market share of each airline, ranging either from 0 to 10,000 for a percentage-based computation, or from 0 to 1 for a computation without percentage consideration (Abdelghany & Abdelghany, 2009, pp. 47–48). The HHI in this study followed the latter technique, which varied from an approximation of 0 as a clue of a heavily competition level to 1 as a clue of a monopolistic market.

Distance was defined as the geographic distance in miles between origin and destination airports.

Route type was defined as either nonstop route market coded with 1, or connecting route market coded with 0. Conventionally, the number of coupons (i.e., referred to the number of boarding pass of a flight) speak to the characteristic of route market (Yuan, 2016). It is very commonly accepted in the literature that in a specific city-pair route, if the number of coupons is 1 and nonstop flights are served, the route is considered nonstop market; otherwise it is considered a connecting market (Coldren, 2005; Coldren, Koppelman, Kasturirangan, & Mukherjee, 2003; Garrow, 2010). For example, the route ABE-ATW was a connecting market due to no nonstop flight being served across airlines.

Existence of low cost carriers was defined as at least one operation of a low-cost carrier on a given route. The variable was coded as 1 if having at least one LCC, and otherwise it was 0.

Business model of airline accounting for the largest share in the given route was partitioned exclusively into low-cost carriers (LCC) and full-service carriers (FSC). The former was coded as 1 and the latter was coded as 0 using dummy coding strategy. A total of 36 commercial airlines reported as ticketing carriers in the 2018 dataset, and 7 of them correspond to the business model of a low-cost carrier, according to reports in their official websites. These LCCs includes Allegiant Air (G4), Frontier Airlines (F9), JetBlue (B6), Spirit Airlines (NK), Southwest Airlines (WN), Sun Country Airlines (SY), and Virgin America (VX) (i.e., Virgin America’s flights continued to be reported in the study timeframe, but the airline will cease its operations as of April 2018 due to the consolidation with Alaska Airlines). The market shares were then calculated for each airline in a given route based on the number of passengers it transported during the Quarter 1 of 2018. Only the airline that had the largest share in each route was reflected in this variable.

Origin hub airport was a dummy variable coded as 1 if the origin airport is a large hub, and 0 if it is a non-large hub. The large hubs were primary commercial service airports and categorized by the Federal Aviation Administration (FAA) as having 1% and more of annual enplanements (“Airport Categories – Airports,” n.d.).

Destination hub airport was a dummy variable coded as 1 if the destination airport is a large hub, and 0 if it was a non-large hub. The large hubs were primary commercial service airports and were categorized by Federal Aviation Administration (FAA) as having 1% and more of annual enplanements (“Airport Categories – Airports,” n.d.).

Method and Data Construction

Method

The research methodology was retrospective, known as ex-post facto, and the corresponding design was cause-type. This methodology was appropriate to answer the research questions because we were determining the extent to which the targeted factors influenced whether airlines entered or exited routes. Furthermore, the effects on the two dependent variables, which were group memberships, had already occurred. As a part of the study, two statistical approaches, descriptive and inference, were utilized to answer the research questions. The latter approach was involved in logit analyses (i.e., logistic regressions), which were appropriate to answer the research problem because the dependent variables were binary nominal (Cameron & Trivedi, 2005; Greene, 2011; Hair, Black, Babin, & Rolph, 2010).

Greene (2011) suggested that if the two binary dependent outcomes are interrelated as opposed to independent and have a significant correlation coefficient, a bivariate logit model should be applied. An example used by (Katchova, 2013) for a bivariate logit model is an investigation on factors influencing the joint outcome of being in an excellent health status (Y_1) and visiting the doctor (Y_2). Another would be a business decision of whether to use marketing contracts or not (Y_1) versus whether to use environment contracts or not (Y_2). Given the current study, the correlation coefficient between entry and exit patterns were $r = -.3520$, $p < .0001$; however, each decision to enter or exit was made independently and the decisions were unrelated to each other. In a route, some airlines choose to enter/increase, while others choose to maintain their frequency, or even exit at the same time. Therefore, two separated binary logit models were more appropriate to estimate the effect of factors on the entry and exit decision patterns.

Table 1

Summary and Description of Independent and Dependent Variables Overall

Variables	Description
Total passengers	Continuous variable represented the aggregated number of passengers carried by all airlines on a route.
Number of departures	Continuous variable represented the aggregated number of flights on a route.
Average market fare	Continuous variable represented the average of price all airlines offered in a route.
Number of carriers	Continuous variable as all airlines operating on a given city-pair route.
HHI score	Continuous variable, range from 0 to 1, measuring market concentration by a summation of squared market share of each airline.
Distance	Continuous variables ad the geographic distance in miles between origin and destination airports.
Route type	Categorical (dichotomous) variable represented having at least one nonstop flight on a given route. Dummy coded with 1 as nonstop market and 0 as connection market (the reference group).
LCC existence	Categorical (dichotomous) variable represented having at least one LCC operation on a given route. Dummy coded with 1 as yes group and 0 as no group (the reference group).
The business model of the largest share airline	Categorical (dichotomous) variable represented the largest share airline is a LCC or FSC. Dummy coded with 1 as a LCC and 0 as a FSC (the reference group).
Origin airport	Categorical (dichotomous) variable represented the origin airport is a large hub or non-large hub. Dummy coded with 1 as a large hub and 0 as a non-large hub (the reference group).
Destination airport	Categorical (dichotomous) variable represented the destination airport is a large hub or non-large hub. Dummy coded with 1 as a large hub and 0 as a non-large hub (the reference group).
Entry Patterns	Categorical (dichotomous) variable represented the airline decisions of entering a new route or increasing the frequency in the existing route. Dummy coded with 1 as an entry decision and 0 as a no-entry decision (the reference group).
Exit Patterns	Categorical (dichotomous) variable represented the airline decisions of exiting from an existing route or reducing the frequency in a given route. Dummy coded with 1 as an exit decision and 0 as a no-exit decision (the reference group).

Note. The order of the variables was arranged for the convenience of the readers based on market-related factors: total passengers, number of departures, average market fare, number of carriers, and HHI; route-related factors: distance, route type, LCC existence, and the business model of the largest share airline; and airport-related factors: origin airport and destination airport.

