Survival in the U.S. Domestic Airline Market:
Strategies for Entry, Exit, and Air Fare Competition

by
Selçuk Baran

Bachelor of Science
Industrial Engineering
Faculty of Engineering, Ankara, Turkey
Bilkent University
2007

Master of Science
Engineering Management
Florida Institute of Technology
2014

A dissertation submitted to the College of Aeronautics
Florida Institute of Technology
as part of the degree requirements for a

Doctor of Philosophy
in
Aviation Sciences

Melbourne, Florida
May, 2018
© Copyright 2017 Selçuk Baran

All Rights Reserved

The author grants permission to make single copies
We the undersigned committee
hereby approve the attached dissertation

Survival in the U.S. Domestic Airline Market:
Strategies for Entry, Exit, and Air Fare Competition

by

Selçuk Baran

__________________________
Michael A. Gallo, Ph.D.
Professor
College of Aeronautics
Committee Chair

__________________________
John Deaton, Ph.D.
Professor
College of Aeronautics

__________________________
Ulreen Jones, Ph.D.
Assistant Professor
College of Aeronautics

__________________________
Barry Webster, Ph.D.
Assistant Professor
Engineering Systems

__________________________
Korhan Oyman, Ph.D., Dean, College of Aeronautics
ABSTRACT

TITLE: Survival in the U.S. Domestic Airline Market: Strategies for Entry, Exit, and Air Fare Competition

AUTHOR: Selçuk Baran

MAJOR ADVISOR: Michael A. Gallo, Ph.D.

The purpose of this study was to identify factors that distinguished between airlines entering new routes or exiting existing routes between 2011–2015, and determine factors related to air fare levels between 2005–2015. Part A employed a cause-type ex post facto design to determine the relationship between the targeted variables and the criterion variable, which distinguished between the “enter” and “exit” groups. Part B employed an explanatory correlational design to measure the relationship between the targeted variables and air fare levels. Research factors included carrier type (full service vs. low-cost), route length, city populations and per capita income at each endpoint of a city-pair market, market concentration, number of competitors, the presence of a hub airport at the origin or destination airports, average air fare level of the airline with the largest average air fare in a city-pair market, and the carrier type of the airline having the largest fare. The sample consisted of 2,111 cases for Part A and 1,082 cases for Part B.

A logistic regression analysis (Part A) found that airlines were: 1.5 times more likely to enter routes at least 850 miles long vs. shorter than 850 miles, 1.5 times more likely to enter a new route at Endpoint 1 with a city population of at
least 2.8 million vs. fewer than 2.8 million, 1.33 times more likely to enter a new route at Endpoint 1 with a city per capita income of at least $47,254 vs. fewer than $47,254, and nearly two times more likely to enter a new route with one or more competitors than no competitors. FSCs also were 1.7 times more likely than LCCs to enter a new route. A multiple regression analysis (Part B) found that all targeted factors had a significant relationship with air fares. Findings suggest there are still distinctions between FSC and LCC business models, and airlines might consider expanding route maps in rural markets instead of major cities or metropolitan areas.
Table of Contents

Abstract ................................................................................................................................. iii

List of Figures ....................................................................................................................... x

List of Tables ........................................................................................................................ xi

Acknowledgements ............................................................................................................. xiii

Dedication ............................................................................................................................. xiv

Chapter 1: Introduction ........................................................................................................ 1

  Background and Purpose ..................................................................................................... 1

    Background ...................................................................................................................... 1

    Purpose .......................................................................................................................... 11

  Definition of Terms ........................................................................................................... 15

  Research Questions and Hypotheses ................................................................................ 19

    Research questions ......................................................................................................... 19

    Research hypotheses ...................................................................................................... 20

  Study Design ..................................................................................................................... 21

  Significance of the Study .................................................................................................. 21

  Study Limitations and Delimitations ................................................................................. 23

    Limitations .................................................................................................................... 23

    Delimitations ................................................................................................................ 25
Chapter 2: Review of Related Literature ................................................................. 27

Introduction .................................................................................................................. 27

Overview of Underlying Theory ............................................................................... 27

Part A: Route entry and exit decision patterns ........................................... 27

Part B: Air fare competition dynamics ................................................................. 31

Review of Past Research Studies .......................................................................... 36

Part A: Route entry and exit decision patterns ........................................... 36

Part B: Air fare competition dynamics ................................................................. 50

Summary and Concluding Remarks ..................................................................... 61

Chapter 3: Methodology ........................................................................................... 62

Population and Sample ......................................................................................... 62

Population .................................................................................................................. 62

Sample ......................................................................................................................... 63

Power analysis ........................................................................................................... 64

Instrumentation ......................................................................................................... 65

Procedures .................................................................................................................. 67

Research methodology ............................................................................................ 67

Human subjects research ......................................................................................... 68

Study implementation ............................................................................................... 68

Threats to internal validity ....................................................................................... 68

Treatment verification and fidelity .......................................................................... 74
Data Analysis ............................................................................................................. 75

Summary of IVs and DVs ............................................................................................. 78

Summary of statistical strategies used ................................................................. 78

Chapter 4: Results ........................................................................................................... 80

Introduction ..................................................................................................................... 80

Descriptive Statistics ................................................................................................... 80

Part A: Route entry and exit decision patterns ............................................. 80

Part B: Air fare competition dynamics ............................................................. 87

Inferential Statistics ..................................................................................................... 90

Preliminary analyses ................................................................................................. 91

Primary analysis 1: Logistic regression .............................................................. 104

Primary analysis 2: Multiple regression ............................................................... 113

Results of Hypotheses Testing ............................................................................... 118

Null hypothesis 1: Independent of set membership, and when examined from a simultaneous perspective, none of the targeted variables will have a significant predictive value in distinguishing between LCCs and FSCs that entered or exited a route ................................................. 119

Null hypothesis 2a: The set of route factors (Set R) will have no significant relationship with airlines’ air fare levels for the targeted city-pair markets ........................................... 119

Null hypothesis 2b: The set of market factors (Set M) will have no significant relationship with airlines’ air fare levels for the targeted city-pair markets when examined in the presence of the set of route factors .......... 120
Chapter 5: Conclusions, Implications, and Recommendations.......................... 122

Summary of Study ......................................................................................... 122

Summary of Findings.................................................................................... 125

Preliminary analyses ................................................................................... 125

Primary analysis 1: Logistic regression ....................................................... 125

Primary analysis 2: Multiple regression ..................................................... 127

Conclusions and Inferences......................................................................... 128

Research question 1: Independent of set membership, what is the relationship between the targeted variables and the dichotomous response variable that distinguished between LCCs and FSCs that entered a new route and those that exited from an existing route? ........................................... 129

Research question 2: When examined from a hierarchical perspective with set entry order R–M, what is the incremental knowledge gained at each step of the analysis relative to airlines’ air fare levels for the targeted city-pair markets? ..................... 133

Implications................................................................................................. 142

Implications relative to theory ................................................................. 142

Implications relative to prior research ..................................................... 147

Implications relative to practice ............................................................... 153

Generalizability, Limitations, and Delimitations ....................................... 159

Generalizability ......................................................................................... 159

Study limitations and delimitations......................................................... 160

viii
Recommendations for Future Research and Practice......................... 164

Recommendations for future research relative to study limitations................................................................. 164

Recommendations for future research relative to study delimitations................................................................ 166

Recommendations for future research relative to implications.............................................................................. 167

Recommendations for practice relative to implications........ 169

Final comments and observations................................................................. 170

References .................................................................................................................................................. 172

Appendix A: IRB Documents......................................................................................................................... 178

Appendix B: Raw Data .................................................................................................................................. 184
List of Figures

Chapter 2

2.1 Bain’s (1956) market structures that are related to barriers to entry (inner circle quadrants) applied to the current study ........ 29

2.2 Determinants of demand and relationship to price .................... 33

2.3 An example of the demand curve relative to air fares .................. 35

2.4 The number of route entries by NWCs (FSCs) and LCCs (all markets, 1996–2009) among U.S. domestic airlines .............. 48

2.5 Survival rates of market entries by NWCs (FSCs) and LCCs (all markets, entry years 1996–2007) .................................. 48

2.6 Percentage of entries by NWCs (FSCs) and LCCs into new markets (1996–2009) ................................................................. 49

2.7 An example of adjacent flight and connecting flights paradigm ................................................................. 58
List of Tables

Chapter 3

3.1 Post Hoc Power Analysis Based on Actual Sample Sizes and Effect Sizes ................................................................................... 66

3.2 Summary and Description of Independent and Dependent Variables Overall .................................................................................... 76

3.3 Summary of the Sets of Independent and Dependent Variables for Each Part of the Study ......................................................... 77

Chapter 4

4.1 Summary of Key Attributes Associated with the Eight Targeted Airlines .......................................................................................... 81

4.2 Descriptive Statistics of Factors Associated with Part A: Route Exit and Entry Strategies (Overall) .................................................. 83

4.3 Descriptive Statistics Relative to Route Entry Strategies by Carrier Type ......................................................................................... 85

4.4 Descriptive Statistics Relative to Route Exit Strategies by Carrier Type ......................................................................................... 87

4.5 Summary and Description of Factors Associated with Part B: Air Fare Competition Dynamics ..................................................... 88

4.6 Assignment of Variables into Dichotomies ........................................... 102

4.7 Significance of the Simultaneous (Full) Model ........................................ 105

4.8 Summary of Logistic Regression Estimates for the Null and Simultaneous (Full) Model ............................................................... 106
4.9 Summary of Odds Ratios for the Independent Variables in the Full (Simultaneous) Model .......................................................... 108

4.10 Summary of Hierarchical Multiple Regression Results............ 115

Chapter 5

5.1 Summary of the Results of Hypothesis Testing......................... 126
Acknowledgments

Merve Baran, thank you so much, my wonderful wife

Burak Efe and Sami, my cute sons. Thanks for not bothering me during my work

Dr. Gallo thank you for your extraordinary contribution to this dissertation and my career development.

Dr. Jones, your dissertation was a great guideline for me. Thank you very much.

Family, friends, and colleagues thank you for being there for me.

Thank you very much my committee members, Dr. Webster, Dr. Jones, and Dr. Deaton for your advice and support.
Dedication

In memory of

Mehmet and Ayse Baran
Osmaniye, Turkey
Chapter 1
Introduction

Background and Purpose

Background. The survival strategies of U.S. airlines relative to route entry and exit decisions and air fares changed radically after the aviation market was deregulated in 1978 as a result of the Airline Deregulation Act (Daraban, 2012). Before deregulation, competition among the airlines was minimal because there were barriers to entry into the market and several price restrictions, which were monitored and controlled by the Civil Aeronautics Board (CAB; Hendricks,Feuille, & Szersen, 1980). The Airline Deregulation Act removed most of these barriers to entry and most of the price restrictions.

The Airline Deregulation Act had a snowballing effect on the airline market resulting in mostly new and innovative scheduling, tariff, network, and pricing strategies. For example, airlines began deciding their own fare levels for each of their routes, and they selected the routes they wanted to enter or exit by themselves without the restrictions that were in place prior to deregulation. Airline passengers also benefitted from these changes. After deregulation, the number of airlines in the U.S. domestic airline market increased substantially. Goetz and Vowles (2009) reported that the period from 1978 to 1983 should be named as the rise of new entrants because the newly deregulated industry witnessed an influx of new entrants such as PEOPLExpress, New York Air, America West and Southwest in
many markets that had previously been protected by regulations. The domestic market share of former trunks in terms of revenue passenger miles declined from 87% to 75% from 1978 to 1983 because of new entrants (Goetz & Sutton, 1997).

As the number of airlines in the U.S. domestic airline market increased, so too did competition among the airlines, requiring them to expand their capacity and increase the number of routes they offered (Goetz & Vowles, 2009). This increased competition also resulted in lower air fare levels, which in turn led to an increase in the penetration of the public selecting air transportation for traveling purposes (Goetz, 2002). These lower price levels, however, did not support large profit margins (International Air Transportation Association, IATA, 2013a). At the same time, though, the relative cost of flying was reduced by 60%, the aircraft themselves became 70% more energy efficient and 75% quieter, and the utilization levels of airports and aircraft also increased (International Civil Aviation Organization, ICAO, 2011). Deregulation all over the world also provoked innovation in the airline industry and new business models such as low-cost carriers emerged. All of these activities led to an increase in the number of passengers and in the number of aircraft in use. Over the past 3 decades, the airline industry has continued to add more direct flights between cities. The number of such flights all over the world went from slightly over 6,000 in 1980 to over 15,000 in 2012, an increase of 250% (IATA, 2013b).
Because deregulation removed restrictions and regulations within the airline industry, many airlines have engaged in games of strategy to help acquire a competitive advantage over their rivals. Before the Airline Deregulation Act, the U.S. government regulated air fares and kept them at particular levels. This resulted in stable, or fixed, ticket prices for each of the city-pair markets. After deregulation, though, the regulations and restrictions related to price levels were removed and the rivalry within the airline industry became intense. This resulted in the airline industry becoming a free market economy. Revenue management concepts came into effect, and the fixed prices for city-pair markets were no longer effective.

Therefore, it became vital to investigate the expansion policies, the entry and exit decision patterns of airlines, and the determinants of average price levels, which are the most important factors affecting profitability in airline markets. The literature is replete with studies that have modeled the strategy game among airlines. Hüschelrath and Müller (2011) explained that the existing literature about the issue could sketchily be split into two strands: ”the determinants of entry” literature and “the effects of entry” literature.

In the “determinants of entry” literature, the research focus was on entry patterns of specific airlines such as Southwest, JetBlue Airways, and Gol Airlines, or the general entry patterns of groups of airlines in terms of business models such as low-cost carriers (LCCs) and full size carriers (FSCs). As an example of the former, Oliveira (2008) examined the entry patterns of LCC Gol Airlines in the
Brazilian domestic market in 2001, and compared Gol Airline’s entry patterns with two well known LCCs, Southwest Airlines and JetBlue Airways. Oliveira concluded that market size, which can be measured by density, and route length and rivals’ presence at the route were primary determinants of profitability. Unlike previous studies, Oliveira used an origin and destination (O&D) markets paradigm, which involved an origin airport (e.g., MLB), a destination airport (e.g., JFK), and a transfer airport (e.g., ATL) located between the O&D markets. It should be noted that most of the previous studies such as Boguslaski, Ito, and Lee (2004) used a non-stop markets paradigm, which ignored routings with flight connections and stops. The transfer passengers are important in determining route entry decision patterns of airlines. Therefore, an O&D markets paradigm is needed to get more meaningful results about entry decision patterns of airlines. Oliveira used the data for O&D markets to detect the determinants of entry decisions of airlines.

As another example that focused on a specific airline, Boguslaski et al. (2004) examined the entry patterns in Southwest Airlines’ route system. The study analyzed the patterns of non-stop entry by Southwest into city-pair markets and detected the changes in entry strategies. The study also evaluated the effect of the Wright and Shelby Amendments, which limited the markets that Southwest could serve and measured the fare saving as a consequence of these restrictions. Lastly, Boguslaski et al. predicted the potential candidate routes for future non-stop entry and assessed the magnitude of threat on other airlines in domestic market because
of Southwest’s expansion. Bogulaski et al. reported that actual and potential passenger density, distance, the hubs of its competitors, and per capita income at the endpoints of a market were the most influential factors when Southwest selects which city-pairs to enter. Bogulaski et al.’s study extended the views of researchers by focusing on potential future routes, the magnitude of threats of airlines on each other, and the expansion strategy of an airline.

In contrast to Oliveira (2008) and Boguslaski (2004), Lederman and Januszewski (2003) examined the entry patterns of a group of low-cost carriers relative to their business models. Lederman and Januszewski analyzed the extent to which product differentiation was a significant element of LCCs’ entry strategies by observing the specialties of the routes that LCCs entered. Lederman and Januszewski reported that the determinants of the entry patterns were similar with the previous studies and these were grouped as a carrier’s own characteristics, exogenous market characteristics, and the presence of competitors at the endpoint airports of the route.

Complementing these previous studies, Ito and Lee (2003) examined the past, present, and future of low-cost carrier growth in the U.S. airline industry by examining the market characteristics that influenced LCC entry in the U.S. airline industry from 1990 to 2002. Ito and Lee analyzed some formalized realities about the growth of LCCs, the uniqueness of the markets that LCCs enter, and the potential for LCC growth and its impact on the network carriers. Ito and Lee
analyzed the five largest LCCs during the period: Southwest, AirTran, ATA (formerly known as American Trans Air), JetBlue, and Frontier. Ito and Lee’s findings refuted the claim that an incumbent’s price or capacity reaction to low-cost carrier entry has a negative impact on the probability of exit of the entrant. Ito and Lee’s study revealed that LCCs started to enter longer routes and they predicted that the threat from LCCs to FSCs would increase in the future, which warrants an analysis of future competition between LCCs and FSCs.

Hüschelrath and Müller (2011) examined the patterns and effects of entry in the U.S. airline markets by analyzing the top 500 non-stop U.S. airline markets from 1996 to 2009. Patterns of entry in the U.S. airline market were analyzed with a descriptive approach, and the effects of entry were analyzed with both a descriptive and econometric approach. Hüschelrath and Müller reported that entry activity of low-cost carriers resulted in substantial fare reductions. Hüschelrath and Müller concluded that although route entry patterns of FSCs do not have similar effects, the entry patterns and expansion policies of low-cost carriers must be considered as the main driver of competition in the domestic U.S. airline industry. Hüschelrath and Müller showed that the survival rates for LCC entries were more than FSC entries, and LCCs started to enter long haul markets. The descriptive approach of Hüschelrath and Müller’s study revealed many important sides of the topic, including a detailed map of all entries by LCCs and FSCs. Different from the
other studies, Hüschelrath and Müller focused on recent aspects of LCC and FSC competition.

Independent of route entry decisions in the “effects of entry” literature, variations in air fare price levels in airline markets is one of the hot topics in the air transport management literature. The literature also is replete with the studies examining the determinants of air fare price levels in airline markets. For example, Vowles (2000) analyzed the effect of LCCs on air fares in the U.S. with an emphasis on geographic and competition issues. The factors included the presence of low fare carriers in a market, hub domination, market share, and type of destination. The data set included 1,000 largest domestic O&D markets in the U.S., and the study period was for the 1st quarter of 1997. Although the narrow study period limited the generalizability of the results, Vowles’ study was one of the first that examined the determinants of air fare levels in U.S. domestic airline markets.

In a separate study, Vowles (2006) examined the air fare pricing determinants in hub-to-hub markets. He reported that route type, presence and type of low fare carrier, and competition in hub-to-hub markets were the primary factors affecting pricing in hub-to-hub markets in the U.S. domestic airline industry. The contribution of Vowles’ study was related to its classification of routes as independent variables. Vowles classified the routes as dominant, dominant-to-secondary, secondary-to-secondary, and primary-to primary in terms of dominance of carriers in endpoints of a route and included these types of routes in the final
model as determinants of air fare pricing. Vowles concluded that the most prominent air fare pricing factors in hub-to-hub markets were route type, low fare carrier, and competition. His study showed that hub-to-hub markets were unique markets, however the factors affecting air fare pricing were similar to other types of markets. Although the novelty of Vowles’ study was his focus on different route types, the types of the routes defined by Vowles were not effective in explaining air fare pricing determinants as expected.

Zhang, Derudder, and Witlox (2013) expanded the studies of Vowles (2000, 2006) by analyzing the determinants of air fare levels of full service carriers as hub hierarchy and market competition in U.S. hub-to-hub markets. O&D markets connected by 17 hubs in the networks of U.S. FSCs were included in Zhang et al.’s model. Their study was not comprehensive, though, because Zhang et al. ignored most of the FSCs and the effect of LCCs on air fare variations. However, Zhang et al.’s study was different from previous studies because they focused on a hub hierarchy paradigm first in U.S. hub-to-hub markets. They reported that an airport’s importance in carriers’ hub hierarchies (primary or secondary), competition from low-cost carriers, and other market structure variables influenced average air fares.

Abda, Belobaba, and Swelbar (2012) examined the impacts of LCC growth on domestic traffic and fares at the largest U.S. airports. Abda et al. surveyed the development of domestic O&D air traffic and fares at the top 200 airports in the U.S. between 1990 and 2008. It should be noted that Abda et al. emphasized the
impacts of LCC entry and growth in the study. The novelty of Abda et al. was
related to their approach in which they considered each airport as a market and the
airlines that had operations to or from the airport could be considered as firms
competing in it. Critical to Abda et al.’s study was the business models of the
carriers, and Abda et al. found a diminishing influence of LCCs on market
parameters of the airline industry.

Brueckner, Lee, and Singer (2013) analyzed airline competition and
domestic U.S. airfares with a comprehensive reappraisal approach. They extended
previous research by utilizing the adjacent airport approach and by focusing on the
comprehensive effects of both low-cost and full service carriers (also referred to as
“legacy” carriers) on air fare level variations. As part of their study, Brueckner et
al. also considered nonstop and connecting markets, and these extensions of their
study were the first in the area. Brueckner et al. reported that although most forms
of legacy carrier competition have weak effects on average fares, low-cost carrier
competition radically affects average air fare levels, whether it occurs on the
airport-pair, at adjacent airports, or as potential competition. Their study was one of
the most comprehensive studies about the topic, and it was the first to analyze
within a single study nonstop and connected markets.

The previous studies cited in the foregoing paragraphs demonstrated there is
still a gap in the published literature relative to understanding the entry and exit
strategies of LCCs and FSCs. Many of these studies also were conducted before
2005. To help bridge the gap in the literature, the current study included the time period between 2010 and 2015, and therefore is the most recent study about the topic. The previous studies also provided guidance to the current study with respect to identifying appropriate factors relative to entry-exit decisions. For example, Oliveira (2008) identified market size, route length, market competition, and the endpoints of a city-pair as having an influence on whether an airline enters a new route or exits an existing route. Similarly, Boguslaski et al. (2004) also concluded that the endpoints of city-pairs are critical as well as whether a competitor has a hub in a route and the per capita income at the endpoints of a city-pair. Lederman and Januszewski (2003) and Ito and Lee (2003) also indicated the importance of considering the presence of competitors with a route. As a result, all of these factors were considered in the current study. With respect to air fare dynamics, most of the studies cited were relatively recent and flagged factors such as market share, route type, market competitiveness as a function of market concentration, and the carrier type (Abda, 2012; Brueckner et al., 2013; Vowles, 2000, 2006; and Zhang et al., 2013). Therefore, these factors also were included in the current study.

Daraban (2012) explained it has become obvious that the modifications on business models as a consequence of the competition between LCCs and FSCs could potentially result in convergence to a state where the apparent distinction between the two models becomes more and more imprecise. The Center For Aviation (2014) emphasized that LCC and FSC models are converging day by day
in terms of various dimensions because some FSCs try to decrease their cost to match LCCs. Daraban reported that the nature of networks of LCCs and FSCs seem to be altering as they operate longer-haul flights between airports they were primarily escaping from (e.g., legacy hubs), with an improved dependence on transfer passengers with connecting flights and sometimes even hub airports.

Given the possibility of an improved bottom line, it is reasonable to assume that the convergence between the LCC and the FSC business models will continue into the future. As a result, it would be prudent to compare the two business models’ entry and exit strategies to gauge how likely it is for one carrier type to enter or exit a route, which factors have the strongest relationship to a carrier’s decision to enter or exit a route, and what factors influence the air fares the two types of carriers assign to a route. This was the focus of the current study.

**Purpose.** The purpose of this study was to examine the survival strategies of U.S. domestic airlines relative to the routes they fly. In the context of the current study survival strategies were defined with respect to two parts: Part A corresponded to route entry and exit decision patterns, and Part B corresponded to air fare competition dynamics. The study was restricted to passenger airline routes that have both origin and destination airports within the borders of the United States, and examined city-pair markets, which are specific markets in which at least one commercial airline is operating. Examples of city-pair markets are New York–Orlando, Miami–Atlanta, and Boston–Houston.
Part A: Route entry and exit decision patterns. Route entry and exit decision patterns are factors that airlines consider when deciding whether or not to begin operating in a new route or stop operating in an existing route to increase market share or to exist in profitable markets. The strategy of an airline determines all these factors. Moreover, the strategy of an airline should be consistent with its mission, vision, and business model. Some of the crucial aspects of the strategy of an airline that determines route entry decision patterns are measurable. These include the route length, population and per capita income of the city-pairs that form the route, market concentration, the number of competitors in the route, and the existence of a hub airport of an airline in the origin or destination airports of the route. Because I was interested in the extent to which these factors are related to airlines’ route entry-exit decisions, the criterion variable was dichotomous in nature: either an airline entered a new route (Yes) or exited from an existing route (No). As a result, my examination of airlines’ route entry-exit decision patterns was to determine the relationship between the targeted variables and the dichotomous response variable that distinguished between the Yes and No groups. In this context the targeted factors related to the two groups were predictor variables that discriminated between airlines that entered a new route and airlines that exited from an existing route. Following Cohen, Cohen, West, and Aiken’s (2003) recommendation, the predictor variables were partitioned into three functional sets:
• Set R = Route Factors consisted of $X_1 = \text{Carrier type}$, which distinguished between airlines that were full-service carriers (FSCs) and low-cost carrier (LCCs); $X_2 = \text{Route length}$, $X_{3a}$ and $X_{3b} = \text{City population at each end-point of a city-pair market, respectively}$; and $X_{4a}$ and $X_{4b} = \text{Per capita income at each end-point of the city-pairs that formed the route, respectively}$.

• Set C = Competitor Factors and consisted of $X_{5a} = \text{Total market concentration of the route}$, $X_{5b} = \text{Market concentration of the incumbent airlines in each route relative to carrier type}$, and $X_7 = \text{Number of competitors within a route}$.

• Set A = Airport Factors consisted of the single factor $X_8 = \text{Existence of a hub airport of an airline in the origin or destination airports of the route (yes or no)}$.

Daraban (2012) reported that after deregulation there was intense competition between LCCs and FSCs, and in recent years there has been a convergence in some key features of these competing airline business models. The airlines included in Part A were the top eight airlines in the U.S. domestic market in terms of market share. The FSCs included in the study were Alaska Airlines, American Airlines, Delta Airlines, SkyWest, and United Airlines, and the LCCs were Jet Blue, Southwest, and Spirit Airlines. The study’s targeted timeframe was
between 2010 and 2015, inclusive, and all metrics were constructed using data from the Bureau of Transportation Statistics (BTS).

**Part B: Air fare competition dynamics.** Air fare competition dynamics refer to factors that influence the average air fare an airline assigns to a city-pair market. The average price level of a city-pair market is one of the most important determinants of an airline’s profitability, and the targeted factors that can influence average price levels are both route and market related.

In the current study, the route factors were exactly the same as those in Set R, which was presented above in Part A. The targeted market factors were assigned to Set $M = \text{Market Factors}$ and consisted of: $X_{5a} = \text{Total market concentration of the route}$, which also was a variable in Set $C$ as presented above in Part A; $X_9 = \text{Market size}$; $X_{10a} - X_{10h} = \text{Market share of eight competing airlines}$; $X_{11} = \text{Largest fare}$, which was the mean air fare of the airline having the largest average air fare in a city-pair market; and $X_{13} = \text{Type of carrier (FSC or LCC) having the largest fare in a route}$.

For the current study, I examined the relationship Sets $R$ and $M$ had with the overall mean air fares in the U.S. airline domestic city-pair market. Thus, in this context, the mean air fares were the criterion variable, the targeted sets of factors were the predictors, and I determined both the collective and individual contributions these variables made in explaining the variability in air fares. I considered the top 1,000 domestic city-pair markets of each of the years for the 11-
year period 2005–2015 in the continental United States in terms of total number of passengers. These city-pairs accounted for approximately 70% of all domestic air travel (U.S. Department of Transportation, 2015a). The airlines targeted for this part of the study were American, Alaska, JetBlue, Delta, Frontier, AirTran, Allegiant, Hawaiian), Spirit, Northwest, SkyWest), US Airways, Virgin, Southwest, and United.

**Definition of Terms**

Key terms and phrases relative to the current study were operationally defined as follows:

1. *Air fare competition dynamics* were defined as the factors affecting average price levels of airlines in various competitive markets. The dynamics of air fare competition included key factors partitioned into two functional sets: Set $R =$ Route factors (route length, population at endpoints of a city-pair market, and per capita income of the city-pairs that form the route), and Set $M =$ Market factors (total market concentration, market shares of specific competitors, market size, the average airfare level of the airline having largest average airfare, and the business model of the airline having the largest air fare in the route). The specific factors of each set are defined separately in this section.
2. *Air fare* was defined as the average price of airlines in the targeted city-pairs markets.

3. *Business models of incumbent airlines* were defined as any airline that carried more than 1,000 passengers and were partitioned into two groups: low-cost carriers (LCCs) and full-service carriers (FSCs).

4. *Hub airport* was defined as a dichotomous variable that represented whether or not an endpoint city was the hub of an entrant airline.

5. *Incumbent airline* was defined as any established airline that carried more than 1,000 passengers per month (both enplaning and deplaning) and flew within a specific route market. This definition was imposed to eliminate nonscheduled flights with less than 1,000 passengers per month.

6. *Largest fare* was defined as the most expensive average air fare among the competing airlines for each route.

7. *Market concentration* was defined as an airline’s “score” on the Herfindahl–Hirschman Index (HHI) of origin and destination passengers for each market. The HHI is a commonly accepted measure of market concentration and can range from 0 to 10,000. An HHI that approaches 0 indicates a heavily concentrated market and consists of many competing firms of equal size whereas an HHI of 10,000 indicates that the market is controlled by a single entity and therefore
has no competition. Moderately concentrated markets generally have HHIs between 1,500 and 2,500, and heavily concentrated markets generally have HHIs of more than 2,500. Thus, a higher HHI indicates a more concentrated market structure and a potentially less competitive environment (Ito & Lee, 2003). For the current study the total HHI and the HHI by business model were included separately for each of the city-pair markets.

