Understanding Airport Leakage through Supply-and-Demand Interaction Models

by

Qian Fu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

TRANSPORTATION ENGINEERING

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Airport leakage is a phenomenon that occurs when air passengers choose to travel longer surface distances to take advantage of better air services at an airport further away (i.e., the substitute airport), instead of, as expected, using their local airport. The overall objective of this research is to investigate what factors affect airport leakage and how they affect airport leakage, in the context of models that consider the two-way interactions between air transportation demand and supply. More specifically, three categories of factors are investigated, including demographic, ground access, and air service factors. Two models have been explored in this regard. The first is a two-stage least squares model which is used to test the hypothesis that airport leakage occurs at 10 medium-size airports in the United States. It was found that the substitute airport, with lower airfare and higher enplanements, may attract passengers that would otherwise use their local, medium-size airport. In addition, passengers travelling in a group of three or more were shown to prefer their local airport even when the substitute airport provides lower airfare. It was also found that airports with higher traffic would attract more passengers. The second model explores the supply-demand equilibrium using a binary logit model to estimate the market shares of two competing (local and substitute) airports. A numerical analysis was performed to explore the sensitivity of equilibrium market share to coefficients, airfare, flight frequency and ground access distance. Results show that passengers will be attracted to the substitute airport to take advantage of lower airfare and higher flight frequency. If the substitute airport reduces its airfare, the airfare at the local airport will also be reduced. As a combination effect of the two airfares, the equilibrium market share changes. Furthermore, it was found that locations will have different market shares even if their ground access distances to the local airport are identical.

Acknowledgement

I would like to express my deepest gratitude to my advisor, Dr. Amy Kim, who has inspired me by her devotion and dedication to research. Without her insights, guidance, patience, and support, my thesis would have been impossible. She also helps me to be professional and to work effectively. I am fortunate to have her as my advisor.

I would like to thank Dr. Karim El-Basyouny and Dr. Ania Ulrich, for their time and guidance as my dissertation committee. I would like to thank Dr. Tony Qiu and Dr. Hui Zhang, for having me in one of their projects. I would also like to thank all the professors in my master's courses. Not only the knowledge, but I have also got a lot of training with respect to critical thinking, quantitative analysis, and academic writing. These courses have provided me a solid foundation for this thesis.

I would like to thank all my colleagues who are full of ideas. Thanks to Rokib SA, Lawrence Lan, and Ran Li, for their help and advice in the data processing. Thanks to Rajib Sikder, Xiaobin Wang, Lin Shao, Sudip Barua, Michael Ge, Cindy Wang, and Ariel Luo, for their help and care in the past two years. I would also like to thank my friend, Monika Ha Tran, for her help in the proofreading of this thesis.

Finally, special thanks to my family for their continuous encouragement, support, and love.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2SLS	Two-Stage Least Squares Model
FGLS	Feasible Generalized Least Squares Model
LCC	Low-cost Carrier
OD	Origin-Destination
GIS	Geographic Information System
ORD	Chicago's O'Hare Airport, IL
JAX	Jacksonville International Airport, FL
TUS	Tucson International Airport, AZ
GRR	Gerald R. Ford International Airport, MI
CAE	Columbia Metropolitan Airport, SC
PWM	Portland International Jetport, ME
BDL	Bradley International Airport, CT
CMH	Port Columbus International Airport, OH
CHS	Charleston International Airport, SC
CHA	Chattanooga Metropolitan Airport, TN
HSV	Huntsville International Airport, AL
PHL	Philadelphia International Airport, PA
SEA	Seattle–Tacoma International Airport, WA
TPA	Tampa International Airport, FL
LGA	LaGuardia Airport, NY
CLT	Charlotte Douglas International Airport, NC
DCA	Ronald Reagan Washington National Airport, VA
MCO	Orlando International Airport, FL
PHX	Phoenix Sky Harbor International Airport, AZ
BOS	Logan International Airport, MA
JFK	John F. Kennedy International Airport, NY
DTW	Detroit Metropolitan Wayne County Airport, MI
ATL	Hartsfield–Jackson Atlanta International Airport, GA

CHAPTER 1. INTRODUCTION

This chapter introduces and defines airport leakage. In the context of the findings and gaps in previous airport leakage studies, the research motivation, objective, and scope of this thesis are discussed. The last section contains an outline of the thesis.

1.1. Background

Airport passenger traffic has a huge impact on local economic development. It was predicted that 185,000 jobs would be created if Chicago's O'Hare Airport (ORD) expands and attracts 50% more passengers (Brueckner, 2003). Thus, a thorough understanding of airport passenger demand is important for urban planners and airport managers. Airport passenger demand has been studied extensively. It is mainly determined by factors in three major categories: demography, airport accessibility, and air services (Zou & Hansen, 2012a). Demographic factors include population and economy, and may cover trip purpose. Airport accessibility is related to location of the airport, ground access origin, the ground access mode, and the ground transportation network. Air services include airline services as well as airport services. Airlines determine the origin and destination airports they will serve, airfares, flight frequencies, and aircraft sizes, which greatly influence passenger demand (Pels, Nijkamp, & Rietveld, 2001; W. Wei & Hansen, 2005). In return, passenger demand for an airport also influences airline services at the airport (Wiltshire, 2013). The characteristics of airport services that impact airport demand include the number of airlines at the airport, customer parking, check-in and retailing services (Gupta, Vovsha, & Donnelly, 2008; Loo, 2008). When more than one airport is available to passengers, demand at one airport is not only impacted by its own air services but also by air services at alternative airports (Zou & Hansen, 2012a). There are some other factors of airport demand that do not belong to any of the three categories, such as deregulation (Ishutkina, 2009).

Airport leakage is a phenomenon that occurs when air passengers choose to travel longer surface distances to take advantage of better air services at an airport further away, instead of, as expected, using their local airport (Suzuki & Audino, 2003). Because airport leakage reduces the local airport's passenger demand, understanding airport leakage is also important for urban planners and airport managers, who attempt to attract more air passengers to the local airport and stimulate economic development. In this thesis, we will call the local airport as such, and call the "leakage" airport the substitute airport. In addition, we assume that the local airport is the only airport in its metropolitan region. This definition distinguishes airport leakage from airport competition in a multi-airport system, where more than one airport is located within a metropolitan region.

1.2. Research Background and Motivation

Our first research question asks whether air passengers that would otherwise use the local medium-size airport serving their metropolitan region leak to major hub (or substitute) airports outside their metropolitan region.

The second research question arises from the fact that most airport leakage studies focus on an airport's catchment area, the geographic service area of an airport. These studies emphasize on the market share distribution around airports, instead of each attribute that affects airport leakage (Fuellhart, 2007; Lieshout, 2012). In consequence, our second research question asks what factors affect airport leakage and how they affect this phenomenon.

In addition, a very limited number of airport leakage studies have accounted for the inherent interactions between supply and demand. A majority of airport leakage studies build discrete choice models based on survey data, and treat supply-side attributes as exogenous explanatory variables for demand (de Luca, 2012; Lian & Rønnevik, 2011; Suzuki, Crum, & Audino, 2003). A As a result, this research will consider supply-and-demand interaction in the study of airport leakage.

1.3. Research Objectives and Scope

The overall objective of this research is to investigate what factors affect airport leakage and how they affect airport leakage, in the context of models that consider the two-way interactions between air transportation demand and supply. More specifically, three categories of factors are investigated, including demographic, ground access, and air service factors.

To accomplish this objective, this research investigates the hypothesis that airport leakage exists when major hub (or substitute) airports provide better air services than medium-size airports. The hypothesis will be tested by assessing how attractive the substitute airport is to passengers who are assumed to use a local airport. If the air services at the substitute airport are shown to have a significant impact on the demand at the local airport, then we may conclude that airport leakage exists.

The research scope has been narrowed down by three considerations. Firstly, we only consider airport leakage from medium-size airports to major hub airports in the U.S.; and each airport is in a distinct metropolitan region. Secondly, passengers' airline choice is excluded from our research scope. All air services, such as airfare and flight frequency, are treated as airport services. Thirdly, all passengers are assumed to use private vehicles to go to the departure airport.

1.4. Structure of Thesis

There are five chapters in this thesis. Chapter 1 introduces the background of the research and gaps in previous studies, followed by the motivation, objective, and scope.

Chapter 2 provides a comprehensive literature review of the air transportation market with an emphasis on airport leakage. Three types of studies are reviewed: studies exploring the one-way impact of air services on air transportation demand, studies exploring the one-way impact of passengers on airfare or airline costs, and studies exploring the two-way interaction between passengers and air services. Models and methodologies that are most common in each of the three categories are discussed.

Chapter 3 explores the impacts of supply-side factors as well as substitute airport attributes on local airport demand. Two-stage least squares models have been specified to capture the endogeneity between airfare and airport passengers. This chapter can be divided into two parts. The first discusses data collection and processing. The second describes the estimation process and results.

Chapter 4 explores airport equilibrium market share using a binary logit model to estimate market share. The variables that are considered in the market share model include airfare, flight frequency, and ground access distance. The airfare variable is based on the airfare function from Chapter 3. A numerical analysis explores the sensitivities of variables and coefficients to airport market share at equilibrium. Chapter 5 provides an overview of the research and major conclusions. Research contributions, limitations, and recommendations for future work are also discussed.

CHAPTER 2. LITERATURE REVIEW

This section provides a review of airport competition studies with respect to study approaches. Previous studies are divided into three categories depending on whether the impact of air services (supply) on air transportation demand is considered, and whether the impact of airport or airline passengers (demand) on air transportation supply is considered. Models and methodologies that are most common in each of the three categories are discussed. Approaches that have been used to study airport leakage are emphasized.

2.1. Air Transportation Demand

There are extensive studies of air transportation demand (airport demand and airline demand) which only consider the one-way impact of supply on demand, and treat supply-side attributes as exogenous. Because we can hardly discuss airport competition without mentioning airline competition, studies that only focus on airline competition are also included. Two methodologies that have been used widely are discrete choice models and linear (or log-linear) regression models (Harvey, 1987; S. Hess, 2004; Hutchinson, 1993).

2.1.1. Discrete Choice Models

Discrete choice models can estimate the probability of choosing an airport among a set of alternative airports for an individual passenger, or they can estimate the market share of an airport among a set of competing airports. The first is considered a disaggregate choice model, and the second an aggregate market share model.

In disaggregate demand studies, passengers' airport choice behaviors are analyzed based on characteristics and attributes, which are specific for each individual and can be obtained through surveys. Two types of survey exist: revealed preference (RP) survey and stated preference (SP) survey. RP survey asks survey respondents about their past experiences regarding travel. SP surveys ask respondents about their choice behaviors in hypothetical situations. RP data reflect real situations but may not capture all factors while SP data is able to control variation but has a risk of underestimating attributes that are not available in the survey (Cherchi & de Dios Ortúzar, 2002; de Luca, 2012).