Data Construction

The dataset used for analyses in this study was directly downloaded from the U.S. Department of Transportation (US DOT) Origin and Destination Data Bank 1B (DB1B). The stored database contains a 10% sample of tickets collected from passengers as they boarded aircraft operated by any of the U.S. airlines. The selected dataset was for Quarter 1 in 2018and it was subsequently imported through *JMP* software. In particular, the raw data provided quarterly demand information on the number of passengers transported between origin-destination pairs, itinerary information (e.g., ticketing carriers, operating carriers, number of coupons, distance…), and quarterly fare charged by each airline for a route that is averaged across all classes of service. Following Garrow's (2010) recommendation, we eliminated 127,429 routes from the total 6,093,175 routes due to missing data on ticketing carriers (i.e., missing ticketing carriers were coded either as "--" or 99 on the original dataset). The data were then reconstructed by sorting out all variables based on origin-destination airport pairs. There were N = 61,024 airport pairs after reconstruction, which were different from the original routes by the unique appearance of an airport pair. An example for the reconstructed data was that the ABE-ATW pair had 27 departures over the quarter from two airlines, while on the raw dataset, it had 27 repeated ABE-ATW pairs. However, on both datasets, the 27 departures were all connecting flights, which indicated a connecting route market as discussed earlier.

To construct the two dependent variables--entry and exit patterns--we sorted out the number of departures performed by each ticketing airline on each O&D airport pair. The same technique was applied to the dataset of Quarter 4 of 2017 to acquire the differences in entry and exit decisions made throughout Quarter 1 of 2018. The number of carriers, the existence of at least one LCC, and the number of departures were counted as nonzero values across sort-outs of ticketing airlines on each airport pair. To construct competition-related factors, total passengers, market shares of each airline, largest share and its respective airline business model, and HHI, we sorted out the number of passengers transported by each ticketing airline on each O&D airport pair. To construct average market fare, we averaged the market fare from all flights performed by all airlines on a given route. To construct distance, we took the minimum of the distance in miles flown because the minimum distances indicate the geographic distances between the O&D airport pairs in nonstop routes.

The size of the sample, N = 61,024 airport pairs, exceeded all recommendations for a minimum sample size of the logistic regression model in the literature. For example, Hosmer Jr, Lemeshow, and Sturdivant (2013) suggested a sample size greater than 400 observations, and Peduzzi, Concato, Kemper, Holford, and Feinstein (1996) called on researchers to obtain at least 10 times the number of independent variables in the model (i.e., the number of independent variables in the study, k = 11). Apart from a large sample size, a logistic analysis requires a sample size for each group membership of at least 10 observations per estimated parameter (Hair et al., 2010, p. 322), and each independent variable consists of a minimum of one cell frequency and no more than 20% of cell frequencies less than five (Tabachnick & Fidell, 2013). The dataset was valid and met all suggestions, thereby ensuring the study's statistical power, and this was confirmed by contingency tables discussed in the sections below.

Findings

Descriptive Analysis

Statistical summary of entry patterns. As reported in Table 2, on itineraries airlines made entry decisions, the mean of the demand was more than 277 passengers, which was higher than that on itineraries airlines had no change in their schedule with approximately 114 passengers. On such routes, airlines supplied more available seat per mile (ASMs) to meet the high demand, and thus had more departures with 128 quarterly flights, which was more than double the non-entry routes at 57 quarterly flights. The market concentration on entry routes was lower than non-entry routes with 0.72 versus 0.83 in the respective HHI scores, which was a sign of a heavier competition with more airlines. Higher average market fare and longer stage length in geographic distance exhibited as long haul flights with higher yield rate for entry routes. Abnormalities or potential outliers were spotted in the variables of average market fare and distance. For average market fare on entry route, the standard deviation (SD) was high at $1,163.19, while the figure for non-entry routes was $151.93. The maximum fare on the range of average market fare was $215,353.35, which probably belonged to charter flights and was identified as an outlier that discussed in the preliminary analysis below. For the variable of distance with the minimum at 11 miles between two airports, we retrieved the route from the dataset and uncovered that the O&D airport pairs are OAK (Metropolitan Oakland International Airport) and SFO (San Francisco International Airport), which the distance was only from East to West boundary of San Francisco Bay.

Table 2

Descriptive Statistics of Continuous Variables in the Model of Entry Patterns

Continuous Variables	Entry			No Entry			Overall		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Total passengers	277.37	1272.29	1 – 34,582	113.75	649.05	1 – 20,560	207.29	1,054.68	1 – 34,582
Departures	128.29	403.12	1 – 9,229	57.01	223.04	1 – 6,040	97.76	339.78	1 – 9,299
Average market fare	324.51	1,163.19	0 – 215,353.35	310.50	151.93	0 – 3,680.84	318.51	885.11	0 –215,353.35
Number of carriers	2.19	1.35	1 – 10	1.61	0.94	1 – 8	1.94	1.22	1 – 10
HHI	0.72	0.26	0.14 – 1	0.83	0.23	0.19 – 1	0.77	0.26	0.13 – 1
Distance	1,495.63	1,015.93	11 – 9,700	1,312.44	913.17	55 – 9,571	1,427.16	977.45	11 – 9,700

As reported in Table 3 and Table 4, entry routes accounted for 57.2% in the total of 61,024 routes. Overall, there were 8,468 routes (13.9%) having at least an operation of a LCC in which entry decisions appeared on 6,029 routes (17.3% of the total entry route) and exit decisions appeared on 2,439 routes (9.3% of the total exit route). By taking advantage of connecting flights to cover all airports in the nation, full-service airlines wholly dominated the overall network with 93.6% in terms of the largest share airline. On routes having entry decision made, there were only 7.3% if the largest share is a LCC, but 92.7% if a FSC. This was also the reason why connection market considerably overweighed nonstop one with 87.8% in comparison with 12.2%. Airlines made 85.9% entry decisions in the total on connection route market as opposed to 14.1% entry decisions made on nonstop route market. This implies that once airlines see a potential growth in connection itineraries, they could either increase their frequencies, or launch single-connecting in place of previous double-connecting flights, or even serve non-stop flights and make the route become nonstop route market. In case of O&D airport pairs are large hubs, approximately 15% entry decisions were made on such routes, which was by far lower than those on routes with non-hub airport pairs.