8. *Market share of competitors* was defined as the ratio of available seats of each incumbent airline in a city-pair market to the total number of available seats in the market. This ratio was determined using the data acquired from the U.S. Department of Transportation’s DB1B database.

9. *Market size* was defined as the total number of available seats in a market. The market size of a city-pair market was determined using the data acquired from the U.S. Department of Transportation’s DB1B database.

10. *Number of competitors* was defined as the number of incumbent airlines in a route market.

11. *Per capita income of the city-pairs* was defined as the average per capita income of the cities at each end of a city-pair market. This was calculated by dividing the total income of each of the cities of a city-pair market by the total population of the corresponding city. For the
current study, max (income) and min (income) were the larger and smaller of the average per capita incomes (measured in thousands of dollars), respectively, at the endpoint cities of each market. The most recent U.S. Census Bureau data for the per capita income of cities were used for the current study.

12. *Population of each city of a city-pair market* was defined as all the inhabitants in the endpoint cities of an airline route market. For the current study, this was determined using the standard approach of calculating the geometric mean of the populations at two end points (Boguslaski et. al., 2004; Brueckner, Dyer, & Spiller, 1992). The most recent U.S. Census Bureau data for the city populations were used for the current study.

13. *Route entry and exit decision patterns* of an airline were defined as the resultant of parameters that airlines took into consideration while deciding to enter or exit a route. The targeted factors were placed into functional sets: Set $R =$ Route related factors (carrier type, route length, city population at each endpoint of a city-pair market, and per capita income of the city-pairs that formed the route), Set $C =$ Competitor related factors (total market concentration, market concentration of the incumbent airlines in the route relative to carrier type, and the number of competitors in the route), and Set $A =$ Airport related factors, which
included the existence of a hub airport of an airline in the origin or destination airports of the route. The specific factors of each set are defined separately in this section.

14. *Route Length* was defined as the distance in miles between the airports at the endpoints of the city-pair markets. This was determined by the great circle mapper method tool (http://www.greatcirclemapper.net/), which measures the shortest distance between two points on the surface of a sphere (Gade, 2010).

15. *Survival strategies* referred to various aspects of airlines’ ways of managing strategic decisions while competing in a competitive market. In general, airlines try to expand their route map while maximizing their profitability. After deregulation, many airlines had to terminate some of their routes or file for bankruptcy. Therefore, survival strategies of airlines became vital especially after deregulation. The aspects of survival strategies of airlines that were analyzed were route entry decision patterns and air fare competition dynamics, both of which are defined separately in this section.

**Research Questions and Hypotheses**

**Research questions.** The research questions that guided the current study were as follows:
Research question 1. Independent of set membership, what is the relationship between the targeted variables and the dichotomous response variable that distinguished between LCCs and FSCs that entered a new route and those that exited from an existing route?

Research question 2. When examined from a hierarchical perspective with set entry order R–M, what is the incremental knowledge gained at each step of the analysis relative to airlines’ air fare levels for the targeted city-pair markets?

Research hypotheses. The corresponding research hypotheses were as follows (the statistical hypotheses are provided in Chapter 4):

Research hypothesis 1. Independent of set membership, and when examined from a simultaneous perspective, at least one of the targeted variables will have predictive value relative to distinguishing between LCCs and FSCs that entered a new route or exited from an existing route.

Research hypothesis 2a. When examined from a hierarchical perspective with set entry order R–M, Set R will have a predictive gain in the relationship with airlines’ air fare levels for the targeted city pair markets, and at least one factor in Set R will have a direct relationship with air fares.

Research hypothesis 2b. When examined from a hierarchical perspective with set entry order R–M, Set M will have a predictive gain in the relationship with airlines’ air fare levels for the targeted city pair markets in the presence of Set R, and at least one factor in Set M will have a direct relationship with air fares.
Study Design

The current study employed two different research methodologies. For Part A, airlines’ route entry and exit decisions, the study was retrospective in nature because the effects on the dependent variable have already occurred. For example, airlines all ready have made a decision whether or not to enter or exit a route. As a result, the research design for Part A was ex post facto (cause-type) and focused on examining the extent to which the targeted factors were related to the preexisting effect of airlines entering or exiting a route. For Part B, air fare competition dynamics, the research methodology was correlational because I examined a single group (15 U.S. airlines in the top 1,000 domestic markets) and the relationship among multiple measures related to this group. The corresponding design included both explanatory and predictive perspectives. For both parts of the study, I used archival data that were accessible from public, U.S. government web sites.

Significance of the Study

The contributions of the current study to the published aviation literature were manifold and related to the study’s two major parts. First, the study investigated recent strategies of market makers of the U.S. domestic airline industry. Although the current literature was replete with similar studies, these past studies were not up-to-date and did not include company level analyses except for Southwest Airlines. The current study extended this knowledge by examining the LCCs and FSCs to clarify recent entry and exit strategies of these different business
models. Second, the current study focused on the real market shares of competitors in the market instead of the existence of various competitors. Most of the studies in the published literature focused on the “Southwest effect,” which simply refers to the entry decision of Southwest to a route dramatically decreases the price levels of incumbent airlines. However, it should be emphasized that a great part of the Southwest effect is related to the unique expansion strategy of Southwest Airlines. In contrast to other airlines, Southwest enters a route with very high frequencies. This also decreases the average price levels in the route market. In reality, the decrease in average price levels is related to the unique expansion strategy of Southwest Airlines, not to its existence in the route market. To observe a more realistic effect, the current study focused on market shares of competitors in the route market. Another contribution of the study was its focus on the probability of entrance of LCCs and FSCs into the route markets. This information currently was absent from the literature. Although there are various methods to define potential entrants in a route, the approach used in current study was seminal and will produce more realistic results. Finally, considering their expansion policy, airlines generally respond to questions such as, “What are the next routes to fly?” and “What are the best frequency increase options in existing routes?” To help them answer these questions, airlines utilize revenue management software to maximize revenue. The results of the current study will help strengthen this software in airlines’ quest to
decide what new markets to enter along with the expected price levels, breakeven prices, and profit margins for these markets.

**Study Limitations and Delimitations**

As is the case with other research studies, the current study was constrained by several limitations and delimitations. Limitations are conditions, events, or circumstances over which a researcher has no control that limit the generalizability of the study’s results whereas delimitations are conditions, events, or circumstances that a researcher imposes to make the study feasible to implement but further limit the generalizability of the results. The limitations and delimitations of the current study are explained below. The reader is advised to consider any conclusions or inferences derived from the study’s results with respect to these limitations and delimitations.

**Limitations.** The limitations of the current study are as follows:

1. **Data integrity.** The primary data sources for the current study was the U.S. Department of Transportation’s (2015b, 2015c) DB1B database and the T-100 domestic segment report from reporting carriers by Office of Airline Information of the Bureau of Transportation Statistics, and the U.S. Census Bureau. Because I had no control over how these data were reported, collected, and stored, overall data integrity is problematic. As a result, the findings of the current study are relative to the data that were available during the study period regardless of the data’s accuracy. This also means that if any changes are made to the database for data
correction purposes subsequent to the current study, then any future studies that replicate the current study might get different results.

2. **Historical events.** The time period of the data used for the current study might include various historical events such as terrorist attacks, social conflicts, airline accidents, or meteorological events such as hurricanes and snowstorms. All of these could have influenced the air transportation demand, which in turn could have impacted airlines’ survival strategies during the current study’s time period. As a result, because I did not account for the presence of any historical events, similar studies that access the same data as the current study but partial out or account for historical events might get different results.

3. **Social trends.** Independent but similar to historical events, trend changes in social and/or cultural values or practices could have an influence on the air transportation market. For example, technological changes such as consumers’ use of smart phones and social media might alter the way in which airlines conduct their business, and how an airline responds to these social trends could impact its route entry-exit decisions and/or air fare competition dynamics. As a result, because I did not account for the presence of any social trends, similar studies that access the same data as the current study but partial out or account for social trends might get different results.

4. **Market reputation of airlines.** Reputation is a very important factor for airlines. Many airlines use advertisements to increase reputation. On the other hand
sometimes the market reputations of airlines decrease because of the reasons such as ongoing delays and operational problems or an unusual event. For example, in 2008, Canadian musician Dave Carrols’ guitar was broken during a trip on United Airlines. He then composed a song named “United Breaks Guitars” and the song became a *YouTube* hit within a few days. The reputation of United Airlines suddenly was diminished and the demand for United Airlines decreased immediately. The stock price of United Airlines also fell 10% and the total cost for stakeholders of United Airlines was $180 million in a few days. A more recent example, which also involved United Airlines, was the involuntary removal of a passenger from United Express Flight 3411 on April 9, 2017, and the corresponding memorandum in which the CEO of United Airlines complemented the employees involved in the incident (Bacon & Mutzabaugh, 2017). The models to define survival strategies throughout the current study did not take into consideration the changes in the market reputation of airlines. As a result, similar studies that take an airline’s market reputation into consideration might get different results.

**Delimitations.** The delimitations of the current study are as follows:

1. **Study period.** The data collection period for the current study was limited to the 5-year period 2011–2015 for Part A, and the 11-year period 2005–2015 for Part B. Therefore, similar studies that use a different data collection period might not get the same results.
2. **Targeted sample.** The sample of the current study was limited to the top eight airlines in the U.S. domestic market in terms of market share for Part A, and the top 1,000 airline markets in the U.S. domestic airline market for Part B. As a result, similar studies that focus on a different sample might get different results.

3. **Data source.** The source of the current study’s data included the U.S. Department of Transportation’s (2015b, 2015c) DB1B database and the T-100 domestic segment report from the Bureau of Transportation Statistics, and the U.S. Census Bureau. As a result, studies that use different data sources might not get the same results.

4. **Use of outliers.** As described in Chapter 4 as part of the data screening discussion, outliers flagged by Jackknife distances were included in the data set that was used for the analysis in Part A, but excluded from the data set that was used for the analysis in Part B. As a result, similar studies that use a different outlier analysis strategy and include or exclude outliers differently than what was done in the current study might get different results.

5. **Creation of dichotomies.** As described in Chapter 4 (Table 4.6) the current study created dichotomies for all continuous factors to facilitate interpretations of the results, and several different strategies were used as the basis for creating these dichotomies. As a result, similar studies that do not use the same strategies for creating dichotomies, or studies that elect not to create dichotomies for continuous factors, might not get the same results.
Chapter 2

Review of Related Literature

Introduction

This chapter includes three main sections. The first section contains information about the theoretical background of the current study. There are two theories on which the study was grounded: demand theory and industrial organization theory. The second section presents a review of past research about both parts of the study: entry and exit patterns in airline markets, and factors affecting air fare price levels. The last section summarizes the implications of the literature to the current study and explains how the study benefitted from past studies.

Overview of Underlying Theory

Part A: Route entry and exit decision patterns. Part A of the current study examined factors related to route entry and exit decision patterns of airlines in the U.S. domestic market. The intent of this part of the study was to determine factors that distinguished between airlines that entered a new route (the Yes group) and airlines that exited from an existing route (the No group). This part of the study was grounded in industrial organization (IO) theory. Without offering a precise definition, suffice it to say that industrial organization is a field of economics that deals with an organization’s strategic behavior and market competition. It involves “the structure and behavior of firms (market strategy and internal organization)”
In short, IO theory is used as the basis for studying how markets function. For the current study, the targeted market was the airline industry.

As a field of economics, IO theory helps to realize the world of perfectly competitive models by adding real life facts such as limited information, barriers to entry of new firms, and transaction costs. Some of the critical topics of IO theory are competition structures in markets, competition and firm behaviors in different competition levels of markets, measures of competition, and competition level as size concentration of firms in a market. Some of the most important subareas of IO theory include market structure, firm strategy, firm objectives, market performance, antitrust issues, regulations, industry studies, primary products, transportation, and utilities.

With respect to the current study, IO theory posits that to understand and explain the basics of strategic decisions of U.S. domestic airlines, the factors influencing entry and exit patterns need to be analyzed. According to Porter (1980), one of the most important factors used to explain why some industries are more profitable than others is related to restrictions that exist in these industries. These restrictions decrease the number of competitors in an industry. Therefore, the average price levels for the industry increase with profit margins of incumbents. According to Bain (1956), as a consequence of barriers to entry, incumbent firms earn very high profits without threat of entry from competitors. Bain identified the specific market structures that keep the profit margin of incumbents high. These
include economies of scale, absolute cost advantages, product-differentiation advantages, and capital requirements. A brief description of each follows, and Figure 2.1 graphically illustrates how these market structures are related to the current study.

**Economies of scale.** Bain (1956) claimed that if the minimum efficient scale covers a noteworthy proportion of the total industry demand, only a small number of firms could exist in the market with high profit margins by discouraging other firms’ entry to the market. For example, in the context of the current study, let’s assume that Delta Airlines is considering which new routes to enter (or exit) next year. From the perspective of IO theory and relative to Figure 2.1, Delta would
need to consider the economies of scale relative to each prospective route. These would include route length, city population at each endpoint of the route, and the per capita income of the city-pairs that form the route.

**Absolute cost advantages.** In a market some of the incumbents may have lower costs than the potential competitors because of efficient cost control techniques or innovative cost reduction methods. Therefore, potential competitors in a market may forgo to enter this market because of low-cost incumbents. Continuing with the previous example and when applied to the current study, before deciding to enter a new route, Delta would have to determine the extent to which the existence of a hub airport of a competing airline in the origin or destination airports could impact its bottom line. For example, Delta might not start operations between Dallas Love Field (DAL) and Orlando (MCO) airports because DAL is a hub airport for Southwest Airlines, a competing LCC. Similarly, Delta might decide to exit an existing route because of a LCC’s decision to establish a hub at an airport at one end of a city-pair.

**Product-differentiation advantages.** Incumbents in a market may have innovative business models, pioneering product specifications, disruptive technologies, and a niche customer segment or customer loyalty. Each of these factors can prevent competitors from encroaching on incumbents’ territory. Continuing with the running example, Delta would need to consider both the
number of competitors in the prospective routes it is considering entering (or exiting), and the business models of the incumbent airlines in the route.

**Capital requirements.** Sometimes entrants in a market need financial support. However, banks or investors generally do not want to fund new entrants while incumbents exist in a market. Thus, trouble in finding financial funds becomes a barrier to entry for new entrants.

The current study endeavored to determine the extent to which these factors were statistically significant barriers to entry by examining them from the perspective of IO theory. For example, with respect to the example given in the “absolute cost advantage” barrier, if the “existence of a hub airport” were a indeed significant factor that distinguished between airlines in that entered a new route (the Yes group and airlines that exited an existing route (the No group), then this result would provide evidence in support of IO theory.

**Part B: Air fare competition dynamics.** The second part of the study focused on factors that influence the average air fare an airline assigns to a city-pair market. This part of the study hypothesized there would be a relationship between the overall average air fares in the U.S. airline domestic city-pair markets (the DV) and the targeted predictors of market size, business model of incumbent airlines, market concentration, specific competitors’ market shares, largest fare, business model of the largest fare airline, route length, per capita income of the city-pairs, and population of end points of a city-pair market. This hypothesized relationship
was deduced from demand theory. Although the hypothesized model is based on average air fare levels or average prices in airline markets, the strong relationship between demand and price was used to explain the determinants of air fare competition via demand theory (Clarkson, 1962).

Demand theory describes the relationship between consumer demand for goods and services and the prices for such goods and services. Figure 2.2 graphically illustrates the determinants of demand and the relationship to price. As a rule of thumb, the relationship between supply and demand is inverse, or negative: As supply increases, demand decreases and the equilibrium price also decreases. As more of a good or service is available, demand drops and therefore so does the equilibrium price. Demand theory is one of the core theories of microeconomics, and answers questions such as: “What is the relationship among income level, substitutes, and supply on the level of demand?” and “In what way do the factors affecting demand impact price levels?” (Riley, 2015).

Demand explains the amount of goods or services consumers are willing and able to purchase at a given price in a given time period. Based on Riley’s (2015) tutorial, there are three main types of demand. Effective demand is when consumers want to purchase a service or goods and this desire is supported by their ability to pay for it. For example, in a robust economy, consumers who want to fly to a destination do so because they are earning an income that enables them to pay for the flight. Latent demand is when consumers have a desire to purchase a service
or goods, but they cannot afford to make the purchase. For example, during an economic downturn, consumers might want to fly to a destination, but choose not to do so because they are unable to pay for the flight. Derived demand is when the demand for a product or service depends on another product or service. For instance, in the airline industry the demand for passenger airplanes like Boeing or Airbus is directly proportional to the demand for air transportation.
Associated with demand theory is the law of demand, which states there is an inverse relationship between the price of a product and the demand for that product (Riley, 2015). Thus, if the price of a product falls, then there will be an increase in demand for the product. Economists refer to this increase as “expansion of demand.” For example, if an airline lowers its air fares, then there will be an increase in the demand for flying on this airline. Conversely, if the price of a product increases, then there will be a decrease in the demand for the product. Economists refer to this decrease as “contraction of demand.” Continuing with the running example of air fares, if an airline increases its air fares, then there will be a decrease in the demand for flying on this airline.

This inverse relationship between price and demand often is depicted as a downward-sloping line and is referred to as the demand curve. A demand curve shows the relationship between price and demand over time. An example of a demand curve is shown in Figure 2.3. The demand for a product can be affected by many different factors. When economists prepare a demand curve for a particular product, they assume that these factors are held constant except for the actual price of the product. This then enables economists to isolate the effect of one variable on another. This is exactly what I did in the current study. For example, consider the average fare level in the JFK (New York)–MCO (Orlando) route in 2015. Based on demand theory, the price for this route is related to the demand of the route, and the factors affecting this demand are manifold. The factors I have targeted, which were
the populations of New York and Orlando, per capita income levels of New York and Orlando, market size, business model of incumbent airlines, market concentration, competitors’ market shares, route length, were based on the corresponding literature (discussed in the next section). Using a multiple regression strategy, I was able to isolate what factors contributed to air fares (and by association, the demand for a route) by holding all factors constant except the one under discussion.

Figure 2.3. An example of the demand curve relative to air fares. When an airline’s air fares decreases, a passenger moves away from a competitor airline and toward the airline with the lower air fare. The decline in air fares also increases a passenger’s willingness and ability to purchase a ticket at this airline, but this also decreases the passenger’s opportunity to purchase the ticket. Adapted from Riley (2015).
Review of Past Research Studies

As noted in Chapter 1, the published research is replete with studies related to route entry and exit decision patterns as well as pricing strategies of airlines. These past studies provided invaluable insight that helped inform the current study. This section provides a review of the most relevant studies and is organized based on the two parts of the study. Part A provides a review of the literature relative to route entry and exit decisions, and Part B contains a review of the literature relative to air fare competition dynamics.

**Part A: Route entry and exit decision patterns.** Of the various past studies related to route entry-exit decision patterns, six were relevant to and helped guide the current study. A separate section for each study is presented.

*Oliveira (2008).* Oliveira examined the entry pattern of a low-cost carrier (LCC), Gol Airlines, in the Brazilian domestic market in 2001. The purpose of his study was two-fold. His first purpose was to design an empirical model for LCC route entry by focusing on the competition between rapidly growing LCCs and full-service carriers (FSCs). His second purpose was to determine how consistent Gol Airline’s entry patterns were to Southwest Airlines and JetBlue Airways, which are two well-recognized LCCs in the U.S. domestic market. Data were collected from the densest 500 routes of the Brazilian domestic airline market and organized in a panel for the 2-year period between 2001 and 2002. To minimize potential heterogeneity across routes, Oliveira removed from the data what he believed were
tourism-related routes—those that had unusually short-haul and/or a high percentage of available weekend seats. The final data set contained 448 routes that corresponded to 27 million passengers. This data set represented 96.66% of all domestic trips during the targeted 2-year period.

Using Amemiya’s (1978) Generalized Least Squares (GLS), Oliveira (2008) found that the correlations between the reported marginal effects and (a) route density, (b) route distance, and (c) rivals’ presence in the route were .078, .464, and -.183, respectively, and all were significant at the 5% significance level. The dummy variable to control the presence of Gol Airlines in any endpoint of the route also was significant with a reported marginal effect of .049 at the 1% significance level. The pseudo $R^2$ was .803 for the entire model. Oliveira concluded that market size, which can be measured with the help of route density, route length, and rivals’ presence at the route, were the primary determinants of profitability. The comparison of Gol Airlines to Southwest Airlines also revealed there was a strong correlation in preference of dense and short-haul routes especially in 2001. However, the results indicated that Gol deviated from the similar paradigm of Southwest and became more consistent with the higher average stage length of the JetBlue paradigm in 2002 when compared to 2001.

Some limitations of Oliveira’s (2008) study included the existence of charter flights, especially in short-haul tourism routes, and the lack of control over governmental restrictions. Although Oliveira removed short-haul tourism routes
from the final data set, he did not take into account charter flights in some of the long-haul routes. Therefore, it is plausible that these charter flights could have influenced Gol Airlines’ decision to enter or exit a route. Oliveira also did not provide any information about governmental restrictions for airlines. Thus, there is a possibility that the entry-exit decision patterns of Gol Airlines could have been influenced by government regulations.

These limitations notwithstanding, Oliveira’s (2008) study included 96.66% of all domestic trips during the research period, which makes the findings compelling and which is why I have included route density, route length, and the number of competitors in a route market as factors for the proposed study. Oliveira’s study also helped inform the proposed study by his use of origin and destination (O&D) markets, which was in contrast with most of the previous studies, particularly Boguslaski et al.’s (2004) study (see below), which analyzed only nonstop markets and ignored routings with flight connections and stops. Another advantage of Oliveira’s study was his definition of the minimum level of operations (MLOs) paradigm, which helps to define entry into a route market. Previous studies used either absolute or relative definitions of MLOs whereas Oliveira used a broader definition of both route and entry, and this is what was used in the current study.

*Boguslaski et al. (2004)*. Boguslaski et al. examined the entry patterns in Southwest Airlines route system. There were three main purposes of the study. The
first was to analyze the patterns of nonstop entry by Southwest into city-pair markets from 1990 to 2000, and to detect the changes in entry strategies. The second purpose was to evaluate the effect of the Wright and Shelby Amendments, which limited the markets that Southwest could serve, and measure the fare savings as a consequence of these restrictions. The third purpose was to predict the potential candidate routes for future nonstop entry and assess the magnitude of threat on other airlines in domestic market because of Southwest expansion.

Traffic data for the study were gathered from the U.S. Department of Transportation’s OD1A Origin and Destination Survey, which consisted of a 10% sample of all tickets reported by U.S. Scheduled Passenger Carriers. All domestic round-trip passengers in the data set and one-way itineraries with three or fewer flight coupons per directional trip from the 48 contiguous U.S. states were included in the study. The largest 2,500 domestic city-pair markets that Southwest did not operate until 1990 were included in the base sample. The endpoints of the city-pair markets were Metropolitan Statistical Areas (MSAs), and there were 18,823 unique city-pair markets within these properties. However, because the majority of them did not have sufficient traffic for a non-stop market, only the city-pair markets with sufficient demand to be a non-stop market were acquired from the sample. This comprised of 2,500 markets corresponding to 90% of all domestic origin and destination passenger traffic. Moreover, because of the unattractiveness of the very long and very short routes for Southwest Airlines, the routes with lengths of more
than 3,000 miles and fewer than 100 miles also were removed from the sample. The base sample included 201 markets in total.

Boguslaski et al. (2004) classified the variables they believed would have an impact on entry patterns of Southwest Airlines into four categories: (a) market characteristics, which included market density and distance; (b) city characteristics, which included population, average per capita income, and the degree to which the city was a vacation destination; (c) Southwest’s pre-existing city/market presence; and (d) market competition and concentration in the market. To identify the general entry strategies of Southwest Airlines, Boguslaski et al. used probit analysis, and to observe the changes in entry strategies, they separated the complete data into two different time periods: from 1991 to 2000, and from 1995 to 2000.

Based on the probit analysis, Boguslaski et al. (2004) reported the following significant predictors at either the 5% or 1% level of significance for entering into a market route:

- Passenger density had a probit coefficient of .481 and a corresponding probability derivative was .01. Thus, for every 1% increase in passenger density, the probability that Southwest would enter a market increased by 1.0%.

- The mean population of the endpoints had a probit coefficient of .563 and corresponding probability derivative was .012. Thus, for every 1%
increase in mean population, the probability that Southwest would enter a market increased by 1.2%.

• The maximum income of the endpoints had a probit coefficient of -.266 and a corresponding probability derivative of -.006. Thus, for every 1% increase in maximum income, the probability that Southwest would enter a market decreased by 0.6%.

• The minimum income of the endpoints had a probit coefficient of -.117 and a corresponding probability derivative of -.003. Thus, for every 1% increase in minimum income, the probability that Southwest would enter a market decreased by 0.3%.

• The maximum number of cities served by Southwest from the endpoints of a market had a probit coefficient of -.044 and a corresponding probability derivative of -.001. Thus, every 1% increase in the maximum number of cities served by Southwest from the endpoints of a market, the probability that Southwest would enter a market decreased by 0.1%.

• Competitors’ carriers having a hub at the endpoints of a route had a probit coefficient -1.149 and a corresponding probability derivative of -.025. Thus, for every 1% increase in the number of rival carriers having a hub at any of the endpoints of a route, the probability that Southwest would enter a market decreased by 2.5%.
The pseudo $R^2$ for the overall model was .52, which means that 52% of the variance in Southwest’s probability to enter a route was explained by the variance in the predictors listed above. One predictor that was not significant was the minimum number of cities served by Southwest from the endpoints of a market (probit coefficient = -.021).

Boguslaski et al. (2004) concluded that passenger density, distance, the hubs of Southwest’s competitors, and per capita income at endpoints of a market were important factors to predict the routes that Southwest will serve. The findings showed that Southwest’s entry strategy changed dramatically in the 1990s. In the first half of 1990s, Southwest expanded with short-haul and dense routes. However in the second half of 1990s, Southwest expanded with mixed types of routes. Boguslaski et al. emphasized that Southwest’s future expansion threatens Alaska, Continental, and American Airlines most. The findings also revealed that the Wright and Shelby Amendments were binding on Southwest.

One of the limitations of Boguslaski et al.’s (2004) study was insufficient sample size for a probit study. The most effective approach to detect the entry pattern changes of Southwest Airlines was using probit study for each of the years from 1990 to 2000 separately. However, the number of entries of Southwest to a route in some of these years was not sufficient to use a probit analysis. Boguslaski et al.’s model also did not take into consideration some of the important historical events that influenced the airline industry such as the Gulf War. Although the
generalizability of the results were limited based on the reasons above, Boguslaski et al.’s approach and model were an improvement in studying route entry and exit strategies of U.S. airlines, and the predictors Boguslaski et al. found significant helped guide the current study.

**Lederman and Januszewski (2003).** Lederman and Januszewski examined the entry patterns of low-cost carriers by examining the properties of the routes that LCCs entered. Lederman and Januszewski analyzed entry decisions of 12 LCCs of U.S. domestic airline market on 3,997 routes in each quarter between 1996 and 2000, which corresponded to 830,989 observations in the final data set. Lederman and Januszewski used probit models to explain the determinants of LCCs to enter a route. The proposed determinants were classified into three broad categories: a carrier’s specific characteristics, exogenous market characteristics, and the presence of competitors at the endpoint airports of the targeted route.

Lederman and Januszewski (2003) provided strong evidence that a carrier’s own presence at the endpoints of a route increased the likelihood of entry ($p < .001$), but the largest share of flights of any competitors at the endpoints was not a significant factor ($p > .001$). These findings indicated that LCCs did not enter airports dominated by other carriers. Another significant factor ($p < .001$) was the presence of FSCs. LCCs entered routes in which FSCs had a presence at any endpoint of a route. Thus, LCCs did not avoid competing with other carriers.
especially with FSCs. Lederman and Januszewski also reported that the type of potential rivals on a route influenced the entry decision patterns of LCCs.

Lederman and Januszewski (2003) concluded that the types of routes that LCCs entered supported the hypothesis that LCCs follow a differentiation strategy. Lederman and Januszewski emphasized that LCCs did not shy away from competition, but to compete with the rivals LCCs should focus on product differentiation or serving with lower fares. The findings suggested that although the LCCs had a cost advantage, they not only provided a similar product but they did so with a lower fare.

A disadvantage of Lederman and Januszewski’s (2003) study is related to the limited number of LCCs in the sample. Only some of the LCCs were represented in the study. Lederman and Januszewski also did not give any detailed information about potential threats to external validity, including the macroeconomic situation of the economy and the potential events that might have had an effect on demand level. Therefore, the generalizability of the results is problematic. Lederman and Januszewski findings, though, were consistent with those of the previous studies presented earlier and therefore helped give greater credibility to the factors targeted in the current study.

Ito and Lee (2003). Ito and Lee analyzed the market characteristics that influenced LCC entry in the U.S. airline industry from 1990 to 2002 by examining three main issues about LCC entry. The first issue was detailing some formalized
realities about the growth of LCCs from 1990 to 2002. The second issue was assessing probit entry models to reveal the specialties of the markets that LCCs entered. The third issue was showing potential for LCC growth and its impact on FSCs. Ito and Lee targeted the five largest LCCs during the period: Southwest, Airtran, ATA, JetBlue, and Frontier. These airlines transported 85% of all passengers traveled with LCCs during the study period. The base sample included 2,500 domestic city-pair market corresponding to 90% of all domestic origin and destination passengers. After removing the markets that were served by only one of the five LCCs in 1991, the final data set consisted of 2,431 markets of which 351 involved a LCC entry until the 2nd quarter of 2002. Of these 351 markets, Southwest Airlines entered more than half (189) of them. The factors Ito and Lee examined included market density, distance, end point city populations and income, hub effects, and competition and concentration variables of LCC entry.