Discrete choice models have been used extensively in estimating airport choice in multi-airport systems, where more than one airport serves a metropolitan area. Most studies do not explore airport choice alone, but joint airport, airline, and ground access mode choices (S. Hess, 2004; S. Hess, 2005; Pels et al., 2001). The results of these studies vary significantly. Some studies find that airport choice is most heavily influenced by ground access or accessibility (Pels, Nijkamp, & Rietveld, 2003) while other studies find that air services attributes like airfare are important (Harvey, 1987). For passengers living in reasonable proximity to two or more airports, impact of access time is not high as flight frequency (Windle & Dresner, 1995). Using survey data from the San Francisco Bay Area, Harvey (1987) built a multinomial logit model, and found that ground access time, airline frequencies, and connections are significant for airport choice for both business and leisure travelers, with the first two variables in a non-linear relationship in airport utility function (Harvey, 1987). As a unique case for the New York Area, whether or not passengers have to make a river crossing to access an airport plays a role on airport choice (Gupta et al., 2008). In summary, significant variables of airport choice in multi-airport system include access time and distance, airfare, frequency, past experience, purpose, car ownership, air trip time, direct or indirect flight, delay, aircraft type, the number of airlines at one airport, and the number of members in travel group. Segmentation of travelers by trip purpose (business or leisure) is commonly done in these models. Different types of discrete choice model have also been applied and compared in previous studies. Hess and Polak applied mixed multinomial logit model for airport choice in the San Francisco Bay Area (S. Hess & Polak, 2005b). A comprehensive literature review of airport choice studies with respect to determinants, survey methods, and discrete choice models can be found in de Luca (2012).

Airport leakage happens more often for leisure travelers than business travelers, and that past experiences at an airport have a significant impact on passengers' airport choice (Suzuki et al., 2003). In a more recent study, joint airport-airline choice has been analyzed in a "two-step" decision process with the first step to screen out choice alternatives that can satisfy passenger's minimum acceptable standards and the second step to build a nested logit model (Suzuki, 2007). In the case of Des Moines International Airport (DSM) competing with Kansas City International Airport (MCI), Minneapolis-St. Paul International Airport (MSP), and Omaha Eppley International Airport (OMA), this modified model shows an improved fit for airline choice but not for airport choice (Suzuki, 2007). In southern Italy, airport choice behaviors among Naples-Capodichino (NAP), Rome Fiumicino (FCO), and Rome Ciampino (CIA) have been studied (de Luca, 2012). FCO and CIA are 20 miles away from each other, and both of them serve Rome. However, NAP serves Naples, and is located 150 miles and 130 miles away from FCO and CIA respectively. In 2013, FCO served about 36 million passengers as a hub airport while NAP and CIA served nearly 6 million and 5 million passengers respectively. In de Luca (2012), however, all of the three airports are treated in a multi-airport system (de Luca, 2012). Based on stated preference survey data, airport choices are analyzed in multinomial logit model, hierarchical logit model, cross nested logit model, and mixed multinomial logit model. It is found that significant factors for airport choice are access time, airfare, age, experience, and income (de Luca, 2012).

In Lieshout (2012), the market share is calculated based on multinomial logit model of airfare, flight frequency, ground access cost, and airside time (Lieshout, 2012). The study assumes that airport demand spreads out around the airport without ground access distance constraint, and areas with market share over 1% are called airport catchment area. It is found that the spatial distribution of airport market share varies with respect to destinations, air service offerings, and the number of competing airports. Understanding airport catchment area is instrumental to understand passengers' airport choice and the competitiveness of alternative airports (Lieshout, 2012).

2.1.2. Linear and Log-linear Regression Model

In studies of airport demand using linear or log-linear regression model, the dependent variable is usually passenger traffic or airport market share (Cohas, Belobaba, & Simpson, 1995; Hutchinson, 1993). The impact of airport or airline competition is reflected by variables of competitors. Canadian domestic air demand has been estimated in log-linear models (Hutchinson, 1993). Aggregate demand model of cross-sectional data is calculated by income at origin, income at destination, airfare, cost of substitute ground access mode, and travel time for the fastest surface mode. By introducing interaction variables or transforming variables, more effects of airport demand can be explored. In this study, interaction effects are counted by using the product of income at origin and income at destination, the ratio of airfare over ground access cost, and the ratio of air travel time over ground access time (Hutchinson, 1993). Improvement of using ratio variables is that air trip is considered comparatively with ground access trip (Hutchinson, 1993). Airport market share in multi-airport system is estimated in a log-linear model of airport dummy variables, the portion of frequency, the average airfare, and the airfare at competing airports (Cohas et al., 1995). Based on ticket-booking data, airport passenger traffic "leaking" from local airports to substitute hub airports is estimated in a two-step regression model. In the first step, the portion of "leakage" passengers is regressed on explanatory variables and time dummy variables. Both the portion of "leakage" passengers and explanatory variables vary with respect to time and routes. In the second step, the residual from the first-step model is regressed on explanatory variables that only vary with respect to routes. These variables include the average airfare from a local airport, the average airfare from a substitute hub airport, the airfare difference, the average flight time from the local airport to a destination, the driving distance between the local airport and the substitute airport, and the portion of available seats per day. Among them, the distance and seats variables are fixed in different time periods, and only vary with respect to routes (Phillips, Weatherford, Mason, & Kunce, 2005).

2.2. Air Transportation Supply

Studies of air transportation supply only consider the one-way impact of demand on supply, and treat demand-side attributes as exogenous. Supply represents airport and airline services, and demand represents airport and airline passengers. Airline service decisions indicate yield, pricing, and seat supply (Ippolito, 1981; Windle & Dresner, 1999; S. Zhang, Derudder, & Witlox, 2013). Because the decision-making of airline services involves assessment of airline costs, the studies of airline cost will also be mentioned. Airline costs consist of capacity cost, traffic-related cost, and overhead cost. In detail, they represent cost for fuel, employees' salary, maintenance, aircraft leasing and landing, advertisement, and administration (O'Connor, 2001). Methodologies in the studies of air transportation supply include linear regression model (Evans & Kessides, 1993; Windle & Dresner, 1999; S. Zhang et al., 2013), log-linear regression model (D. W. Gillen, Oum, & Tretheway, 1990; W. Wei & Hansen, 2003; Zou & Hansen, 2012b) , and other non-linear regression models (Hansen & Kanafani, 1989; Ippolito, 1981; Swan & Adler, 2006).

The passenger traffic variable in airline pricing models shows the one-way impact of airline passengers on supply (Evans & Kessides, 1993; Ippolito, 1981; Windle & Dresner, 1999; S. Zhang et al., 2013). Flight distance, vacation route, and flight connection are also important for pricing (Evans & Kessides, 1993; Windle & Dresner, 1999; S. Zhang et al., 2013). Other pricing variables are slot control, time trend, presence of low-cost carriers, indexes, and market share (Evans & Kessides, 1993; Windle & Dresner, 1999; S. Zhang et al., 2013). Seat supply is in a function of passengers, average airfare, carrier concentration at airport, commuter competition, airport departures, local carrier indicators, and eligibility of subsidy (Ippolito, 1981). Yield, which is the weighted average airfare, is regressed on distance, squared-distance, the product of population on two ends of the route, vacation dummy variable, slot control dummy variable, and quarter dummy variables (Windle & Dresner, 1999).

The variables in airline cost model include airline output, unit fuel price, labor price per employee, material price indicator, capital stock, load factor, stage length, delay, and the number of points served (Zou & Hansen, 2012b). In another study, several log-linear models of airline cost have been compared when using different variables (Hansen & Kanafani, 1989). These variables include quantity of labor, quantity of non-labor inputs, indicator of airline operating characteristics, network concentration, the number of points served, labor cost, trip length, load factor, aircraft seat capacity, and year dummy variables (Hansen & Kanafani, 1989).

It is found that for a specific flight distance, there is an optimal aircraft size that minimizes aircraft operating cost. The optimal aircraft size increases when flight distance increases. In addition, because larger aircraft size usually leads to higher pilot cost, the pilot cost variable is endogenous. By excluding the pilot cost variable in the model, the optimal aircraft size minimizing the aircraft cost becomes smaller (W. Wei & Hansen, 2003).

2.3. Air Transportation Demand and Supply

Studies of air transportation demand and supply refer to studies that have considered the two-way interaction between air services (supply) and air passengers (demand). The most common methodologies in these studies include two-stage least squares model, three-stage least squares model, mathematical optimization, game theory, and spatial competition model.

2.3.1. Two-stage and Three-stage Least Squares Models

The two-way interaction between airfare (supply) and air passengers (demand) can be represented by two simultaneous equations. Two-stage least squares (2SLS) and three-stage least squares (3SLS) are two estimation methods of simultaneous equations model. The 2SLS introduces an instrumental variable to replace the endogenous variable, which is correlated with the error term (Dougherty, 2011; Pindyck & Rubinfeld, 1998). More specifically for a demand model, the endogenous airfare variable is correlated with the error term. In solution, the 2SLS model replaces the endogenous variable with an instrumental variable. The endogenous airfare variable is estimated by the passenger variable and other exogenous variables in the first stage, and the predicted airfare variable (i.e., instrumental variable) is used in the second-stage demand model (Dougherty, 2011; Pindyck & Rubinfeld, 1998). The 3SLS model is based on 2SLS model but assumes that the error terms of simultaneous equations are correlated (Zellner & Theil, 1962).

Two-stage least squares model is built for 14 airports in the United States to capture the endogeneity of the supply-side and demand-side attributes to study airport leakage (Suzuki & Audino, 2003). It estimates the airfare in the first stage, and then uses the predicted airfare variable (instrumental variable) into the second-stage demand model. The variables in the first-stage airfare model include the route dummy variables, quarter dummy variables, flight legs, freight, passengers, the airfare at the substitute airport, and the interaction variable of driving distance and the airfare at the substitute airport. Besides the predicted airfare values, the second-stage airport passengers model is estimated by seasonality, the flight legs at the substitute airport, interaction effect of the flight legs at the substitute airport and the driving distance besides most of the variables in the first-stage model. Four models are compared in the study. Results show that the model is improved by using log-linear model form and by considering the substitute airport attributes. The interaction variable of airfare and driving distance also shows that air passengers may be attracted to a substitute airport that is 250 miles away (Suzuki & Audino, 2003).

Another two-stage least squares model has been used with the first-stage airline demand model and the second-stage seat supply model (Ippolito, 1981). In its log-linear airline demand model, variables include the number of flights, load factor, elasticity of flight frequency, squares of airfare, fare elasticity at mean fare level, distance, income, population, short-haul trip dummy variable, and three attractive-city dummy variables. The short-haul trip dummy variable is an implement of the distance variable to indicate possibility of car driving rather than air travel. On the log-linear seat supply model, variables are the fitted demand value from the first-stage model divided by enplanement, fare, ramp-to-ramp time, enplanement, carrier airport concentration, commuter competition, a dummy variable indicating whether airport departures is larger than 100,000, local carrier dummy variable, and subsidy dummy variable (Ippolito, 1981).

A three-stage least squares model has been built to explore the impact of competition from the United States-Canada transborder cities (Elwakil, Windle, & Dresner, 2013). On the supply side, the average airfare is regressed on log-form of passengers, log-form of great circle distance, an index, and year dummy variable. On the demand side, the number of passengers is in the log-linear model of variables including the average airfare, population in metropolitan area, the product of per capita incomes at origin and destination, year dummy variables, origin dummy variables, destination dummy variables, and border city dummy variables. Border city dummy variables are indicators of competitors (Elwakil et al., 2013). It concludes that the airfare difference is the major cause of airport leakage for the United States-Canada transborder market.

2.3.2. *Mathematical Optimization Studies of Airport and Airline Competition*

In airport or airline competition, mathematical optimization usually combines with game theory to explore the supply-demand equilibrium. Airport and airline competitors optimize their objectives under certain constraints. Only objectives and outputs (i.e., optimal solutions) of mathematical optimization will be reviewed.