Table 3

Descriptive Statistics Relative to LCC Existence and Business Model of the Largest Share Airlines in the Model of Entry Patterns

			LCC existence				Business model of the largest share airline			
			Yes		No		LCC		FSC	
	N	%	N	%	N	%	N	%	N	%
Entry	34,886	57.2	6,029	17.3	28,857	82.7	2,545	7.3	32,341	92.7
No Entry	26,138	42.8	2,439	9.3	23,699	90.7	1,370	5.2	24,768	94.8
Overall	61,024	100	8,468	13.9	52,556	86.1	3,915	6.4	57,109	93.6

Table 4

Descriptive Statistics Relative to Origin, Destination Airport, and Route Type in the Model of Entry Patterns

			Origin Airport				Destination Airport				Route Type			
			Hub		Non-Hub		Hub		Non-Hub		Nonstop		Connection	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Entry	34,886	57.2	5,404	15.5	29,482	84.5	5,223	15.0	29,663	85.0	4,919	14.1	29,967	85.9
No Entry	26,138	42.8	3,798	14.5	22,340	85.5	3,946	15.1	22,192	84.9	2,516	9.6	23,622	90.4
Overall	61,024	100	9,202	15.1	51,822	84.9	9,169	15.0	51,855	85.0	7,435	12.2	53,589	87.8

Statistical summary of exit patterns. As reported in Table 5, exit decisions were made on routes that had higher demand with the average of 298 passengers, but lower fare at $301.40 compared to 46 passengers and $348.69 on routes no exit decisions were made. The competition level in exit routes was stiffer and fiercer with at least two players and 0.7 HHI, and that in non-exit routes were easier with one operation of an airline and 0.89 HHI. When examining the distance between two groups, airlines tended to exit on shorter routes (mean at 1,340.74 miles) and make no exit decision on long routes (mean at 1,551.99 miles).

Table 5

Descriptive Statistics of Continuous Variables in the Model of Exit Patterns

Continuous Variables	Exit			No Exit			Overall		
	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Total passengers	298.38	1,277.40	1 – 34,582	46.60	394.55	1 – 18,249	207.29	1,054.69	1 – 34,582
Departures	138.86	409.36	1 – 9,299	25.26	129.26	1 – 5,141	97.76	339.78	1 – 9,229
Average market fare	301.40	124.40	0 – 3,680.84	348.69	1461	0 – 215,353.35	318.51	885.11	0 – 215,353.35
Number of carriers	2.28	1.33	1 – 10	1.35	0.71	1 – 8	1.94	1.23	1 – 10
HHI	0.70	0.26	0.13 – 1	0.89	0.20	0.21 – 1	0.77	0.26	0.13 – 1
Distance	1,340.74	896.45	54 – 9,571	1,551.99	1,093	11 – 9,700	1,417.17	977.45	11 – 9,700

As reported in Table 6 and Table 7, exit decisions were made on 38,947 routes, accounting 63.8% of the total observations of the study. There were 19.1% of the exit routes in conjunction with the appearance of at least one LCC operation. On routes that exit decisions were made, there were only 8.2% if the largest share is a LCC, but 91.8% if a FSC. Also, on routes that exit decisions made, nonstop market occupied only 16.7%, while the figure for connection market was 83.3%. Similar to entry pattern analysis, approximately 18% exit decisions were made on routes with departures and arrivals at large hubs.

Table 6

Descriptive Statistics Relative to LCC Existence and Business Model of the Largest Share Airlines in the Model of Exit Patterns

			LCC existence				Business model of the largest share airline			
			Yes		No		LCC		FSC	
	N	%	N	%	N	%	N	%	N	%
Exit	38,947	63.8	7,432	19.1	31,515	80.9	3,195	8.2	35,752	91.8
No Exit	22,077	36.2	1,036	4.7	21,041	95.3	720	3.2	21,357	96.8
Overall	61,024	100	8,468	13.9	52,556	86.1	3,915	6.4	57,109	93.6

Table 7

Descriptive Statistics Relative to Origin, Destination Airport, and Route Type in the Model of Exit Patterns

			Original				Destination				Route Type			
			Hub		Non-Hub		Hub		Non-Hub		Nonstop		Connection	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Exit	38,947	63.8	7,042	18.1	31,905	81.9	7,080	18.2	31,867	81.8	6,520	16.7	32,427	83.3
No Exit	22,077	36.2	2,160	9.8	19,917	90.2	2,089	9.5	19,988	90.5	915	4.1	21,162	95.9
Overall	61,024	100	9,202	15.1	51,822	84.9	9,169	15.0	51,855	84.9	7,435	12.2	53,589	87.8

Preliminary analysis

The targeted independent and dependent variables were tested for compliance with the assumptions of logistic regression. First, the assumption of a dichotomous dependent variable was obtained through group memberships of both dependent variables. Entry patterns were coded either 1 for entering/increasing, or 0 for not entering/increasing flights in a route. Exit patterns were coded either 1 for exiting/reducing, or 0 for not exiting/reducing flights in a route. Second, the assumption of mutually exclusive categories on the dependent variables was fulfilled. Each route of airport pairs was a member of one group or the other, but not both, and therefore was exhaustive and mutually exclusive. Third, the assumption of independence of scores on the dependent variables was presumed to be compliant because the study's dataset was not the result of repeated measures or matched data. Instead, the data were acquired from an archival database of U.S. DOT, and thus the data for each targeted variable associated with each route of airport pairs were unrelated. Lastly, the assumption of correct specification of the model, which requires the hypothesized model only include independent variables that are relevant, was met. The reason was that the inclusion of the targeted variables in the hypothesized model was based on prior research, and the significant chi-square test for the fit of the null and convergent models discussed in primary analysis.

Although not required for a logistic regression, outlier analysis and the absence of multicollinearity in the independent variables were also addressed because these issues could be indicative of a poor predictive model. Outliers are extreme data points inconsistent with others, and it potentially could produce results that are not representative of the relationships in the remaining data. Outliers can be labeled as either contaminants or rare cases (Cohen J., Cohen P., West, & Aiken, 2003). We conducted a statistical analysis on the dataset with $N = 61{,}024$ routes, and determined 3,536 routes (5.8%) as potential outliers based on Jackknife distances. "The distance for each observation is calculated with estimates of the mean, standard deviation, and correlation matrix that do not include the observation itself. The jack-knifed distances are useful when there is an outlier." (SAS Institute Inc., 2016, pp. 50–51). Examination of these outliers revealed several instances in which there was an inconsistency; for example, the average fare in BGM-HNL was extremely high at $215,353.35 with only one carrier operating on the route, which was most likely a charter flight rather than a commercially-scheduled flight. To determine the impact of these outliers, we ran analyses before and after excluding the outliers, and then we compared the results of the models. All overall models yielded significant results with few differences; however, the estimations were impacted considerably. Therefore, we decided to eliminate all the flagged outliers from the primary models, and the sample size reduced to N = 57,488 routes.