Using a probit analysis strategy involving cross-section market data, Ito and Lee (2003) reported the following significant predictors at the 1% level of significance for LCCs entering into a market route:

- Market density had a probit coefficient of 1.315.
- Pre-entry price markups had a probit coefficient of 3.718.
- Maximum and minimum populations had respective probit coefficients of -.71 and -.253.
• The existence of a hub FSC or a LCC had respective probit coefficients of - .06 and - .121.

The existence of a hub with multiple carriers in one endpoint of the route, route concentration, and dominance of largest carrier in a route were not significant factors affecting LCC entry. The pseudo $R^2$ for the overall model was .41.

Ito and Lee’s (2003) findings showed that pre-entry passenger density was the most important factor in determining LCC entry. Ito and Lee explained that FSCs were operating in most of these dense routes. Therefore, they concluded that in the future the threat from LCCs to FSCs would increase. The results also revealed that United Airways was under the highest pressure from LCCs. Ito and Lee concluded that LCC entry was not limited to only short and medium haul markets, but that LCCs also were beginning to enter longer haul routes.

One limitation of Ito and Lee’s (2003) study is related to the capabilities of probit models. Because of the low profit margins in the airline industry, especially in the U.S. domestic airline market, most airlines do not have an expansion strategy because the high level of competition has resulted in less expansion. On the other hand, probit models need a minimum number of entries to gather meaningful results, but the number of entries in the U.S. domestic airlines in recent years was not high. To compensate for this, Ito and Lee combined all entry cases. One problem in doing this is that because more than half of the entries were from Southwest, the results are biased toward Southwest. Another problem in combining
all entry cases is that it makes it difficult to get information about entry patterns of LCCs separately. Nevertheless, Ito and Lee’s study was informative because it alerted me to the need to analyze a broader range period to gather meaningful results with probit models.

**Hüschelrath and Müller (2011).** Hüschelrath and Müller examined the patterns and effects of entry of the top 500 nonstop routes in the U.S. airline market from 1996 to 2009. The study focused on two main issues within the U.S. airline market: patterns of entry and the effects of entry. Patterns of entry were analyzed from a descriptive approach whereas the effects of entry were analyzed from both a descriptive and an econometric approach.

The patterns of entry into non-stop U.S. airline markets from 1996 to 2009 demonstrated a considerably high amount of activity for both full-service and low-cost carriers. However, as depicted in Figure 2.4, the overall pattern in the number of route entries after 2003 showed that LCCs have been more active in increasing the number of route entries than network carriers. Hüschelrath and Müller (2011) also reported that the survival rates for LCC market entries were higher than the survival rates of FSCs. This is illustrated in Figure 2.5. The findings also revealed that since 2005, LCCs had a higher percentage of new market entries than network carriers (see Figure 2.6) and that these entries have been into long-haul markets.

With respect to the effects of entry, the results from a descriptive approach showed that only LCC entries resulted in significant decreases in market yields in
Figure 2.4. The number of route entries by NWCs (FSCs) and LCCs (all markets, 1996–2009) among U.S. domestic airlines. (Source: Hüschelrath & Müller, 2011, p. 6).

Figure 2.5. Survival rates of market entries by NWCs (FSCs) and LCCs (all markets, entry years 1996–2007). (Source: Hüschelrath & Müller, 2011, p. 8).
the top 500 U.S. airline markets. However, when Hüschelrath and Müller (2011) examined the data from an econometric approach, they found that in addition to the different effects of entry among monopoly, oligopoly without LCC, oligopoly with LCC market structures, and FSC and LCC carrier structures, there also was a significant within-group variation that resulted from these business models. Hüschelrath and Müller concluded that the growth of low-cost carriers after deregulation was still one of the most important facets of the U.S. airline industry. During the early period of deregulation, LCCs only existed in short-haul routes. Today, however, LCCs appears in long-haul routes as well. According to Hüschelrath and Müller, their findings suggest that the competition level between full-service and low-cost carriers will continue to increase in the future.
Hüschelrath and Müller’s (2011) study was one of the most comprehensive studies about the patterns of entry and effects of entry. The descriptive approach of their study revealed many important aspects of the topic and informed the current study by presenting a detailed map of all entries by LCCs and FSCs. This inspired me to focus on route entry decisions of airlines grouped relative to their business models (FSC vs. LCC) instead of examining each of the airlines separately. This strategy in turn enabled me to acquire a sufficient sample size needed for a logarithmic regression model.

**Part B: Air fare competition dynamics.** Five studies related to air fare competition dynamics were instrumental in helping guide the second part of the current study. A separate section for each study is presented.

*Zhang, Derudder, and Witlox (2013).* Zhang et al. examined the determinants of air fare levels of full-service carriers in U.S. hub-to-hub (HH) markets. They posited two hypotheses: (a) the presence of hub hierarchies could be relevant because there might be differential impacts based on the level of “hubness” of both points in a route, and (b) there might be duopolistic effects or intensive competition in routes connecting hubs of different carriers. Zhang et al. classified transfer passenger hubs into two categories—primary hubs and secondary hubs—and they acquired their data set from origin and destination markets connected by 17 hubs in the networks of U.S. FSCs. The data collection period was May 2009, and their study included 131 routes and 2.2 million passengers in total.
Zhang et al. (2013) reported that ordinary least squares regression results showed that 10 of the 12 targeted predictors were significant at the 10% level, but none was significant at the preset alpha level of .05. Zhang et al. also reported that the overall $R^2 = .694$, which indicates that the 12 IVs collectively explained about 70% of the variability in average price levels in the U.S. FSC HH markets. One significant IV ($p < .10$) was the total number of passengers in the route, $B = -0.042$, which indicated that a 1% increase in the total number of passengers decreased the price levels by .04% in the U.S. FSC HH markets. A second significant variable was distance, $B = 0.255$, $p < .10$, which indicated that a 1% increase in distance increased the price levels by 25.5% in the U.S. FSC HH markets. Two other significant factors were vacation routes, $B = -0.318$, $p < .10$, and slot controlled hub, $B = -0.162$, $p < .10$. Thus, hub hierarchies were a significant factor in the model. With respect to air fares, Zhang et al. reported that air fares on routes between a carrier’s primary hubs were on average 44% higher than general fares and 39% higher for primary-secondary routes. Zhang et al. concluded that (a) the monopolistic effects of airfares at hubs became less critical in a carrier’s network, and (b) duopolies in routes connecting the hubs of different carriers have no or very little impact on air fares relative to pricing variation in HH markets.

The biggest limitation of Zhang et al.’s (2013) study is they used only the data from U.S. FSCs in 17 hubs. They neither considered the effect of many of the other FSCs nor the effect of LCCs on the average price levels. The competition
with LCCs and the existence of alternative origin and destination routes with
different transfers could have affected the model. Therefore, the generalizability of
their results was limited. On the other hand, the design of the model, hub hierarchy
paradigm, and using OLS regression as a method helped inform the current study.
Thus, the current study will use OLS regression methods for second part of the
study (Part B: Air Fare Competition Dynamics), although the model for the current
study did not include a hub hierarchy paradigm as a research factor.

Vowles (2000). Vowles examined factors affecting average air fare levels of
low-cost carriers in the U.S. airline markets by focusing on geographic and
competition issues. The factors emphasized in Vowles’ model were (a) the presence
of LCCs in a market, (b) hub domination, (c) market share, and (d) type of
destination served. The data included 1,000 of the largest domestic origin and
destination markets in the U.S. excluding Alaska and Hawaii. These 1,000 markets
included 75% of the entire traffic during the study’s period along with 117
additional city pairs that were in the previous 1,000 largest domestic markets. The
study period was the Q1 1997, and Vowles used multiple regression as his
statistical strategy.

Vowles (2000) reported a significant overall model, $R^2 = .716$, $F = 402.60$,
and found that all variables were significant for $\alpha = .05$. Thus, all the targeted
factors collectively explained about 72% of the variability in average air fares. The
significant factors included route distance, $B = 0.068$, resort city in an endpoint of
the route, $B = -42.19$, the presence of Southwest Airlines in the route, $B = -77.61$, the existence of a hub in endpoints of a route, $B = 16.89$, the presence of a LCC in the route, $B = -45.46$, the market share of lowest fare carrier, $B = -0.429$, and the market share of the largest carrier in the route, $B = 0.427$. These regression coefficients should be interpreted similar to the way they were interpreted in the Zhang et al. (2013) discussion presented earlier. For instance, in markets where Southwest was present, the predicted price of a one-way ticket was forecasted to decline by $77.61$ when holding all other variables constant.

One disadvantage of Vowles’ (2000) study is related to the study period, which was Q1 1997. There may be many other factors that could have impacted average air fares such as seasonality that were not considered. The external validity of the study also was limited because of the small sample size. Nevertheless, Vowles’ study helped inform the current study by identifying specific factors to consider that could have an influence on air fares. Examples include route length (distance), market share of competitors, and existence of a hub airport.

**Abda, Belobaba, and Swelbar (2012).** Abda et al. examined the impact of LCC growth on domestic origin and destination air traffic and air fares at the top 200 airports in the U.S.—ranked by the number of passengers in 2005—between 1990 and 2008. The same airports from 1990, 1995, 2000, 2005, and 2008 were selected and identified as five milestones. Abda et al. reported that each airport
could be considered as a market and the airlines that had operations to or from the airports could be considered as firms competing in it.

Abda et al. (2012) analyzed the data set with a descriptive approach and concluded that LCCs’ market share was still increasing, but LCCs’ market stimulation effect decreased in recent years. Abda et al. indicated that the average number of LCCs per airport primarily increased from 0.5 in 1990 to 2.8 in 2005, but has fallen since. Abda et al. concluded that LCCs expanded rapidly and started existing in large airports and as a consequence of this trend LCCs started competing with each other in large airports in recent years. Abda et al. also reported that although many people thought LCCs would avoid major airports and direct competition with FSCs, as LCCs have evolved they have concentrated on the major airports and highest density origin and destination markets. This evolution, however, has not impacted the LCC business model. According to Abda et al., LCCs still have lower average air fares and higher passenger volumes, but trends suggest that rapid development of the LCC paradigm might end.

Although Abda et al.'s (2012) study was descriptive and lacked detailed statistical data analysis, the study uncovered an important point about LCCs, which helped inform the current study: Abda et al. indicated that the diminishing influence of LCC on market parameters of the airline industry was becoming a critical fact day-by-day. This comment made me realize that the business model of carriers could impact air fare dynamics. As a result, the current study included the
business models (FSC vs. LCC) of the carriers as factors to see whether or the same effect was detected.

**Vowles (2006).** Separate from his 2000 study, Vowles (2006) examined the air fare pricing determinants in hub-to-hub markets by focusing on factors that influence pricing in hub-to-hub markets in the U.S. on a carrier-by-carrier basis. The study included 20 hub airports of six airlines and involved 185 markets. No information was given about the study period. A distinguishing characteristic of the study was the classification of routes as independent variables. Vowles focused on four types of hypothetical routes: dominant, dominant-to-secondary, secondary-to-secondary, and primary-to-primary.

Vowles defined a dominant route as one in which the same carrier controlled both endpoints of the route. He then hypothesized that the dominant carrier would charge higher expected fares because of its dominant presence at both airports in the market. Vowles defined a dominant-to-secondary route as a route in which one endpoint was the dominant hub of a carrier, but the second endpoint was a secondary hub for a different carrier. He then hypothesized that the fare leader would be the carrier controlling the dominant hub with the carrier serving the secondary hub having similar but cheaper fares. Vowles defined a secondary-to-secondary route as one in which both endpoints were secondary hubs of an airline. He then hypothesized that this type of route would have a more justifiable distribution of customers between the carriers and as the distance between hubs
increased, the number of carriers gaining market share in a particular market also
would increase. Vowles defined a primary-to-primary route as one in which both
endpoints of the route were a primary hub for two different carriers. He then
hypothesized that this type of route would have an equal distribution of market
share between the hub airlines with none of the other airlines having any
measurable market share.

To identify the factors of air fare pricing in hub-to-hub markets, Vowles
(2006) used a multiple regression strategy, which resulted in seven separate
regression models. One was for the overall model, and the others were for each of
the largest hub carriers in the U.S. domestic airline industry. Vowles reported that
for the overall model, \( R^2 = .62 \), which indicates that the factors collectively
explained 62% of the variability in air fares. Two route types, dominant-to-
secondary and primary-to-primary, were not included in this model because they
were not statistically significant in the preliminary analysis. Two significant factors
were the total number of passengers in the route, \( B = -7.46 \times 10^{-5} \), and route
distance, \( B = 0.054 \). Thus, with respect to the former result, for every 1,000-unit
increase in total passengers for a route, air fares decreased on average by $7.46, and
with respect to the latter result, for every 100-mile increase in route distance, air
fares increased on average by $5.40. Other significant factors included: (a) the
relationship between air fares and all LCCs sans Southwest Airlines, \( B = -45.340; \)
(b) the occurrence of multiple airports within a region, $B = -14.53$; (c) the presence of Southwest Airlines, $B = -55.7$; and (d) the dominant route, $B = -11.63$.

Secondary-to-secondary routes, however, were not significant. Vowles concluded that the most influential air fare pricing factors in hub-to-hub markets were route type, LCCs, and competition. Vowles also indicated that hub-to-hub markets were unique markets, but the factors affecting air fare pricing were similar with other types of markets. The types of the routes defined by Vowles were not effective in explaining air fare pricing determinants as expected.

One limitation to Vowles’ (2006) study is related to the lack of sufficient parameters in the overall model. As described by Vowles, the number of flights in a particular market, the number of seats made available in the market versus the number of seats made available for passengers continuing outside the market, competition from one-stop service in the market, and the role of intervening opportunity should be included to gather more meaningful results. Although the study period was not provided, the number of routes examined was not sufficient. To have more generalizable results, Vowles should have collected data for more routes. The current study addressed this limitation.

**Brueckner, Lee, and Singer (2013).** Brueckner et al. extended past studies on air fare pricing determinants by using an adjacent airport approach and competitive effects of both LCCs and FSCs. One of the purposes of their study was to measure the effect of airport pair competition and adjacent competition for both
types of carriers. Brueckner et al. explained that adjacent competition in a market consists of service on an airport-pair with at least one alternative end point in the same metro area. For example, there is an adjacent competition between Delta’s flight from Orlando to JFK airports and JetBlue’s flight from Orlando to La Guardia. Thus, Brueckner et al.’s approach to air fare dynamics was extended by considering nonstop and connecting markets. An example of these extensions is presented in Figure 2.7. Brueckner et al.’s sample included last 2 quarters of 2007 and the first 2 quarters of 2008. Nonstop markets and connecting markets were included in the data set as separate markets.

Figure 2.7. An example of adjacent flight and connecting flights paradigm. (Source: Brueckner, Dyer, & Spiller, 1992, p. 5).
Brueckner et al. (2013) reported that findings for the base nonstop model was $R^2 = .32$, but the effects of nonstop legacy competition were not significantly different from zero. The presence of nonstop competition from the leading LCC, Southwest, however, was significant and reduced fares by 26% ($B = -.26, p < .001$). Similarly, nonstop competition from other LCCs also reduced air fares but at a smaller effect of 12% ($B = -0.116, p < .001$). Brueckner et al. concluded that adjacent legacy nonstop competition had no effect on air fares, but adjacent nonstop competition from Southwest reduced air fares by 11%, matching the findings of prior studies. For the market-level model, the findings showed (a) the overall $R^2 = .73$, (b) a second in-market nonstop legacy competitor increased air fare levels 5.3%, and (c) the impact of legacy connecting competition became negative, with an additional competitor reducing fares by 2.4%. The air fare impact of nonstop in-market competition from Southwest was 25%, and the effect of the other LCCs was 21%. Brueckner et al. explained that the impact of Southwest’s connecting competition was significant ($p < .001$) again for the market-level model, and the effect of adjacent potential competition from other LCCs was significant at a slightly higher alpha rate of $\alpha = .075$. The results for the connecting markets main model had an overall $R^2 = .56$, and revealed that a second legacy connecting competitor did not have any effect on airfare levels. However, the presence of a third competitor decreased fares by 2.7% and that an additional competitor beyond three reduced fares by 1.5%. Nevertheless, Brueckner et al. emphasized that the
presence of a Southwest connecting competition decreased air fares by 8% and connecting competition from other LCCs resulted in a 6% fare reduction.

Brueckner et al. added that the adjacent connecting competition from both FSCs and LCCs did not influence air fare levels. Based on these findings, Brueckner et al. concluded:

- FSC competition had a weak effect on average air fares for most of the cases, but LCC competition had a large effect on average air fare levels in each of the cases of airport-pair, at adjacent airports, or as potential competition.

- With respect to the base model, Brueckner et al. (2013) emphasized that potential competition from Southwest, both in-market and at adjacent airports, had critical impacts on air fare levels. The impact of other connecting LCC competition, however, although smaller than that of Southwest, was still considerable.

- With respect to the market level model, Brueckner et al. (2013) indicated that the market level regression analyses supported the results of the regression of the base model. They also reported that legacy competition applied only to uncertain descending pressure on air fares, but the effect of LCC nonstop competition, both in-market and adjacent, was important.

One of the distinguishing features of Brueckner et al.’s (2013) study was its focus on different types and levels of market. The number of factors for analyzing
air fare pricing determinants was sufficient and helped inform the current study by including some of the same factors from Brueckner et al. The strength of their study was related to its comprehensive approach: they included an adjacent airport approach, they focused on the competition impact of both FSCs and LCCs, and they included both nonstop and connected markets. Brueckner et al. reminded the reader that these two markets had not been analyzed simultaneously before within a single study.

**Summary and Concluding Remarks**

The literature reported in this chapter included studies that provided the most salient information about the targeted topics of the current study, namely, route entry-exit decisions and air fare competition dynamics. With respect to the former, the literature demonstrated there is still a gap about what factors influence airlines’ decisions to expand (or reduce) their network. The corresponding literature review helped identify what factors to consider, including the determinants of entry and exit patterns relative to an airline’s business model (LCC vs. FSC), which enabled me to measure the probability of route entry (or exit) of these airlines in the U.S. domestic market. As for the literature on air fare competition dynamics, the studies reviewed helped identify and understand the effects of various factors on average air fare levels in different settings. As noted earlier, many of these factors were incorporated into the current study.
Chapter 3
Methodology

Population and Sample

Population. The target population for Part A (route entry and exit decisions) was all U.S. domestic airline route markets for the 5-year period 2011–2015. The accessible population was all U.S. domestic airline route markets for the 5-year period 2011–2015 that were reported in the U.S. Department of Transportation’s (2015b) Airline Origin and Destination Survey (DB1B). This survey contains a 10% sample of airline ticket data, which are reported by the airlines and collected quarterly by the Office of Airline Information of the Bureau of Transportation Statistics. These data are stored in a publicly accessible database and are used to determine air traffic patterns, air carrier market shares, and passenger flows. This accessible population was further delimited to the route markets associated with the top eight airlines in the U.S. domestic airline market in terms of market share in 2016. Five of these airlines—Alaska, American, Delta, SkyWest, and United—are full-service carriers, and three airlines—JetBlue, Southwest, and Spirit—are low-cost carriers. Collectively, these eight airlines represented approximately 85% of the total U.S. domestic market during the targeted 5-year period. The data from the DB1B database were augmented by data from the Bureau of Transportation Statistics’ Air Carrier Statistics database—known as the T-100 data bank—which is a publicly accessible database that contains domestic market
and segment data. U.S. air carriers report these data monthly, and the aviation industry, media organizations, and U.S. government officials use this database to inform their respective audiences about the state of the domestic airline industry.

The target population for Part B (air fare dynamics) was all city-pair markets in the U.S. domestic airline industry for the 11-year period 2005–2015. The accessible population was the 1,000 largest city-pair markets in the 48 contiguous states reported in the U.S. Department of Transportation’s Airline Origin and Destination Survey (DB1B), which was described in the previous paragraph. Included among the data contained in the DB1B database is the average market fares, which are reported quarterly and represent the average prices paid by all fare paying passengers. These data represented approximately 75% of all 48-state passengers and 70% of total domestic passengers. The targeted airlines for Part B were AirTran, Alaska, Allegiant, American, Delta, Frontier, Hawaiian, JetBlue, Northwest, SkyWest, Spirit, Southwest, United, US Airways, and Virgin.

Sample. The sampling strategy used for Part A was purposive (or judgment), which is a non-probability sampling procedure that relies on a researcher’s prior experiences and personal judgment in selecting a sample. Unlike a convenience sample where participants are selected because they are conveniently available, participants of a purposive sample are selected because the researcher judges them to be representative of a specific group. In the current study, my primary focus for Part A was to select airlines that (a) were representative of both
full-service and low-cost carriers and (b) accounted for at least 85% of the total
U.S. domestic air travel in terms of market share and city-pair markets. As a result,
I selected all entry and exit data associated with the targeted eight airlines. In this
sense, the sample represented a census because it contained all the reported cases
for the eight airlines during the targeted timeframe. The total sample size for Part A
was \( N = 2,111 \) of which \( n = 1,590 \) cases reflected routes the airlines entered and \( n = 521 \) cases represented routes the airlines exited.

The sampling strategy for Part B was simple random. I first randomized the
data from the accessible population for each targeted year separately, and then
combined these randomized data sets into a single data set. Using this single
randomized data set, I then randomly selected \( N = 1,300 \) city-pair markets of which
\( N = 1,082 \) were related to the 15 targeted airlines. Thus, the total sample size for
Part B was \( N = 1,082 \).

**Power analysis.** Prior to implementing the current study, I performed an a
priori power analysis to determine the minimum number of cases I would need for
each part of the study. For Part A, which involved logistic regression, I considered
minimum sample sizes from three perspectives: (a) Agresti (2007) and Peduzzi,
Concato, Kemper, Holford, and Feinstein (1996) suggested the sample size be at
least 10 times the number of predictors, which yielded a minimum sample size of
150 cases; (b) The power tables from Hsieh (1989, p. 799), in conjunction with an
assumed \( R^2 = .13 \), \( ES = .15 \), event proportion of .25, odds ratios of 1.3, 1.5, and 1.7,
yielded minimum sample sizes 1606, 691, and 421, respectively; and (c) The computer program G*Power, relative to the parameters used with Hsieh’s table, yielded minimum sample sizes of 1600, 681, and 406, respectively. For Part B, which involved hierarchical multiple regression, the overall minimum sample size was \( N = 131 \) based on \( \alpha = .05 \), number of IVs = 13, \( \beta = .20 \), and \( ES = .15 \). As indicated earlier, the actual sample sizes for Part A and Part B were \( N = 2,111 \) and \( N = 1,082 \), respectively, and both sample sizes exceeded the minimum needed.

In addition to an a priori power analysis, I also performed a post hoc power analysis. Unlike the former, which is designed to help determine the minimum sample size needed relative to a set of assumed parameters, the latter is designed to determine the actual power of the study based on the results from data analysis. As summarized in Table 3.1, the power for all aspects of the study was greater than .96. Thus, the power of the current study was greater than the minimum a priori threshold of .80.

**Instrumentation**

The current study did not employ any formal data collection instruments. Instead, the data were acquired from three publicly accessible U.S. government databases: the U.S. Department of Transportation’s (2015a/b) DB1B database, the T-100 data bank, and the U.S. Census Bureau. The first two databases were described above. The U.S. Census Bureau database contains a variety of statistics related to each state’s voting-age population and industries.
**Table 3.1**  
*Post Hoc Power Analysis Based on Actual Sample Sizes and Effect Sizes*

### Part A: Route Exit and Entry Decisions (N = 2,111)  
(Logistic Regression)

<table>
<thead>
<tr>
<th>Model</th>
<th>Actual Value</th>
<th>Actual Effect Size(^b)</th>
<th>Number of Predictors</th>
<th>Estimated Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>1.08 ≤ OR ≤ 1.89</td>
<td>0.08 ≤ ΔOR ≤ 0.89</td>
<td>1</td>
<td>.96–1.00</td>
</tr>
</tbody>
</table>

### Part B: Air Fare Competition Dynamics (N = 916)  
(Hierarchical Multiple Regression)

<table>
<thead>
<tr>
<th>(Overall Model)(^d)</th>
<th>(R^2 = .95)</th>
<th>(ES = 19.00)</th>
<th>13</th>
<th>&gt; .99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set R = Route Factors(^e)</td>
<td>(R^2 = .49)</td>
<td>(ES = 0.96)</td>
<td>5</td>
<td>&gt; .99</td>
</tr>
<tr>
<td>Set M = Market Factors(^f)</td>
<td>(sR^2 = .46)</td>
<td>(ES = 0.85)</td>
<td>8</td>
<td>&gt; .99</td>
</tr>
</tbody>
</table>

Note.  
\(^a\)Odds ratios (OR) were determined for any single research factor while controlling for other factors in the model. In logistic regression, effect size is synonymous with the size of the treatment’s effect on the odds ratio. Thus, the change in odds above or below 1.00 produced by any single factor represents its effect size with respect to the dependent measure. For the current study, the actual odds ratios varied from 1.08 to 1.89 (see Chapter 4, Table 4.9), which means that the respective changes in these ORs from 1 vary between 0.08 and 0.89 and represent the corresponding effect sizes. For the hierarchical multiple regression model, the effect sizes were calculated by \(ES = R^2 / (1-R^2)\).  
\(^b\)The initial total sample size for Part B was \(N = 1,082\). However, after preliminary data screening, 166 cases were flagged as outliers and removed from the final data set used for the hierarchical multiple regression analysis (see Chapter 4 for additional information).  
\(^c\)The overall model consisted of 13 independent variables that were partitioned into two functional sets, R and M.  
\(^d\)Set R = Route Factors consisted of \(X_2 = \) Route length, \(X_{3a} = \) City population at Endpoint 1, \(X_{3b} = \) City population at Endpoint 2, \(X_{4a} = \) Per capita income of city at Endpoint 1, and \(X_{4b} = \) Per capita income of city at Endpoint 2.  
\(^e\)Set M = Market Factors consisted of \(X_9 = \) Market size, which was the total number of available seats in a market; \(X_{10a} = \) Market share of Alaska Airlines, \(X_{10b} = \) Market share of American Airlines, \(X_{10c} = \) Market share of Delta Airlines, \(X_{10d} = \) Market share of SkyWest Airlines, \(X_{10e} = \) Market share of United Airlines, where market share was the ratio of available seats to the total number of available seats in the market relative to the eight targeted airlines; \(X_{11} = \) Largest fare; and \(X_{13} = \) Business model of largest fare (full-service carrier or low-cost carrier).

Although the use of archival data provided rapid and unobtrusive access to large amounts of data, there are several disadvantages to using these data. Primary among them is related to validity and reliability issues. Because I did not know how the data were acquired, I was unable to assess or give attention to the corresponding validity and reliability related to any inferences, conclusions, and recommendations that result from data analysis. This disadvantage was mitigated, though, because of the presumed integrity of the data based on its usage. For example: (a) the U.S.
government uses the DB1B to determine air traffic patterns, air carrier market
shares, and passenger flows; (b) the aviation industry, the media, and the U.S.
legislature use the T-100 to “produce reports and analyses on air traffic patterns,
carrier market shares, as well as passenger, freight, and mail cargo flow within the
aviation mode” (U.S. Department of Transportation, 2015a, “Overview”); and (c)
the U.S. government relies on census data to make critical decisions, including state
apportionments. Data in the DB1B and T-100 databases are from 1990–2015, and
the U.S. Census Bureau database maintains current as well previous census data.

Procedures

Research methodology. The current study employed two different research
methodologies, one for each part. In Part A, the methodology was ex post facto and
the corresponding design was cause-type. This methodology-design was
appropriate because I was determining the extent to which the targeted factors led
to whether an airline entered a new route (the Yes group) or exited an existing route
(the No group). Because I was working with archival data, the effect on the
dependent variable, which was group membership and corresponded to the decision
to enter or exit a route, had already occurred. In Part B, the methodology was
correlational and the corresponding design was both explanatory and predictive.
This methodology-design was appropriate because the focus of Part B was to
determine the relationship the targeted variables had with air fares.
Human subject research. The current study involved the collection of existing data stored in publicly accessible databases and included airline, airline passenger, and market related data. With respect to passenger data, all identifying information was removed before the data became available publicly and therefore it was impossible to data to data provider. Although using publicly accessible archival data is not considered human subjects research as defined at 45 CFR 46.102, I still submitted an application to FIT’s IRB for review. A copy of the application, which was approved as an “exempt” status, is provided in Appendix A.

Study implementation. The implementation of the study essentially consisted of accessing the targeted databases and downloading the corresponding data. Because the data sets were very large, they automatically were downloaded in “csv” format and opened in Microsoft Excel. I imported the data into JMP and then modified the columns of the data table to reflect the targeted variables and prepared the data for analysis. (See also the description provided in Population and Sample.)

Threats to internal validity. Internal validity addresses the issue of whether changes in a dependent variable are directly related to the independent variables. In the context of the current study, internal validity depended on the extent to which entry-exit decision patterns and airfare pricing dynamics were related to the corresponding targeted independent variables and not to some other exogenous or confounding variables. Campbell and Stanley (1963) initially identified several possible threats to internal validity. Since then, others have
extended this list of threats. According to Ary, Jacobs, and Sorenson (2010), there are 12 threats to internal validity that should be considered: history, maturation, testing, instrumentation, statistical regression, selection bias, experimental mortality (attrition), selection-maturation interaction, experimenter effect, subject effects, diffusion, and location. Because these threats were considered in the context of research involving humans, many of the threats were not applicable to the current study because I used publicly accessible archival data and therefore by definition the study did not constitute human subjects research. Nevertheless, a discussion of each threat follows, including information about their relevance to the current study and how the effects of those that are relevant were controlled.