Three optimization objectives exist in previous studies including profit maximization, welfare maximization, and cost minimization. Profit maximization is the most common for airline while both profit and welfare maximization are commonly used for airport (Barbot, 2009; Brueckner & Flores-Fillol, 2006; D. Gillen & Morrison, 2003; A. Zhang, Fu, & Yang, 2010). Profit equals to revenue minus cost. Airline revenue is the product of the number of passengers and airfare while airport revenue is divided into two parts: revenue from aviation operation, such as runways, aircraft landing and parking, terminals, and the revenue from commercial activities such as advertisement, car parking, and retailing. The commercial activities become increasingly important recently, thus, it is essential to have the two revenues in airport profit function (Barbot, 2009; D. Gillen & Morrison, 2003; A. Zhang et al., 2010; A. Zhang & Czerny, 2012). Welfare maximization represents social benefits maximization and is usually assumed to be the objective of publically funded airports. Welfare equals to airport tax revenue minus passenger costs, capital cost, and external cost (Pels, Nijkamp, & Rietveld, 1997; Pels, Nijkamp, & Rietveld, 1998). It is the sum of total utility and airline profit (Brueckner & Flores-Fillol, 2006). In a study (Adler, Pels, & Nash, 2010), social welfare contains environmental cost and out-of-pocket cost which is paid by government. Thirdly, airline cost includes passenger cost, flight cost and fixed cost while airport cost consists of capacity cost, passenger cost, airport operation cost, and external cost. When airline performance is taken into account, delay cost is also in its cost function (Hsu & Wu, 1997). In addition, airline's costs in fully connected and hub-and-spoke network are treated differently (Pels et al., 1997). The demand part in the profit can be linear demand function or market share in discrete choice models, as discussed in the first section of this chapter.

When airline and airport objectives are considered simultaneously, two-stage or three-stage models are adopted, with one stage to satisfy airline objective and another stage to satisfy airport objective. Usually, airport profit maximization is on the first stage, and airline profit maximization is on the second stage (Barbot, 2009; A. Zhang et al., 2010). Multi-stage model is also applicable to obtain airline network based on airline's optimal services and airport charges (Pels et al., 1997). For instance, the first stage is targeting at distance minimization to obtain an optimized network and the second stage is airline's profit maximization problem (Adler, 2001). In this model, airlines are able to determine their routes and whether to serve routes concurrently (Adler, 2001).

Outputs of optimization are the decisions that airlines and airports are making. For airlines, the outputs include airfare, frequency, and aircraft size. The radius of market size is also the output in a catchment area study (Hsu & Wu, 1997). For airports, the outputs are airports' charges (or taxes) to airlines. Other derivative outputs may include passengers' generalized cost, demand for each airline, aircraft size, flight operating cost, and the traffic/capacity ratio (Zou & Hansen, 2012a).

2.3.3. Other Airport and Airline Competition Studies

Compared to two-stage least squares (2SLS) models, three-stage least squares (3SLS) models, and mathematical optimization, the other methods used in airport competition studies are game theory and spatial competition model. As stated earlier, game theory studies usually combine with mathematical optimization method (Adler, 2001; W. Wei & Hansen, 2007; W. Wei, 2006). In this section, the various types of "game" in previous studies are reviewed.

Airline and airport competition mainly deals with how each "player" in the "game" makes decisions. Most of previous studies assume that players make decisions simultaneously and independently (Adler, 2001; Brueckner & Flores-Fillol, 2006; W. Wei, 2006; A. Zhang et al., 2010; Zou & Hansen, 2012a). For example, each airline makes airfare or frequency decisions to maximize its own profit across all available routes under conditions of knowing, partly knowing or not knowing competitors' information. There are also sequential game and accommodating game (Basso & Zhang, 2007; Hsu & Wen, 2003). Sequential game represents that decision marker makes decisions one by one; and accommodating game represents a phenomenon that one airline decreases flight frequency, and its competitor increases its flight frequency as a response to accommodate market from the other airline.

Wei and Hansen (2007) explored how duopoly airlines determine aircraft size and flight frequency in three game scenarios: one-shot simultaneous game, leader-and-follower Stackelberg game, and two-level hierarchical game. In the one-shot simultaneous game, airlines maximize their own profits by determining aircraft size and flight frequency at the same time. Airlines are assumed to have perfect information of their competitors' decisions. In the second game which is a leader-and-follower Stackelberg game, one airline makes a decision and then based on this, the other airline makes a decision. One airline in the game acts as a leader. In the two-level hierarchical game, two airlines determine their flight frequencies at the same time, and after knowing competitors' flight frequency decisions, airlines simultaneously determine their aircraft size decisions (W. Wei & Hansen, 2007).

Airport-airline collusion is a cooperative relationship between an airport and an airline in pursuit of larger objectives respectively or larger combined airport-airline objective. The objective may be profit or market share. There are studies that assume airlines at one airport provide the same air services; thus, airport-airline collusion in this condition reflects the decision power of airport for airline service attributes (Basso & Zhang, 2007; D. Gillen & Morrison, 2003). However, airport and airline may also decide to collude or not before making price decisions, as shown in a three-stage game (Barbot, 2009). However, Zhang et al (2010) derived a different conclusion from a two-stage competition model for airport-airline vertical cooperation focusing on the impact of revenue sharing. It was found that airport competition stimulates airport to cooperate with airlines, leading to a reduced joint profit but an increased social welfare (A. Zhang et al., 2010). Besides, Pels et al (1997) found there is no exact airport and airline equilibrium (Pels et al., 1997).

Spatial competition models have been well-studied, but their applications to airport competition are limited (Dmitry, 2012). The theorem of spatial competition model is that "transportation costs have the effect of creating different demand elasticities in spatially separated markets" (Fröhlich & Niemeier, 2011). As a pioneering study, the Hotelling model has been used to show how two airports in two locations compete with each other when they offer homogeneous services (Fröhlich & Niemeier, 2011). Airport catchment area, including the overlapping catchment area, depends on airport pricing, transportation cost and passengers' utility of taking advantage of air service. The underlying assumption of the model is that market is distributed evenly within the whole area. The baseline of airport's pricing decision is to prevent passengers to withdraw from the market. It was shown in the Hotelling model that if two airports are within a multi-airport system and passengers' costs to airports are low, the overlapping catchment area of the two airports will be large. If there are airport price differentiation and unit transportation cost differentiation for two competing airports, one airport will attract passengers from the hinterland of the other airport. For multi-airport systems like Greater London and the New York Area, even though primary airports mainly serve full service carriers and smaller airports serve low-cost carriers, the airfare in one airport would still be constrained to the airfare in the competing airport (Fröhlich & Niemeier, 2011). In addition, spatial competition model can also account for access time, delay, and cooperation or non-cooperation between airports (Basso & Zhang, 2007; Fröhlich & Niemeier, 2011).

2.4. Summary

In airport competition and demand studies that do not consider supply-and-demand interaction and treat supply-side attributes as exogenous, three types of models have been discussed including discrete choice models, linear and log-linear regression models. Discrete choice model has been used to estimate disaggregate airport choice or aggregate market share for multi-airport system and for airport leakage (Hansen, 1995; Harvey, 1987; Lieshout, 2012). The basis of discrete choice model is passengers' utility maximization. Discrete choice model has also been applied in combination with the geographic information system (GIS) to study airport catchment area (Fuellhart, 2007; Lieshout, 2012). Linear and log-linear regression models are able to estimate demand or supply for airports in competition by including attributes of competitors. Although attributes of competing airlines or competing airports can be included, the competition or cooperation pattern cannot be reflected.

There are mainly four types of airport competition study methods considering supply-and-demand interaction, including two-stage and three-stage least squares models, mathematical optimization, game theory, and spatial competition model. Linear or log-linear models of demand and supply can be estimated simultaneously by two-stage or three-stage least squares estimation method to account for supply-and-demand interaction. Mathematical optimization studies assume the decision-making of airline or airport is based on profit maximization, social benefit maximization, or airline cost minimization. They usually combine with discrete choice models (Barbot, 2009; Hansen, 1990; Pels et al., 1998; Suzuki, Crum, & Audino, 2004). Classical game theory models account for the decision-making process of competing airports and competing airlines (Barbot, 2009; Hansen, 1990). It includes the sequence, information known, and decision variables in decision-making. Meanwhile, the objective of decision is usually profit maximization or welfare maximization, which implicates that normally classical game theory associates with mathematical optimization. If both airport choice and airline choice are considered, it is important to show the relationship between airport and airline in analysis of airport competition (Barbot, 2009). Output of spatial competition model on the demand side is airport catchment area, and that on the supply side is airport-airline relationship or airport pricing (Fröhlich & Niemeier, 2011). However, the basic spatial competition model, Hotelling model, cannot reflect the impact of this factor. Among all the methodologies, only two-stage least squares model and three-stage least squares model are based on real data and meanwhile can account for supply-and-demand interaction.

Based on the findings in previous studies (Harvey, 1987; S. Hess, 2004; S. Hess & Polak, 2005b; Pels et al., 2003; Windle & Dresner, 1995), variables that are deemed important to airport demand include ground access time and distance, airfare, flight frequency, air trip time, direct or indirect flight, delay, aircraft type, the number of airlines at one airport, group size, and characteristics of passengers. The characteristics of passengers contain past experience, trip purpose, car ownership, income in disaggregate study, and contain trip purposes, population, employment, and income in aggregate study.

In supply-side studies, the dependent variables are normally airline cost, yield (Windle & Dresner, 1999), pricing (S. Zhang et al., 2013) , and seat supply (Ippolito, 1981). No matter what the dependent variable is, the number of passengers is a variable in the function (Evans & Kessides, 1993; Ippolito, 1981; Windle & Dresner, 1999; S. Zhang et al., 2013). Other variables that have been used in supply-side models include revenue, unit fuel price, labor price per employee, material price indicator, capital stock, load factor, stage length, the number of points served, delay (Zou & Hansen, 2012b), flight distance, vacation route dummy variables, flight connection, slot control, time trend, presence of low-cost carriers, indexes of market share (Evans & Kessides, 1993; Windle & Dresner, 1999; S. Zhang et al., 2013), carrier concentration at airport, commuter competition, airport departures, local carrier indicators, and eligibility of subsidy (Ippolito, 1981).

In Table 2.1, studies categorized by their focus on demand, supply, or demand and supply interaction are summarized, along with methodology and focus.

In conclusion, two gaps were found in previous airport leakage studies. One is that the studies specifically exploring how major hub airports affect airport leakage at local airports are limited. The other gap is that so few leakage studies have accounted for the inherent interactions between supply and demand. Based on the two gaps, this research explores whether major hub airports affect airport leakage at local airports, and if so, how they affect airport leakage, in the context of models that consider the two-way interactions between demand and supply.