Multicollinearity can occur if the independent variables in a model are highly-correlated, and it can be examined through a correlation matrix. Cohen et al. (2003) suggested an existence of multicollinearity if the correlation coefficient is $r > .8$ between two independent variables. As reported in Table 8, compared to the threshold, the number of carriers and HHI were labeled as multicollinearity, $r_{\text{Carriers vs. HHI}} = -.89$, which indicates the negative relationship that the more carriers, the smaller HHI score is. Evidence of serious multicollinearity also was found between the total passengers and the number of departures, $r_{\text{Pax vs. Departures}} = .91$, indicating a positive relationship when airlines have more flights, and thus, can carry more passengers. For the sake of interpretations in later sections, we decided to retain the number of carriers and total passengers in the model and to exclude HHI and number of departures from the final model. At this point, the number of independent variables was reduced, k = 9 in total.

Table 8

Correlation Matrix between Continuous Variables

	Carriers	Market Fare	Total Passengers	HHI	Departures	Distance
Carriers	1.0000	-0.1129	0.3835	-0.8906	0.5523	-0.0119
Market Fare	-0.1129	1.0000	-0.1659	0.0808	-0.1723	0.3811
Total Passengers	0.3835	-0.1659	1.0000	-0.2017	0.9142	-0.1093
HHI	-0.8906	0.0808	-0.2017	1.0000	-0.3398	0.0123
Departures	0.5523	-0.1723	0.9142	-0.3398	1.0000	-0.1053
Distance	-0.0119	0.3811	-0.1093	0.0123	-0.1053	1.0000

Primary Analysis

Two separate simultaneous models were developed by regressing two independent variables: Entry versus No entry and Exit versus No Exit, on the nine independent variables simultaneously. Following Warner's (2008) recommendations, the overall goodness of fit of null models, which regressed group memberships in the absence of nine independent variables, were compared to that of the full models. The assessments point to the log likelihood (LL) function and the chi-square statistic. The former is comparable to the sum of the squared residuals in multiple regression, while the latter is the difference between $-2LL$ for the full model and $-2LL$ for the null model (Warner, 2008).

Entry patterns models. As reported in Table 9, the full model was statistically significant, $\chi^2(9) = 3441.82$, $p < .0001$. In addition, Cohen et al. (2003) recommended reporting the Pseudo-R^2 ($R_L{}^2$) as a gain in prediction obtained from adding variables to a model. The full model provided a predictive gain of 4.36% over the null model ($R_{Lfull}{}^2 = .0436$, df = 9).

As reported in Table 10, the null model was significant, $\chi^2(0) = 785.17$, $p < .0001$. The null model's logit in the Entry group was $B_{\text{Constant}} = 0.235$, which means that in the absence of information provided the independent variables, the odds of entering/increasing flights in a route was $e^{0.235} = 1.26$. When applied the mathematical expression, $e^{0.235} / (1 + e^{0.235}) = 0.5575$, it indicated that 55.75% of the observations associated with entering/increasing flights in Quarter 4 of 2018. Because the omnibus test yielded a significant result, we examined the relationship between each IV and the DV—Entry patterns. In the full model, seven of nine IVs were significantly related to the group membership, Entry versus No Entry, in the presence of the other predictors.

Table 9

Significance of the Simultaneous Model of Entry Patterns

Model	Log Likelihood	*df*	χ^2
Null	39452.34		
Full	37731.43		
Difference	1720.91	9	3441.82***

Note. $N = 57{,}488$. $R_L^2 = .0436$, ***$p < .001$

Directions of relationships. The original logistic coefficient for total passengers was $B_{\text{Pax}} = 0.0003$, which indicates a positive relationship between total passenger variable and the group membership. As passengers increased, airlines were more likely to enter/increase their operations in the market route. Similarly, $B_{\text{Fare}} = 0.0002$, $B_{\text{Carriers}} = 0.5057$, and $B_{\text{Distance}} = 0.0003$ showed positive signs, which signified a likelihood of making entry decision in the market route when market fare, number of carriers, and the distance between origin and destination airports increased. For dummy nominal coded variables, $B_{\text{LCC}} = -0.2462$, $B_{\text{Origin}} = -0.3469$, and $B_{\text{Dest}} = -0.4281$ implied that airlines were less likely to enter/increase their flights in the market route in which there had been the appearance of at least one LCC operation, or either the origin or destination airport was a large hub. Two other nonsignificant variables, route type and business model of the largest share airline, were not interpreted.

Magnitude interpretations. To interpret the magnitude of the relationships, we turned our attention to the exponential coefficients (odds ratio) that are calculated by raising e to the original coefficients (B_{i}), $e^{0.0003} = 1.0003$ and its reciprocal $e^{-0.0003} = 0.9997$. In the case, $e^{Bi} > 1$ indicates a positive relationship, $e^{Bi} < 1$ indicates a negative relationship, and $e^{Bi} = 1$ indicates a no change in the odds for the membership relative to the discussion IV. In our study, it means that with an increase of one passenger, airlines were 1.0003 times more likely to be involved in entry decisions than maintaining their schedule or considering no new entry. Also, $e^{Bi} - 1$ equals the percentage change in odds in which the odds increased by 0.03% if the demand increased by one passenger, holding all other variables constant. In the same way, airlines were 1.0002, 1.6581, and 1.0003 times more likely to enter the route or increase their flights in the route as an increase of market fare by $1, or one more competitor, or 1 mile in the distance between the origin and destination airports. The corresponding odds ratios of making entry decisions also increased 0.02%, 65.81%, and 0.03% respectively. For negative signs in the relationships (i.e., in our entry pattern study, all significant dummy coded nominal variables were negative) reciprocals were reported as opposed to its odds ratios. Particularly, airlines were 1.2791, 1.4147, and 1.5343 times more likely to enter/increase the frequency in the route in which no LCC had existed, and either origin or destination airport was not a large hub. The corresponding odds ratios of making entry decisions in routes characterized by LCC existence, origin large hub, and destination large hub decreased by 21.82% (= 0.7818 – 1), 29.32% (= 0.7068 – 1), and 34.83% (= 0.6517 – 1), respectively.