**History.** Some events or conditions during the period of a study could affect the dependent variable and therefore provide an alternative explanation for a study’s findings (Ary et al., 2010). These events or conditions may be very general and macro issues such as an economic crisis, financial bankruptcies, wars, and terrorist attacks. Nevertheless, these events may also be very specific to the sector of the study. For the current study, there is the possibility of a history threat because the overall study period (2005–2015) could have been impacted by events such as: (a) the September 11, 2001 attacks, which sharply decreased air transportation demand; (b) the invasion of Iraq (2003–2011), which had a very crucial impact on the U.S. economy; and (c) the “great recession” of 2008, which triggered a financial crisis. Moreover, specifically for the airline industry,
were several mergers, including AirTran and Southwest (2010) and American Airlines and US Airways (2013). Because I had no control over such historical events, the reader is informed that it is possible that the data and corresponding results could have been impacted by a history threat to internal validity.

**Maturation.** A maturation threat refers to whether the findings of a study are impacted by any biological or psychological changes participants might experience during the course of the study (Ary et al., 2010). Examples include but are not limited to age, wisdom, experience, and motivation, all of which can lead to different outcomes on a dependent measure. Because the current study used archival data, the maturation threat was not applicable.

**Testing.** A testing threat refers to instances when an assessment is administered both prior to and after an intervention. Because participants are exposed to the post-intervention assessment prior to treatment, it is possible that the participants became aware of or more sensitive to the test items, test format, and testing environment. Thus, participants’ performance on the post-assessment might be a function of their pre-assessment exposure and not the treatment. Because the current study used archival data and did not include the administration of any pre- or post-assessments, the testing threat was not applicable.

**Instrumentation.** According to Ary et al. (2010), an instrumentation threat could lead to inconsistencies in performance as a consequence of three main issues: instrumentation decay, data collector characteristics, and data collector bias.
Instrument decay refers to changes made to a data collection instrument during the course of a study, including how the instrument is scored. Data collector characteristics refer to changes in the characteristics of the person collecting the data, including age, gender, and ethnicity. Data collector bias refers to the unconscious distortion of data by the data collector or scorer. In the current study, I did not know of any changes to the instruments used to collect the data stored in the databases I consulted, or to what extent the data might have been impacted by the manner in which they data were collected. As a result, the reader is informed that it is possible—although unlikely given that these data are used by the media, government organizations, and airlines for public presentations, policy decisions, and business decisions—that the data and corresponding results could have been impacted by an instrumentation threat to internal validity.

**Statistical regression.** A statistical regression threat is related to the phenomenon of “regression toward the mean,” which refers to situations in which participants who score extremely high (or low) on an assessment to have a tendency to regress toward the mean in a subsequent assessment. Because the current study used archival data and did not involve any types of assessments, the statistical regression threat was not applicable.

**Selection bias.** A selection bias threat refers to situations when nonrandom factors such as sex, weight, hair, eye, skin color, personality, mental capabilities, and physical abilities, but also attitudes such as motivation or willingness to
participate influence the selection of participants. Therefore, as a consequence of selection problems, there could be differences between control and treatment groups (Ary et al., 2010). Because the current study used archival data consisting of routes or flights and did not involve human participants being placed into control or treatment groups, the selection bias threat was not applicable.

**Mortality.** A mortality threat, also known as attrition, refers to the loss of participants during the implementation of a study. This loss of participants could lead to biased outcomes because it has the potential to alter the characteristics of the sample. Thus, the sample might no longer be representative of the parent population, and there also could be a reduction in statistical power depending on the magnitude of the loss. Because the current study used archival data, there was no risk of participant attrition. In other words, no one dropped out of the study. As a result, the mortality threat from the perspective of “participant loss” was not applicable to the current study. However, as discussed in Chapter 4, \( N = 166 \) outlier cases (18%) were removed from the final analysis in Part B. To determine the extent to which these cases were different from the rest of the Part B data set (other than reflecting extremely low or high scores relative to the rest of the data), I compared the mean air fares between the outliers \( (N = 166) \) and the final cases \( (N = 916) \). The mean air fare of the outliers was $9.30 higher than the mean air fare of the final sample, but this difference was not significant, \( t(1080) = -1.91, p = .0558 \). Therefore, the mortality threat was not applicable to the current study.
**Selection-maturation interaction.** The selection-maturation interaction threat refers to the combination of both selection and maturation threats. The interaction between these two phenomena applies when treatment group participants have different maturation rates and therefore this interaction could be mistaken for a treatment effect. Because neither the selection threat nor the maturation threat was applicable to the current study, the resultant interaction of these two threats also was not applicable.

**Experimenter effect.** An experimenter effect refers to any unintentional behavior or bias a researcher might exhibit or have that could cause the researcher to unconsciously influence the participants of a study. Because the current study relied on archival data and did not involve an experimenter implementing any type of treatment, the experimenter effect was not applicable to the current study.

**Subject effects.** A subject effect threat refers to any changes that occur to participants’ attitudes during the course of a study. Such changes could include: (a) the Hawthorne effect in which participants are elated to be participating in a study and therefore want to do well regardless of treatment; (b) the John Henry effect in which participants want to intentionally perform poorly because they perceive they have been placed in an inferior group and therefore feel resentful; and (c) the novelty effect, which can occur when participants are presented with something new and exciting. Because the current study did not involve human participants, but instead used archived data, the subject effects threat was not applicable.
**Diffusion.** Diffusion refers to a situation when the participants of one group share information with participants in another group with respect to the type of treatment they are receiving. This correspondence could influence the second group’s performance. Because the current study did not involve human participants or groups of participants (Part B of the study was correlational), the diffusion threat was not applicable.

**Location.** A location threat refers to the environment in which an assessment is rendered. An example would be pilot candidates who are taking simulator sessions in different flight schools and then assessed on their performance. If the conditions at the different flight schools are not exactly (or nearly) the same, then it becomes difficult to determine if the changes in the dependent variable are directly and solely related to the independent variables or if they were a function of the different locations. Because I had no control over the location in which the data were collected and stored, I presumed that the location was comfortable and a stress-free environment that resulted in accurate entries. As a result, I do not believe the location threat had any impact on the current study.

**Treatment verification and fidelity.** The concept of treatment fidelity refers to what a researcher does to ensure that a study is implemented exactly as planned (Shaver, 1983). If a study is not implemented as intended, then subsequent replications might not be possible. To promote treatment fidelity, Shaver indicated that it is critical for researchers to verify the independent variables by describing
them fully and providing rationale for including them in a study. Such verification enables valid interpretations of effects and enables researchers to estimate the population and ecological generalizability of the results. Applying Shaver’s guidelines to the current study, I provided a detailed description of the IVs in both narrative and table forms (Tables 3.2 and 3.3).

**Data Analysis**

**Summary of IVs and DVs.** As summarized in Table 3.2, the current study involved 21 independent and two dependent variables. Following Cohen et al.’s (2003) “less is more” edict, the IVs were partitioned into functional sets as described in Table 3.3. Sets R, C, and A were relative to Part A and involved 10 IVs and one DV ($Y_1$). Sets R and M were relative to Part B and involved 17 IVs and one DV ($Y_2$). A brief description of each set follows:

**Set R = route factors.** Set R contained six factors: $X_1 = \text{Carrier type}$, a dichotomous variable that distinguished between low-cost and full-service carriers; $X_2 = \text{Route length}$, a continuous variable that represented the distance in miles between the airports at the endpoints of the city-pair markets; $X_{3a}$ and $X_{3b} = \text{City populations}$ of the endpoint cities that formed a city-pair market; and $X_{4a}$ and $X_{4b} = \text{Per capita income}$ in U.S. dollars of the cities that formed the endpoints of a city-pair market.
Table 3.2
Summary and Description of Independent and Dependent Variables Overall

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1 = \text{Carrier type}$</td>
<td>$X_1$ was a categorical (dichotomous) variable that represented the targeted airlines’ business model (LCC or FSC). It was dummy coded with FSC as the reference group.</td>
</tr>
<tr>
<td>$X_2 = \text{Route length}$</td>
<td>$X_2$ was a continuous variable that represented the distance in miles between the airports at the endpoints of the city-pair markets as determined by the Great Circle Mapper tool (<a href="http://www.greatcirclemapper.net/">http://www.greatcirclemapper.net/</a>)</td>
</tr>
<tr>
<td>$X_{3a}$ and $X_{3b} = \text{City populations}$</td>
<td>$X_{3a}$ and $X_{3b}$ were continuous variables that represented the city populations of the endpoints of a city-pair market. $X_{3a}$ = city population city at first endpoint, and $X_{3b}$ = city population at second endpoint.</td>
</tr>
<tr>
<td>$X_{4a}$ and $X_{4b} = \text{Per capita income}$</td>
<td>$X_{4a}$ and $X_{4b}$ were continuous variables that represented the per capita income in U.S. dollars ($) of the cities that formed the endpoints of a city-pair market. $X_{4a}$ = per capita income of the city at first endpoint, and $X_{4b}$ = per capita income of the city at second endpoint.</td>
</tr>
<tr>
<td>$X_{5a}$ and $X_{5b} = \text{Market concentration}$</td>
<td>$X_{5a}$ and $X_{5b}$ were continuous variables that represented an airline’s “score” on the Herfindahl–Hirschman Index (HHI) of origin and destination passengers for each market. $X_{5a}$ = total HHI, and $X_{5b}$ = HHI by carrier type (LCC or FSC).</td>
</tr>
<tr>
<td>$X_7 = \text{Number of competitors}$</td>
<td>$X_7$ was a continuous (discrete) variable that represented the number of incumbent airlines in a route market.</td>
</tr>
<tr>
<td>$X_8 = \text{Existence of a hub airport}$</td>
<td>$X_8$ was a categorical (dichotomous) variable that represented if the airport at an endpoint city was the hub of an entrant airline. It was dummy coded with No as the reference group.</td>
</tr>
<tr>
<td>$X_9 = \text{Market size}$</td>
<td>$X_9$ was a continuous variable that represented the total number of available passenger seats carried by airlines in a market.</td>
</tr>
<tr>
<td>$X_{10a}$–$X_{10h} = \text{Market share of eight competing airlines}$</td>
<td>$X_{10a}$–$X_{10h}$ were continuous variables that represented the ratio of available seats to the total number of available seats in the market relative to the targeted eight airlines, respectively.</td>
</tr>
<tr>
<td>$X_{11} = \text{Largest fare}$</td>
<td>$X_{11}$ was a continuous variable that represented the mean air fare of the airline with the largest mean air fare in a city-pair market.</td>
</tr>
<tr>
<td>$X_{13} = \text{Business model of largest fare}$</td>
<td>$X_{13}$ was a categorical (dichotomous) variable that represented the largest fare ($X_{11}$) airline’s business model (LCC or FSC). It was dummy coded with FSC as the reference group.</td>
</tr>
<tr>
<td>$Y_1 = \text{Route entry-exit decision}$</td>
<td>$Y_1$ was a dichotomous dependent variable for Part A of the study and represented whether or not an airline decided to enter a route (Yes group, coded 1) or to exit from a route (No group, coded 0).</td>
</tr>
<tr>
<td>$Y_2 = \text{Air fare}$</td>
<td>$Y_2$ was a continuous dependent variable for Part B of the study and represented airlines’ mean air fares in a city-pair market.</td>
</tr>
</tbody>
</table>
Table 3.3
Summary of the Sets of Independent Variables for Each Part of the Study

Part A: Route Exit and Entry Decisions
(Y1 = Route exit-entry decision: Yes vs. No)

Set R = Route Factors
X1 = Carrier type
X2 = Route length
X3a and X3b = City populations
X4a and X4b = Per capita income

Set C = Competitor Factors
X5a = Total market HHI
X5b = HHI by carrier type
X7 = Number of competitors

Set A = Airport Factors
X8 = Existence of a hub airport

Part B: Air Fare Competition Dynamics
(Y2 = Mean air fares in a city-pair market)

Set R = Route Factors
X2 = Route length
X3a and X3b = City populations
X4a and X4b = Per capita income

Set M = Market Factors
X5a = Total market HHI (market concentration)
X9 = Market size
X10a–X10h = Market share of eight competing airlines
X11 = Largest fare
X13 = Business model of largest fare

Note. See Table 3.2 for descriptions of the variables.

Set C = competitor factors. Set C contained three factors: X5a = Market concentration overall, which was measured by the Herfindahl–Hirschman Index (HHI; Chin, 2010); X5b = Market concentration (HHI scores) for each of the targeted airlines; and X7 = Number of competitors, which represented the number of incumbent airlines in a route market.
Set $A = \text{airport factors.}$ Set $A$ contained the single variable $X_8 = \text{Existence of a hub airport, which indicated if the airport at an endpoint city was (Yes) or was not (No) the hub of an entrant airline.}$

Set $M = \text{market factors.}$ Set $M$ contained 12 factors: $X_{5a} = \text{Market concentration overall as presented earlier in Set C;}$ $X_9 = \text{Market size, which represented the total number of available passenger seats carried by airlines in a market;}$ $X_{10a} - X_{10h} = \text{Market share of the respective eight competing airlines that were targeted (Alaska, American, Delta, JetBlue, SkyWest, Southwest, Spirit, and United), which was the ratio of available seats to the total number of available seats in the market relative to the targeted airlines;}$ $X_{11} = \text{Largest fare, which was the mean air fare of the airline having the largest average air fare in a city-pair market;}$ and $X_{13} = \text{Type of carrier (FSC or LCC) having the largest fare in a route.}$

Summary of statistical strategies used. As noted earlier in this chapter as well as in Chapter 1, the current study used two primary inferential statistical procedures. For Part A, because the DV was dichotomous, I performed a logistic regression analysis to determine the likelihood each IV makes in distinguishing between airlines that entered a new route (Yes group) and those that exited from an existing route (No group). To facilitate the interpretations of the results, I followed MacKinnon and Dwyer’s (1993) suggestion to transform the continuous IVs to corresponding dichotomies and expressed them as binary data. This transformation was with respect to each IV’s corresponding measure of central tendency. For Part
B, I performed a hierarchical multiple regression analysis with the set entry order of R–M to determine the amount of knowledge each set of IVs provided in explaining the variability in air fares. For both parts, I also performed preliminary data screening including: (a) outlier analysis, (b) multicollinearity checks, (c) missing data analysis, and (d) confirming that the assumptions for binary logistic regression and multiple regression were met. I also performed omnibus tests before conducting any pairwise comparisons to control for the possibility of inflated Type I and Type II error rates. The reader is directed to Chapter 4 for complete results of these analyses.
Chapter 4

Results

Introduction

This chapter summarizes the result of data analysis and is organized into three main sections. The first section contains a summary of the descriptive statistics results relative to the archival data on (a) route entry and exit decision patterns of low cost carriers and full service carriers, and (b) air fare competition dynamics for the targeted airlines in the U.S. domestic market. The second section consists of the results of inferential statistics and is separated into three parts: preliminary analyses and two primary analyses. The preliminary analyses involved various strategies commonly used for producing a “clean” data set. These included invalid/missing data analysis, outlier analysis, and checking for compliance with the assumptions of logistic and multiple regression. The two primary data analyses focused on answering the two research questions and involved logistic regression and multiple regression, respectively. The last section contains the results of hypothesis testing that corresponded to the research questions and hypotheses given in Chapter 1, and are based on the respective results of the two primary analyses.

Descriptive Statistics

Part A: Route entry and exit decision patterns. Part A examined route entry and exit patterns of the top eight airlines in the U.S. domestic airline market in terms of revenue passenger miles for the 5-year period 2011–2015.
Table 4.1
Summary of Key Attributes Associated with the Eight Targeted Airlines

<table>
<thead>
<tr>
<th>Airline</th>
<th>Type</th>
<th>Route Length (Miles)</th>
<th>Market Rank(^a)</th>
<th>Air Fares ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alaska</td>
<td>FSC</td>
<td>1182 663</td>
<td>6</td>
<td>185.47 47.26</td>
</tr>
<tr>
<td>American</td>
<td>FSC</td>
<td>1112 542</td>
<td>1</td>
<td>214.00 43.00</td>
</tr>
<tr>
<td>Delta</td>
<td>FSC</td>
<td>1081 672</td>
<td>3</td>
<td>226.26 64.28</td>
</tr>
<tr>
<td>JetBlue</td>
<td>LCC</td>
<td>1280 788</td>
<td>5</td>
<td>197.00 60.00</td>
</tr>
<tr>
<td>SkyWest(^b)</td>
<td>FSC</td>
<td>991 544</td>
<td>8</td>
<td>201.21 56.82</td>
</tr>
<tr>
<td>Southwest</td>
<td>LCC</td>
<td>906 520</td>
<td>2</td>
<td>176.97 46.05</td>
</tr>
<tr>
<td>Spirit</td>
<td>LCC</td>
<td>927 388</td>
<td>7</td>
<td>131.00 53.00</td>
</tr>
<tr>
<td>United</td>
<td>FSC</td>
<td>1372 749</td>
<td>4</td>
<td>240.66 55.32</td>
</tr>
<tr>
<td>Overall</td>
<td>—</td>
<td>975 682</td>
<td>—</td>
<td>198.53 57.63</td>
</tr>
</tbody>
</table>

Note. \(N = 2,111\). All entries rounded to nearest whole number and are for the 5-year period 2011–2015.

\(^a\)Based on revenue passenger miles. \(^b\)Although SkyWest is a regional carrier, its business model is similar to that of a FSC and therefore it is listed here as a FSC.

**Overall results.** As reported in Table 4.1, which contains an alphabetical listing of the targeted airlines, the overall mean route length between any two city-pairs was \(M = 975\) miles (\(SD = 682\) miles) and ranged from 11 miles (Delta Airlines) to 4,184 miles (United Airlines). Of the four full-service carriers, United Airlines had the longest mean route length (\(M = 1,372\) miles, \(SD = 739\) miles) and Delta Airlines had the shortest mean route length (\(M = 1,081\) miles, \(SD = 672\) miles). Of the four low-cost carriers, JetBlue Airlines had the longest mean route length (\(M = 1,280\) miles, \(SD = 788\) miles), and Southwest Airlines had the shortest mean route length (\(M = 906\) miles, \(SD = 520\) miles). With respect to market size based on revenue passenger miles for the targeted 5-year period, American Airlines was ranked number one overall and was the top full-service carrier, and Southwest
Airlines was the top low-cost carrier. When air fare data were disaggregated with respect to the targeted airlines, the overall mean air fare was $M = 198.53$ ($SD = 57.63$). Among the full-service carriers, United Airlines had the highest mean air fare ($M = 240.66$, $SD = 55.32$), and American Airlines had the lowest mean air fare ($M = 214.00$, $SD = 43.00$). Among the four low-cost carriers, JetBlue Airlines had the highest mean airfare ($M = 197.00$, $SD = 60.00$) and Spirit Airlines had the lowest mean air fare ($M = 131.00$, $SD = 53.00$).

Table 4.2 contains a summary of the continuous independent variables of Part A, which focused on route entry and exit strategies. Augmenting the data reported in Table 4.1 with respect to route lengths, the median overall route length was $Mdn = 850$ miles and the route lengths ranged overall between 11 miles and 4,184 miles, which indicates a mix of both short- and long-haul routes.

For the population of the cities at the first ($X_{3a}$) and second ($X_{3b}$) endpoints, the respective means, medians, standard deviations, and ranges were approximately the same. The city population at the first endpoint was $M = 4,720,161$, $Mdn = 2,814,330$, $SD = 4,734,835$, range = 92,663 to 20,182,305, the city population at the second endpoint was $M = 5,153,394$, $Mdn = 3,498,362$, $SD = 5,098,365$ range = 99,371 to 20,182,305.

For the per capita income of the city at the first endpoint ($X_{4a}$), the mean was $M = 48,494$ ($SD = 8,385$), the median was $Mdn = 47,254$, and the range was from $29,329$ to $107,887$. For the per capita income of the city at the second
Table 4.2
Descriptive Statistics of Factors Associated with Part A: Route Exit and Entry Strategies (Overall)

<table>
<thead>
<tr>
<th>Factor</th>
<th>M</th>
<th>Mdn</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_2$  = Route length (miles)</td>
<td>975</td>
<td>850</td>
<td>682</td>
<td>11–4184</td>
</tr>
<tr>
<td>$X_{3a}$  = City population at 1st endpoint</td>
<td>4720161</td>
<td>2814330</td>
<td>4734835</td>
<td>92663–20182305</td>
</tr>
<tr>
<td>$X_{3b}$  = City population at 2nd endpoint</td>
<td>5153394</td>
<td>3498362</td>
<td>5098365</td>
<td>99371–20182305</td>
</tr>
<tr>
<td>$X_{4a}$  = Per capita income of city at 1st endpoint ($)</td>
<td>48494</td>
<td>47254</td>
<td>8385</td>
<td>29329–107887</td>
</tr>
<tr>
<td>$X_{4b}$  = Per capita income of city at 2nd endpoint ($)</td>
<td>49339</td>
<td>47894</td>
<td>8659</td>
<td>29329–107887</td>
</tr>
<tr>
<td>$X_{5a}$  = Total HHI$^a$</td>
<td>7652</td>
<td>8610</td>
<td>2445</td>
<td>846–10000</td>
</tr>
<tr>
<td>$X_{5b}$  = HHI for LCCs$^b$</td>
<td>5082</td>
<td>5029</td>
<td>3944</td>
<td>0–10000</td>
</tr>
<tr>
<td>$X_{5c}$  = HHI for FSCs$^c$</td>
<td>2570</td>
<td>31</td>
<td>3914</td>
<td>0–10000</td>
</tr>
<tr>
<td>$X_6$  = Number of competitors</td>
<td>0.8</td>
<td>1</td>
<td>0.9</td>
<td>0–5</td>
</tr>
</tbody>
</table>

Note. $N = 2,111$. All entries rounded to nearest whole number and are for the 5-year period 2011–2015. $^a$HHI = Herfindahl–Hirschman Index of origin and destination passengers for each market. The HHI index is a measure of the competitive pressures from rivals. A HHI that approaches zero indicates a more concentrated market structure and a potentially less competitive environment (Ito & Lee, 2003). $^b$LCC = Low-cost carrier. $^c$FSC = Full-service carrier.

endpoint ($X_{4b}$), the mean was $M = $49,339 ($SD = $8,659), the median was $Mdn = $47,894, and the range was from $29,329 to $107,887. Given the diverse sample of very short- to very long-haul routes throughout the United States, these levels of variability for population and per capita income values were expected.

With respect to the overall HHI ($X_{5a}$), the reader is reminded that the Herfindahl–Hirschman Index is a commonly accepted measure of market concentration and can range from 0 to 10,000. An HHI that approaches zero indicates a heavily concentrated market and consists of many competing firms of equal size whereas an HHI of 10,000 indicates that the market is controlled by a single entity and therefore has no competition. The U.S. Department of Justice
(2015) considers moderately concentrated markets to have HHIs between 1,500 and 2,500, and heavily concentrated markets to have HHIs of more than 2,500. In the context of the current study, the HHI was used as a measure of competitiveness. As reported in Table 4.2, the overall mean HHI was $M = 7,652$ ($SD = 2,445$), which indicates that the overall market during the targeted 5-year period was being controlled by one or two airlines and there was little competition among the airlines. The median HHI was $Mdn = 8,610$ and ranged from 846 to 10,000. Among the low-cost carriers ($X_{5b}$), the mean HHI was $M = 5082$ ($SD = 3,944$), the median was $Mdn = 5029$, and the range was 0–10,000. Among the full-service carriers ($X_{5c}$), the mean was $M = 2,570$ ($SD = 3,914$), the median was $Mdn = 31$, and the range was 0–10,000. These latter statistics indicate that the HHI data for the full-service carriers were considerably skewed, and that when compared to the low-cost carriers, the full-service carriers had a less concentrated market structure and concomitantly a less competitive environment.

Lastly, as for the number of competitors, the mean was $M = 0.8$ ($SD = 0.9$), the median was $Mdn = 1$, and the range was 0–5. Thus, on average, there was less than one competitor in the routes the targeted eight airlines entered or exited.

**Route entry strategies by carrier type.** As reported in Table 4.3, when the route entry data were disaggregated by carrier type, there was little difference in mean route lengths between low-cost carriers ($M = 992$, miles, $SD = 528$ miles) and full-service carriers ($M = 1,045$ miles, $SD = 745$ miles). This small difference
Table 4.3
Descriptive Statistics Relative to Route Entry Strategies by Carrier Type

<table>
<thead>
<tr>
<th>Factor</th>
<th>LCC Enter (N = 638)</th>
<th>FSC Enter (N = 952)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>$X_2$ = Route length (miles)</td>
<td>992</td>
<td>528</td>
</tr>
<tr>
<td>$X_{3a}$ = City population at 1st endpoint</td>
<td>4487420</td>
<td>3764381</td>
</tr>
<tr>
<td>$X_{3b}$ = City population at 2nd endpoint</td>
<td>4606634</td>
<td>3863005</td>
</tr>
<tr>
<td>$X_{4a}$ = Per capita income of city at 1st endpoint ($)</td>
<td>48427</td>
<td>7954</td>
</tr>
<tr>
<td>$X_{4b}$ = Per capita income of city at 2nd endpoint ($)</td>
<td>48896</td>
<td>8428</td>
</tr>
<tr>
<td>$X_5$ = Total HHI*</td>
<td>8060</td>
<td>2175</td>
</tr>
<tr>
<td>$X_6$ = Number of competitors</td>
<td>0.8</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Note. N = 1,590. All entries rounded to nearest whole number and are for the 5-year period 2011–2015.

supports recent arguments that the route entry strategies of LCCs and FSCs have been converging. With respect to the city populations at the first ($X_{3a}$) and second ($X_{3b}$) endpoints: (a) the mean population sizes at the first endpoint were $M = 4,487,420 (SD = 3,764,381)$ for LCCs and $M = 5,068,143 (SD = 5,175,178)$ for FSCs, and (b) the mean population sizes at the second endpoint were $M = 4,606,634 (SD = 3,863,005)$ for LCCs and $M = 5,507,865 (SD = 5,497,650)$ for FSCs. Based on these data, it appears that full-service carriers entered more crowded origins and destination cities than low-cost carriers during the target 5-year period.

For the per capita income of the cities at the first ($X_{4a}$) and second ($X_{4b}$) endpoints: (a) the mean incomes at the first endpoint were $M = $48,427 (SD = $7,954) for LCCs and $M = $49,523 (SD = $8,286) for FSCs; and (b) the mean incomes
at the second endpoint were $M = 48,896 (SD = 8,428)$ for LCCs and $50,336 (SD = 8,476)$ for FSCs. These results suggest that the per capita income levels of the origins and destinations of the routes FSCs entered were larger than that of LCCs. This could be due to the presumably higher service quality FSCs provide, which would then yield more demand from higher income level cities. Finally, there was not much difference in the number of competitors ($X_6$) between low-cost and full-service carriers relative to the routes they entered. The respective means for LCCs and FSCs were $M = 0.8 (SD = 0.9)$ and $M = 0.9 (SD = 1.0)$.

**Route exit strategies by carrier type.** As reported in Table 4.4, when the route exit data were disaggregated by carrier type, the mean route length for LCCs was $M = 671$ miles ($SD = 613$ miles) whereas the mean route length for FSCs was $M = 887$ miles ($SD = 732$ miles). Thus, full-service carriers exited longer routes on average than low-cost carriers during the 5-year targeted period 2011–2015.

With respect to the city populations at the first ($X_{3a}$) and second ($X_{3b}$) endpoints: (a) the mean population sizes at the first endpoint were $M = 2,988,194 (SD = 3,256,701)$ for LCCs and $M = 4,970,462 (SD = 5,384,369)$ for FSCs, and (b) the mean population sizes at the second endpoint were $M = 3,915,478 (SD = 4,554,318)$ for LCCs and $M = 5,723,545 (SD = 5,916,074)$ for FSCs. Based on these data, it appears that full-service carriers exited more crowded origins and destinations than LCCs.
Table 4.4
Descriptive Statistics Relative to Route Exit Strategies by Carrier Type

<table>
<thead>
<tr>
<th>Factor</th>
<th>LCC Exit (N = 158)</th>
<th>FSC Exit (N = 363)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>$X_2 =$ Route length (miles)</td>
<td>671</td>
<td>613</td>
</tr>
<tr>
<td>$X_{3a} =$ City population at 1st endpoint</td>
<td>2988194</td>
<td>3256701</td>
</tr>
<tr>
<td>$X_{3b} =$ City population at 2nd endpoint</td>
<td>3915478</td>
<td>4554318</td>
</tr>
<tr>
<td>$X_{4a} =$ Per capita income of city at 1st endpoint ($)</td>
<td>45059</td>
<td>9199</td>
</tr>
<tr>
<td>$X_{4b} =$ Per capita income of city at 2nd endpoint ($)</td>
<td>46256</td>
<td>9865</td>
</tr>
<tr>
<td>$X_{5a} =$ Total HHIa</td>
<td>8258</td>
<td>2187</td>
</tr>
<tr>
<td>$X_6 =$ Number of competitors</td>
<td>0.5</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Note. N = 521. All entries rounded to nearest whole number and are for the 5-year period 2011–2015.

For the per capita income of the cities at the first ($X_{4a}$) and second ($X_{4b}$) endpoints: (a) the mean incomes at the first endpoint were $M = 45,059$ ($SD = 9,199$) for LCCs and $47,407$ ($SD = 8,503$) for FSCs, and (b) the mean incomes at the second endpoint were $M = 46,256$ ($SD = 9,865$) for LCCs and $48,841$ ($SD = 8,585$) for FSCs. Thus, unlike the corresponding route entry per capita income (Table 4.3), there was little difference between LCCs and FSCs in the city-pair income levels of the routes the airlines exited during the targeted 5-year period. Finally, there was not much difference in the number of competitors ($X_6$) between low-cost and full-service carriers relative to the routes they exited. The respective means for LCCs and FSCs were $M = 0.5$ ($SD = 0.7$) and $M = 0.6$ ($SD = 0.8$).