Table 2.1 List of Studies and Other Information in Categorization of Demand, Supply and Interaction

Categorization	Study	Methodology	Focus
Demand-side Studies	Harvey (1987)	Discrete Choice Model	Airport Competition in Multi-Airport
	Pels et al. (2013)	Discrete Choice Model	Airport Competition in Multi-Airport
	Suzuki et al. (2003)	Discrete Choice Model	Airport Leakage
	Lieshout (2012)	Discrete Choice Model	Airport Leakage
	Hutchinson (1993)	Log-linear Model	Airport Demand
Supply-side Studies	Windle and Dresner (1999)	Linear Regression Model	Airline Competition
	Zou and Hansen (2012b)	Log-linear Regression Model	Airline Cost
Supply and Demand Interaction Studies	Suzuki and Audino (2003)	Two-stage Least Squares Model	Airport Leakage
	Elwakil et al. (2013)	Three-stage Least Squares Model	Airport Leakage
	Suzuki et al. (2004)	Mathematical Optimization	Airport Leakage
	Pels et al. (1998)	Mathematical Optimization	Airport Competition in Multi-Airport
	Hansen (1990)	Game Theory	Airline Competition
	Zhang et al. (2010)	Game Theory	Airport Competition
	Fröhlich and Niemeier (2011)	Spatial Competition Model	Airport Competition

CHAPTER 3. AIRFARE AND AIRPORT DEMAND INTERACTION MODEL

The objective of this chapter is to explore variables that influence airport demand under the hypothesis of airport leakage. There are two sections in this chapter. The first section includes data collection, origin-destination (OD) selection, data processing, and descriptive statistics of dataset. In the second section, a two-stage least squares model has been developed to capture the interaction of airfare and airport demand. To eliminate the bias of first-order autocorrelation and heteroskedasticity in the two-stage least squares model, the feasible generalized least squares models are established and compared.

3.1. Data Preparation

3.1.1. Data Sources

Data on airport passenger traffic, airline services, census, aviation fuel cost and distance were gathered from five online sources. Airport passenger traffic and airline services data in the United States are from the Airline Origin and Destination Survey (DB1B) (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014d) and the Air Carrier Statistics (T-100) (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014a), both of which are available from the U.S. Department of Transportation (DOT). Census data is from the U.S. Census, Department of Commerce (Census Bureau, U.S. Department of Commerce, 2014a). Aviation fuel cost data is also available from the U.S. DOT (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014c). Driving distances between airports are from the Travel Math website (Travelmath, 2014). The first four data sources will be described in the following sections.

3.1.1.1. Airline Origin and Destination Survey (DB1B)

The Airline Origin and Destination Survey (DB1B) takes information from 10% of domestic air tickets sold in the U.S., including airfare, coupons (i.e., flight legs), origin, destination, quarter, ticket carrier, market distance, market miles flown, non-stop market miles, and others. DB1B provided airfare, flight legs, and distance information for this research.

There are three types of tables in the DB1B dataset. The DB1B ticket table contains information of every domestic itinerary which may be a round-trip itinerary. The DB1B market table contains information of every trip for which a stop is made for purposes other than changing planes. The DB1B coupon table contains information for every trip segment for which the flight number does not change (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014d). For a trip with a layover in between, two trip segments are recorded. We do not consider round trip or trip segments, so we used the DB1B market table for our modeling purposes.

One record in the DB1B market table provides the information for a single ticket, but this single ticket may have bookings for more than one passenger as an air ticket can be booked for a group. As a result, we are able to obtain group size information. Meanwhile, the airfare and flight leg variables represent the average airfare and the average flight leg per passenger. The non-stop market miles variable, which is the distance of direct flight between origin and destination airports, has been chosen as a distance variable in this thesis. The DB1B data is available aggregated into quarter-years. Data from 2004 quarter 1 to 2014 quarter 1 were used, leading to over 21 million observations for all U.S. airports from the DB1B market dataset.

3.1.1.2. Air Carrier Statistics, U.S. Carriers (T-100)

The Air Carrier Statistics (U.S. Carriers), which is also called T-100 dataset, provides aggregated data about air carriers, enplaned passengers, and freight per month. Two kinds of tables are available for domestic air traffic. The T-100 domestic market table is based on travelers' origin and destination (i.e., trip) including direct and indirect flights. The T-100 domestic segment table is based on trip segment including passengers on direct flight and passengers transferring at origin or destination airport of the trip segment (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014a; Bureau of Transportation Statistics, U.S. Department of Transportation, 2014b). As mentioned above, we do not consider trip segment for transfer passengers, so the market table is used. However, flight departure is only available in the segment table, which is an indicator of flight frequency.

From 2004 quarter 1 to 2014 quarter 1, there are more than 2.5 million observations in T-100 domestic market table, and more than 3.6 million observations in T-100 domestic segment table for all the U.S. airports and all the U.S. carriers.

3.1.1.3. U.S. Census

Demographic information, such as age, race, and income, are available from the Annual Community Survey by U.S. Census, Department of Commerce (Census Bureau, U.S. Department of Commerce, 2014a). The Annual Community Survey is a nationwide survey of around 3.5 million households (Census Bureau, U.S. Department of Commerce, 2014b). Census data are available at different geographic levels, such as county, metropolitan area, division, and state. Population and per capita income for metropolitan areas are used in this research. Metropolitan areas are defined by the Office of Management and Budget (Nussle, 2008). Although specific criteria have been used, a brief definition of metropolitan area is that "Metropolitan Statistical Areas have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties" (Nussle, 2008).

3.1.1.4. Aviation Fuel Cost and Consumption

Aviation fuel cost and consumption were also found on the U.S. DOT website (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014c). Both total fuel cost and unit fuel cost per gallon are available per month. This information is categorized by U.S. carriers or international carriers, scheduled services or unscheduled services, and domestic services or international services. For this research, aviation fuel cost and consumption of U.S. carriers with respect to scheduled domestic services were used.

3.1.2. Origin-Destination (OD) Selection

The purpose of origin-destination (OD) selection is to find metropolitan regions from which airport leakage is hypothesized to occur. Each route - from an origin airport to a destination airport - is called an origin-destination (OD) pair. In this research, we will use "local OD pair" and "substitute OD pair" to differentiate routes originating from a (candidate) local airport and from a (candidate) substitute airport to a given destination, respectively. The process of OD selection involves identifying the local airport in the area from which passengers are leaking, the (substitute) major hub airport to which passengers "leak", and the destination airport (thereby identifying the OD trip). The identification procedure is shown in Figure 3.1.

The first step in Figure 3.1 involves the selection of 25 candidate local airports. The selection is based on literature review or geographic features. When choosing each candidate local airport, their corresponding substitute airports are also chosen. For instance, based on passenger survey data, airport leakage was observed from Des Moines International Airport (DSM) to Kansas City International Airport (MCI), Minneapolis–Saint Paul International Airport (MSP), and Eppley Airfield (OMA) (Iowa Department of Transportation & Iowa Department of Economic Development, 2001; Suzuki et al., 2003). Ten of the 25 candidate local airports are from a previous study of airport leakage (Suzuki & Audino, 2003). In that study, 14 airports were identified as local airports because they were "airports classified as 'medium' by the U.S. General Accounting Office (U.S. GAO) report" without other airports in radius of 70 miles (Suzuki & Audino, 2003). Four out of the 14 airports were excluded in our selection due to the fact that the passenger traffic is too small or the airport is close to a multi-airport region. The airport leakage in the hypothesis of this research is occurring where one airport is expected to serve one metropolitan region. Thus, multi-airport region was excluded.

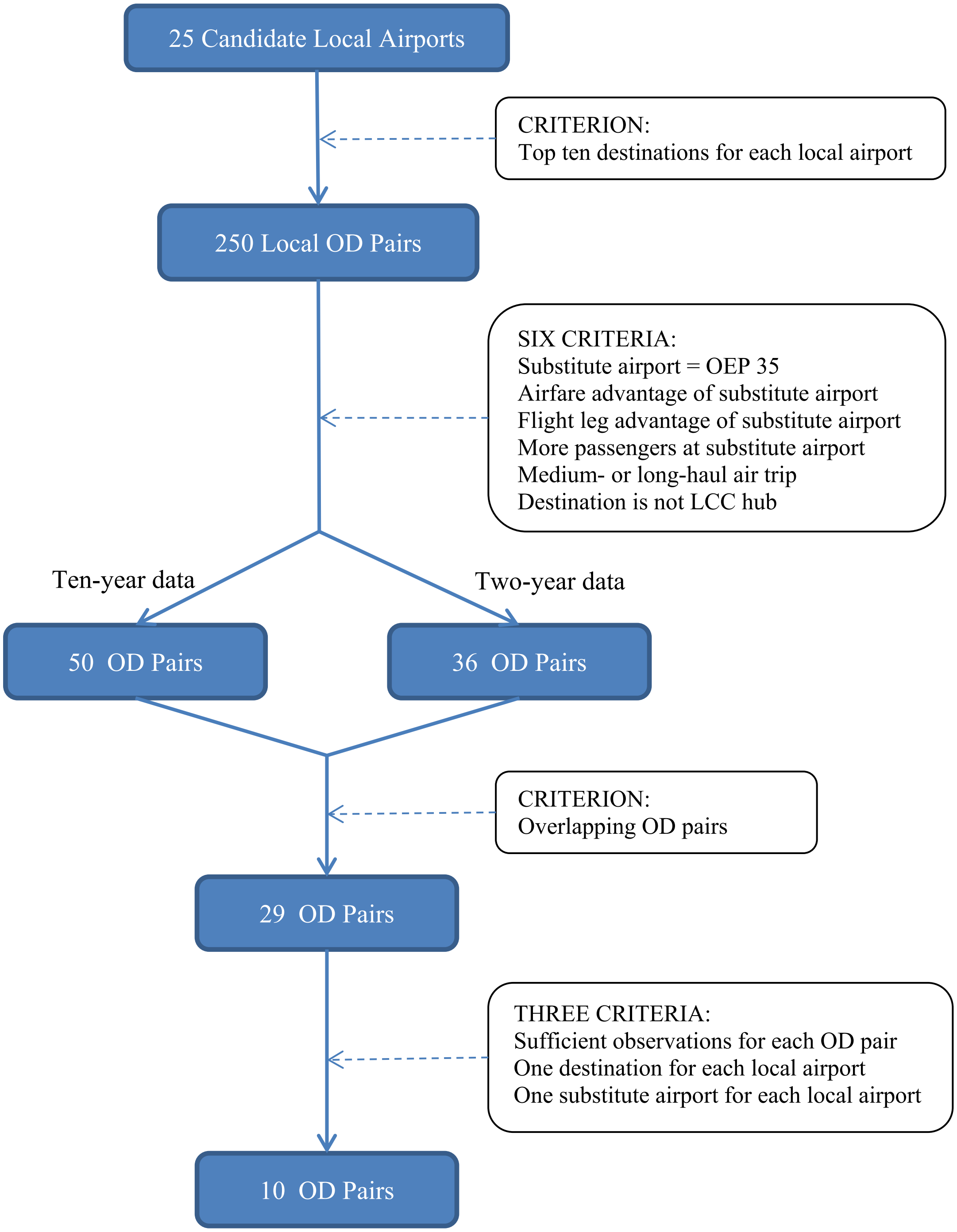

Figure 3.1 Procedure of origin-destination (OD) selection

The remaining 15 of the 25 local airport candidates were selected from “medium” or “small” airports categorized by the U.S. Federal Aviation Administration (FAA) in a single-airport metropolitan area but with hub airports nearby (Federal Aviation Administration, 2014a). Each hub airport holds a functional importance for airlines and serves more significant passengers than the local airport (Ryerson & Kim, 2013). The driving distance between a candidate local airport and the corresponding substitute airport ranges from 45 miles to 237 miles, which is within the distance that air passengers may be willing to drive to a substitute airport (Suzuki & Audino, 2003). It should be noted that one candidate local airport may have more than one candidate substitute airport, because the driving distances to these candidate substitute airports are comparable as well as the number of passengers at each substitute airport.