Table 10

Summary of Logistic Regression Estimates for the Null and Simultaneous Model of Entry Patterns

	B_i	*SE*	χ^2	*p*	Odds Ratio	95% CI	Reciprocal
Null Model							
Constant	0.235	0.0084	785.17	<.0001***			
Full Model							
Constant	-0.9381	0.0302	960.50	<.0001***			
Total Passengers	0.0003	0.00006	25.54	<.0001***	1.0003	[1.0001, 1.0004]	0.9997
Market fare	0.0002	0.00007	5.81	<.0001***	1.0002	[1.0000, 1.0003]	0.9998
Number of carriers	0.5057	0.0112	2038.8	<.0001***	1.6581	[1.6221, 1.6949]	0.6031
Distance	0.0003	0.00001	429.87	<.0001***	1.0003	[1.0002, 1.0003]	0.9997
Route type	-0.0468	0.0448	1.09	.2963	0.9542	[0.8740, 1.0419]	1.0479
LCC existence	-0.2462	0.0499	24.27	<.0001***	0.7818	[0.7089, 0.8622]	1.2791
Business model of the largest share	0.051	0.060	0.71	.3998	1.0523	[0.9345, 1.1850]	0.9503
Origin hub	-0.3469	0.0271	163.28	<.0001***	0.7068	[0.6702, 0.7455]	1.4147
Destination hub	-0.4281	0.0271	248.20	<.0001***	0.6517	[0.6179, 0.6874]	1.5343

Note. N = 57,488. R_L^2 = .0436, df = 9 for the full model, *p < .05. **p < .01. ***p < .001

Another approach in understanding the magnitude of the relationships is to calculate marginal effects for the independent variables. Indeed, reporting marginal effects instead of odds ratio is more popular in econometrics (Cameron & Trivedi, 2005; Greene, 2011). Following Greene's (2011) instructions, we informed readers of two types: marginal effects at the mean and average marginal effects. The former is estimated for the average observation ($\bar{x}$) in the sample, while the latter is estimated as the average of the individual marginal effects. In both ways, the marginal effects reported in Table 11 were almost identical, so we only interpreted the marginal effects at the mean in this study. In our study, for an additional passenger in demand, $1 increase in market fare, one more competitor, and 1-mile increase in distance, airlines were 0.007%, 0.005%, 12.31%, and 0.007% more likely to make an entry decision, respectively. In contrast, under independent conditions, at least one LCC operation, origin large hub, or destination large hub, airlines were 6%, 8.45%, and 10.43% less likely to make an entry decision into the routes.

Table 11

Summary of Marginal Effects for the Logistic Model of Entry Patterns

	Mean (*M*)	Logistic Coefficient for Entry Patterns	Marginal Effects at the Mean[a]	Average Marginal Effects[b]
Total Passengers	68.32	0.0003***	0.00007	0.00007
Market Fare	310.80	0.0002***	0.00005	0.00005
Number of Carriers	1.80	0.5057***	0.1231	0.1246
Distance	1345.95	0.0003***	0.00007	0.00007
Route Type	0.09	-0.0468	-0.0114	-0.0115
LCC Existence	0.11	-0.2462***	-0.06	-0.0607
Business Model of the Largest Share	0.05	0.051	0.0124	0.0126
Origin Hub	0.13	-0.3469***	-0.0845	-0.0855
Destination Hub	0.13	-0.4281***	-0.1043	-0.1055

Note. N = 57,488. Logit equation = $0.0003X_{\text{Total Pax}} + 0.0002X_{\text{Fare}} + 0.5057X_{\text{Carriers}} + 0.0003X_{\text{Distance}} - 0.0468X_{\text{Route Type}} - 0.2462X_{\text{LCC}} + 0.051X_{\text{Largest Share}} - 0.3469X_{\text{Origin}} - 0.4281X_{\text{Dest}} - 0.9381$.

[a]Logit Value at the Mean was calculated by substituting the means of regressors into the Logit equation. Logit Value = 0.329. Odds = $e^{0.329} = 1.389$. Probability Pr (Y = 1 | X) = $e^{0.329} / (1 + e^{0.329}) = 0.58$, Pr (Y = 0 | X) = 1 - 0.58 = 0.42. Marginal Effect at the Means, $\delta_p/\delta_{xj} = f(\overline{x}'\beta) * (1 - f(\overline{x}'\beta)) * \beta_j$ (Greene, 2011), which in our study equals 0.58 x 0.42 x Logistic coefficients. [b]Average Predicted Probability were obtained from *JMP* output for each case before taking an average, Pr (Y = 1 | X) = 0.56, Pr (Y = 0 | X) = 0.44. Average Marginal Effects, $\delta_p/\delta_{xj} = \frac{\Sigma f(x'\beta)}{n} * (1 - \frac{\Sigma f(x'\beta)}{n}) * \beta_j$ (Greene, 2011), which in our study equals 0.56 x 0.44 x Logistic coefficients, *$p < .05$. **$p < .01$. ***$p < .001$

Exit patterns model. As reported in Table 12, the full model was statistically significant, $\chi^2(9) = 10249.31$, $p < .0001$. $R_{Lfull}^2 = .1348$, $df = 9$ indicated that the full model provided a predictive gain of 13.48% over the null model. As reported in Table 13, the null model was significant, $\chi^2(0) = 3564.8$, $p < .0001$. The logistic constant coefficient of the null model in the Exit group was $B_{\text{Constant}} = 0.5156$, which means that in the absence of information provided the independent variables, the odds of exiting/reducing flights in a route was $e^{0.5156} = 1.67$. When applied the mathematical expression, $e^{0.5156} / (1 + e^{0.5156}) = 0.6255$, it indicated that 62.55% of the observations associated with exiting/reducing flights in Quarter 4 of 2018. Because the omnibus test yielded a significant result, we examined the relationship between each IV and the DV—Exit patterns. In the full model, all nine IVs were significantly related to group membership, Exit versus No Exit, in the presence of the other predictors.

Table 12

Significance of the Simultaneous Model of Exit Patterns

Model	Log Likelihood	*df*	χ^2
Null	38005.80		
Full	32881.14		
Difference	5124.66	9	10249.31***

Note. N = 57,488. R_L^2 = .1348
*** $p < .0001$

Directions of relationships. The original logistic coefficient for total passengers (B_{Pax} = -0.0002), Market fare (B_{Fare} = -0.0003), Distance (B_{Distance} = -0.0003), and LCC existence (B_{LCC} = -0.2351) showed a negative relationship with the group membership. As each of the variables increased, airlines were less likely to exit/reduce their operations in a given route. Conversely, the logistic coefficients for the number of carriers (B_{Carriers} = 1.0272), route type ($B_{\text{Route type}}$ = 0.1908), the business model of the largest share airline ($B_{\text{Largest share}}$ = 0.2997), origin airport (B_{Origin} = 0.4350), and destination airport (B_{Dest} = 0.5029) showed positive signs. It signified a likelihood of making exit decisions in the market route when the number of carrier increase, when it is a nonstop market, when the largest share airline is a LCC, and when the origin and destination airports are large hubs.