Part B: Air fare competition dynamics. The second part of the study examined air fare competition dynamics. As noted in Chapter 3, the data for Part B
Table 4.5

Summary and Description of Factors Associated with Part B: Air Fare Competition Dynamics

<table>
<thead>
<tr>
<th>Factor</th>
<th>$M$</th>
<th>$Mdn$</th>
<th>$SD$</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_2$  = Route length (miles)</td>
<td>1053</td>
<td>871</td>
<td>621</td>
<td>129–2724</td>
</tr>
<tr>
<td>$X_3a$ = City population at 1st endpoint</td>
<td>4128206</td>
<td>2509417</td>
<td>3790702</td>
<td>92663–20182305</td>
</tr>
<tr>
<td>$X_3b$ = City population at 2nd endpoint</td>
<td>3819012</td>
<td>2403957</td>
<td>4245193</td>
<td>109210–20182305</td>
</tr>
<tr>
<td>$X_4a$ = Per capita income of city at 1st endpoint ($)</td>
<td>43936</td>
<td>43245</td>
<td>7052</td>
<td>29329–107887</td>
</tr>
<tr>
<td>$X_4b$ = Per capita income of city at 2nd endpoint ($)</td>
<td>43971</td>
<td>43229</td>
<td>8308</td>
<td>23503–118695</td>
</tr>
<tr>
<td>$X_5a$ = Total HHI&lt;sup&gt;a&lt;/sup&gt;</td>
<td>6619</td>
<td>7785</td>
<td>3571</td>
<td>0–10000</td>
</tr>
<tr>
<td>$X_5b$ = HHI for LCCs&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6542</td>
<td>9248</td>
<td>3883</td>
<td>0–10000</td>
</tr>
<tr>
<td>$X_5c$ = HHI for FSCs&lt;sup&gt;c&lt;/sup&gt;</td>
<td>6692</td>
<td>6943</td>
<td>3253</td>
<td>0–10000</td>
</tr>
<tr>
<td>$X_8$  = Market size</td>
<td>1062</td>
<td>550</td>
<td>1460</td>
<td>149–18858</td>
</tr>
<tr>
<td>$X_{11}$ = Largest air fare</td>
<td>202</td>
<td>193</td>
<td>68.2</td>
<td>71–419</td>
</tr>
<tr>
<td>$Y_2$  = Air fare</td>
<td>198</td>
<td>192</td>
<td>57.6</td>
<td>81–383</td>
</tr>
</tbody>
</table>

Note. $N = 1,082$. Results are in the presence of 154 outliers (see the Outliers discussion in Chapter 4).

Consisted of $N = 1,082$ randomly selected U.S. city-pair markets that corresponded to 15 airlines: AirTran, Alaska, Allegiant, American, Delta, Frontier, Hawaiian, JetBlue, Northwest, SkyWest, Southwest, Spirit, United, US Airways, and Virgin. As reported in Table 4.5, the average route length was $M = 1,053$ miles ($SD = 621$ miles), the median was $Mdn = 871$ miles, and the routes ranged in length from 129 miles to 2,724 miles. For the city population at the first endpoint, the average was $M = 4,128,206$ ($SD = 3,790,702$), the median was $Mdn = 2,509,417$, and the population ranged from 92,663 people to 20,182,305 people. For the city population at the second endpoint, the average was $M = 3,819,012$ ($SD = 4,245,193$), the median was $Mdn = 2,403,957$, and the population ranged from 109,210 people to 20,182,305 people.
For the per capita income of the city at the first endpoint, the mean was $M = \$43,936$ ($SD = \$7,052$), the median was $Mdn = \$43,245$, and income ranged from $\$29,329$ to $\$107,887$. For the per capita income of the city at the second endpoint, the mean was $M = \$43,971$ ($SD = \$8,308$), the median was $Mdn = \$43,229$, and income ranged from $\$23,503$ to $\$118,695$.

With respect to market concentration/competitiveness, the overall mean HHI for the 15 targeted airlines was $M = 6,619$ ($SD = 3,571$), the median HHI was 7,785, and HHIs ranged from 0 to 10,000. When HHI data were disaggregated by carrier type (low-cost vs. full-service), the corresponding means, standard deviations, and ranges were nearly identical to those of the overall HHI. The only difference was the median HHI, which was not affected by outliers (LCC $Mdn = 9,248$, FSC $Mdn = 6,943$). Nevertheless, all HHI values indicate that the 15 airlines were in a fairly concentrated market with little competition.

As for the market size, which was defined as the total number of available seats in a market, the mean was $M = 1,062$ ($SD = 1,460$), the median was $Mdn = 550$, and the overall number of available seats ranged from 149 to 18,858. Also, the mean air fare of largest fare carrier of each route was $M = \$202$ ($SD = \$68.20$), the median was $Mdn = \$193$, and the fares ranged from $\$71$ to $\$419$. Lastly, the overall mean air fare, which was the dependent variable for Part B, was $M = \$198$ ($SD = \$57.60$), the median was $Mdn = \$192$, and the fares ranged from $\$81$ to $\$383$. 
Inferential Statistics

As a reminder to the reader, the purpose of the current study was twofold. In Part A, the purpose was to determine the relationship between the targeted sets variables and the dichotomous response variable that distinguished between airline route entry-exit decisions, and the primary data analytic strategy was logistic regression. The sets of variables included: Set R = Route Factors, which consisted of $X_1 = \text{Carrier type (low-cost vs. full-service)}$, $X_2 = \text{Route length}$, $X_3 = \text{City populations at each end-point}$, and $X_4 = \text{Per capita income at each end-point}$; Set C = Competitor Factors, which consisted of $X_5 = \text{Market concentration (HHI points)}$ and $X_7 = \text{Number of competitors}$; and Set A = Airport Factors, which consisted of $X_8 = \text{Existence of a hub airport (yes or no)}$. The dichotomous dependent variable ($Y_1$) was whether an airline entered or exited a particular route (yes or no).

In Part B, the purpose was to examine the relationship among the targeted sets of variables and airlines’ mean air fare in a city-pair market, and the primary data analytic strategy was hierarchical multiple regression. The sets of variables included: Set R = Route Factors, which consisted of $X_2$, $X_3$, and $X_4$ from Part A, and Set M = Market Factors, which consisted of $X_5$ from Part A, and $X_9 = \text{Market size}$, $X_{11} = \text{Largest fare}$, and $X_{13} = \text{Business model of largest fare (LCC or FSC)}$. The dependent variable ($Y_2$) was airlines’ mean air fares in a city-pair market. A summary and descriptions of these variables were provided in Chapter 3 (see Table 3.2 and Table 3.3).
**Preliminary analyses.** Prior to conducting the primary analyses associated with each part of the study, the overall data set was examined for invalid data, missing data, and outliers. In addition, because both logistic and multiple regression strategies were used, the final data set was examined for compliance with the respective assumptions of these two strategies. A discussion of these preliminary analyses follows.

*Invalid data for logistic regression.* The data were examined to ensure validity for conducting logistic regression analysis. According to Tabachnick (2013), each independent variable should include a minimum of one cell frequency and no more than 20% of cell frequencies should be less than five. This guideline was confirmed by contingency analyses and is discussed later in the section on data set modifications. Thus, there were no circumstances in which a cell frequency fell below five and Tabachnick’s recommendation that no more than 20% of cell frequencies should be fewer than five was confirmed.

*Missing data analysis.* Missing or incomplete data generally occur for various reasons. For example, participants might opt not to respond to an item, forget to respond to an item, or are unclear on how to respond to an item. According to Cohen, Cohen, West, and Atkins (2003), it is incumbent for the researcher to determine if data are missing randomly or systematically, and to then treat missing data as information. In the current study, there was little chance for missing data because the primary data source was archived data and not individuals
responding to items on a questionnaire. Nevertheless, I examined the data set for any missing cells and found none. Thus, the data set was complete and had no missing data.

**Outlier analysis.** Outliers are extreme data points relative to the other observations within a data set. The presence of outliers can confound the results of a study because outliers generally inflate significance or mask significance. According to Cohen et al. (2003), outliers are either contaminants or rare cases. With respect to the former, contaminants could be due to inaccurate measurements, data entry/keying errors, errors from data analysis, and inattentive participants (Cohen et al., 2003). For example, an analyst who records total revenue on a U.S. domestic flight as $27.20 most likely meant $27,200 ($27.2K). As for the latter, a rare case is a correct but highly unusual or atypical measurement. For example, a rare case would be a pilot who has 45,000 flight hours, or in the context of the current study, an end city with a per capita income of $20,000 (extremely low) or $80,000 (extremely high). To examine the data set for outliers, I conducted two independent outlier analyses via Jackknife distances: one for each part.

For Part A, 151 of the 2,111 cases (7%) were flagged as outliers. Of these cases, 98 were associated with full-service carriers, 41 were associated with low-cost carriers, and all the outliers appeared to be rare cases (not contaminants). For example: (a) 94 cases (62%) had relatively small route lengths that were fewer than 100 miles, (b) 110 cases (73%) had relatively small city population sizes of fewer
than 10 million people, and (c) 103 cases (68%) had relatively high per capita income of more than $50,000. To determine the impact of these outliers, I ran two logistic regression analyses, one each in the presence and absence of the outliers. The results were nearly identical. Therefore, I included the outliers in the final data set so the data would reflect real-world situations. This decision is reflected in Chapters 1 and 5 as a delimitation to the studies results.

For Part B, 166 of the 1,082 (18%) cases were flagged as outliers. Similar to the outliers from Part A, the cases flagged contained either extremely low or high values relative to route lengths, city populations, and market sizes. To determine the impact of these outliers, I ran two simultaneous regression analyses, one each in the presence and absence of outliers. Although the results were approximately the same, the model without outliers yielded a larger $F$ value (1059.28 vs. 871.92), a 2% increase in $R^2$ (.95 vs. .93), and a smaller root mean square error (12.35 vs. 15.03). Therefore, I removed the outliers from the final data set. This decision is reflected in Chapters 1 and 5 as a delimitation to the studies results.

**Logistic regression assumptions (Part A).** According to Tabachnick (2013), the assumptions of logistic regression include a dichotomous dependent variable, mutually exclusive categories on the dependent variable, independent observations on the dependent measure, and correct specification of the hypothesized model. In addition to these assumptions, Cohen et al. (2003) also recommended that the data set be screened for multicollinearity among the
independent variables, outliers in the solution, and linearity of the logit. To test these assumptions, the data corresponding to Part A of the study were used (see Chapter 3, Tables 3.2 and 3.3). Following is a brief discussion of each assumption and the corresponding results relative to these data.

*Dichotomous DV.* To address the assumption for a dichotomous dependent variable, \( Y_1 \) = Route entry-exit decision was expressed as a dichotomy to represent whether or not an airline decided to enter a route (Yes group, coded 1) or to exit from a route (No group, coded 0). Therefore, the requirement for a dichotomous dependent measure was satisfied.

*Mutually exclusive categories on the DV.* In logistic regression, the categories, or groups, in the dependent measure must be exhaustive and independent of each other such that each case belongs to one group or the other, but not both. In the current study, this assumption was met because each case was assigned either to the “Exit” or the “Entry” group. In other words, routes in which an airline started operations were assigned to the “Entry” group, and routes in which an airline stopped operations were assigned to the “Exit” group. Because no cases were assigned to both groups, the data set was compliant with this assumption.

*Independence of scores on the dependent measure.* According to Tabachnick (2013), “Logistic regression assumes that responses of different cases are independent of each other. That is, it is assumed that each response comes from
a different, unrelated case” (p. 445). In other words, the observations should not be the result of repeated measures or matched data. For the current study, because the data were acquired from an archival data source of U.S. domestic airline traffic, it was assumed that the data for each targeted variable associated with each case were unrelated. Therefore, the data set was presumed to be compliant with the independence assumption.

**Correct specification of the model.** This assumption requires that the independent variables of the model are “relevant” to the research context. As noted in Chapter 2, the targeted variables that were included in the model were based on prior research and theory and hence were assumed to be “relevant.” Nevertheless, following Warner’s (2008) guidance, I developed a baseline (i.e., null) model in the absence of the targeted variables. I then compared the chi-square test for the fit of the null model to the fit of the primary model, which included the targeted sets of independent variables. Because this latter chi-square produced a significantly better fit than the null model, I concluded that the model was correctly specified. The results of this analysis are summarized in Table 4.7, which is provided later in this chapter as part of the discussion involving the first primary analysis.

**Absence of multicollinearity among the independent variables.** Multicollinearity refers to the situation where two or more variables are highly correlated with each other ($r > .80$). In such instances, the results of an analysis can be difficult to interpret or even useless. According to Cohen et al. (2003), the
presence of multicollinearity in a data set can lead to unreliable estimates of
dividual regression coefficients resulting in large standard errors, and can make it
difficult to compute the individual regression coefficients correctly. To determine
the presence of multicollinearity I examined the variance inflation factors (VIFs) of
the IVs. The VIF is an index that identifies multicollinearity among independent
variables by comparing the increase in variance in the regression coefficients of the
model to those of a model in which all of the variables are uncorrelated. According
to Cohen et al., the square root of the VIF represents the amount that the standard
error of a regression coefficient would increase relative to the situation in which all
of the predictor variables are uncorrelated. For example, if VIF = 4, then this would
indicate $\sqrt{4} = 2$, or a two-fold increase in standard error from the case where the
corresponding variable is not correlated with any of the other variables in the
model. Cohen et al.’s threshold for the presence of multicollinearity is VIF = 10.
Because the VIFs for the targeted variables were less than 10, there was no serious
evidence of multicollinearity among the independent variables and hence the data
set was compliant with this assumption.

Absence of outliers in the solution. Tabachnick (2013) indicated that the
solution of a logistic regression model is the predicted probability of each case
belonging to a specific group. This implies if a model contains several cases that
are poorly predicted, then this could suggest a poor model fit. For example, if the
solution of the logistic regression model for Part A predicted that an actual “Exit”
group case had a high probability of being in the “Entry” group, then the case would be considered an outlier. To address this assumption, I followed Tabachnick’s recommendation by generating a histogram and outlier box plot of the residuals. Because none of the residuals were outliers, I deduced that the model provided a good fit of the data and therefore the data set was deemed to be compliant with this assumption.

*Linearity of the IVs and logit.* According to Cohen et al. (2003), logit, which is the natural logarithm (ln) of odds of the dependent variable, is a function of the predicted probability of the dependent variable that is linearly related to the independent variables. Because logistic regression assumes a linear relationship between the continuous independent variables and logit of the dependent variable, the assumption of linearity of IVs and log odds is violated when the interaction between a continuous independent variable and its natural logarithm is statistically significant (Tabachnick, 2013). One way to address this assumption is to eliminate the continuous independent variables by transforming them to dichotomies and expressed as binary data. This is discussed in the “Data Set Modifications” section later in this chapter. Because the linearity of the logit cannot be violated by binary data, the assumption of linearity of the logit was met.

*Multiple regression assumptions (Part B).* Similar to that of Part A, the data set that corresponded to Part B of the current study was examined for compliance with the assumptions for multiple regression, which was the statistical
strategy for Part B. The reader is reminded that although some of the same independent variables from Part A also were used in Part B, the dependent variable was $Y_2 = \text{Air fare}$, which was continuous and represented airlines’ mean air fares in a city-pair market (see Chapter 3, Tables 3.2 and 3.3). Furthermore, as noted in the outlier discussion presented earlier, the data set used for this part was in the absence of outliers (total $N = 916$). Following is a brief discussion of each assumption and the corresponding results.

**Linearity.** Regression analysis assumes a linear relationship between the dependent measure and each of the independent variables in the population. In other words, the slope of the predictor equation should be constant over the full range of the IVs. Following Cohen et al.’s (2003) guidelines, to determine if the linearity assumption was satisfied I examined the plot between the residuals and predicted values. This plot yielded no discernable pattern, which was confirmed by examining the Kernel smoother line against the linear fit (i.e., zero line). The Kernel smoother line followed the trend of the zero line to the point where the two lines were nearly coincidental. As a result, the data set was compliant with the linearity assumption.

**Correct specification of the independent variables.** This assumption implies that the targeted variables are appropriate to the research context and have been identified via theory, the literature, preliminary studies, and/or personal experience. The assumption also implies that the targeted variables are properly measured, and
that the form of the relationship between each IV and DV has been properly specified. If these three conditions are met, then the IVs and residuals will be independent in the population and the regression coefficient estimates will be unbiased. Although it is impossible to know if all the “correct” variables have been targeted, this assumption essentially answers the question, “Do all the targeted IVs really belong in the final model?”

To determine if the IVs were correctly specified, I examined each variable’s respective leverage plot, which tests the relationship between the residuals of the dependent measure (i.e., what remains after all of the factors’ contribution to the DV has been accounted for) and the residuals of the targeted IV (i.e., what remains after all of the other factors’ relationship with the targeted IV has been accounted for). The results of these plots showed that all the variables were correctly specified except for $X_5 = \text{HHI}$ and $X_{10} = \text{Market share of specific competitors}$, which represented the ratio of available seats to the total number of available seats in the market relative to the eight targeted airlines: Alaska, American, Delta, JetBlue, SkyWest, Southwest, Spirit, and United. Of these airlines, JetBlue, Southwest, and Spirit were not correctly specified. As a result, $X_5$ and these three airlines were removed from the final data set.

*Reliable measurement of the independent variables.* This assumption indicates that the independent variables in a regression model are measured without error, which can only occur with a perfectly reliable instrument. According to
Cohen et al. (2003), measurement error is examined via a measure of reliability, and a reliability coefficient between .70 and .90 is acceptable when using instruments that measure attitudes and personality traits. Furthermore, Worthen et al. (1999) posited that reliability coefficients greater than .50 are acceptable in practice when used to make decisions about a group of individuals. In the current study, I did not administer any type of physical measuring instrument but instead accessed archival data from the Bureau of Transportation Statistics. These data were stored directly from actual flights with the help of analysts, computer systems, and software. Because these data are used by policymakers from both the public and private sectors, I assumed the instruments were used to collect these data were reliable, and therefore deemed the data set was compliant with this assumption.

*Constant variance of the residuals.* This assumption, which is commonly referred to as the assumption of homoscedasticity, implies that the variance of $Y$ is the same for any $X$ value. If $Y$-variance changes as the value of $X$ changes, then this assumption is violated and the statistics from the regression analysis will be incorrect. Thus, for the data set to be compliant with this assumption, the variance of the residuals should not be related to any of the IVs or to the predicted value of the DV. Violations of the homoscedasticity of the residuals assumption can be detected via the same residual analysis used to confirm the linearity assumption. Because there was no discernable pattern in this plot, the data set satisfied the homoscedasticity of residuals assumption.
Independence of residuals. This assumption indicates that the residuals of the observations must be independent of one another. In other words, there must be no relationship among the residuals for any subset of cases in the analysis. To determine if this assumption was met, a residual analysis was performed in which the residuals were plotted against the case numbers. This plot yielded no discernable pattern, which was confirmed by examining the Kernel smoother line relative to the fit line. Similar to the linearity assumption, the two lines were nearly coincidental. Thus, the data set was compliant with the independence of the residuals assumption.

Normality of residuals. This assumption implies that for any value of the IV, the residuals around the regression line are assumed to have a normal distribution. The normality of residuals assumption generally is achieved by constructing a normal q-q plot, which consists of a scatter plot of the residuals with a superimposed straight line and a 95% confidence band. A visual inspection of this plot showed the majority of the residuals “hugging” the normal line and falling within the 95% corresponding confidence band. As a result, the normality assumption was deemed to be satisfied.

Data set modifications. To prepare the two data sets for analyses, modifications were made to the data set for Part A only, and involved transforming each independent variable into a dichotomy, which assisted in the interpretation of the odds ratios in the logistic regression analyses. Thus, the likelihood of an airline
### Table 4.6
*Assignment of Variables into Dichotomies*

<table>
<thead>
<tr>
<th>Sets and Corresponding Variables</th>
<th>Assigned Dichotomy</th>
<th>Basis for Dichotomy&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set R = Route Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( X_1 ) = Airline</td>
<td>LCC vs. FSC</td>
<td>—</td>
</tr>
<tr>
<td>( X_2 ) = Route Length</td>
<td>Fewer than 850 miles vs. 850 miles or more</td>
<td>Median</td>
</tr>
<tr>
<td>( X_{3a} ) = City Population of Endpoint 1</td>
<td>Fewer than 2,814,330 vs. 2,814,330 or more people</td>
<td>Median</td>
</tr>
<tr>
<td>( X_{3b} ) = City Population of Endpoint 2</td>
<td>Fewer than 3,498,362 vs. 3,498,362 or more people</td>
<td>Median</td>
</tr>
<tr>
<td>( X_{4a} ) = Per Capita Income of Endpoint 1</td>
<td>Fewer than $47,254 vs. $47,254 or more</td>
<td>Median</td>
</tr>
<tr>
<td>( X_{4b} ) = Per Capita Income of Endpoint 2</td>
<td>Fewer than $47,894 vs. $47,894 or more</td>
<td>Median</td>
</tr>
<tr>
<td><strong>Set C = Competitor Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( X_{5a} ) = Total market HHI</td>
<td>Lower than 8609 vs. 8609 or higher</td>
<td>Median</td>
</tr>
<tr>
<td>( X_{5b} ) = HHI by Carrier Type</td>
<td>None vs. 1 or more</td>
<td>Median</td>
</tr>
<tr>
<td>( X_7 ) = Number of Competitors</td>
<td>0 competitors vs. 1 or more</td>
<td>—</td>
</tr>
<tr>
<td><strong>Set A = Airport Factors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( X_8 ) = Existence of Hub Airport</td>
<td>A hub airport exists vs. does not exit</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>a</sup>Where appropriate, the splits for the dichotomies were tested and validated through the development of contingency tables.

Entering or exiting a route for one dichotomy could be compared directly to the other, rather than comparing odds ratios over a range of values that are more difficult to understand. For example, as shown in Table 4.6, the independent variable \( X_2 = \) Route length was split at the variable’s median of 850 miles. This dichotomy facilitated a comparison of odds ratios between route lengths fewer than 850 miles and those 850 miles or more. This modified data set also served a second purpose. According to MacKinnon and Dwyer (1993), because the regression
coefficients in each step of a logistic regression analysis are measured on different scales, the interpretation of any mediation effects can be problematic if the predictor variables are not dichotomous. Thus, splitting the independent variables into dichotomies also provided a format suitable for measuring mediation effects of the targeted variables on group membership.

With the exception $X_1 = \text{Airline (LCC vs. FSC)}$, $X_8 = \text{Existence of a hub airport (yes vs. no)}$, and $X_7 = \text{Number of competitors}$, a measure of central tendency was used to produce each variable’s dichotomy. To help determine which measure of central tendency was appropriate, I ran a contingency analysis for each IV. This analysis also confirmed that each cell had at least one frequency and there was a relatively balanced number of cases in each group (“1” vs. “0”) for each dichotomy. This confirmation was necessary to ensure the logistic regression yielded maximum likelihood estimates that were not the result of numerous iterations (Tabachnick, 2013). The choice of using the mean, median, or midrange as shown in Table 4.6 was predicated on these objectives. With respect to $X_7 = \text{Number of competitors}$, as shown in Table 4.6 the split for this variable was based on existence of at least one competitor. The routes without competition were labeled as “0” and the others were “1.” The decision to create dichotomies and the corresponding strategies used are reflected in Chapters 1 and 5 as a delimitation to the studies results.
Summary of preliminary analyses. As a result of the foregoing preliminary data screening, the final data set used for Part A consisted of \( N = 2,111 \) cases, 10 independent variables expressed as dichotomies, and the dichotomous dependent variable, which was group membership (route entry-exit decision of Yes or No). The final data set for Part B consisted on \( N = 916 \) cases, 13 independent variables, and the dependent variable, which was air fares. The results of the primary analysis corresponding to each part are discussed separately below.

Primary analysis 1: Logistic regression. Part A of the current study examined the relationship between the targeted sets of variables and the dependent variable (i.e., group membership) via logistic regression by regressing the dichotomous group membership variable on the targeted sets of IVs simultaneously. A simultaneous strategy was used because of the dearth of prior research or theory to guide a set entry order. Prior to this analysis, though, I followed Warner’s (2008) recommendations and developed a baseline or null model by regressing the group membership variable in the absence of the independent variables. I then compared the overall goodness of fit of the null model to the full model. The result of this analysis is summarized in Table 4.7.

According to Warner (2008), a measure that can be used to assess the overall goodness of fit of the logistic regression model is the log likelihood (LL) function, which is comparable to the sum of the squared residuals in multiple regression. As reported in Table 4.7, the full model was statistically significant,
Table 4.7

Significance of the Simultaneous (Full) Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Log Likelihood</th>
<th>df</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null$^b$</td>
<td>1179.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full$^c$</td>
<td>1119.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>60.30</td>
<td>10</td>
<td>120.59*</td>
</tr>
</tbody>
</table>

Note. $N = 2,111$. $R^2 = .0511$.

$^a$Log Likelihood (LL) indicates the agreement between the probabilities of group membership generated by the logistic regression model and the actual group membership within the sample. Larger absolute LL values represent a worse model fit. $\chi^2 = -2(\text{LL}_{\text{null model}} - \text{LL}_{\text{full model}})$. $^b$The null model represents the baseline model without information about the predictor variables. $^c$The full model represents the hypothesized model with the independent variables entered into the model simultaneously. The IVs were $X_1, X_2, X_3a, X_3b, X_4a, X_4b, X_5a, X_5b, X_7$, and $X_8$.

$\chi^2(10) = 120.59, p < .0001$. The reader should note that the chi-square statistic is the difference between -2LL for the full model and -2LL for the null model. According to Warner, this difference should be large for the full model to be considered statistically significant. In addition to the chi-square statistic, Cohen et al. (2003) recommended reporting the Pseudo-$R^2$ ($R^2_{\text{L}}$) for logistic regression, which is the similar to $R^2$ in multiple regression. As reported in the general table note of Table 4.7, $R^2_{\text{L}} = .0511$, which represents the gain in prediction obtained from adding variables to a model. Therefore, the full model provided a predictive gain of approximately 5% over the null model ($R^2_{\text{L}} = .0511, df = 10$).

A summary of the logistic regression estimates for the full and null models is provided in Table 4.8. The reader will note that the null model was significant, $\chi^2(0) = 488.51, p < .0001$, and the null model’s logit for membership in the “Enter”
Table 4.8
Summary of Logistic Regression Estimates for the Null and Simultaneous (Full) Model

<table>
<thead>
<tr>
<th></th>
<th>$B^a$</th>
<th>SE</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Null Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.116</td>
<td>0.504</td>
<td>488.51</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td><strong>Full Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.085</td>
<td>0.182</td>
<td>0.22</td>
<td>.6403</td>
</tr>
<tr>
<td>$X_1 = \text{Airline}$</td>
<td>-0.519</td>
<td>0.116</td>
<td>20.13</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td>$X_2 = \text{Route Length}$</td>
<td>-0.421</td>
<td>0.107</td>
<td>15.46</td>
<td>&lt; .0001***</td>
</tr>
<tr>
<td>$X_{3a} = \text{City Pop of Endpoint 1}$</td>
<td>-0.315</td>
<td>0.123</td>
<td>6.57</td>
<td>.0104*</td>
</tr>
<tr>
<td>$X_{3b} = \text{City Pop of Endpoint 2}$</td>
<td>-0.15</td>
<td>0.127</td>
<td>1.50</td>
<td>.2213</td>
</tr>
<tr>
<td>$X_{4a} = \text{Per Capita Income Endpoint 1}$</td>
<td>-0.283</td>
<td>0.117</td>
<td>5.87</td>
<td>.0154*</td>
</tr>
<tr>
<td>$X_{4b} = \text{Per Capita Income Endpoint 2}$</td>
<td>-0.222</td>
<td>0.118</td>
<td>3.56</td>
<td>.0592</td>
</tr>
<tr>
<td>$X_5 = \text{HHI}$</td>
<td>0.079</td>
<td>0.146</td>
<td>0.30</td>
<td>.5865</td>
</tr>
<tr>
<td>$X_{5b} = \text{HHI by Carrier Type}$</td>
<td>0.200</td>
<td>0.166</td>
<td>1.46</td>
<td>.2275</td>
</tr>
<tr>
<td>$X_7 = \text{Number of Competitors}$</td>
<td>-0.639</td>
<td>0.180</td>
<td>12.59</td>
<td>.0004***</td>
</tr>
<tr>
<td>$X_8 = \text{Hub Existence}$</td>
<td>-0.214</td>
<td>0.122</td>
<td>3.10</td>
<td>.0785</td>
</tr>
</tbody>
</table>

Note. $N = 2,111$. $R^2 = .0511$, $df = 10$ for the full model. Number of correctly classified cases = 1,598 (76%) at a predicted probability cut of .5.

$a$Logistic regression estimates are the natural logarithm of the odds ratio. Therefore, the exponential of the estimate ($e^{B}$) yields the odds ratio. $b$The null model represents the baseline model and predicts the odds for exiting or entering a route without information provided by the independent variables. In the sample, these odds differed significantly from 1; that is, the probability of exiting or entering a route differed significantly from 0.5. For example, in the null model, $B_0 = -1.116$ and therefore $e^{-1.116} = 0.33$, which means that the odds of exiting a route was 0.33 (approximately one third odds). This indicates that approximately 1/3 of the cases in the sample exited a route during the targeted period. $c$The full model represents the hypothesized model and predicts the odds of exiting or entering a route with the independent variables entered into the model simultaneously. In the sample, these predicted odds differed significantly from the null model as demonstrated in Table 4.19.

$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$. 

...
result as a test of the omnibus, I next examined the relationship between each IV and group membership.