Top 10 destinations with the highest passengers were identified for each candidate local airport based on the 10-year DB1B market data. As a result, there are 250 local OD pairs after this step. For each local OD pair, the average airfare per passenger, average flight legs per passenger, and non-stop miles were also obtained. Meanwhile, these three values have also been obtained for each corresponding substitute OD pair. In order to strengthen the hypothesis of airport leakage, six criteria were used to select ODs.

1. Substitute airports are included in the Operational Evolution Partnership (OEP) 35 airports by the U.S. Federal Aviation Administration (FAA) (Federal Aviation Administration, 2009).

   OEP 35 airports are the 35 busiest commercial airports in the U.S., taking on over 70% of air passenger movements in U.S. (Federal Aviation Administration, 2009; Federal Aviation Administration, 2014b). They serve major metropolitan areas as airlines’ hub airports to transfer traffic volume (Ryerson & Kim, 2013).

2. Average airfare per passenger for local OD pair < average airfare per passenger for substitute OD pair.

3. Average flight legs per passenger for local OD pair < average flight legs per passenger for substitute OD pair.

   Based on our assumption that passengers travel to the substitute airport to take advantage of better air services, the average airfare and flight leg for the substitute OD pair should be lower than the local OD pair.

4. Passengers for substitute OD pair ≥ 150% * passengers for local OD pair

   This criterion is based on economies of density in aviation industry, meaning that the more passengers an airport serves, the lower the airport cost per passenger will be. To guarantee that a substitute airport serves more passengers to the destination than the local airport, we assume that passengers for the substitute OD pair will be least 50% more than passengers for the local OD pair. Thus, the substitute airport is more likely to provide lower airfare and attract passengers from the local airport.

5. Distance from local airport to destination airport is greater than 500 miles.

   Because airport leakage is less likely to occur for short-haul air trips (Hsu & Wu, 1997), the OD pairs that are less than 500 miles were eliminated.

6. Destination airport is not a low-cost carrier (LCC) hub airport.

   If the destination airport is served by at least one LCC, the local OD pair may also be served by LCC, regardless of whether the local airport is a LCC hub airport. However, we only consider LCC as an attribute of the local airport. Thus, the destination airport cannot be a LCC hub.

The six criteria have been used to filter the OD pairs in the 10-year DB1B data and 2-year DB1B data. The 2-year DB1B data is from 2012 quarter 2 to 2014 quarter 1 (8 quarters in total). The reason to use the 2-year DB1B data is to exclude the potential impact of the 2008 economic downturn on airport leakage. As shown in Figure 3.1, based on the 10-year DB1B data, there are 50 results that satisfy the above six criteria. Based on the 2-year DB1B data, there are

36 results that satisfy the six criteria. Together, there are only 29 overlapping OD pairs for the two periods. Due to the fact that many OD pairs share the same origin airport, the 29 OD pairs were filtered again using the following three criteria.

1. Only OD pairs with more than 30 quarterly observations of passenger enplanement in T-100 dataset are retained in the dataset. All other variables from DB1B dataset have 41 quarterly observations without missing values.

2. For each local airport, only the OD pair with the highest number of passengers is selected. 12 OD pairs are left.

3. Only one substitute airport with the highest passengers is selected for each local airport.

As shown in Figure 3.1, finally, there are 10 local OD pairs left after filtering. The local airports, their corresponding substitute airports and destination airports are displayed in Table 3.1. Their detailed airfare, flight leg, passengers and distance information from 2004 to 2014 (with only one quarter in 2014) are contained in Appendix A.

Table 3.1 Result of OD Selection

Local Airport	Substitute Airport	Destination Airport	Local OD Pair (Code)
Jacksonville International Airport, FL	Orlando International Airport, FL	Philadelphia International Airport, PA	JAX - PHL
Tucson International Airport, AZ	Phoenix Sky Harbor International Airport, AZ	Seattle–Tacoma International Airport, WA	TUS - SEA
Gerald R. Ford International Airport, MI	Detroit Metropolitan Wayne County Airport, MI	Tampa International Airport, FL	GRR - TPA
Columbia Metropolitan Airport, SC	Charlotte Douglas International Airport, NC	LaGuardia Airport, NY	CAE - LGA
Portland International Jetport, ME	Logan International Airport, MA	Charlotte Douglas International Airport, NC	PWM - CLT
Bradley International Airport, CT	John F. Kennedy International Airport, NY	Tampa International Airport, FL	BDL - TPA
Port Columbus International Airport, OH	Detroit Metropolitan Wayne County Airport, MI	Tampa International Airport, FL	CMH - TPA
Charleston International Airport, SC	Charlotte Douglas International Airport, NC	LaGuardia Airport, NY	CHS - LGA
Chattanooga Metropolitan Airport, TN	Hartsfield–Jackson Atlanta International Airport, GA	Ronald Reagan Washington National Airport, VA	CHA - DCA
Huntsville International Airport, AL	Hartsfield–Jackson Atlanta International Airport, GA	Ronald Reagan Washington National Airport, VA	HSV - DCA

3.1.3. Description of Dataset

After origin-destination selection, data from different sources were processed to create the final dataset. Because the DB1B data is presented in a quarterly format, all variables will be in quarters except census data which are in years.

The average airfare per passenger, average group size, flight legs per passenger, and non-stop miles are from the DB1B market dataset as mentioned in Section 3.1.1. Traffic data from the T-100 dataset have been processed and organized to derive more variables related to traffic volume. The total passenger enplanement per quarter from the local airport to all the U.S. airports except the subject destination airport is set as the total enplanement variable for the local airport. The reason to exclude the subject destination is to eliminate the endogeneity between this enplanement variable and the passenger variable. The passenger variable represents the number of passengers from the local airport to the subject destination (Suzuki & Audino, 2003). Total passenger enplanement from the substitute airport to all the U.S. airports is set as the enplanement variable for the substitute airport to show traffic volume of the substitute airport in a certain quarter. The total passenger enplanement from all the U.S. airports to the destination per quarter, excluding that from the local airport and from the substitute airport, has been used as the seasonality variable to reflect seasonal fluctuation of air passengers to the destination (Suzuki & Audino, 2003). The reason of excluding the local airport and the substitute airport is also to eliminate the endogeneity between the passenger variable and the seasonality variable, or the endogeneity between the enplanement variable and the seasonality variable. In addition, the number of passengers served by low-cost carriers (LCC) in each quarter has been assessed for the 10 local OD pairs. Identified LCCs are Southwest Airlines, AirTran Airways, Allegiant Air, Frontier Airlines, JetBlue Airways, Spirit Airlines, Sun County Airlines, and Virgin America.

Yearly population and per capita income in metropolitan areas are only available in the years between 2005 and 2013 (Census Bureau, U.S. Department of Commerce, 2014a). The population in 2004 was estimated by using population change rate from 2004 to 2005 for each state. When one metropolitan area covers more than one state, the average population change rate was used. Per capita income in 2004 was estimated by using per capita income growth rate from 2004 to 2005 in the United States. Population and income in quarter 1, 2014 was set as the same values as 2013. In addition, freight enplanement for each local OD pair per quarter from T-100 has been included in the dataset which may also be able to reflect economy (Suzuki & Audino, 2003).

Aviation fuel cost and consumption are in months from Air Carrier Financial Reports (or Form 41 Financial Data), U.S. DOT (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014c). Quarterly fuel cost per gallon is not the mean value of fuel cost per gallon in three months; instead, it is weighed by fuel consumption. Fuel cost is time specific, meaning it does not change for different routes. The values are shown in Figure 3.2. From this figure, the economic crisis in 2008 caused a significant decrease in fuel cost. Since 2009, fuel cost has been increasing and remained relatively stable after 2011.

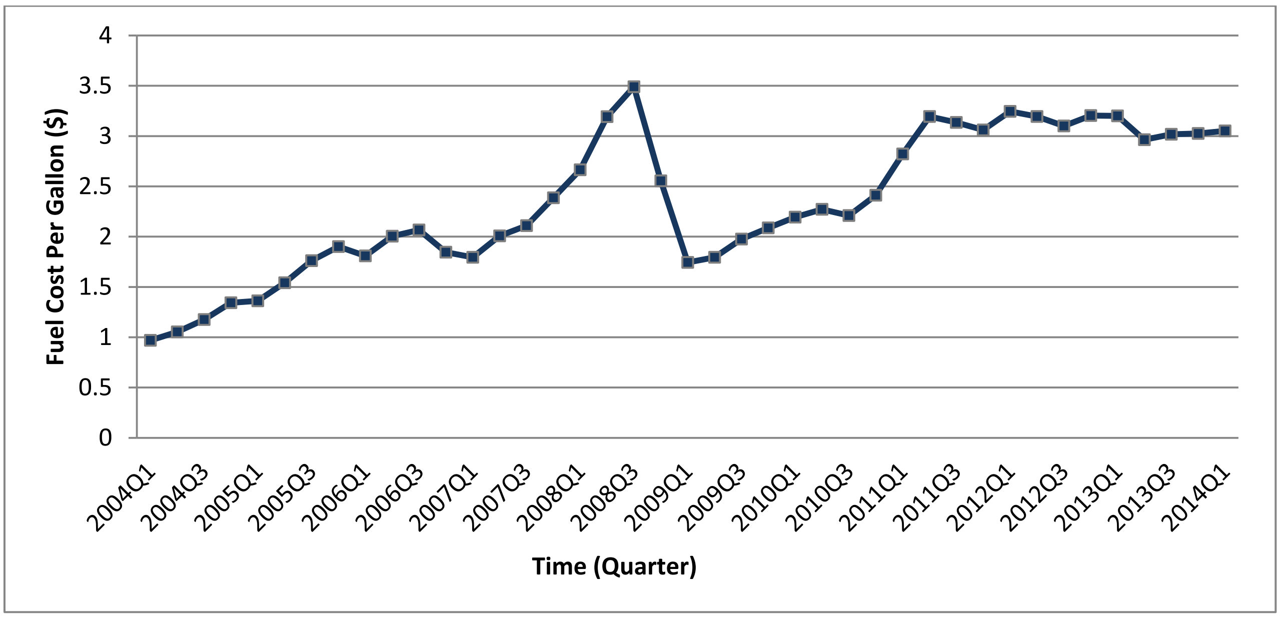

Figure 3.2 Quarterly aviation fuel cost per gallon for domestic services (U.S. carriers)

Descriptive statistics, including the number of observations, mean, standard error, minimum, median, and maximum of all variables in the dataset are shown in Table 3.2.