Magnitude interpretations. As reported in Table 13, airlines were 1.0002, 1.0003, 1.0003, and 1.2651 times more likely to be involved in exit decisions than maintaining their schedule if there is a decrease by one passenger in demand, by $1 in market fare, by 1 mile in distance, and there is an operation of at least one LCC in a given route. The corresponding odds ratios of making exit decisions also decreased by 0.02% (= 0.9998 – 1), 0.03% (= 0.9997 – 1), 0.03% (= 0.9997 – 1), and 20.95% (= 0.7905 – 1), respectively. On the other hand, airlines were 2.7933, 1.2102, 1.3495, 1.5450, and 1.6534 times more likely to exit/reduce the frequency in the route if one more carrier enters the competition, if a nonstop market, if the largest share airline is a LCC, and if either origin or destination airport are large hubs. The corresponding odds ratios of making exit decisions in routes increased by 179.33%, 21.02%, 34.95%, 54.50%, and 65.34%, respectively.

Table 13

Summary of Logistic Regression Estimates for the Null and Simultaneous Model of Exit Patterns

	B_i	*SE*	χ^2	*p*	Odds Ratio	95% CI	Reciprocal
Null Model							
Constant	0.5146	0.0086	3564.8	<.0001***			
Full Model							
Constant	-0.8049	0.0324	616.05	<.0001***			
Total passengers	-0.0002	0.00008	5.52	.0188*	0.9998	[0.9996, 0.9999]	1.0002
Market fare	-0.0003	0.00007	16.79	<.0001***	0.9997	[0.9995, 0.9998]	1.0003
Number of carriers	1.0272	0.0140	5337.5	<.0001***	2.7933	[2.7717, 2.8713]	0.3580
Distance	-0.0003	0.00001	480.69	<.0001***	0.9997	[0.9996, 0.9997]	1.0003
Route type	0.1908	0.0592	10.40	.0013**	1.2102	[1.0777, 1.3590]	0.8263
LCC existence	-0.2351	0.0742	10.05	.0015**	0.7905	[0.6835, 0.9142]	1.2651
Business model of the largest share	0.2997	0.0890	11.33	.0008**	1.3495	[1.1334, 1.6068]	0.7410
Origin hub	0.4350	0.0310	197.04	<.0001***	1.5450	[1.4540, 1.6418]	0.6472
Destination hub	0.5029	0.0312	259.99	<.0001***	1.6534	[1.5554, 1.7576]	0.6048

Note. N = 57,488. R_L^2 = .1348
*** p < .0001

Alternatively, marginal effects for the independent variables were reported and interpreted in Table 14. Again, the marginal effects calculated in both ways were almost identical, and marginal effects at the mean were used for interpretations. For an additional passenger in demand, $1 increase in market fare, 1-mile increase in distance, and an existing LCC operation in a given route, airlines were 0.005%, 0.007%, 0.007%, and 5.28% less likely to make an exit decision, respectively. In contrast, under independent conditions, one more competitor, nonstop market, a LCC holding the largest share, origin large hub, or destination large hub, airlines were 23.05%, 4.28%, 6.73%, 9.76%, and 11.29% more likely to make an exit decision from the routes.

Table 14

Summary of Marginal Effects for the Logistic Model of Exit Patterns

	Mean (*M*)	Logistic Coefficient for Exit Patterns	Marginal Effects at the Mean[a]	Average Marginal Effects[b]
Total Passengers	68.32	-0.0002***	-0.00005	-0.00005
Market Fare	310.80	-0.0003***	-0.00007	-0.00007
Number of Carriers	1.80	1.0272***	0.2305	0.2394
Distance	1345.95	-0.0003***	-0.00007	-0.00007
Route Type	0.09	0.1908***	0.0428	0.0445
LCC Existence	0.11	-0.2351***	-0.0528	-0.0548
Business Model of the Largest Share	0.05	0.2997***	0.0673	0.0699
Origin Hub	0.13	0.4350***	0.0976	0.1014
Destination Hub	0.13	0.5029***	0.1129	0.1172

Note. N = 57,488. Logit equation = $-0.0002X_{\text{Pax}} - 0.0003X_{\text{Fare}} + 1.0272X_{\text{Carriers}} - 0.0003X_{\text{Distance}} + 0.1908X_{\text{Route Type}} - 0.2351X_{\text{LCC}} + 0.2997X_{\text{Largest Share}} + 0.4350X_{\text{Origin}} + 0.5029X_{\text{Dest}} - 0.8049$.

[a]Logit Value at the Mean was calculated by substituting the means of regressors into the Logit equation. Logit Value = 0.675. Odds = $e^{0.675} = 1.965$. Probability Pr (Y = 1 | X) = $e^{0.675} / (1 + e^{0.675}) = 0.66$, Pr (Y = 0 | X) = 1 - 0.66 = 0.34. Marginal Effect at the Means, $\delta_p/\delta_{xj} = f(\overline{x}'\beta) * (1 - f(\overline{x}'\beta)) * \beta_j$ (Greene, 2011), which in our study equals 0.66 x 0.34 x Logistic coefficients. [b]Average predicted probabilities were obtained from *JMP* output for each case before taking an average, Pr (Y = 1 | X) = 0.63, Pr (Y = 0 | X) = 0.37. Average Marginal Effects, $\delta_p/\delta_{xj} = \frac{\Sigma f(x'\beta)}{n} * (1 - \frac{\Sigma f(x'\beta)}{n}) * \beta_j$ (Greene, 2011), which in our study equals 0.63 x 0.37 x Logistic coefficients, *$p < .05$. **$p < .01$. ***$p < .001$

Classification accuracy. Classifications also can be used as supplementary analyses to determine the goodness of fit of a logistic regression model (Cohen et al., 2003; Hair et al., 2010). We compared the statistical classifications of group memberships in the full models to actual group memberships by determining predicted probabilities for each case and developing contingency tables of predicted versus actual group membership. With respect to the entry pattern model as reported in Table 15, 35201 (= 22416 + 12785) cases were classified as belonging to the Entry group, and 22287 (= 9695 + 12592) cases to the No Entry group in the full model. There were 35008 out of 57488 cases, which was 61%, correctly classified in the full model at the predicted probability cut of 0.5. These correctly classified cases consisted of 22416 cases (70% hit rate) and 12592 cases (50% correct rejection rate). With respect to the exit pattern model as reported in Table 16, 35154 (= 26456 + 8698) cases were classified as belonging to the Exit group, and 22334 (= 9525 + 12809) cases to the No Exit group in the full model. There were 39265 out of 57488 cases, which was 68%, correctly classified in the full model at the predicted probability cut of 0.5. These correctly classified cases consisted of 26456 cases (74% hit rate) and 12809 (60% correct rejection rate).