As summarized in Table 4.8, the full model logit \( (L_i) \) for group membership was predicted by the equation
\[
L_i = -0.519X_1 - 0.421X_2 - 0.315X_{3a} - 0.15X_{3b} - 0.283X_{4a} - 0.222X_{4b} + 0.079X_{5a} + 0.2X_{5b} - 0.639X_7 - 0.214X_8 + 0.085.
\]
Furthermore, five IVs were significantly related to group membership in the presence of the other predictors: \( X_1 = \text{Airline}, X_2 = \text{Route length}, X_{3a} = \text{City population of endpoint 1}, X_{4a} = \text{Per capita income of endpoint 1}, \) and \( X_7 = \text{Number of competitors}. \) The exponent of each regression coefficient \( (e^{Bi}) \) in the prediction equation specifies the change in odds relative to the independent variable \( (X_i) \) while controlling for the other predictors in the model. In context of the current study: (a) if \( e^{Bi} < 1.00 \), then the odds decrease for the membership in the “Enter” group relative to \( X_i \); (b) if \( e^{Bi} > 1.00 \), then the odds increase for membership in the “Enter” group relative to \( X_i \); and (c) if \( e^{Bi} = 1.00 \), then there is no change in the odds for the membership in the “Enter” group relative to \( X_i \). (The reader is reminded that the “Enter” group refers to routes of flights in which the targeted airlines entered or began operations.) A discussion of the interpretation of the five statistically significant factors shown in Table 4.8 follows.

\( X_1 = \text{Airline}. \) As reported in Table 4.8, the regression coefficient for \( X_1 = \text{Airline} \) was \( B_2 = -0.519 \). When \( e \) is raised to this coefficient, \( e^{-0.519} = 0.60 \) and its reciprocal \( e^{0.519} = 1.68 \). These exponential results represent the odds ratios for
### Table 4.9

**Summary of Odds Ratios for the Independent Variables in the Full (Simultaneous) Model**

<table>
<thead>
<tr>
<th>Independent Variables&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Odds Ratios (OR)</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1 = \text{Airline}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LCC vs. FSC</td>
<td>0.60</td>
<td>[0.47, 0.75]</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>FSC vs. LCC</td>
<td>1.68</td>
<td>[1.34, 2.11]</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>$X_2 = \text{Route Length}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fewer than 850 miles</td>
<td>0.66</td>
<td>[0.53, 0.81]</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>850 miles or more</td>
<td>1.52</td>
<td>[1.23, 1.88]</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>$X_3a = \text{City Pop of Endpoint 1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fewer than 2,814,330</td>
<td>0.73</td>
<td>[0.57, 0.93]</td>
<td>.0103*</td>
</tr>
<tr>
<td>2,814,330 or more</td>
<td>1.37</td>
<td>[1.08, 1.74]</td>
<td>.0103*</td>
</tr>
<tr>
<td>$X_3b = \text{City Pop of Endpoint 2}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fewer than 3,498,362</td>
<td>0.86</td>
<td>[0.66, 1.09]</td>
<td>.2214</td>
</tr>
<tr>
<td>3,498,362 or more</td>
<td>1.16</td>
<td>[0.91, 1.48]</td>
<td>.2214</td>
</tr>
<tr>
<td>$X_4a = \text{Per Capita Income Endpoint 1}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fewer than $47,254</td>
<td>0.75</td>
<td>[0.60, 0.95]</td>
<td>.0152*</td>
</tr>
<tr>
<td>$47,254 or more</td>
<td>1.33</td>
<td>[1.06, 1.67]</td>
<td>.0152*</td>
</tr>
<tr>
<td>$X_4b = \text{Per Capita Income Endpoint 2}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fewer than $47,894</td>
<td>0.80</td>
<td>[0.64, 1.00]</td>
<td>.0590</td>
</tr>
<tr>
<td>$47,894 or more</td>
<td>1.25</td>
<td>[0.99, 1.57]</td>
<td>.0590</td>
</tr>
<tr>
<td>$X_5a = \text{HHI}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower than 8609</td>
<td>1.08</td>
<td>[0.81, 1.44]</td>
<td>.5868</td>
</tr>
<tr>
<td>8609 or higher</td>
<td>0.92</td>
<td>[0.69, 1.23]</td>
<td>.5868</td>
</tr>
<tr>
<td>$X_5b = \text{HHI by Carrier Type}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero</td>
<td>1.22</td>
<td>[0.88, 1.69]</td>
<td>.2256</td>
</tr>
<tr>
<td>One or more</td>
<td>0.82</td>
<td>[0.59, 1.13]</td>
<td>.2256</td>
</tr>
<tr>
<td>$X_7 = \text{Number of Competitors}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None vs. 1 or more</td>
<td>0.53</td>
<td>[0.37, 0.75]</td>
<td>.0003**</td>
</tr>
<tr>
<td>1 or more vs. None</td>
<td>1.89</td>
<td>[1.34, 2.71]</td>
<td>.0003**</td>
</tr>
<tr>
<td>$X_8 = \text{Hub Existence}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A hub exists</td>
<td>0.81</td>
<td>[0.64, 1.03]</td>
<td>.0796</td>
</tr>
<tr>
<td>A hub does not exist</td>
<td>1.24</td>
<td>[0.97, 1.57]</td>
<td>.0796</td>
</tr>
</tbody>
</table>

**Note.** $N = 2,111.$

<sup>a</sup>Significance tests and confidence intervals (CI) on odds ratios for the independent variables are likelihood ratio ($\chi^2$) based.

*p < .05. **p < .01. ***p < .001.

Airline type (FSC vs. LSC) and are summarized in Table 4.9. Interpreting these results in the context of the current study, $e^{-0.519} = 0.60$ indicates that the odds of belonging to the “Enter” group decreased significantly for routes operated by
LCCs, and the reciprocal, $e^{0.519} = 1.68$, means that the odds of belonging to the “Enter” group increased significantly for routes operated by FSCs. More concretely, this latter result indicates that FSCs were 1.68 times more likely than LCCs to enter (i.e., begin operations) a new route. This result also suggests that the odds of entering a new route decreased by approximately 40% for LCCs compared to FSCs, holding all other variables constant. It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time the odds ratio in the population would vary anywhere between 1.34 and 2.11. In other words, 95% of the time FSCs would be between 1.34 and 2.11 times more likely than LCCs to enter a new route. Given the relatively narrow width of this interval, the corresponding accuracy in parameter estimation is relatively high.

$X_2 = \text{Route length}$. As reported in Table 4.8, the regression coefficient for $X_2 = \text{Route length}$ was $B_2 = -0.421$, and the corresponding odds ratio was $e^{-0.421} = 0.66$ and its reciprocal $e^{0.421} = 1.52$ (Table 4.9). Thus, the odds of belonging to the “Enter” group decreased significantly for routes fewer than 850 miles, and the odds of belonging to the “Enter” group increased significantly for routes with length 850 miles or more. More concretely, this latter result indicates that routes with lengths of 850 miles or more were 1.52 times more likely to be “entered” than routes with lengths of fewer than 850 miles. This result also suggests that the odds of entering a new route decreased by 34% for routes that were fewer than 850 miles in length when compared to routes that were 850 miles or more in length, holding all other
variables constant. It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time the odds ratio in the population would vary anywhere between 1.23 and 1.88. In other words, 95% of the time routes of 850 miles or more in length would be between 1.23 and 1.88 times more likely to be “entered” when compared to routes fewer than 850 miles in length. The width of this interval is fairly narrow, which makes the accuracy in estimating these odds in the population relatively high.

\( X_{3a} = \text{City population of endpoint 1} \). As reported in Table 4.8, the regression coefficient for \( X_{3a} = \text{City population of endpoint 1} \) was \( B_{3a} = -0.315 \), and the odds ratio was \( e^{-0.315} = 0.73 \) and its reciprocal \( e^{0.315} = 1.37 \) (Table 4.9). Thus, the odds of belonging to the “Enter” group decreased significantly for routes with a population at Endpoint 1 of fewer than 2,814,330 people, and the odds of belonging to the “Enter” group increased significantly for routes with a population at Endpoint 1 of 2,814,330 or more people. More concretely, this latter result indicates that prospective new routes with a population at Endpoint 1 of 2,814,330 or more people were nearly 1.5 times more likely \((OR = 1.37)\) to be entered than routes with a population at Endpoint 1 of fewer than 2,814,330 people. This result also suggests that the odds of entering a new route decreased 27% for routes with a population at Endpoint 1 of fewer than 2,814,330 people, holding all other variables constant. It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time the odds ratio in the population
would vary anywhere between 1.08 and 1.74. In other words, 95% of the time routes with a population at Endpoint 1 of 2,814,330 or more people would be between 1.08 and 1.74 times more likely to be “entered” when compared to routes with a population at Endpoint 1 of fewer than 2,814,330 people. Given the width of this interval, the accuracy in estimating these odds in the population is moderate.

\( X_{4a} = \text{Per capita income endpoint 1} \). The regression coefficient for \( X_{4a} \) was \( B_2 = -0.283 \) (Table 4.8), and the corresponding odds ratio was \( e^{-0.283} = 0.75 \) and its reciprocal \( e^{0.283} = 1.33 \) (Table 4.9). Thus, the odds of belonging to the “Enter” group decreased significantly for routes with a per capita income of fewer than $47,254 at Endpoint 1, and the odds of belonging to the “Enter” group increased significantly for routes with a per capita income of $47,254 or more at Endpoint 1. More concretely, this latter result indicates that prospective new routes with a per capita income of $47,254 or more at Endpoint 1 were 1.33 times more likely to be entered than routes with a per capita income of fewer than $47,254 at Endpoint 1. This result also suggests that the odds of entering a new route decreased by 25% for routes with a per capita income of fewer than $47,254 at Endpoint 1 when compared to routes with a per capita income $47,254 or more at Endpoint 1, holding all other variables constant. It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time the odds ratio in the population would vary anywhere between 1.06 and 1.67. In other words, 95% of the time routes with a per capita income of
$47,254 or more at Endpoint 1 would be between 1.06 and 1.67 times more likely to be entered when compared to routes with a per capita income of fewer than $47,254 at Endpoint 1. Given the width of this interval, the accuracy in estimating these odds in the population is moderate.

\[ X_2 = \text{Number of competitors}. \] The regression coefficient for \( X_2 = \text{Number of competitors} \) was \( B_2 = -0.639 \) (Table 4.8), and the corresponding odds ratios were \( e^{-0.639} = 0.53 \) and its reciprocal \( e^{0.639} = 1.89 \) (Table 4.9). Thus, the odds of belonging to the “Enter” group decreased significantly for routes with no competitors, and the odds of belonging to the “Enter” group increased significantly for routes with one or more competitors. More concretely, this latter result indicates that prospective new routes with one or more competitors were nearly two times more likely \( (OR = 1.89) \) to be entered than routes with no competitors. This result also suggests that the odds of entering a new route decreased by 47% for routes with no competitors when compared with routes with one or more competitors, holding all other variables constant. It also should be noted that the corresponding 95% confidence interval (Table 4.9) indicates that 95% of the time the odds ratio would vary anywhere between 1.34 and 2.71. In other words, 95% of the time routes with one or more competitors would be between 1.34 and 2.71 times more likely to be entered when compared to routes with no competitors. Given the width of this interval, the accuracy in estimating these odds in the population is moderate.
Nonsignificant variables. Although not statistically significant at the preset alpha level of $\alpha = .05$, the results reported in Table 4.9 show that $X_{4b} = \text{Per Capita Income Endpoint 2}$ was significant for $\alpha = .06$ ($p = .0590$), and $X_8 = \text{Hub Existence}$ was significant for $\alpha = .08$ ($p = .0796$). Thus, if the reader is willing to accept a slightly higher increase in Type I error, then (a) prospective new routes with a per capita income of $47,894$ or more at Endpoint 2 were 1.25 times more likely to be entered than routes with a per capita income of fewer than $47,894$ at Endpoint 2, and (b) prospective new routes that did not have an existing hub were 1.24 times more likely to be entered than routes that did have an existing hub. Thus, the odds of entering a new route decreased where the per capita income was fewer than $47,894$ at Endpoint 2, and where a hub airport in any of the endpoints of the route existed.

Primary analysis 2: Multiple regression. Part B of the current study examined the relationship two sets of variables, $R = \text{Route Factors}$ and $M = \text{Market Factors}$, had with the dependent variable (air fare) via a hierarchical multiple regression strategy. This was done by regressing air fares on the two sets of IVs via the set entry order of $R–M$. Set $R$ consisted of $X_2 = \text{Route length}$, $X_{3a}$ and $X_{3b}$, which consisted of the populations at city endpoints 1 and 2, respectively, and $X_{4a}$ and $X_{4b}$, which consisted of the per capita income at city endpoints 1 and 2, respectively. Set $M$ consisted of $X_9 = \text{Market size}$, $X_{10} = \text{Market share of five competing airlines}$ ($X_{10a} = \text{Alaska}$, $X_{10b} = \text{American}$, $X_{10c} = \text{Delta}$, $X_{10d} = \text{SkyWest}$,}
and $X_{10e} = \text{United}$, $X_{11} = \text{Largest fare}$, and $X_{13} = \text{Business model of largest fare}$.

The results are summarized in Table 4.10, and a discussion of these results follows.

**Set R: Route factors.** As reported in Table 4.10, when the five variables comprising Set R entered the model together and in the presence of no other variables other than themselves, the collective contribution they made in explaining the variance in air fares was significant, $R^2 = .49$, $F(5, 910) = 175.10$, $p < .0001$.

Given a significant omnibus test, inspection of the individual factors within Set R showed that four of the five factors were significant. The exception was $X_{3a} = \text{City population at Endpoint 1}$, $B_{3a} = 0.0000004$, $t(910) = 0.89$, $p = .3724$. With respect to the four significant factors:

- $X_2 = \text{Route length}$, $B_2 = 0.052$, $t(910) = 22.46$, $p < .0001$. This result indicates that for every 100-mile increase in route length, air fares increased on average by $5.24$.

- $X_{3b} = \text{City population at endpoint 2}$, $B_{3b} = -0.0000010$, $t(910) = -2.27$, $p = .0232$. This result indicates that for every 1-million increase in the population of the city at Endpoint 2 of a city-pair route, air fares decreased on average by $1.00$.

- $X_{4a} = \text{Per capita income of city at endpoint 1}$, $B_{4a} = 0.00160$, $t(910) = 6.65$, $p < .0001$. This result indicates that for every $10,000$-increase in the per capita income of the city at Endpoint 1 of a city-pair route, air fares increased on average by $16.00$. 


Table 4.10

Summary of Hierarchical Multiple Regression Results

<table>
<thead>
<tr>
<th>Factor</th>
<th>Model 1 $B^0$</th>
<th>Model 2c $B$</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-12.27</td>
<td>-17.35***</td>
</tr>
<tr>
<td></td>
<td>$X_2$</td>
<td>0.0524***</td>
<td>0.0068***</td>
</tr>
<tr>
<td></td>
<td>$X_3a$</td>
<td>0.0000004</td>
<td>0.0000008***</td>
</tr>
<tr>
<td></td>
<td>$X_3b$</td>
<td>-0.0000010*</td>
<td>0.0000005***</td>
</tr>
<tr>
<td></td>
<td>$X_4a$</td>
<td>0.00160***</td>
<td>0.000459***</td>
</tr>
<tr>
<td></td>
<td>$X_4b$</td>
<td>0.00203***</td>
<td>0.000486***</td>
</tr>
<tr>
<td></td>
<td>$X_9$</td>
<td>-0.0055***</td>
<td>[-0.0065, -0.0045]</td>
</tr>
<tr>
<td></td>
<td>$X_{10a}$</td>
<td>23.98*</td>
<td>[1.98, 45.99]</td>
</tr>
<tr>
<td></td>
<td>$X_{10b}$</td>
<td>10.73***</td>
<td>[7.11, 14.36]</td>
</tr>
<tr>
<td></td>
<td>$X_{10c}$</td>
<td>6.30***</td>
<td>[3.22, 9.38]</td>
</tr>
<tr>
<td></td>
<td>$X_{10d}$</td>
<td>11.53*</td>
<td>[2.53, 20.52]</td>
</tr>
<tr>
<td></td>
<td>$X_{10e}$</td>
<td>6.31**</td>
<td>[1.65, 10.98]</td>
</tr>
<tr>
<td></td>
<td>$X_{11}$</td>
<td>0.77***</td>
<td>[0.75, 0.79]</td>
</tr>
<tr>
<td></td>
<td>$X_{13}$</td>
<td>15.24***</td>
<td>[12.98, 17.49]</td>
</tr>
</tbody>
</table>

Statistical Results

- $R^2$: .49
- $F$: 174.10***
- $\Delta R^2$ ($sR^2$): .46
- $\Delta F$: 1037.30***

**Note.** $N = 916$. Set entry order was R–M.

aSet R = Route Factors: $X_2$ = Route length, $X_3a$ = City population at Endpoint 1, $X_3b$ = City population at Endpoint 2, $X_4a$ = Per capita income of city at Endpoint 1, and $X_4b$ = Per capita income of city at Endpoint 2. Set M = Market Factors: $X_9$ = Market size, which was the total number of available seats in a market; $X_{10a}$ = Market share of Alaska Airlines, $X_{10b}$ = Market share of American Airlines, $X_{10c}$ = Market share of Delta Airlines, $X_{10d}$ = Market share of SkyWest Airlines, $X_{10e}$ = Market share of United Airlines, where market share was the ratio of available seats to the total number of available seats in the market relative to the eight targeted airlines; $X_{11}$ = Largest fare; and $X_{13}$ = Business model of largest fare (full-service carrier or low-cost carrier).
bModel 1 corresponds to $Y =$ Air Fare regressed on Set R. cModel 2 corresponds to $Y =$ Air Fare regressed on Set M in the presence of Set R. Model 2 also is the final model, which would result from a simultaneous multiple regression strategy.

*p < .05. **p < .01. ***p < .001.
\[ X_{4b} = \text{Per capita income of city at endpoint 2}, \ B_{4b} = 0.00203, \ t(910) = 9.38, \ p < .0001. \] This result indicates that for every $10,000-increase in the per capita income of the city at Endpoint 2 of a city-pair route, air fares increased on average by $20.30.

**Set M: Market factors.** As reported in Table 4.10, when the eight variables comprising Set M entered the model in the presence of Set R, the 13 factors collectively explained 95% of the variance in air fares, which was significant, \( R^2 = .95, \ F(13, 902) = 1385.55, \ p < .0001. \) Furthermore, in the presence of the five factors of Set R, the eight factors of Set M uniquely explained an additional 46% of the variance in air fares, which also was significant, \( sR^2 = .46, \ F(8, 902) = 1037.30, \ p < .0001. \) Given a significant omnibus test, inspection of the individual factors within Set M in the presence of Set R showed that all eight factors were significant:

\[ X_9 = \text{Market size}, \ B_9 = -0.0055, \ t(910) = -10.63, \ p < .0001. \] This result indicates that for every 100-unit increase in the total number of available seats in a market, air fares decreased on average by $0.55.

\[ X_{10a} = \text{Alaska Airlines}, \ B_{10a} = 23.98, \ t(910) = 2.14, \ p = .0327. \] This result indicates that for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $23.98.

\[ X_{10b} = \text{American Airlines}, \ B_{10b} = 10.73, \ t(910) = 5.81, \ p < .0001. \] This result indicates that for every 1-unit increase in the ratio between available seats to
the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $10.73.

\[ X_{10c} = Delta Airlines, B_{10c} = 6.30, t(910) = 4.01, p < .0001. \] This result indicates that for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $6.30.

\[ X_{10d} = SkyWest Airlines, B_{10d} = 11.53, t(910) = 2.51, p = .0121. \] This result indicates that for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $11.53.

\[ X_{10e} = United Airlines, B_{10e} = 6.31, t(910) = 2.66, p = .0081. \] This result indicates that for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $6.31.

\[ X_{11} = Largest fare, B_{11} = 0.77, t(910) = 75.03, p < .0001. \] This result indicates that for every $1-increase in the most expensive air fare reported for each route among the competing airlines, air fares increased on average by $0.77.

\[ X_{13} = Business model of largest fare, B_{13} = 15.23, t(910) = 13.25, p < .0001. \] This result indicates that when the business model of the airline (low-cost carrier vs. full-service carrier) with the largest fare for each route is compared relative to overall air fares, the overall air fare of LCCs was $15.23 more than the
overall air fare of FSCs in the presence of all of the other factors. The reader is cautioned not to confuse this result with the difference in largest fare between LCCs and FSCs. The dependent variable in the current situation is $Y_2 = \text{Air Fare}$, not $X_{11} = \text{Largest fare}$. However, for the interested reader, when $X_{11} = \text{Largest fare}$ was regressed on $X_{13} = \text{Business model of largest fare}$ in the absence of any other factor, LCCs’ largest fare for each route was $74$ less on average than FSCs’ largest fare.

As reported in Table 4.10, with the exception of the variables associated with market share ($X_{10a} - X_{10e}$), the corresponding 95% CIs for the other factors in the final model were relatively narrow, which indicates that the results provided relatively high accuracy in parameter estimation. For example, with respect to $X_2 = \text{Route length}$, the 95% CI = [0.005, 0.009]. This indicates that 95% of the time we expect air fares in the population to increase on average anywhere between $0.49$ and $0.86$ for every 100-mile increase in route length when considered in the presence of the other 12 factors. The reader also should observe for Table 4.10 that all 13 factors in the final model are statistically significant, and that the model explains 95% of the variability in air fares.

**Results of Hypotheses Testing**

The research hypotheses of the current study, which were presented in Chapter 1, are restated here in null form for testing purposes. The decision to reject or fail to reject a null hypothesis was based on the results of the respective primary
analyses reported in this chapter. A discussion of the decisions made with respect to each null hypothesis follows.

**Null hypothesis 1: Independent of set membership, and when examined from a simultaneous perspective, none of the targeted variables will have a significant predictive value in distinguishing between LCCs and FSCs that entered or exited a route.** As reported in Table 4.7, the simultaneous (full) model was significant, $\chi^2(10) = 120.59, p < .0001$. Given a significant omnibus, the individual factors within this model were examined for significance. As reported in Table 4.8, five variables were significant at the preset alpha level of $\alpha = .05$: $X_1 =$ Airline ($p < .0001$), $X_2 =$ Route length ($p < .0001$), $X_{3a} =$ City population of Endpoint 1 ($p < .0104$), $X_{4a} =$ Per capita income of city at Endpoint 1 ($p < .0154$), and $X_7 =$ Number of competitors ($p = .0004$). Because the corresponding odds ratios ($OR$) as reported in Table 4.9 differed significantly from 1.00 for $X_1 =$ Airline ($OR = 1.68$), $X_2 =$ Route length ($OR = 1.52$), $X_{3a} =$ City population of Endpoint 1 ($OR = 1.37$), $X_{4a} =$ Per capita income of city at Endpoint 1 ($OR = 1.33$), and $X_7 =$ Number of competitors ($OR = 1.89$), I rejected Hypothesis 1.

**Null hypothesis 2a: The set of route factors (Set R) will have no significant relationship with airlines’ air fare levels for the targeted city-pair markets.** As reported in Table 4.10, a hierarchical regression analysis in which Set R = Route Factors was entered into the model first resulted in statistical significance, $R^2 = .49, F(5, 910) = 174.10, p < .0001$. A follow-up analysis showed
that four of the five factors had a significant relationship with air fares: $X_2 = \text{Route length (} p < .0001 \text{)}, X_{3b} = \text{City population at Endpoint 2 (} p = .0232 \text{), } X_{4a} = \text{Per capita income of city at Endpoint 1 (} p < .0001 \text{), and } X_{4b} = \text{Per capita income of city at Endpoint 2 (} p < .0001 \text{)}. More specifically: (a) as route lengths increased, air fares increased; (b) as the population of the city at Endpoint 2 increased, air fares decreased; and (c) as the per capita income of the respective cities at Endpoints 1 and 2 increased, air fares also increased. As a result, Hypothesis B1 was rejected.

**Null hypothesis 2b: The set of market factors (Set M) will have no significant relationship with airlines’ air fare levels for the targeted city-pair markets when examined in the presence of the set of route factors.** As reported in Table 4.10, when the market factors set entered the analysis in the presence of the route factors set, the increment was significant, $sR^2 = .46, F(8, 902) = 1037.30, p < .0001$. A follow-up analysis showed that all eight factors were significant: $X_9 = \text{Market size (} p < .0001 \text{), } X_{10a} = \text{Market share of Alaska Airlines (} p = .0327 \text{), } X_{10b} = \text{Market share of American Airlines (} p < .0001 \text{), } X_{10c} = \text{Market share of Delta Airlines (} p < .0001 \text{), } X_{10d} = \text{Market share of SkyWest Airlines (} p = .0121 \text{), } X_{10e} = \text{Market share of United Airlines (} p = .0081 \text{), } X_{11} = \text{Largest air fare within a particular route (} p < .0001 \text{); and } X_{13} = \text{Business model (LCC vs. FSC) of largest fare within a particular route (} p < .0001 \text{). More specifically: (a) as market size increased, air fares increased; (b) as the market share of each of the competing airlines increased, air fares increased; (c) as the most expensive average air fare
reported for each route among the competing airlines increased, overall air fares increased; and (d) the overall air fare of LCCs was significantly higher than the overall air fare of FSCs when the most expensive average fares for each route were compared. As a result, Hypothesis B2 was rejected.
Chapter 5

Conclusions, Implications, and Recommendations

Summary of Study

The current study focused on two survival strategies of U.S. domestic airlines: (a) exit and entry strategies and (b) air fare competition strategies. In Part A I analyzed factors relevant to airlines’ route entry and exit decisions. The dichotomous criterion variable was whether an airline entered a new a route or exited from an existing route. In Part B I examined the relationship various sets of factors had with average air fare, which was a continuous criterion variable.

The targeted research factors, which were derived from both theory and the literature, were grouped into four functional sets. Set R = Route Factors was comprised of six variables: carrier type, which distinguished between airlines that were full-service carriers (FSCs) and low-cost carrier (LCCs), route length, city populations at each end-point of a city-pair market; and per capita income at each end-point of the city-pairs that formed the route. Set C = Competitor Factors was comprised of three variables: the total market concentration of a route, the market concentration of the incumbent airlines in each route relative to carrier type, and the number of competitors in the route. Set A = Airport Factors was comprised of a single dichotomous variable that represented whether there was a hub airport of an airline in the origin or destination airports of the route. Set M = Market Factors was comprised of 12 factors: total market concentration, market concentration of
specific competitors, the average air fare level of the airline having largest average air fare in a city-pair market, and the carrier type of the airline having the largest fare in the route. In the context of the current study, which was relative to the U.S. domestic airline market, I analyzed the relationship sets R, C, A had with the entry and exit strategies of FSCs and LCCs, and I examined the relationship sets R and M had with the average air fare level in the city-pair markets.

I employed two different research methodologies. For Part A I employed a cause-type ex post facto design because the grouping was on the dependent variable (whether an airline entered or exited a route) and I analyzed the extent to which the targeted factors were linked to the preexisting effect. For Part B I employed both an explanatory and predictive correlational design because I analyzed a single group (all U.S. airlines in the top 1,000 domestic markets) and the extent to which the targeted factors were related to this group. Data for both parts of the study were archived in publicly accessible databases maintained by the U.S. government. As a result, I did not use any formal data collection instrument, and was not able to give attention to instrumentation validity and reliability.

The target population for Part A was all U.S. domestic airline route markets for the 5-year period 2011–2015. The accessible population was all U.S. domestic airline route markets for the same 5-year period that were reported in the U.S. Department of Transportation’s (2015b) Airline Origin and Destination Survey (DB1B) and the U.S. Department of Transportation’s (2015c) Bureau of
Transportation Statistics’ T-100 data bank. The sampling strategy was purposive, and I selected all entry and exit data associated with the targeted eight airlines from the databases. Thus, the sample for Part A represented a census because it contained all the reported cases for the targeted airlines during the given timeframe. The total sample size for Part A was \( N = 2,111 \): \( n = 1,590 \) cases reflected routes the airlines entered and \( n = 521 \) cases represented routes the airlines exited. The eight airlines represented approximately 85% of the total U.S. domestic market during the targeted 5-year period and included five FSCs (Alaska, American, Delta, SkyWest, and United) and three LCCs (JetBlue, Southwest, and Spirit).

The target population for Part B was all city-pair markets in the U.S. domestic airline industry for the 11-year period 2005–2015. The accessible population was the 1,000 largest city-pair markets in the 48 contiguous states reported in the DB1B database. The targeted airlines for Part B were AirTran, Alaska, Allegiant, American, Delta, Frontier, Hawaiian, JetBlue, Northwest, SkyWest, Spirit, Southwest, United, US Airways, and Virgin. The data associated with these airlines represented approximately 70% of total U.S. domestic passengers. The sampling strategy for Part B was simple random. I randomized the data from the accessible population by each year separately, combined these randomized data sets into a single data set, and then randomly selected \( N = 1,300 \) city-pair markets of which \( N = 1,082 \) were related to the 15 targeted airlines. Thus, the total sample size for Part B was \( N = 1,082 \).
Summary of Findings

Preliminary analyses. Prior to testing the study’s hypotheses, I performed various preliminary data screening activities that included examining the respective data sets for invalid data, missing data, outliers, and confirming the data sets were compliant with the assumptions of logistic regression (Part A) and multiple regression (Part B). As result of these preliminary analyses, the final sample size for Part A was maintained at $N = 2,111$. This data set, however, included 151 outliers (7%), which were flagged by Jackknife distances. They were kept in the final data set because they had little-to-no impact on the findings and they reflected real-world situations. For Part B, though, the initial data set was reduced by 166 outlier cases, and the final sample size was $N = 916$. The sample sizes for both parts were adequate relative to the recommendations from Agresti (2007), Peduzzi et al. (1996), and Hsieh (1989, p. 799), and both data sets were compliant with the assumptions of logistic and multiple regression, respectively. A brief discussion of the significant findings from each part of the study follows, and a summary of the results of the corresponding hypothesis tests that were performed relative to these findings are provided in Table 5.1.