Table 3.2 Descriptive Statistics of Variables in the Dataset

Variable	Explanation	Obs.	Mean	Std. Error	Min.	Median	Max.
$FARE$	Airfare from local airport, per passenger per quarter	410	177.00	46.37	97.88	171.36	372.67
LEG	Flight leg from local airport, per passenger per quarter	410	1.40	0.25	1.06	1.37	2.07
$MILES$	Non-stop miles for local OD pair	410	814.53	224.01	523.00	777.00	1,216.00
$PASS$	Passengers for local OD pair per quarter	395	14,829.31	11,153.00	190.00	13,170.00	53,726.00
$FREIGHT$	Freight for local OD pair per quarter (pounds)	395	9,622.15	31,630.00	0.00	538.00	304,816.00
$FREQ$	Performed departure for local OD pair	393	229.55	156.10	2.00	185.00	661.00
SEA	Seasonality; total passenger enplanement per quarter from all U.S. airports excluding local airport and substitute airport to the destination	410	2,586,606.00	724,544.00	1,545,803.00	2,337,216.00	5,001,127.00
$PASS_{LCC}$	Passengers served by LCC for local OD pair per quarter	410	4,498.10	7,452.00	0.00	0.00	26,017.00
P_{LCC}	Portion of passengers served by LCC for local OD pair per quarter	395	0.22	0.36	0.00	0.00	1.00
$SIZE$	Average group size for local OD pair per quarter	410	1.89	0.62	1.03	1.75	4.15
ENP	Passengers from local airport to all U.S. destinations excluding passengers for local OD pair per quarter	410	365,565.20	256,617.00	51,795.00	257,806.50	1,007,612.00
POP	Population in the metropolitan area	410	882,339.60	415,970.00	348,211.00	773,619.00	19,990,193.00

	of local airport per quarter						
INC	Per capita income in the metropolitan area of local airport per quarter	410	6,611.70	761.84	5,451.00	6,460.00	8,833.00
FS	Airfare from substitute airport, per passenger per quarter	410	152.02	31.20	88.66	149.63	308.12
LS	Flight leg from substitute airport, per passenger per quarter	410	1.10	0.06	1.01	1.08	1.36
$FREQS$	Performed departure for substitute OD pair	410	1,193.07	523.35	260.00	1,218.00	2,141.00
ES	Passengers from substitute airport to all U.S. destinations	410	4,912,370.00	2,486,211.00	2,183,002.00	3,944,060.00	10,822,651.00
DIS	Driving distance between local airport and substitute airport	410	130.90	28.71	88.00	139.00	181.00
$FUEL$	Unit aviation fuel cost per gallon per quarter in the U.S.	410	2.34	0.70	0.97	2.21	3.49

3.2. Model Estimation and Results

The two-stage least square (2SLS) model has been chosen for use. The reason is that the 2SLS model is able to estimate simultaneous equations model. More specifically for this research, the 2SLS model is able to estimate how airfare and other variables impact airport demand, and, meanwhile, how airport passengers impact airfare. In the first stage of the model, the airfare is estimated by the passenger variable and other exogenous variables. In the second stage, the demand is estimated by the predicted airfare variable from the first stage (i.e., instrumental variable) and other variables (Dougherty, 2011; Pindyck & Rubinfeld, 1998).

The 2SLS model in this research is based on model form and findings in Suzuki and Audino (2003). Log-linear form is used for the first-stage model and second-stage model, because it performed better than linear form (Suzuki & Audino, 2003). One principle of variable selection is to keep as many variables as possible in each model so as to explore their impacts on demand (Suzuki & Audino, 2003). Variables that have been tested include route indicator variables, flight leg, seasonality variable, quarter indicator variable, freight, airfare at the substitute airport, flight leg at the substitute airport, and interaction variables of the driving distance between the local airport and the substitute airport with the airfare at the substitute airport and with the flight leg at the substitute airport, travel group size, fuel cost per gallon, the enplanement at the local airport and the substitute airport, low-cost carrier (LCC) indicator variables, non-stop miles, flight frequency, population, and income (Suzuki & Audino, 2003).

3.2.1. Model 1-a: Two-Stage Least Squares (2SLS) Model

Two-stage least squares model, which is also called Model 1-a in this thesis, can be estimated in the Statistical Analysis System (SAS) software. All the variables mentioned above including their interaction terms have been tested in the airfare model and the demand model. The final model form obtained is below.

First-stage Model

$$\ln(FARE_{it}) = \sum_{i} \lambda_i \cdot \mathrm{I}(i=1) + \alpha_1 \ln(PASS_{it}) + \alpha_2 \cdot \mathrm{I}(LCC_{it} = 1) + \alpha_3 \ln(FS_{it}) + \alpha_4 \ln(FUEL_t * MILES_i) + \mu_t \qquad \textbf{(3-1)}$$

Where

The subscript i denotes the local OD pair, particular to each of the 10 OD pairs represented in the dataset.

The subscript t denotes time or quarter.

$FARE_{it}$ is the average airfare per passenger for the local OD pair i at quarter t.

$\mathrm{I}(i = 1)$ is the route indicator variable. $\mathrm{I} = 1$ if the route is for the local OD pair i; and $\mathrm{I} = 0$ otherwise.

$PASS_{it}$ is the number of passengers for the local OD pair i at quarter t.

$\mathrm{I}(LCC_{it} = 1)$ is the low-cost carrier (LCC) indicator variable. $\mathrm{I} = 1$ if 25% or more passengers used low-cost carriers (LCC) for the local OD pair i at quarter t; and $\mathrm{I} = 0$ otherwise.

FS_{it} is the average airfare per passenger for the substitute OD pair corresponding to the local OD pair i at quarter t.

$FUEL_t$ is the unit aviation fuel cost per gallon for U.S. domestic services provided by U.S. carriers at quarter t.

$MILES_i$ is the non-stop miles of from origin airport to destination airport for the local OD pair i.

There are two reasons to use the interaction term $FUEL_t * MILES_i$. Firstly, total fuel cost, as one of the major airline costs, depends on miles flown. Secondly, $Fuel_t$ is time-specific and $Miles^L$ is route-specific, so their combination is specific for every observation.

μ_t is the error term in the airfare model at quarter t.

$\alpha_1 \dots \alpha_4$, and λ_i are parameters.

Second-stage model

The predicted airfare from the first-stage model is used in the second-stage demand model.

$$\begin{aligned}\ln(PASS_{it}) = \sum_i \delta_i \cdot \mathrm{I}(i=1) + \beta_1 \ln(\widehat{FARE}_{it}) + \beta_2 \ln(LEG_{it}) + \ \beta_3 \ln(SEA_{it}) \\ + \beta_4 \ln(POP_{it}) + \beta_5 \ln(FS_{it}) + \beta_6 \ln(SIZE_{it}) + \beta_7 * SIZE_{it} \\ * \ln(FS_{it}) + \beta_8 DIS_i + \beta_9 \ln(ENP_{it}) + \varepsilon_t\end{aligned} \tag{3-2}$$

Where

$PASS_{it}$ is the number of passengers for the local OD pair i at quarter t.

$\mathrm{I}(i = 1)$ is the route indicator variable. I = 1 if the route is for the local OD pair i; and I = 0 otherwise.

$ln(\widehat{FARE}_{it})$ is the fitted log value of airfare per passenger for the local OD pair i at quarter t.

LEG_{it} is the average flight leg per passenger for the local OD pair i at quarter t. If all passengers take direct flights from the origin airport to the destination airport, the average flight leg is one. If all passengers transfer once between origin and destination, the average flight leg is two.

SEA_{it} is the seasonality variable, represented by total number of passengers from all U.S. airports except the local airport and the substitute airport to the destination airport for the local OD pair i at quarter t.

POP_{it} is the annual population in the year of quarter t in the metropolitan area served by the local airport (i.e., origin airport) of local OD pair i.

FS_{it} is the average airfare per passenger for the substitute OD pair corresponding to the local OD pair i at quarter t.

$SIZE_{it}$ is the average group size of passengers for the local OD pair i at quarter t.

$SIZE_{it} * \ln(FS_{it})$ is to show how the airfare at the substitute airport impact the demand when group size changes.

DIS_i is the driving distance between the local airport and the corresponding substitute airport for the local OD pair i; in miles.

ENP_{it} is the total passenger enplanement from the local airport to all U.S. destination airports minus the number of passengers of the local OD pair i at quarter t.

ε_t is the error term.

$\beta_1 \ldots \beta_9$, and δ_i are parameters.

3.2.1.1. Estimation Results

Parameter estimation results and goodness of fit for Model 1-a are in Table 3.3.

Table 3.3 Estimation Result of Two-stage Least Squares Model (Model 1-a)

	Parameter Notation	Variable	Coefficient	Std. Error	t-value	Pr>t
First Stage	$\lambda_{JAX-PHL}$	$\mathrm{I}(JAX-PHL=1)$	2.90	0.25	11.66	<.0001
	$\lambda_{TUS-SEA}$	$\mathrm{I}(TUS-SEA=1)$	2.76	0.25	10.93	<.0001
	$\lambda_{GRR-TPA}$	$\mathrm{I}(GRR-TPA=1)$	2.64	0.23	11.49	<.0001
	$\lambda_{CAE-LGA}$	$\mathrm{I}(CAE-LGA=1)$	2.83	0.24	11.82	<.0001
	$\lambda_{PWM-CLT}$	$\mathrm{I}(PWM-CLT=1)$	2.87	0.25	11.61	<.0001
	$\lambda_{BDL-TPA}$	$\mathrm{I}(BDL-TPA=1)$	2.83	0.25	11.35	<.0001
	$\lambda_{CMH-TPA}$	$\mathrm{I}(CMH-TPA=1)$	2.85	0.24	11.63	<.0001
	$\lambda_{CHS-LGA}$	$\mathrm{I}(CHS-LGA=1)$	2.93	0.25	11.67	<.0001
	$\lambda_{CHA-DCA}$	$\mathrm{I}(CHA-DCA=1)$	2.87	0.23	12.26	<.0001
	$\lambda_{HSV-DCA}$	$\mathrm{I}(HSV-DCA=1)$	3.40	0.25	13.64	<.0001
	α_1	$PASS_{it}$	-0.09	0.01	-6.8	<.0001
	α_2	$\mathrm{I}(LCC_{it}=1)$	-0.13	0.02	-5.18	<.0001
	α_3	FS_{it}	0.24	0.03	6.92	<.0001
	α_4	$FUEL_t * MILES_i$	0.26	0.02	15.29	<.0001
	Model fit statistics	Sum of squared residual	19.349			
		Mean squared Error	0.010			
		R-square	0.832			
		Adjusted R-square	0.827			
Second Stage	**Parameter Notation**	**Variable**	**Coefficient**	**Std. Error**	**t-value**	**Pr>t**
	$\delta_{JAX-PHL}$	$\mathrm{I}(JAX-PHL=1)$	-6.31	1.12	-5.65	<.0001
	$\delta_{TUS-SEA}$	$\mathrm{I}(TUS-SEA=1)$	0.00	.	.	.
	$\delta_{GRR-TPA}$	$\mathrm{I}(GRR-TPA=1)$	-7.07	1.04	-6.8	<.0001
	$\delta_{CAE-LGA}$	$\mathrm{I}(CAE-LGA=1)$	-13.97	2.26	-6.18	<.0001
	$\delta_{PWM-CLT}$	$\mathrm{I}(PWM-CLT=1)$	-11.28	1.99	-5.68	<.0001
	$\delta_{BDL-TPA}$	$\mathrm{I}(BDL-TPA=1)$	-11.07	1.96	-5.66	<.0001
	$\delta_{CMH-TPA}$	$\mathrm{I}(CMH-TPA=1)$	-4.73	0.85	-5.56	<.0001

	$\delta_{CHS-LGA}$	$\mathrm{I}(CHS - LGA = 1)$	-4.22	0.71	-5.95	<.0001
	$\delta_{CHA-DCA}$	$\mathrm{I}(CHA - DCA = 1)$	-10.63	1.67	-6.35	<.0001
	$\delta_{HSV-DCA}$	$\mathrm{I}(HSV - DCA = 1)$	-1.99	0.47	-4.27	<.0001
	β_1	$FARE_{it}$	-1.60	0.24	-6.7	<.0001
	β_2	LEG_{it}	-3.05	0.29	-10.54	<.0001
	β_3	SEA_{it}	1.00	0.16	6.34	<.0001
	β_4	POP_{it}	1.89	0.33	5.74	<.0001
	β_5	FS_{it}	0.38	0.13	2.96	0.0033
	β_6	$SIZE_{it}$	1.03	0.40	2.55	0.011
	β_7	$SIZE_{it} * \ln(FS_{it})$	-0.12	0.04	-3.16	0.0017
	β_8	DIS_i	-0.15	0.03	-5.9	<.0001
	β_9	ENP_{it}	0.27	0.12	2.23	0.0265
	Model fit statistics	Sum of squared residual	380.194			
		Mean squared Error	0.083			
		R-square	0.924			
		Adjusted R-square	0.920			

In Model 1-a, almost all variables are significant at the 99% confidence level while only the enplanement at the local airport is significant at the 95% confidence level. Demand and airfare have a negative relationship, meaning higher demand leads to lower airfare while higher airfare leads to lower demand. When the airfare increases, passengers are less willing to choose the local airport. Thus the demand at the local airport reduces. The negative impact of demand on airfare occurs when airline competition exists. If more than one airline serves for the same route, increasing demand will intensify their competition which will eventually reduce the average airfare.