Table 15

Classification Matrix for Entry Pattern Model

Actual Group Membership	Predicted Group Membership	
	Entry	No Entry
Entry	22416	9695
	(Hits[a] = 70%)	(Misses[b] = 30%)
No Entry	12785	12592
	(False Alarms[c] = 50%)	(Correct Rejections[d] = 50%)

Note. N = 57,488. The probability cut was equal to $p_i = 0.5$. Actual group membership was 32111 Entry cases (56%) and 25377 No Entry cases (44%).

[a]Hits were the accurate classification of Entry cases to membership in the Entry group. [b]Misses were the misclassification of Entry cases to membership in the No Entry group. [c]False alarms were the misclassification of No Entry cases to membership in the Entry group. [d]Correct rejections were the accurate classification of No Entry cases to membership in the No Entry group.

Table 16

Classification Matrix for Exit Pattern Model

Actual Group Membership	Predicted Group Membership	
	Exit	No Exit
Exit	26456	9525
	(Hits[a] = 74%)	(Misses[b] = 26%)
No Exit	8698	12809
	(False Alarms[c] = 40%)	(Correct Rejections[d] = 60%)

Note. N = 57,488. The probability cut was equal to $p_i = 0.5$. Actual group membership was 35981 cases (63%) and 21507 No Entry cases (37%).

[a]Hits were the accurate classification of Exit cases to membership in the Exit group. [b]Misses were the misclassification of Exit cases to membership in the No Exit group. [c]False alarms were the misclassification of No Exit cases to membership in the Exit group. [d]Correct rejections were the accurate classification of No Exit cases to membership in the No Exit group.

Results of Hypotheses Testing

The research hypotheses of the current study are restated here in null form for testing purposes. The decision to reject or fail to reject a null hypothesis relied on the results of the respective primary analyses.

Null hypothesis 1: When examined from the entry pattern model, none of the targeted variables will have significant predictive value relative to distinguishing between airline route entry and non-entry decisions. As reported in Table 9, the simultaneous model was statistically significant, $\chi^2(9) = 3441.82$, $p < .0001$. Given a significant overall model, the individual variables within this model were examined for significance. As reported in Table 10, seven of nine variables were significant at the preset alpha level of .05: Total passengers ($p < .0001$), Market fare ($p < .0001$), Number of carriers ($p < .0001$), Distance ($p < .0001$), LCC existence ($p < .0001$), Origin hub ($p < .0001$), and Destination hub ($p < .0001$). Therefore, the decision was to reject the null hypothesis 1, and to accept the alternative hypothesis 1 that stated when examined from the entry pattern model, at least one of the targeted variables will have significant predictive value relative to distinguishing between airline route entry and non-entry decisions.

Null hypothesis 2: When examined from the exit pattern model, none of the targeted variables will have significant predictive value relative to distinguishing between airline route exit and non-exit decisions. As reported in Table 12, the simultaneous model was statistically significant, $\chi^2(9) = 10249.31$, $p < .0001$. Given a significant overall model, the individual variables within this model were examined for significance. As reported in Table 13, all nine variables were significant at the preset alpha level of .05: Total passengers ($p = .0188$), Market fare ($p < .0001$), Number of carriers ($p < .0001$), Distance ($p < .0001$), Route type ($p = .0013$), LCC existence ($p = .0015$), Business model of the largest share ($p = .0008$), Origin hub ($p < .0001$), and Destination hub ($p < .0001$). Therefore, the decision was to reject the null hypothesis 2, and to accept the alternative hypothesis 2 that stated when examined from the exit pattern model, at least one of the targeted variables will have significant predictive value relative to distinguishing between airline route exit and non-exit decisions.

Conclusions

With respect to entry pattern decisions, the simultaneous logistic regression yielded seven significant factors that distinguished between entry and non-entry decisions in a given U.S. route. For a 100-passenger increase, airlines were 7% more likely to enter a new route or increase the frequency in their existing routes. In this case, the purpose of making entry decisions is to increase available seats per mile (ASM) to meet the increasing demand on the given route. For a 100-dollar increase in market fare, airlines were 5% more likely to enter/increase their operations. A high air fare would be a clue of a profitable market, which are appealing on an eye of airline network planners. With an appearance of one new competitor on a given route, airlines were 12.31% more likely to increase their operations. The probable reason is that in order to maintain the current market share on the O&D airport pair, airlines are likely to compete against others by increasing the frequency, which would provide passengers a less total trip time (Belobaba, Odoni, & Barnhart, 2015). For a 100-mile increase in distance between O&D airports, airlines were 7% more likely to enter/increase the number of departures on that route. On long haul routes, airlines could leverage the economies of scale, low operating costs, high aircraft utilization. Additionally, passengers on long range flights usually have a high willingness-to-pay for ticket fares as well as additional on-board services.

On routes with at least one operations of a LCC, airlines were 6% less likely to make an entry decision. Indeed, competition on route market with the appearance of LCCs is stiffer and fiercer due to its large effect on average air fare (Brueckner, Lee, & Singer, 2013). Furthermore, airlines were 8.45% less likely to enter/increase their operations on the route in which origin airport is a large hub, and 10.43% less likely to enter/increase their operations in which destination airport is a large hub. Airlines, especially low-cost carriers, have a tendency to move their operations away of large hubs to avoid higher landing fees, terminal congestions, or self-uncontrollable delays. Instead, routes with departures and arrivals performed at secondary airports within the targeted airport's catchment area are strategically taken into consideration.

With respect to exit pattern decision, the simultaneous logistic regression yielded all nine significant factors that distinguished between exit and non-exit decisions in a given U.S. route. For an additional 100 passengers in demand, airlines were 5% less likely to make exit/reduce their operations in the given route. It was consistent between entry and exit patterns, which airlines tend to either enter the route or increase the frequency or at least maintain their frequency rather than making exit decision if the demand is growing. Moreover, the consistency in entry and exit patterns also were reflected through affective factors: market fare, distance between origin and destination airports, and origin and destination large hubs. For either 100-dollar increase in market fare and 100-mile increase in distance, airlines were 7% less likely to make an exit decision, but 9.76%, and 11.29% more likely to exit/reduce their flights if the origin and destination airports are large hubs in the given route. Obviously, airlines are considering long haul routes or those with high yield to be profitable market, at the same time, avoiding large hubs that might potentially cost airlines the most compared to medium or small hubs.