Primary analysis 1: Logistic regression. In Part A of the current study I examined the relationship between the targeted variables and the dichotomous dependent variable (whether an airline entered or exited a route) via logistic regression by regressing the dichotomous group membership variable on the
targeted sets of IVs simultaneously. The results of this analysis revealed that the full model was significant, and five of the eight factors were significantly related to group membership in the presence of other predictors. These significant factors were $X_1 = (FSC \text{ vs. LCC}), X_2 = \text{Route length}, X_{3a} = \text{City population at Endpoint 1},$ $X_{4a} = \text{Per capita income of city at Endpoint 1},$ and $X_7 = \text{Number of competitors}.$ The results were as follows: (a) FSCs were 1.68 times more likely than LCCs to enter (i.e., begin operations) a new route, (b) the airlines were 1.52 times more likely to enter routes with lengths of 850 miles or more than routes with lengths of fewer than 850 miles, (c) the airlines were nearly 1.5 times more likely to enter a new route with a city population at Endpoint 1 of 2,814,330 or more people than routes with a city population at Endpoint 1 of fewer than 2,814,330 people, (d) the airlines
were 1.33 times more likely to enter a new route with a city per capita income of $47,254 or more at Endpoint 1 than routes with a city per capita income of fewer than $47,254 at Endpoint 1, and (e) the airlines were nearly two times more likely to enter a new route with one or more competitors than routes with no competitors.

**Primary analysis 2: Multiple regression.** In Part B I examined the relationship between the targeted variables and average air fare levels in city-pair markets via multiple regression by regressing air fares on the targeted sets of IVs hierarchically using the set entry order R-M. The results of this analysis revealed that when Set R = Route Factors entered the model, the overall set was significant, and four of the five factors in this set were significant: $X_2 = \text{Route length}$, $X_{4a} = \text{Per capita income of city at Endpoint 1}$, and $X_{4b} = \text{Per capita income of city at Endpoint 2}$, had a positive relationship with air fares, and $X_{3b} = \text{City population at Endpoint 2}$ had a negative relationship with air fares.

When Set M = Market Factors entered the analysis in the presence of Set R, the overall model was significant, the increment Set M made in the presence of Set R was significant, and a follow-up analysis showed that all eight factors in Set M were significant. The overall final model indicated: (a) for every 10-mile increase in route length, average air fares on average increased by 68 cents; (b) for every 1-million increase in the population of the city at Endpoint 1 of a city-pair route, air fares increased on average by 80 cents; (c) for every 1-million increase in the population of the city at Endpoint 2 of a city-pair route, air fares increased on
average by 50 cents; (d) for every $10,000-increase in per capita income at Endpoint 1, average air fares on average increased by $4.60; (e) for every $10,000-increase in per capita income at Endpoint 2, average air fares on average increased by $4.80; (f) for every 100-unit increase in the total number of available seats in a market, air fares decreased on average by $55 cents; (g) for every $1-increase in the most expensive air fare reported for each route among the competing airlines, air fares increased on average by 77 cents; and (h) when the business model of the airline (LCC vs. FSC) with the largest fare for each route is compared relative to overall air fares, the overall air fares of LCCs was $15.24 more than the overall air fares of FSCs in the presence of all of the other factors. Finally, for every 1-unit increase in market share, which was the ratio between available seats to the total number of available seats in the market relative to the competing airlines, the average air fares for the competing airlines increased. More specifically, Alaska Airlines’ air fares increased on average by $23.98, American Airlines’ air fares increased on average by $10.73, Delta Airlines’ air fares increased on average by $6.30, SkyWest Airlines’ air fares increased on average by $11.53, and United Airlines’ air fares increased on average by $6.31.

Conclusions and Inferences

In this section, I summarize the study’s findings and discuss these findings relative to the research questions presented in Chapter 1. A separate presentation is given for each question. In addition to the summary of findings, I also include
interpretations of the results in the context of the research setting, and plausible explanations for the results obtained. Moreover, if there is a necessity, I included anecdotal information and additional analysis related to each of the results.

**Research question 1: Independent of set membership, what is the relationship between the targeted variables and the dichotomous response variable that distinguished between LCCs and FSCs that entered a new route and those that exited from an existing route?** All IVs relevant to Part A were included in the model and the results of the simultaneous logistic regression revealed that the overall model was significant, $\chi^2(10) = 120.59, p < .0001$. This result revealed that the relationship between the targeted factors and the dichotomous variable that distinguished between the entry and exit groups was unlikely due to chance. There were eight factors in the overall model and five were significant: $X_1 = \text{Airline (FSC vs. LCC)}$, $X_2 = \text{Route length}$, $X_{3a} = \text{City population at Endpoint 1}$, $X_{4a} = \text{Per capita income of city at Endpoint 1}$, and $X_7 = \text{Number of competitors}$.

$X_1 = \text{Airline}$. This factor represented the business model of the airline examined in the overall model. FSCs were 1.68 times more likely than LCCs to enter (i.e., begin operations) a new route. One plausible explanation for this finding is related to the main characteristic differences between FSCs and LCCs. The purpose of the LCC business model is to decrease the cost of flying so as to increase the penetration of air transportation. LCCs are more cost sensitive than
FSCs. Therefore, LCCs revised their expansion policies with increasing operating costs between 2011 and 2015. During that time period one of the most important cost drivers of airlines—jet fuel costs—increased a lot. Besides that, many other cost drivers such as airport fees, employee costs, and ground handling costs increased. Moreover, after the 2008–2009 economic recession in United States, the purchasing power of people decreased. These issues mostly encouraged airlines to revise their expansion strategies. During that time, LCCs such as Southwest and Jet Blue were more conservative than FSCs in terms of network growth.

\[ X_2 = \text{Route length}. \] The results revealed that routes with lengths of 850 miles or more were 1.52 times more likely to be “entered” than routes with lengths of fewer than 850 miles. One plausible explanation for this finding is similar to the previous explanation about the airlines’ business model. One of the main differences between FSCs and LCCs is length of the flight routes. The main driver of entering a flight route was increasing fuel cost in recent years. It is obvious that the longer the flight route, the higher the fuel cost. The anecdotal information in here is the aplomb between the previous result and that result. It is obvious that one of the most important differences between LCCs and FSCs is the difference between average route lengths. In general, LCCs fly between close city-pairs. However, FSCs fly between very far city-pairs. Although the business model of LCCs is based on point-to-point short haul markets, the business model of FSCs is
based on hub-and-spoke long haul markets. Thus, it is reasonable that FSCs were more likely to enter new routes that were greater in length than LCCs.

\[ X_{3a} = \text{City population of endpoint 1.} \] The results showed that prospective new routes with a city population at Endpoint 1 of 2,814,330 or more people were nearly 1.5 times more likely to be entered than routes with a population at Endpoint 1 of fewer than 2,814,330 people. One plausible explanation for this finding is related with the type of the markets that FSCs mostly serve. As explained before, increasing costs, decreasing purchasing power of passengers, and high level of competition affected the dynamics of expansion strategies of airlines. FSCs’ business models are based on hub-and-spoke and medium and long-haul markets. In general, these types of airlines have a hub airport in major cities such as New York, Atlanta, and Dallas. For example, Atlanta Airport is the major hub of Delta Airlines, and Dallas Love Field Airport is the major hub of Southwest Airlines. Delta is one of the most well known FSCs of the U.S. domestic market, and Southwest is one of the most well known LCCs of the U.S. domestic market. Through the presence of these hub airports and the hub-and-spoke network model, FSCs are able to provide service between major cities.

A second plausible explanation is related to another very important distinction between LCCs’ and FSCs’ business models, namely, product differentiation. The product that airlines sell is the seat. LCCs mostly have only economy class seats, but FSCs have business, first, and economy class seats. To
compensate for the increasing costs, airlines tend to expand their network map with high population cities. Moreover, the FSCs use these major cities as hubs in their hub-and-spoke network model. LCCs mostly operate in secondary airports, which serve low population cities. Therefore, network growth of FSCs, especially in major cities, increased the likelihood of airlines’ expansion in high population cities.

\[ X_{4a} = \text{Per capita income of endpoint 1.} \] The results indicated that prospective new routes with a per capita income of $47,254 or more at Endpoint 1 were 1.33 times more likely to be entered than routes with a per capita income of fewer than $47,254 at Endpoint 1. One plausible explanation for this finding is similar to the explanation of the previous factor. Major cities with high population mostly have higher per capita income. The higher the per capita income level of customers the more they will pay for a ticket. Many managers, directors, or C-level executives live in major cities with high per capita income level. All these factors stimulate airlines to expand their network map in cities with high per capita income level. Thus, it makes sense to establish a presence where the money is located.

\[ X_7 = \text{Number of competitors.} \] The results revealed that prospective new routes with one or more competitors were nearly two times more likely to be entered than routes with no competitors. One plausible explanation for this finding is higher competition levels represent high revenue potential flight routes. Especially after an economic downturn and an increase in jet fuel prices, the
attitudes of airlines were toward increasing network growth in flight routes with high revenue potential.

A second plausible explanation is related to the concept of competition dynamics. Given the voluminous amount of passenger data the airlines collect and the sophisticated data analyses they conduct, there must be a reason why a particular route has no competition. One possible reason could be because it is not profitable or it is untested. Thus, it makes sense for an airline to pursue routes with existing competition (there must be a reason why other airlines are present) than to pursue routes absent of any competition.

Research question 2: When examined from a hierarchical perspective with set entry order R–M, what is the incremental knowledge gained at each step of the analysis relative to airlines’ air fare levels for the targeted city-pair markets? When Set R = Route Factors entered the model in the absence of any other variables other than those in the set, the overall model was significant. The five factors of Set R collectively provided a 49% predictive gain over the null model, $R^2_{\text{Set } R} = .49$, $F(5, 910) = 174.10$, $p < .0001$. Furthermore, four of the five factors were significant at this stage: route length, city population at Endpoint 2, per capita income Endpoint 1, and per capita income Endpoint 2.

When the eight variables of Set M = Market Factors entered the analysis in the presence of the five variables from Set R, the overall model was still significant, and the 13 factors collectively provided a 95% predictive gain over the
null model, $R^2_{SetsRM} = .95, F(13, 910) = 1385.55, p < .0001$. Furthermore, the unique contribution the eight factors of Set M made in the presence of Set R also was significant, $sR^2 = .46, F(8, 902) = 1037.30, p < .0001$. There was a 46% additional knowledge over the model with only set R. At this final stage, all 13 factors were significant. Plausible explanations for these significant variables are as follows:

$X_2 = \text{Route length}$. This factor represented the distance between origin and destination in a city pair market, and its relationship with average air fare was positive: for every 10-mile increase in route length, air fares on average increased by 68 cents. Thus, it appears that higher route lengths resulted in higher average air fare levels. One plausible explanation for this finding is increasing fuel cost and maintenance cost with increasing route length. Airplanes have various types of periodic maintenance checks. The periods of these checks are mostly related with the utilization of aircraft. The more the aircraft are utilized the higher the total distances that the aircraft fly. Therefore, increasing route length increases the utilization of aircraft, and the number of periodic maintenance checks in a constant time range increases with increasing route length and utilization.

On the other hand, as mentioned earlier, fuel costs increase with increasing route length. In recent years, on average between 30% and 40% of the total cost of an airline is due to fuel cost. In general, air fares for long haul routes are higher when compared to short haul routes. The anecdotal information here is related to
the difference in properties of city-pair markets that FSCs and LCCs have operations and the difference in product differentiation styles of FSCs and LCCs. As was previously mentioned, average air fares for FSCs are higher than LCCs. One reason for that is, in general, FSCs operate in long and medium haul routes, whereas LCCs operate in short and medium haul routes. The other reason for that is FSCs have business or first-class compartments with higher fares whereas LCCs have only economy class compartments. It is clear that FSCs operate in longer routes and they have higher average air fares. This is why average air fares increase with increasing route length. It should be noted that during the study’s time period, one of the most important cost drivers of airlines, jet fuel cost, increased a lot.

\[ X_{3a} = \text{City population of endpoint 1}. \] This factor represented the population of origin city in a city-pair market, and its relationship with average air fares was positive: for every 1-million increase in the population of the city at Endpoint 1 of a city-pair route, air fares increased on average by 80 cents. Thus, it appears that higher population of origin cities result in higher average air fare levels. However, the size of the impact seems slight. One plausible explanation for this finding is increasing demand with increasing population. Especially in recent decades, cities have continued to become gradually more urbanized. Therefore, with the help of urbanization there were high populations in certain “megacities” (defined as cities with populations of greater than 10 million), some of which are business centers and some of which are tourism centers. It is obvious that airlines
want to connect these cities to tap into the massive conceivable market for passengers there. It should be noted there also are 62 cities with populations of upwards of 5 million all around the world, and these cities account for 40% of air travel worldwide (International Air Transportation Association (IATA), 2011). All these clues show that the focus of the airlines is serving the areas where most people live and where most people want to go. It also is the rule of economy. If the demand increases when supply is constant, then the prices also increase.

\[ X_{3b} = \text{City population of endpoint 2}. \]

This factor represented the population of destination city in a city-pair market, and its relationship with average air fare was positive: for every 1-million increase in the population of the city at Endpoint 2 of a city-pair route, air fares increased on average by 50 cents. Thus, it appears that higher population of destination cities result in higher average air fare levels. One plausible explanation for this finding is exactly the same as the one given above for cities at Endpoint 1 of a city-pair market.

The anecdotal point that should be mentioned here is that impact of the population of the destination city on average air fare levels is smaller than the impact of the population of origin city. This is because most of the time origin cities of city-pair markets are hubs of major airlines and these hubs are transfer points for passengers living in hubs’ feeding areas. The hub-and-spokes network model used mostly by FSCs is intended to produce a high rate of transfer passengers, and this network strategy helps airlines maximize transfer passenger
rates by rising connectivity. The reader is reminded there are two types of passengers in airline industry: origin and destination passengers and transfer passengers. Origin and destination passengers are those who are traveling only from a spoke to the hub, or the hub to a spoke. These types of passengers do not transfer to another flight at the hub to go to someplace else. However, transfer passengers travel from somewhere to the hub, and change flights to travel from the hub to somewhere else. As the number of flight routes increase, the number of transfer passengers also increases. Every new flight route that is added into the network means new potential transfer passengers for existing flights. If the number of passengers of a single flight increases, this signifies a rise in the market size of that flight, which may increase the average air fare level of that flight. Briefly, especially most of the origin cities for highly populated cities are hub points of the major airlines, and the impact of the population on average air fare levels increase with the transfer passengers.

\[ X_{4a} = \text{Per capita income of endpoint 1}. \] This factor represented per capita income of origin city in a city-pair market, and its relationship with average air fare was positive: for every $10,000-increase in per capita income at Endpoint 1, average air fare levels on average increased by $4.60. Thus, it appears that higher per capita income of origin cities result in higher average air fare levels. There are two main plausible explanations for this finding. One of them is that the higher purchasing power of customers attracts airlines to increase average air fare levels.
for higher profit margins. The other plausible explanation is about airlines being a central dynamic in increasing tourism and economic connections through passenger and freight transportation. Approximately, 3 billion people and roughly 47 million metric tons of cargo were transported by air in 2012, and all of these activities relevant to the transportation industry have supported some 57 million jobs and $2.2 trillion in economic activity, amounting to approximately 3.5% of the global Gross Domestic Product (GDP) (IATA, 2013b). It is obvious there is a mutual relationship between air transportation and economic wealth. Airlines trigger economic activity. Therefore, the per capita income of the people who live in the corresponding city increases, which in turn increases the demand for airlines. The people who have higher per capita income tend to buy business- or first-class air fares. Therefore, average air fare levels increase in the markets that have higher per capita income.

\[ X_{4b} = \text{Per capita income of endpoint 2}. \] This factor represented per capita income of destination city in a city-pair market, and its relationship with average air fare was positive: for every $10,000-increase in per capita income at Endpoint 2, average air fare levels on average increased by $4.80. Thus, it appears that higher per capita income of destination cities result in higher average air fare levels. A plausible explanation for this finding is exactly the same as the one given for the previous finding.
$X_9 = \text{Market size.}$ This factor represented total available seats of all incumbent airlines in a city-pair market, and its relationship with average air fare was negative: for every 100-unit increase in the total number of available seats in a market, air fares decreased on average by 55 cents. Thus, it appears that higher market size of a city-pair market results in lower average air fare levels.

A plausible explanation for this finding is totally related to the basics of economic theory. The essential rule of a liberal economy—that price is the crossing point of supply and demand—is also relevant to airlines. It should be noted here that in the early years of commercial aviation, most governments took an active role in regulating air fare levels and capacity planning of airlines in most regions of the world. This meant that the airline industry was not a free market economy. In this manner, available seats or capacity of airlines controlled and kept at specific levels. After the Deregulation Act of the airline industry in the U.S. at the end of the 1970s, all rules of the game were defined from the beginning. After deregulation, the competition level within the airline industry suddenly increased. Airlines began to fly the same routes. The side effect of this was extremely high level of capacity because of high level of competition in some of the city-pair markets. Therefore, because of increasing supply the average air fare levels in some of the city-pair markets and profit margins of airlines decreased. It is obvious there is competitive pressure for the airlines to be assertive when regarding capacity increase. Therefore, airlines set their capacity too high and allocate capacity to
routes that hardly regain the airlines’ costs, much less earn them high profit margins (IATA, 2011).

$X_{11} = \text{Largest fare.}$ This factor represented the largest average air fare in a city pair market, and its relationship with average air fare was positive: for every $1-$increase in the most expensive air fare reported for each route among the competing airlines, air fares increased on average by 77 cents. Thus, it appears that higher largest fare in a city-pair market results in higher average air fare levels. A plausible explanation for this finding is related to market dynamics. If the average air fare level is high in a city-pair market, this is because of the high average air fare of incumbent airlines in the city-pair market. In some of the city-pair markets, there is only one incumbent airline. If there is no competitor in a city-pair market, then because of low level of competition the incumbent airline will have more room to increase the average air fare level to maximize profit margin.

$X_{13} = \text{Business model of largest fare.}$ This factor represented the business model of largest average air fare in a city-pair market, and its relationship with average air fare was positive: when the business model of the airline (LCC vs. FSC) with the largest fare for each route is compared relative to overall air fares, the overall air fare of LCCs was $15.23$ more than the overall air fare of FSCs in the presence of all of the other factors.

A plausible explanation for this finding is related to the market conditions of each of the cases. If the business model of the largest fare airline in a route
market is LCC, then this might mean there is a propensity of high air fare levels in the market. The tendency of the airlines toward higher prices changed the market dynamics and LCCs also serve with high prices.

\[ X_{10a-10e} = \text{Market share of competing airlines.} \]  These five factors represented the market share of Alaska, American, Delta, SkyWest, and United Airlines, respectively, in a city-pair market, and their relationships with average air fare were positive: for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to the five competing airlines, air fares increased on average by $23.98, $10.73, $6.30, $11.53, and $6.31, respectively. A plausible explanation for these findings is related to the previous explanations about the FSC business model. The average air fare levels of FSCs are higher than the average air fare levels of LCCs. All five competing airlines are full-service carriers.

The reader will note the absence of any low-cost carriers from the final model as well as the absence of the remaining 10 of the initial set of 15 airlines targeted for Part B (AirTran, Allegiant, Frontier, Hawaiian, JetBlue, Northwest, Southwest, Spirit, US Airways, and Virgin). This is because these airlines were eliminated from the final dataset during preliminary data screening. When included in the dataset, these airlines generated “singularity” error, which indicated they were providing redundant information with other variables, and/or had high variable inflation factor (VIF) scores. After examining various models with
involving these airlines, the best model was the one that resulted with the remaining five airlines.

However, a comparable model, which included the absence of exactly the same outliers, also was significant, $R^2 = .95, F(16, 899), p < .0001$. Of the 15 competing airlines, seven remained in the model after preliminary data screening: American, JetBlue, Frontier, AirTran, Spirit, US Airways, and Southwest. The reader will note that all these airlines except for American and US Airways are LCCs. What is interesting about this model is that not only were the regression coefficients for the other factors similar to those of the final model reported here, but there was a negative relationship between the LCCs’ market share and air fare levels. More specifically, for every 1-unit increase in the ratio between available seats to the total number of available seats in the market relative to these competing airlines, air fares decreased on average by $11.07, $32.51, $19.22, $35.22, and $6.27, respectively, for JetBlue, Frontier, AirTran, Spirit, and Southwest.

**Implications**

The implications of the results of the current study relative to the theoretical grounding, prior research, and aviation practice are examined below.

**Implications relative to theory.** The current study was grounded in two theories. Part A (entry-exit decisions) was grounded in industrial organization (IO) theory, and Part B (air fare dynamics) was grounded in economic (demand) theory.
These theories are summarized below along with a discussion of the implications of the study’s findings relative to each of these theories.

**Industrial organization theory.** Part A of the current study was grounded in industrial organization theory. This theory is a branch of economics and the main focus areas are an organization’s strategic behavior and market competition. Throughout the literature, IO theory was mostly used to understand how markets work. In the context of the current study, IO theory was used to understand entry and exit strategies of airlines in the U.S. domestic airline market. For this aim, Bain’s (1956) model was used to explain the market structures that are related to barriers to entry. Bain’s model states that barriers to entry result in incumbent firms’ earning very high profits without threat of entry from competitors because of economies of scale, absolute cost advantages, product-differentiation advantages, and capital requirements. Bain’s theory clarifies that all these market specialties make the incumbent firm more powerful and prevents competitor’s entering to the market or forces competitors to exit from the market. Briefly, having economies of scale advantage, cost advantage, disruptive or innovative business model, or capital requirement advantage makes an airline more powerful in a market. Besides that, the existing competitors in the market generally cannot sustain being in the market and therefore exit from the market. These factors also could lead to other potential competitors being discouraged from entering a market. Thus, having economies of scale advantage increase the probability of an airline sustaining itself in a route.
market. In the context of the current study, these economies of scale would be assisted by: a long route market, high population or high per capita income of origin and/or destination points of a route market, having absolute cost advantage via the presence of a hub airport at an origin or destination, or having a product differentiation advantage with the help of airline being a LCC.

In the current study, I posited that the dynamics of IO theory would shape the entry and exit strategies in the airline industry. The findings of the current study, however, partially supported this theory. The relationship between route length, city population, and per capita income, and the likelihood of an airline entering a route when compared to the likelihood of an airline exiting from an existing route was supported by the data. The results showed that an airline’s likelihood of entering a route was higher if the route length, per capita income, and population of the city-pair market also were higher. According to Bain (1956) these factors make it easier for a firm to sustain itself in a market, and therefore it made sense for an airline to want to enter the market.

However, the current study also revealed that the presence of a hub airport at an origin or destination for an airline did not have a significant impact on an airline’s exit or entry strategies. This was contrary to IO theory because according to Bain the presence of a hub airport should provide an absolute cost advantage, which should increase the likelihood of the airline’s presence in a particular route market. This was not the case, though, as the results showed an nonsignificant
relationship between having a hub and the decision to enter or exit a route. The findings of the current study relative to product differentiation advantage also did not support IO theory. The overall model’s results revealed that FSCs were more likely to enter a route than LCCs, which implies that LCCs’ product differentiation advantage did not help them to survive better than FSCs as expected.

**Demand theory.** Part B of the current study was grounded in demand theory. Throughout the current study, the interest area of demand theory, which describes the relationship between demand and price, was used to clarify the determinants of airlines’ air fare pricing strategies. Demand theory is a branch of economics that examines the relation between the demand for goods or services and prices of these goods or services. Demand theory considers various factors that could have impact on demand level. These factors include but are not limited to income level, substitutes of the goods or services, competition level, complements, preferences, and population. Demand theory focuses on how these factors affecting demand level also impact price levels. The law of demand states there is an inverse relation between demand and price levels.

In the context of the current study, if an airline increases price levels in a route market then the demand for the flights of the airline in that route should decrease. This implies there should be a strong connection between air fare levels in route markets and demand theory. Demand theory can be used to clarify the air fare level setting strategies of airlines. In general, economists use demand curves to
visualize the relationship between demand and price level. There are multiple factors affecting demand levels and also price levels. Specifically for the airline industry market, these would include size, business model of incumbent airlines, route length, number of competitors, market share of competitors, population of origin and destination cities, and per capita income of origin and destination cities. All these factors should have an impact on demand levels and also on air fare levels.

The findings of the current study mostly supported demand theory. As expected, the relationships between all of the targeted factors and the average air fare level in a route market were significant and positive. These factors included: route length, city population of origin and destination cities, per capita income of both origin and destination cities, air fare of airline with the largest market share, business model of the airline with the largest air fare level, and the market shares of Alaska, American, Delta, SkyWest, and United Airlines in a route market. It is self-evident that as route lengths increase, so too do expectations and needs so the price levels increase as was shown in the simultaneous model. It also is self-evident that as the city population of an origin city and the per capita income level of both origin and destination cities increase, the demand for the flights and therefore, average air fare levels also will increase. The reader also is reminded that the business model of the airlines with the largest fare was FSC. Thus, by increasing the level of the largest air fare level in a route market and increasing the market
shares of the five FSC airlines—which are the largest FSCs in U.S. domestic airline market—would also increase the average price levels because of the increase in needs and expectations. For example, FSCs serve their customers with a high-quality service level and one of the unique differences between FSCs and LCCs is FSCs having business or first-class compartments. This implies that FSCs’ existence in a route market increases the expectations and needs of the marketplace, and therefore the price levels also increase in these markets. What’s more, the relationship between market size, which specifically represents supply level in the context of the current study, was significant and negative as expected. As a rule of thumb if the supply level increases when demand is constant, then price levels decrease. The results of the study showed that increasing market size decreased the price levels.

**Implications relative to prior research.** In this section, the findings of the current study were compared with the findings of the published literature. Prior research studies were analyzed in Chapter 2 in two parts. Part A included the review of studies related with route entry and exit decisions. Part B included the review of studies related with air fare competition dynamics. In this section, the findings of the current study are compared to the previous studies in two subsections that correspond to each part of the study. It should be noted that the comparison of the current study’s findings to that of prior research was limited because of the seminal nature of the current study. Previous literature mostly
focused on entry or exit strategies of airlines separately whereas the current study examined entry and exit cases in the same model and also included a comparison between FSCs and LCCs. On the other hand, some of the determinants of air fare levels were not considered in most of the previous studies relevant to air fare competition dynamics. Although the current study benefited from different aspects of past studies that were included in the literature review section, there are no past similar studies to the current one to make comprehensive comparisons.

**Route entry and exit decision patterns.** Similar to the findings of the current study, Oliveira (2008) detected that route density, route distance, and rivals’ presence in a route were significant factors in explaining the entry pattern of a low-cost carrier, Gol Airlines, in the Brazilian domestic market in 2001. The current study also revealed that route length, having more than one competitor in a route market, per capita income of Endpoint 1 of the route market, and population of Endpoint 1 of the route market were significant factors in explaining the difference between entry and exit strategies of LCCs and FSCs. Although the current study did not include route density as a parameter, in the context of current study endpoint population or income level represented the same construct as route density in terms of explaining demand level. Oliveira’s study also analyzed a LCC in the Brazilian domestic market. That means the sample size of Oliveira’s study was very limited. It also should be noted there are strong similarities between market dynamics in North America and South America. Therefore, it was not surprising to
detect similar factors as significant factors in explaining entry or exit decision patterns.

The findings of the current study also were, to a degree, consistent with Boguslaski et al.’s (2004) study on entry patterns in Southwest Airlines route system. Boguslaski et al. (2004) revealed that (a) passenger density, (b) route distance, (c) the hubs of Southwest’s competitors, and (d) per capita income at the endpoints of a market were important factors to predict the entry strategies of Southwest Airlines. The current study also revealed that route length, per capita income, and population of origin city of the route market were significant factors in explaining the difference between entry and exit strategies of LCCs and FSCs. The main difference between the results was relevant to passenger density, income level of destination cities, and the existence of hubs. The inconsistency between these results was because of the difference of the nature of the current study and the study of Boguslaski et al. The latter study analyzed the entry patterns of Southwest Airlines, however, the current study examined the distinction between LCCs and FSCs groupings in terms of entry and exit decision patterns. On the other hand, Boguslaski et al. used probit for each of the years from 1990 to 2000, separately. However, the number of entries of Southwest to a route in some of these years was not sufficient to use a probit analysis. The most important lesson that the current study learned from Boguslaski et al. was relevant to that issue. The current study did not analyze the entry and exit patterns of airlines separately. The current study
used LCC and FSC groupings to meet the number of entry and exit requirements of probit model and included a sufficiently large sample size.

On the other hand, there was an important contradiction between the findings of the current study and those of Lederman and Januszewski (2003), which analyzed the entry patterns of LCCs. Lederman and Januszewski reported there was convincing evidence that a carrier’s own presence at the endpoints of a route increased the likelihood of entry. Existence of hub airport as a parameter for the current study was not same but similar to the parameter that represented a carrier’s own presence at the endpoints of a route. The findings of the current study revealed that the existence of a hub airport at any endpoint of a route market was not a significant factor in explaining entry or exit strategies of airlines.

In one of the other studies, Ito and Lee (2003) examined factors that influenced entry strategies of LCCs in the U.S. airline industry from 1990 to 2002. The findings of the current study were, to a degree, consistent with Ito and Lee’s findings. Ito and Lee reported that market density, distance, end point city populations and income, and the existence of a hub airport at any endpoint of the route market were significant factors that influenced LCC entry. The findings of the current study were similar to those of Ito and Lee in terms of the impact of market density, distance, population, and income level of origin cities on entry or exit strategies. Moreover, similar Ito and Lee, the current study also showed that the existence of a hub with multiple carriers at one of the endpoints of the route and
route concentration were not significant factors affecting LCC entry. The current study also revealed that the HHI index, which represented market concentration, and the existence of a hub airport were not significant factors in explaining entry and exit strategies.