As expected, the presence of LCC on the route will reduce the average airfare. The airfare at the local airport will decrease if the airfare at the substitute airport decreases. This may result from competition of airlines serving the two airports. The interaction effect of unit fuel cost and non-stop miles is positive on the airfare because increase of both unit fuel cost and non-stop miles will increase airline cost. The longer distance is or the higher fuel cost is, the more expensive air ticket will be. Comparing the absolute values of the parameters, the passenger variable is small because the digits of passenger values are more than other variables. The impact of low-cost carrier on the airfare at the local airport is smaller than that of the airfare at the substitute airport. Parameters of route indicator variables are close to each other, but they are able to reflect characteristics of local OD pairs which have not been explained by other variables such as non-stop miles itself and driving distance to the substitute airport. The goodness of fit for the first-stage airfare model is not good as the second-stage model because the R-square value is 0.832. The goodness of fit for the first-stage airfare model can also be shown by Figure 4, which is the plot of observed $\ln(FARE_{it})$ against predicted $\ln(\widehat{FARE_{it}})$.

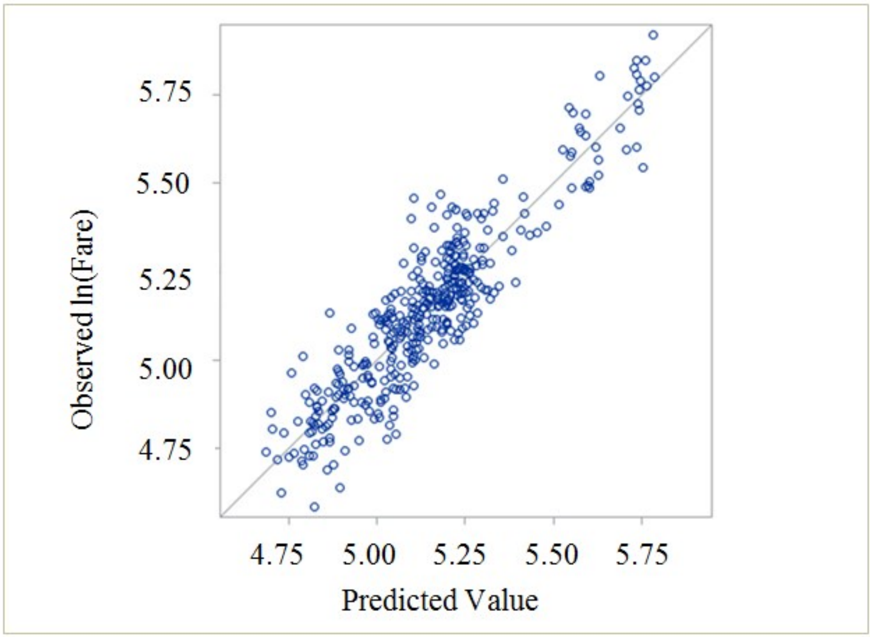

Figure 3.3 Observed airfare vs. predicted airfare

Comparing with TUS-SEA route indicator variable, all other route indicator variables have negative impacts on demand. The characteristics of these routes that have not been captured by other explanatory variables, such as existence of a hub airport nearby and non-stop miles, actually decrease demand comparing with the Route TUS-SEA. The airfare and flight legs at the local airport have negative impacts on the demand, showing passengers are less willing to choose the local airport if airfare increases or more transfers are needed. The seasonality variable indicates seasonal fluctuation of traffic to the destination and contributes positively to the demand at the local airport. In other words, more passengers going to the destination in a season/quarter means more passengers for the local OD pair. Population in metropolitan area has also a positive impact on demand. Normally, more activities exist when population increase, which leads to higher passenger demand at the local airport. Looking at the airfare at the substitute airport alone, its parameter is positive. This means if the substitute airport provides lower airfare, the demand at the local airport will decrease. On the contrary, if the airfare at the substitute airport increases, there will be more passengers using the local airport. This supports our hypothesis of airport leakage in the process of origin-destination selection. The positive impact of travelers' group size can be interpreted to mean that for a larger travel group, passengers would prefer to use the local airport. Dividing the parameter of airfare at the substitute airport by the parameter of the interaction variable of group size and airfare at the substitute airport, it shows that the positive impact of airfare at the substitute airport will be eliminated when group size is more than three. In other words, lower airfare at the substitute airport does not have attraction to passengers from the local airport when passengers travel in a group of more than three people. Driving distance to the substitute airport impacts the demand at the local airport negatively, showing that more passengers would use the local airport if the substitute airport is farther. The enplanement at the local airport contributes positively to demand. Higher traffic at the local airport would attract more passengers, which is a positive feedback effect found in the previous study (Hansen, 1995). More variables, such as interaction effect of the driving distance and airfare at the substitute airport, are tested in the demand model, but they are insignificant. Goodness of fit for the second-stage model can be shown be R-squared value, which is $R^2 = 0.924$. The goodness of fit for the second-stage model can also be shown by plotting observed $\ln(PASS_{it})$ against predicted $\ln(\widehat{PASS}_{it})$ in Figure 3.4.

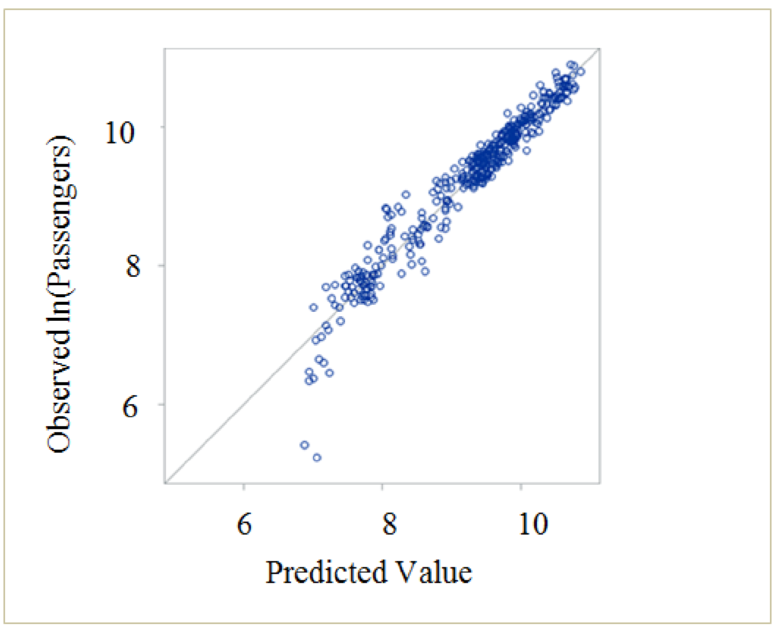

Figure 3.4 Observed demand vs. predicted demand

3.2.1.2. Test for Autocorrelation

Hypotheses of airfare model and demand model in Model 1 are that values of the error terms (or residuals) are independent with time periods (Dougherty, 2011). That is, in the first-stage model, μ_t is independent with $\mu_{t'}$ when $t \neq t'$. In the second-stage model, ε_t is independent with $\varepsilon_{t'}$ when $t \neq t'$. If these hypotheses do not meet, time serial autocorrelation exists, which is also called autocorrelation. Two types of tests have done for autocorrelation. The first one is Durbin-Watson test based on linearity assumption of error term and lagged error term; and the second test is Lagrange Multiplier General test by adding lagged residual into regression. The purpose of conducting more than one test is to eliminate the impact of test assumptions on the result and to validate the results (Ayyangar, 2007).

- Durbin-Watson Test

The standard test for first-order autocorrelation is Durbin-Watson d statistic.

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

Where e_t is residual, and e_{t-1} is lagged residual (Dougherty, 2011).

Whether to reject the null hypothesis of no autocorrelation is based on value of d, lower-level threshold d_L and upper-level threshold d_U. Values of the two thresholds depend on the number of explanatory variables in the model and the number of observations (Dougherty, 2011).

First-order Durbin-Watson test is available for the 2SLS models in SAS software. The result for Model 1-a is in Table 3.4.

Table 3.4 Result of Durbin-Watson Test

First-stage model	Durbin-Watson (DW)	0.921
	Number of Observations	395
	First-Order Autocorrelation	0.530
Second-stage model	Durbin-Watson (DW)	1.069
	Number of Observations	395
	First-Order Autocorrelation	0.462

Durbin-Watson test shows that $d = 0.921 < d_L(Var. = 14, Obs. = 390) = 1.754$ for the first-stage model and $d = 1.069 < d_L(Var. = 19, Obs. = 390) = 1.738$ for the second-stage model, so we can reject the null hypothesis of no autocorrelation and conclude that that there is positive autocorrelation in both the first-stage model and second-stage model. To have a better understanding of how residual correlates with lagged residual, two plots of residual against the time-dependent (i.e., quarter) variable for the first-stage model and second-stage model are shown below respectively.