However, the two models also produced conflicting results that pointed to factors: number of carriers, route type, LCC existence, and the airline business model with the largest share. For one more competitor joining the competition, airlines were 12.31% more likely to increase their operations, and 23.05% more likely to reduce the operations or stop their service on the given route. The magnitude of making exit decisions were nearly double over that of making entry decisions. It indicates that although airlines could proactively increase their frequency to compete with others, they are still preparing exit strategies once the competition becomes heavier and fiercer and leads to a "thin-razor" profit margin. The operation of at least one incumbent LCC could put newcomer airlines on initial reluctance to enter their flights, but existing airlines were 5.28% less likely to make exit/reduce their flights. Although LCC operations could make the competition more difficult by introducing cheaper fares, the existing airlines do not consider it to be serious threats because of using other strategies such as higher frequencies, on-board entertainment, baggage lost-and-found services to offset their higher fares. Route type and the business model of the largest share airline were the two factors that not significantly affected entry decisions, but exit decisions. For nonstop route markets, airlines were 4.28% more likely to reduce their flights or cease their operations on the route. It is commonly accepted in the literature that the level of service with nonstop flights in a nonstop market is a the most important and significant factor in attracting the attention of passengers' choice (Coldren, 2005; Coldren et al., 2003; Garrow, 2010). Therefore, airlines having connecting flights are highly likely to lose their passengers to those having nonstop flights, and thus the former would cut down the frequency of connecting flights. If the largest share of a given route market is accounted by a LCC, airlines were 6.73% more likely to exit their operations out of the market or reduce their flights. Passengers on route markets with a LCC competition are mostly leisure travelers who are sensitive to price and thus are simply attracted by affordable fares of LCCs (Belobaba et al., 2015). Hence, once a LCC holds the largest share and dominates the route market, the market structure will probably be fixed, and consequently it is difficult for the remaining airlines to overturn the situation. The remaining airlines would reduce their frequency to avoid directly the competition with the largest LCC, or in worst scenarios, completely cease their operations on the route market.

Finally, the findings of the study make an implication to airline planners in understanding key drivers as well as barriers to entry and exit decisions. The tasks of airline planners in network planning are to draw the network and route map of airlines in general, and constantly evaluate the efficiency and profitability of each route on the network. Therefore, in aid of significant affective factors found in this study, they could gain predictive insights before making right decisions, and could assess their competitors' decisions in the same routes of the network. Further implication speaks to airport operators at large hubs in more and more airlines moving their operations to secondary airports in which the airlines have low operating costs.

Generalizability, Limitations, and Delimitations

The sample used in the study was also the accessible population that is indeed a census of the target population—10% random sample of all U.S. itineraries reported in the 2018 dataset for Quarter 1. For this reason, the sample analyzed in the study was somehow highly representative to the target population, and thus the results could be generalizable to the target population. However, the ecological generalizability could be limited to only the U.S. domestic market because of the unique characteristics of the market that make it difficult to transcend to other market. For example, on international itineraries, flights are predominantly operated in long haul routes, and these transcontinental flights usually place their operations at large hubs for connections to spoke cities.

One limitation of the study pointed to the data integrity, which means that we had no control over how the data were reported and stored in the DB1B database. Therefore, any changes are made to the dataset subsequent to the current study, then any replication studies could yield different results. In the meanwhile, the delimitation of the study was relative to the data collection period that was the 2018 dataset for Quarter 1; therefore, similar studies that use a different data collection period might not get the same results. Other minor delimitations were our choice of eliminating all outliers in the dataset and using dummy coding strategy for categorical variables in the study. Other studies that decide to keep outliers in the dataset for analyses and use other coding techniques such as effects coding or contrast coding might yield different results.

References

Abdelghany, A., & Abdelghany, K. (2009). *Modeling applications in the airline industry*. Abingdon, U.K.: Routledge.

Abdelghany, A., & Guzhva, V. S. (2010). Analyzing airlines market service using panel data. *Journal of Air Transport Management*, *16*(1), 20–25.

Airport Categories – Airports. (n.d.). Retrieved November 21, 2017, from https://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/categories/

Baran, S. (2018). *Survival in the U.S. domestic airline market: Strategies for entry, exit, and air fare competition*. Florida Institute of Technology.

Belobaba, P., Odoni, A. R., & Barnhart, C. (2015). *The global airline industry* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Boguslaski, C., Ito, H., & Lee, D. (2004). Entry patterns in the southwest airlines route system. *Review of Industrial Organization*, *25*(3), 317–350.

Brueckner, J. K., Lee, D., & Singer, E. S. (2013). Airline competition and domestic US airfares: A comprehensive reappraisal. *Economics of Transportation*, *2*(1), 1–17.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erblaum Associates.

Coldren, G. M. (2005). *Modeling the competitive dynamic among air-travel itineraries with generalized extreme value models*. Northwestern University.

Coldren, G. M., Koppelman, F. S., Kasturirangan, K., & Mukherjee, A. (2003). Modeling aggregate air-travel itinerary shares: logit model development at a major US airline. *Journal of Air Transport Management*, *9*(6), 361–369.

Garrow, L. A. (2010). *Discrete Choice Modelling and Air Travel Demand*. Farnham, England: Ashgate.

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Hair, J. F., Black, W. C., Babin, B. J., & Rolph, A. E. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Pearson Education.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). New York, NY: John Wiley & Sons.

Ito, H., & Lee, D. (2003). Low cost carrier growth in the U.S. airline industry: Past, present, and future. *Brown University Department of Economics Paper*.

Katchova, A. L. (2013). *Bivariate probit and logit models example*. Retrieved from https://sites.google.com/site/econometricsacademy/

Oliveira, A. V. (2008). An empirical model of low-cost carrier entry. *Transportation Research Part A: Policy and Practice, 42*(4), 673–695.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12), 1373–1379.

SAS Institute Inc. (2016). *JMP® 13 Multivariate Methods*. Cary, NC: SAS Institute Inc.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Upper Saddle River, NJ: Pearson Education.

Warner, R. M. (2008). *Applied statistics: From bivariate through multivariate techniques*. Thousand Oaks, CA: Sage Publications Inc.

Yuan, Z. (2016). *Essays on Network Competition in the Airline Industry* (Doctoral dissertation). University of Toronto.