There also were important contradictions between the findings of the two studies. For instance, Ito and Lee’s (2003) study emphasized that existence of a hub airport of a LCC or FSC and destination population and income were important factors that influenced entry and exit strategies. However, the findings of the current study revealed that the existence of a hub airport and population and income of destination cities were not significant in explaining entry and exit patterns. As a disadvantage of the model of Ito and Lee, it should be noted that probit models need a minimum number of entries to gather meaningful results, but the number of entries in the U.S. domestic airlines during the study period of Ito and Lee was very low. To compensate for this, Ito and Lee combined all entry cases and the most important problem with the study of Ito and Lee emerged because of that. Most of the entries during the study period of Ito and Lee were from Southwest Airlines and the results of their study were biased. Therefore, the findings of their study were biased toward Southwest entry cases, which makes it difficult to fully compare to the current study.

Air fare competition dynamics. The findings of the current study were, to a degree, consistent with the two studies of Vowles (2000, 2006). In one of these
studies, Vowles (2000) analyzed determinants affecting average air fare levels of low-cost carriers (LCCs) in the U.S. by focusing on geographic and competition issues. Separate from this study, Vowles (2006) also examined factors affecting air fare pricing in hub-to-hub markets. In both these studies route length, market size, and market share of the largest fare carrier were significant factors in explaining average air fare levels in route markets. This was exactly the case for the current study as well. In the current study and both of Vowles’ studies, increasing route length and largest fare airline’s being a FSC increased the average air fare levels, but increasing market size decreased average air fare levels. There is an important reason behind the differences between Vowles’ studies and the current study. For example, it should be noted that both of Vowles’ studies focused only on LCCs. Moreover, the samples of both of Vowles’ studies were very limited in size. Vowles’ (2000) study was related to the Q1 1997, and Vowles’ (2006) study only analyzed 185 hub-to-hub route markets. One advantage of current study to both of Vowles’ studies was that the current study focused on most recent data set and employed a very large sample size.

The findings of the current study also were very similar to the findings of Zhang et al. (2013) who examined the determinants of air fare levels of full service carriers in U.S. hub-to-hub markets. Both studies showed that with increasing market size, average air fare levels in route markets decreased, and with increasing route length, average air fare levels in route markets increased.
Furthermore, Brueckner et al. (2013) proposed a new model of air fare pricing determinants by using an adjacent airport approach and competitive effects of both LCCs and FSCs. Similar to Brueckner et al., the current study considered the existence of LCC or FSC incumbents on average air fare levels in airline route markets. There were two differences between the approaches of the current study and Brueckner et al.’s study, however. Brueckner et al.’s study analyzed the impact of LCCs and FSCs on average air fare levels by the help of their existence in a route market. The current study, though, analyzed the impact of market shares of major FSCs Alaska, American, Delta Airlines, SkyWest, and United Airlines on average air fare levels in airline route markets. The findings of the current findings about the impact of FSCs supported those of Brueckner et al.’s. According to Brueckner et al., the existence of FSCs in a route market increased the average air fare levels, and this is exactly what was found in the current study relative to the five competing FSC airlines. Although the findings of these two studies were similar, it should be noted that the Brueckner et al.’s sample included the last two quarters of 2007 and the first two quarters of 2008. This study period effectively was a subset of the sample used in the current study, which included the domestic flights between 2005 and 2015.

**Implications relative to practice.** Understanding what the results of current study mean for practice within the aviation industry requires some information about the air transport management in U.S. domestic market. The
reader should note that the U.S. government at the end of the 1970s deregulated the airline industry, which became a starting point for a new era for the airline industry. Before deregulation, entry and exit strategies of airlines and air fare setting strategies of airlines were regulated by government. For example, it was impossible to enter a route without the permission of government, and the government regulated ticket prices. With deregulation, which liberalized the airline industry, rivalry within the airline industry became intense. It turned out to be ordinary for many airlines to fly the same routes, and airlines started experimenting with different business models such as LCCs and innovative strategies and services in an effort to gain a competitive edge.

Especially after deregulation, the two most important departments within the organizational structure of airlines were the network and capacity planning department and the yield management department. Relevant to the current study, the network and capacity planning department of an airline develops entry and exit strategies of an airline, and the yield management department works on air fare setting strategies. These two departments are the owners of the effort to produce high profit returns because the two essential parameters of profitability, namely, supply and demand, are formed by the choices of these departments. Briefly, network and capacity planning departments define the supply of airline resources by entering into a new route or by exiting from an existing route, and yield management departments try to maximize the total revenue by predicting the
demand and setting optimal air fare levels. Thus, “airlines are constantly tweaking their schedules, trying to find profitable new routes or pulling the plug on ones that have underperformed” (Mutzabaugh, 2017, para. 1). For example, in January 2017, American Airlines announced a shuttle operation at Chicago O'Hare, Delta Airlines opened three new routes from Seattle, Icelandair opened a nonstop route between Kansas City and Reykjavik making it the first regularly scheduled trans-Atlantic route in the Midwest, and Allegiant Airlines is adding two new routes to Charleston, SC and Sarasota, FL (Mutzabaugh, 2017).

The first part of the current study, which examined the entry and exit strategies of airlines, was mostly relevant to air transport management professionals who work at network and capacity planning department of airlines as well as and government authorities such as thee FAA, which regulates entry and exit decisions of airlines. The second part of the current study, which examined the air fare setting strategies of airlines, was mostly relevant to air transport management professionals who work at yield management departments of airlines and government authorities like FAA who also regulate air fare level setting policies of airlines.

Traditionally, by using various information technology systems and databases that include historical and recent data, yield management and network and capacity planning departments shape their strategies for many years. In recent years and similar to many other industries, the airline industry began widely using advanced statistical techniques, supervised machine learning methods like linear or
logistic regression, and artificial intelligence to increase the efficiency of their predictions. Thus, the first main implication of the current study to practice relates to the way in which the advanced statistical model built for detecting entry and exit strategies of airlines can be used as a reference to predict the potential behaviors of competitors. As explained above, the main purpose of a network and capacity planning department is drawing the network and route map of an airline. The duty of this department is strategic, and the decisions of that department shape the strategies of yield management department. Network and capacity planning departments continuously work on master capacity planning, which generally includes summer and winter capacity plans for most of the airlines. Efficient master capacity planning requires efficient prediction methodologies. The professionals of network and capacity planning departments analyze market conditions, forecast potential demand, and predict competitors’ strategies to shape their own strategies. The most vital aspect of master capacity planning is the prediction of competitors’ strategies. If a network and capacity planning professional can predict the future entry or exit strategies of competitors than he or she can maximize the efficiency of capacity planning decisions of an airline. Furthermore, by using the multiple regression model that was developed to understand the dynamics of air fare level setting strategies of airlines, a yield management professional may shape pricing and yield maximizations strategies. Think about a yield management professional working for American Airlines. She can use this multiple regression model to
predict average air fare levels while deciding about dynamic pricing strategies of American Airlines in the U.S. domestic market. Yield management is one of the most vital issues in airline industry. In general, yield management professionals working for an airline have designated flight routes, which are city-pair markets. Yield management professionals have two main duties. One is managing short-term dynamic pricing daily and/or weekly, and the other is long-term dynamic pricing monthly and/or seasonally. Mostly, daily and weekly dynamic pricing strategies are shaped by competitors’ pricing strategies and unexpected last-minute demand. However, budget targets, profitability targets, and general market conditions shape long-term dynamic pricing strategies. Airlines easily can use both the multiple regression model of the current study in conjunction with the logistic regression model of the current study to come identify monthly and/or seasonal dynamic pricing strategies. For example, by using the logistic regression model an airline professional may detect the potential total capacity changes in route markets of the airline. Then, by using the multiple regression model she can predict the average air fare level in the market. Therefore, it will be much easier to build long-term dynamic pricing strategies.

Another implications of the study’s results to practice relates to the way in which the LCC and FSC business models converge or diverge. There are ongoing discussions about this issue. For example, Ferrer-Rosell and Coenders (2017) reported that low-cost carriers are becoming less and less low-cost-like, and full-
service airlines are becoming less and less full-service-like. The findings of the current study do not support this observation, but instead found that FSCs were 1.68 times more likely than LCCs to enter a new route. The current study also found that for every 1% increase in the market share of Alaska, American, Delta, SkyWest, and United Airlines in a city pair market, average air fares in city pair market increased average air fare levels in the U.S. domestic airline market. Moreover, the business model of largest fare airline in a market being LCC increased the average air fare levels by $15.24. All these findings support the idea that there are still significant differences between LCCs and FSCs. Thus, airline professionals should still consider the distinction between these two business models when shaping their strategies.

The current study also found that airlines were 1.52 times more likely to enter routes of 850 miles or more than routes of fewer than 850 miles, and for every 10-mile increase in route length, average air fare level increased by 5 cents on average. These findings imply that increasing jet fuel prices has a potential to disturb the airline industry. These findings and their corresponding implications also are timely because they support the FAA’s recent working to empower the U.S. use of 1 billion gallons per year of “drop-in” sustainable alternative jet fuels by 2018. The FAA works on increasing the usage of sustainable alternative jet fuels created from renewable sources to minimize the harm of increasing jet fuel prices on commercial aviation and to solve problems relevant to environmental, and
energy security challenges that arise from petroleum based jet fuel use. Thus, these findings of the current study supported the motivation of the FAA to work sustainable alternative jet fuels.

**Generalizability, Limitations, and Delimitations**

**Generalizability.** The generalizability of the study’s results was examined from both a population and ecological perspective. In terms of population generalizability, the findings of the current study with respect to Part A are generalizable to the accessible population because the sample essentially was a census of this population. However, generalizing the results to the target population is problematic because the data source, the DB1B database, contains only a 10% sample of all airline ticket data. However, if these data are representative of the overall industry, then a case could be made that the results from the current study are generalizable to the target population. For example, the reader is reminded that the targeted airlines represented 85% of the total U.S. domestic airline market during the study period. As for Part B, because the data were randomly selected from a randomized database, the results are generalizable to both the accessible and target populations. Furthermore, the airlines targeted represented 70% of the total U.S. domestic airline market during the study period.

In terms of ecological generalizability, it will be difficult to generalize the current study’s results from both Part A and Part B to other markets outside the U.S. domestic market such as the European or Asian domestic markets, or even the
U.S. international market. This because each domestic airline market has its own unique characteristics that make it difficult to transcend. To help the reader assess the extent to which the current study’s results could be applied to other airline markets, I have provided detailed information in Chapters 3 and 4 of this dissertation about how the current study was implemented and how the data were analyzed.

**Study limitations and delimitations.** In this section, the limitations and delimitations from Chapter 1 are restated here as a convenience to the reader to help facilitate a transition to the next section that includes recommendations for future research relative to these limitations and delimitations.

**Limitations.** Limitations are conditions, events, or circumstances over which a researcher has no control that will limit the generalizability of the study’s results. The limitations of the current study were as follows:

1. **Data integrity.** The primary data sources for the current study was the U.S. Department of Transportation’s (2015b, 2015c) DB1B database and the T-100 domestic segment report from reporting carriers by Office of Airline Information of the Bureau of Transportation Statistics, and the U.S. Census Bureau. Because I had no control over how these data were reported, collected, and stored, overall data integrity is problematic. As a result, the findings of the current study were relative to the data that are available during the study period regardless of the data’s accuracy. This also means that if any changes are made to the database for data
correction purposes subsequent to the current study, then any future studies that replicate the current study might get different results.

2. **Historical events.** The time period of the data that I used for the current study might include various historical social events such as terrorist attacks, social conflicts, airline accidents, or meteorological events such as hurricanes and snowstorms. All of these could have had an influence on the air transportation demand, which in turn could have impacted airlines’ survival strategies during the current study’s time period. As a result, because I did not account for the presence of any historical events, similar studies that access the same data as the current study but partial out or account for historical events might get different results.

3. **Social trends.** Independent but similar to historical events, trend changes in social and/or cultural values or practices could have an influence on the air transportation market. For example, technological changes such as consumers’ use of smart phones and social media might alter the way in which airlines conduct their business, and how an airline responds to these social trends could impact its route entry decisions and/or air fare competition dynamics. As a result, because I did not account for the presence of any social trends, similar studies that access the same data as the current study but partial out or account for social trends might get different results.

4. **Market reputation of airlines.** Reputation is a very important factor for airlines. Many airlines use advertisements to increase reputation. On the other hand
sometimes the market reputations of airlines decrease because of the reasons such as ongoing delays and operational problems or an unusual event. For example, in 2008, Canadian musician Dave Carrols’ guitar was broken during a trip on United Airlines. He then composed a song named “United Breaks Guitars” and the song became a *YouTube* hit within a few days. The reputation of United Airlines suddenly was diminished and the demand for United Airlines decreased immediately. The stock price of United Airlines also fell 10% and the total cost for stakeholders of United Airlines was $180 million in a few days. A more recent example, which also involved United Airlines, was the involuntary removal of a passenger from United Express Flight 3411 on April 9, 2017, and the corresponding memorandum in which the CEO of United Airlines complemented the employees involved in the incident (Bacon & Mutzabaugh, 2017). The models to define survival strategies throughout the current study did not take into consideration the changes in the market reputation of airlines. As a result, similar studies that take an airline’s market reputation into consideration might get different results.

**Delimitations.** Delimitations are conditions, events, or circumstances that a researcher imposes to make the study feasible to implement but further limit the generalizability of the results. The delimitations of the current study are as follows:

1. **Study period.** The data collection period for the current study was limited to the 5-year period 2011–2015 for Part A, and the 11-year period 2005–2015 for
Part B. Therefore, similar studies that use a different data collection period might not get the same results.

2. **Targeted sample.** The sample of the current study was limited to the top eight airlines in the U.S. domestic market in terms of market share for Part A, and the top 1,000 airline markets in the U.S. domestic airline market for Part B. As a result, similar studies that focus on a different sample might get different results.

3. **Data source.** The source of the current study’s data included the U.S. Department of Transportation’s (2015b, 2015c) DB1B database and the T-100 domestic segment report from the Bureau of Transportation Statistics, and the U.S. Census Bureau. As a result, studies that use different data sources might not get the same results.

4. **Use of outliers.** As described in Chapter 4 as part of the data screening discussion, outliers flagged by Jackknife distances were included in the data set that was used for the analysis in Part A, but excluded from the data set that was used for the analysis in Part B. As a result, similar studies that use a different outlier analysis strategy and include or exclude outliers differently than what was done in the current study might get different results.

5. **Creation of dichotomies.** As described in Chapter 4 (Table 4.6) the current study created dichotomies for all continuous factors to facilitate interpretations of the results, and several different strategies were used as the basis for creating these dichotomies. As a result, similar studies that do not use the same
strategies for creating dichotomies, or studies that elect not to create dichotomies for continuous factors, might not get the same results.

Recommendations for Future Research and Practice

The purpose of the current study was twofold: to determine what factors distinguished between airlines entering a new route or exiting from an existing route, and to determine what factors were related to air fare levels. In earlier sections of this chapter, I presented and discussed inferences and implications about the findings, and I restated the study’s limitations and delimitations from Chapter 1. In this section, I present a numerical listing of recommendations for future research relative to the study’s limitations, delimitations, and implications. I then conclude this section with a set of recommendations for practice relative to the study’s implications.

**Recommendations for future research relative to study limitations.**

Following is a set of recommendations for future research based on the current study’s limitations.

1. I had no control over the integrity of the data stored in the primary data sources used for the current study. Therefore, a recommendation for future research is to replicate the study with the same statistical models and parameters but use data from other third-party data sources such as air travel intelligence companies like OAG, Flight Stats, or Routes Online. Alternatively, subsequent
studies could attempt to acquire the data directly from the airlines, although this might be hard to obtain.

2. The time periods of the current study was between 2011 and 2015 for Part A, and between 2005 and 2015 for Part B. Both periods are very long and include various historical social events such as Hurricane Katrina in 2005, the economic crisis of 2008, and Ebola disease in 2014. Therefore, a recommendation for future research relative to this limitation is to replicate the study during a time period that is absent any historical events that could impact the primary focus of the current study.

3. The manner in which prospective airline passengers use smart phones and social media has been instrumental in forcing airlines to review their business practices. As a result, a recommendation for future research is to incorporate factors related to the use of social media and contemporary technologies such as smart phones to measure the effect they have on airlines’ survival strategies of route entry-exit decisions and air fare levels.

4. Sometimes the market reputations of airlines decrease because of reasons such as ongoing delays, operational problems, or an unusual event. A good example of this is the United Airlines debacle in which a passenger was forcibly removed from a flight (Bacon & Mutzabaugh, 2017). Events such as this could damage the reputation of an airline and lead to reduction in seat demand. As a result, a recommendation for future research is to include airlines’ reputation as
a factor to determine its impact on route entry-exit decisions and air fare levels. One suggestion is to consult companies such as SKYTRAX, which are compile airline ratings, to acquire the necessary data.

**Recommendations for future research relative to study delimitations.**

Following is a set of recommendations for future research based on the current study’s delimitations.

1. Data collection for the current study was restricted to the 5-year period 2011-2015 for Part A, and the 11-year period 2005–2015 for Part B. Therefore, it is recommended that future research focus on the most recent period beginning after the end of 2015 to reveal the most recent impacts of the determinants of the models.

2. The current study was delimited to a sample of the top eight airlines in terms of market share for Part A and the top 1,000 airline markets in the U.S. domestic airline market for Part B. The study also focused exclusively on LCCs and FSCs in terms of business models. However, there are other types of business models such as regional and charter airlines. Although the total market shares of these airlines are not as high as the U.S. domestic airline market, it may be beneficial to include these business models in the targeted sample. Therefore, a recommendation for future research is to consider these business models.

3. The data sources used for the current study were restricted to the U.S. Department of Transportation’s (2015b, 2015c) DB1B database and the T-100
domestic segment report from the Bureau of Transportation Statistics, and the U.S. Census Bureau. Therefore, it is recommended that future research try to acquire data from other sources. Suggestions include (a) air travel intelligence platforms like OAG and FlightStats, (b) alliances like Oneworld, Star, or SkyTeam, and (c) authorities such as IATA and ICAO. A more challenging data source is the targeted airlines.

4. As part of preliminary data screening, outlier analyses were conducted using Jackknife distances, and the results from these analyses flagged 151 outliers in Part A and 166 outliers in Part B. Based on follow-up analyses, the outliers in Part A remained in the final data set but the outliers in Part B were excluded from the final data set. As a result, a recommendation for future research is to (a) use a different outlier analysis strategy (e.g., Mahalanobis distances, $T^2$, univariate box plots, or data that falls outside $\pm 3 SD$ for normally distributed data), and (b) examine the extent to which the outliers impact the results.

5. For Part A, all continuous variables were transformed into dichotomies to facilitate interpreting the logistic regression results. As result, a recommendation for future research is to use different strategies than what was used in the current study to create these dichotomies.

**Recommendations for future research relative to implications.** Specific recommendations for future research based on the implications to prior research as well as theory are discussed as follows:
1. One of the significant implications of the current study was related to the finding that showed that an airline’s propensity of exiting from a route is higher if the route length, per capita income, and population of the city-pair market are higher. Another finding revealed that origin’s or destination’s being the hub airport of an airline did not have a significant impact on the airline’s exit or entry strategies. Therefore, it is recommended that future research focus on determining why higher route lengths, per capita income, and population of the city-pair market increase the propensity of exiting from a route. This result was expected because these characteristics are mostly the characteristics of FSCs, which have decreased capacity especially in recent years. Is there an inverse relation? Is the relationship because being a FSC requires these characteristics? Do these characteristics stimulate an airline’s exit from a route?

2. Having a hub airport in an origin or destination of a city-pair market, which means absolute cost advantage for an airline, did not have significant relationship with an airline exiting a route. This result was not an expected and was contrary to IO theory. Therefore, it is recommended that future research analyze why having a hub airport at the endpoints of a city-pair market, which is a great absolute cost advantage, did not significantly affect exit behaviors of airlines. Future research also might analyze the difference between exit and entry strategies of airlines in the markets including hub airports and in the markets not including hub airports.
3. One of the other findings of the current study revealed that the relationships between route length, city population of origin city, per capita income of both origin and destination cities, air fare of airline with the largest market share, business model of the airline with the largest air fare level, and market share of American Airlines in a route market and the average air fare level in a route market were significant and positive. Therefore, it is recommended that future research focus on detecting new factors affecting the average air fare levels positively. Only the relationship between market shares of Alaska Airlines, American Airlines, Delta Airlines, SkyWest Airlines, and United Airlines and the average air fare levels were examined throughout the current study. A researcher may also include the market shares of other major airlines such as Hawaiian, Virgin, and Southwest Airlines in the model and detect their correlation with average air fare levels.

**Recommendations for practice relative to implications.** In addition to recommendations for future research, following is a set of recommendations for practice based on the implications of the study:

1. The study’s findings strongly emphasize there are still significant distinctions between the FSC and LCC business models. Therefore, it is recommended that airline professionals consider this distinction when developing advanced statistical models to predict network growth strategies and air fare strategies of competitors.
2. The current study’s findings strongly suggest that increasing fuel prices disrupt the airline industry. The propensity of airlines exiting from a market and average air fare levels in markets increase with increasing route length. It is recommended that FAA, aircraft manufacturers, and the airlines themselves focus on projects to increase the penetration of sustainable alternative jet fuels in the U.S. domestic market.

3. The current study’s findings suggest that increasing the population and per capita income of endpoints of a route market both increase average air fare levels and the propensity of airlines exiting from route markets. Therefore, it is recommended that airlines might consider expanding their route maps in rural markets instead of major cities or metropolitan areas.

**Final Comments and Observations**

Although the results of the current study were consistent with what was expected, the results were based on legacy data and assumed statistical stationarity, which presumes that statistical properties such as the mean and variance are constant over time. However, the reader will note from Chapter 3 there was the possibility of a history threat, which suggests the data represented a complex system filled with changing distributions. This concept of changing distributions is supported from the following observations of Tables 4.1–4.5: (a) the standard deviations are 50% or more of the mean, which indicates extremely high variability in the reported results; and (b) in nearly every instance, the mean was either far to
the left or right of the median, which indicates extreme skewness. When considered collectively, these are indicators of a complex system, which is comprised of four parameters: diversity, connectiveness, interdependency, and robustness. Therefore, future research involving these data might benefit by approaching such studies from a complex system perspective and manipulating all four corresponding parameters.
References


176


Appendix A

IRB Documentation
Notice of Exempt Review Status

From: Florida Tech Institutional Review Board
      FA00014319, IRB00001690

To: Selcuk Baran

Date: April 4, 2016

IRB Number: 16-089

Study Title: Survival in the US domestic airline market: Strategies for entry, exit and air fare competition

Dear Researcher:

Your research protocol was reviewed and approved by the IRB Chairperson. Per federal regulations, 45 CFR 46.101, your study has been determined to be minimal risk for human subjects and exempt from 45 CFR 46 federal regulations and further IRB review or renewal unless you change the protocol or add the use of participant identifiers. This study is approved for one year from the above date. If data collection continues past this date, a Continuing Review Form must be submitted.

All data, which may include signed consent form documents, must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Access to data is limited to authorized individuals listed as key study personnel.

The category for which exempt status has been determined for this protocol is as follows:

1) Research involving the collection or study of existing data, documents, records, or specimens if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, indirectly or through identifiers linked to the subjects.
RESEARCH INVOLVING HUMAN PARTICIPANTS
Exempt Application

This form shall be used if there is minimal risk to human subjects and one or more of the conditions below apply: If there is more than minimal risk associated with the research (none of the conditions below apply) or if the research utilizes a special population (children, prisoners, institutionalized individuals, etc.), please use the expedited/full application form found on the IBS website.

You should consult the university's document "Principles, Policy, and Applicability for Research Involving Human Subjects" prior to completion of this form. Copies may be obtained from the Office of Sponsored Programs and on the IBS website.

IRB Contact Information:
Dr. Lisa Blackmon, IRB Chairperson
315x89
John Politanos, Associate Vice President for Research
3074-7318
joelpolitanos@fit.edu 374-7230

Investigator Information:

Title of Project: Survival In the U.S. Domestic Airlines Market: Strategies for Entry, Exit, and Air Fares Competition
Date of Submission: March 23, 2016
Expected Project Start Date: May 2016
Expected Project Duration: 1 month

Principal Investigator: Sedaer Ramaz
Title: PhD Student
Academic Unit: College of Aeronautics
Phone:
Email: sramaz2013@my.fit.edu

Co Investigator: 
Title:
Academic Unit:
Phone:
Email:

Co Investigator:
Title:
Academic Unit:
Phone:
Email:
Categories of Exempt Research

Researcher must choose one:

☐ 1) Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as:
   a. research on regular and special education instruction strategies, or
   b. research on the effectiveness of or the comparison among instruction techniques, curricula, or classroom management methods.

☐ 2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior unless:
   a. the subjects can be identified, directly or through identifiers linked to the subjects and
   b. any disclosure of information that the research could reasonably place the subject at risk of criminal or civil liability or be damaging to the subject's financial standing, employability, or reputation.

Note: This exemption does not apply to surveys, questionnaires, or interviews involving minors.

☐ 3) Research involving the use of educational tests, surveys, or interview procedures, or observation of public behavior if:
   a. the subjects are elected or appointed public officials or candidates for public office or
   b. the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter.

☐ 4) Research involving the collection or study of existing data, documents, records, or specimens if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

☐ 5) Research and demonstration projects that are conducted by or subject to the approval of Department or Agency heads and that are designed to study, evaluate, or otherwise examine:
   a. public benefit or service programs,
   b. procedures for obtaining benefits or services under those programs,
   c. possible changes in or alternatives to those programs or procedures, or
   d. possible changes in methods or levels of payment for benefits or services under those programs.

☐ 6) Tests and studies of educational achievement and consumer acceptance studies if:
   a. wholesome foods without additives are consumed or
   b. foods are consumed that contain food ingredients found to be safe by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture.

If any part of this study will be funded by an external funding source, you must note the funding source and award/solicitation number below:

N/A
1. List the objectives of the proposed project.

The purpose of the IS is to examine the survival strategies of U.S. domestic airlines relative to their peers. In the context of the proposed study, survival strategies are defined with respect to two basic parts. Part A corresponds to market risk and entry decision patterns and Part B corresponds to airfare competition dynamics. The study will be restricted to passenger airline routes that have both origin and destination airports within the borders of the United States, and will examine city-pair markets, which are specific markets in which at least one commercial airline is operating. Examples of city-pair markets are New York-Orlando, Miami-Atlanta, and Houston-Houston.

2. Describe the research project design/methodology. Discuss how you will conduct your study, and what measurement instruments you are using. Attach all research materials to this application. Please describe your study in enough detail so the IRB can identify what you are doing and why.

The proposed study will employ two different research methodologies. Part A: The methodology for the first part of the study is on post facts and the corresponding design is case-type. This methodology is appropriate because I will be determining the extent to which the targeted factors led to whether an airline decided to institute a route (the Yes group) or decided not to institute a route (the No group). Because I will be working with recorded data, the effect on the dependent variable, which is group membership and corresponds to the decision to establish or not establish a route, has already occurred. Part B: The methodology for the second part of the study is comparative and the corresponding design is both explanatory and descriptive. This methodology is appropriate because the focus of the second part of the study is to determine the relationship the targeted variables have with air fares.

3. Describe the characteristics of the participant population, including number, age, sex, and recruitment strategy (attach actual recruitment email text, recruitment flyers etc.).

The proposed study involves the collection of existing data stored in publicly available databases and will include flights, airline passengers, and market-related data. With respect to passenger data, all identifying information was removed before the data became available publicly and therefore it is impossible to identify the data source and it is impossible to describe the characteristics of the passengers. Because the proposed will use de-identified, publicly available data, it does not constitute human subjects research as defined at 45 CFR 46.102.

4. Describe any potential risks to the participants (physical, psychological, social, legal etc.) and assess their likelihood and seriousness. Describe steps that will be taken to mitigate each risk.

N/A (See Item 3)

5. Describe the procedures you will use to maintain the confidentiality and privacy of your research participants and project data. If video or audio recordings will be made, you must review the video/audio recording policy found on the IRB website and address precautions you will take in this section.

N/A (See Item 3)

6. Describe your plan for informed consent (attach proposed form).

N/A (See Item 3)
Signature Assurance

I understand Florida Institute of Technology's policy concerning research involving human participants and I agree:

1. To accept responsibility for the scientific and ethical conduct of this research study.
2. To obtain prior approval from the Institutional Review Board before amending or altering the research protocol or implementing changes in the approved consent form.
3. To immediately report to the IRB any serious adverse reactions and/or unanticipated effects on participants which may occur as a result of this study.

PI Signature

Date 03/23/2016

Advisor Assurance: If primary investigator is a student
This is to certify that I have reviewed this research protocol and that I attest to the scientific merit of the study, the necessity for the use of human subjects in the study to the student's academic program, and the competency of the student to conduct the project.

Major Advisor  Michael A. Gallo
Major Advisor (print)  Michael A. Gallo

Date 3/25/2016

Academic Unit Head: It is the PI's responsibility to obtain this signature
This is to certify that I have reviewed this research protocol and that I attest to the scientific merit of this study and the competency of the investigator(s) to conduct the study.

Academic Unit Head

Date 3/28/16

FOR IRB USE ONLY

IRB Approval

Name

Date 4/4/16

IRB #
Appendix B

Raw Data
Given the voluminous amount of raw data, the data files for each part of the study proved too unwieldy to present here and occupied over 50 pages. As a result, readers who are interested in acquiring the raw data may contact the author directly: Selçuk Baran, Ph.D., sbaran1985@gmail.com.