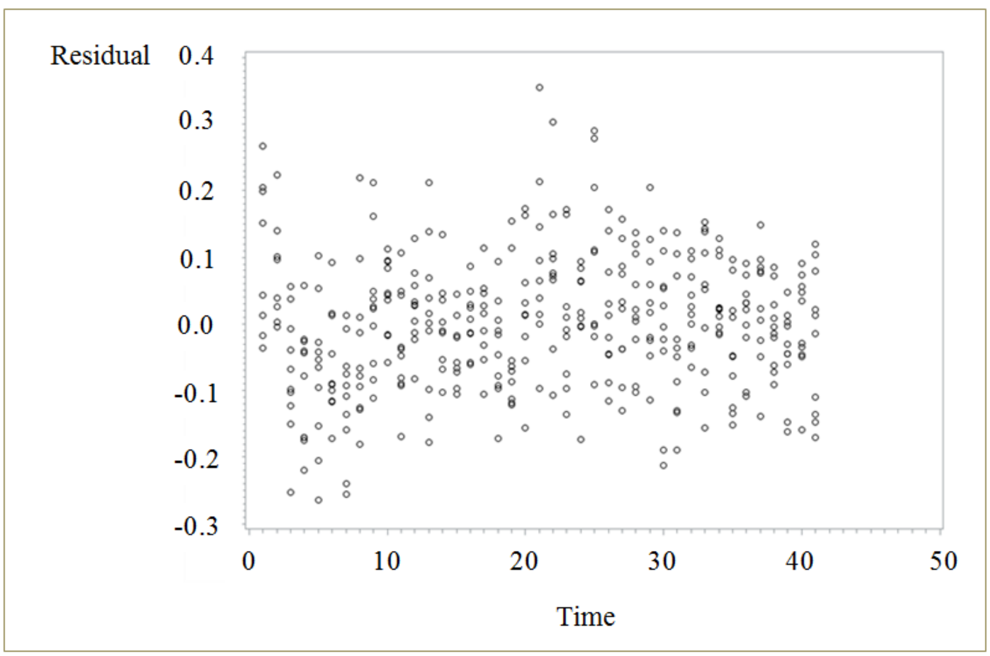

Figure 3.5 Residual against time-dependent variable in the first-stage model (Model 1-a)

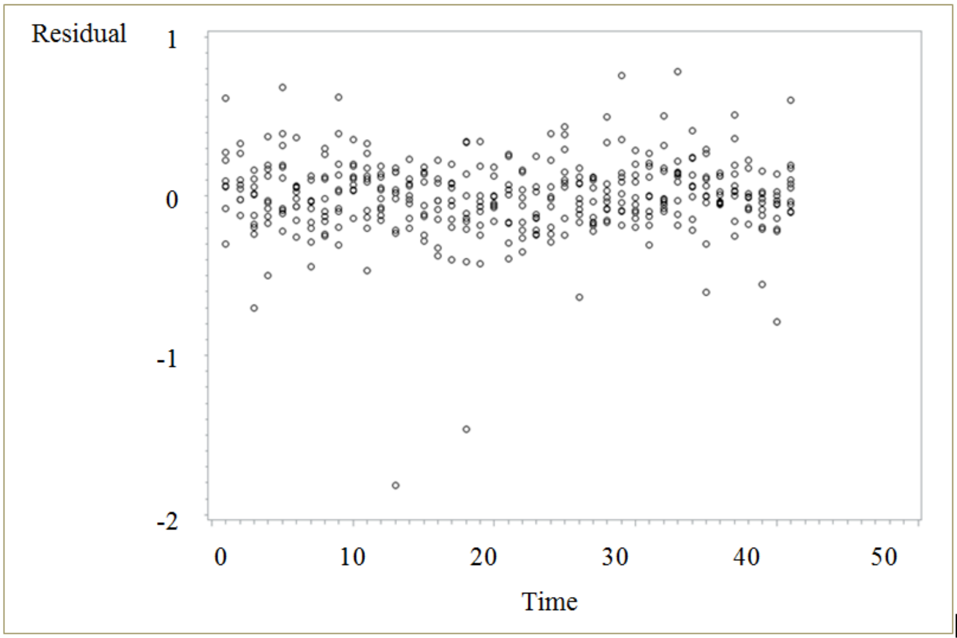

Figure 3.6 Residual against time-dependent variable in the second-stage model (Model 1-a)

Residuals in the two figures are not randomly distributed against time; instead, they have formed a wave-like curve along $residual = 0$ axis. Such formation suggests that there is positive autocorrelation, meaning lagged residuals have positive relationship with residual (Dougherty, 2011).

- Lagrange Multiplier General Test

The Lagrange Multiplier General test checks model performance by adding a lagged residual as a variable in the regression. Lagrange Multiplier General test includes Breusch–Godfrey test and Durbin alternative test. Criterion of Breusch–Godfrey test is to check whether the assumption of $nR^2 \sim \chi_p^2$ holds, where p is the number of lagged residual and n is the number of observations minus number of lagged residual (Godfrey, 1978; SAS, 2014). The Durbin alternative test checks whether lagged residuals are not equal to zero (Park, 2006; Wooldridge, 2012) .

A lagged residual was added to Model 1-a, and only results that are related to the two tests are shown in Table 3.5.

Table 3.5 Result of Lagrange Multiplier General Tests

<table>
<tr><td rowspan="4">Breusch–Godfrey test</td><td colspan="3">Number of observations</td><td colspan="2">390</td></tr>
<tr><td colspan="3">Number of lagged residual</td><td colspan="2">1</td></tr>
<tr><td colspan="2" rowspan="2">R-Square</td><td>First-stage model</td><td colspan="2">0.284</td></tr>
<tr><td>Second-stage model</td><td colspan="2">0.104</td></tr>
<tr><td rowspan="3">Durbin alternative test</td><td>Variable</td><td></td><td>Parameter Estimate</td><td>t Value</td><td>Pr > |t|</td></tr>
<tr><td rowspan="2">Lagged residual</td><td>First-stage model</td><td>0.531</td><td>12.12</td><td><.0001</td></tr>
<tr><td>Second-stage model</td><td>0.320</td><td>5.94</td><td><.0001</td></tr>
</table>

In the Breusch–Godfrey test, nR^2=(390-1)*0.284=110.476>$\chi^2_{1,5\%}$= 3.841 for the first-stage model and R^2=(390-1)*0.104=40.456>$\chi^2_{1,5\%}$= 3.841 for the second-stage model. Thus, we can reject null hypothesis of no autocorrelation and conclude that autocorrelation is present in the first-stage model and second-stage model.

For the Durbin alternative test, both parameters of lagged residual are significantly positive in the first-stage model and second-stage model, indicating that positive first-order autocorrelation is present. In conclusion of all the tests above, there is positive first-order autocorrelation in Model 1-a.

3.2.1.3. Test for Heteroskedasticity

Test for heteroskedasticity is to check whether distribution of error term is homogenous, meaning whether the variance of error term is fixed with respect to different time periods. White test is able to detect heteroskedasticity (Dougherty, 2011), and thus has been conducted for the first-stage model and second-stage model respectively. Results of the White tests are shown in Table 3.6.

Table 3.6 Result of White Test for Heteroskedasticity

White Test	First-stage model	Chi-Square	76.92
		Pr > ChiSq	0.011
	Second-stage model	Chi-Square	102.51
		Pr > ChiSq	0.930

Under the hypothesis of no heteroskedasticity, nR^2 is distributed as chi-squared statistic, i.e., $nR^2 \sim \chi_p^2$ where p is number of regressors minus one and n is number of observations minus p. If the P-value as shwon in the table is smaller than 5%, we can reject null hypothesis of no heteroskedasticity. In the first-stage model, P value is 0.011 and smaller than 0.05, so we can reject the null hypothesis, and conclude that heteroskedasticity is present. In the second-stage model, P value is 0.930 and larger than 0.05, so we cannot reject the null hypothesis of no heteroskedasticity. In conclusion, heteroskedasticity has been detected in the first-stage model of Model 1.

3.2.2. Model 1-b: Feasible Generalized Least Squares (FGLS) Model

Due to the detection of first-order autocorrelation and heteroskedasticity in Model 1-a, it will be estimated using feasible generalized least squares. Feasible generalized least squares is able to estimate parameters in the model when first-order autocorrelation and heteroskedasticity are present (Wooldridge, 2012). To compare with the 2SLS estimation method for Model 1-a, we call this model as feasible generalized least squares (FGLS) model or Model 1-b. As Model 1-b uses the same explanatory variables as Model 1-a, the variables will not be explained again in this section. The variables that have appeared in the previous section are also listed in Appendix C.

First-stage Model

$$\ln(FARE_{it}) = \sum_i \lambda_i \cdot \mathrm{I}(i = 1) + \alpha_1 \ln(PASS_{it}) + \alpha_2 \cdot \mathrm{I}(LCC_{it} = 1) + \alpha_3 \ln(FS_{it}) + \alpha_4 \ln(FUEL_t * MILES_i) + \mu_t \tag{3-3}$$

$$\mu_t = \rho_1 \mu_{t-1} + \epsilon_t$$

Where

μ_t is the error term in airfare model at quarter t.

μ_{t-1} is the error term in airfare model at quarter $t-1$.

$\mu_t = \rho_1 \mu_{t-1} + \epsilon_t$ is the autoregressive error model.

ρ_1 is the first-order autoregressive parameter.

ϵ_t is the error term in the autoregressive error model, which is assumed to be normally and independently distributed with mean 0 and variance σ^2, $\epsilon_t \sim N(0, \sigma^2)$.

Second-stage model

The predicted airfare from the first-stage model is used in the second-stage demand model.

$$\ln(PASS_{it}) = \sum_i \delta_i \cdot \mathrm{I}(i=1) + \beta_1 \ln(\widehat{FARE}_{it}) + \beta_2 \ln(LEG_{it}) + \beta_3 \ln(SEA_{it}) + \beta_4 \ln(POP_{it}) + \beta_5 \ln(FS_{it}) + \beta_6 \ln(SIZE_{it}) + \beta_7 * SIZE_{it} * \ln(FS_{it}) + \beta_8 DIS_i + \beta_9 \ln(ENP_{it}) + \varepsilon_t \tag{3-4}$$

$$\varepsilon_t = \rho_2 \varepsilon_{t-1} + v_t$$

Where

ε_t is the error term in the demand model at quarter t.

ε_{t-1} is the error term in the demand model at quarter $t-1$.

$\varepsilon_t = \rho_2 \varepsilon_{t-1} + v_t$ is autoregressive error model.

ρ_2 is the first-order autoregressive parameter.

v_t is the error term in the autoregressive error model, which is assumed to be normally and independently distributed with mean 0 and variance σ^2, $v_t \sim N(0, \sigma^2)$.

3.2.3.1. Estimation Results

Parameter estimation results and goodness of fit for Model 1-b are in Table 3.7.

Table 3.7 Estimation Result of Feasible Generalized Least Squares Model (Model 1-b)

Stage	Parameter Notation	Variable	Coefficient	Std. Error	t-value	Pr>t
First Stage	$\lambda_{JAX-PHL}$	$\mathrm{I}(JAX - PHL = 1)$	2.75	0.29	9.36	<.0001
	$\lambda_{TUS-SEA}$	$\mathrm{I}(TUS - SEA = 1)$	2.63	0.30	8.83	<.0001
	$\lambda_{GRR-TPA}$	$\mathrm{I}(GRR - TPA = 1)$	2.58	0.28	9.26	<.0001
	$\lambda_{CAE-LGA}$	$\mathrm{I}(CAE - LGA = 1)$	2.72	0.29	9.5	<.0001
	$\lambda_{PWM-CLT}$	$\mathrm{I}(PWM - CLT = 1)$	2.76	0.30	9.34	<.0001
	$\lambda_{BDL-TPA}$	$\mathrm{I}(BDL - TPA = 1)$	2.62	0.30	8.85	<.0001
	$\lambda_{CMH-TPA}$	$\mathrm{I}(CMH - TPA = 1)$	2.70	0.29	9.35	<.0001
	$\lambda_{CHS-LGA}$	$\mathrm{I}(CHS - LGA = 1)$	2.75	0.30	9.29	<.0001
	$\lambda_{CHA-DCA}$	$\mathrm{I}(CHA - DCA = 1)$	2.74	0.28	9.66	<.0001
	$\lambda_{HSV-DCA}$	$\mathrm{I}(HSV - DCA = 1)$	3.25	0.29	11.14	<.0001
	α_1	$PASS_{it}$	-0.06	0.01	-4.4	<.0001
	α_2	$\mathrm{I}(LCC_{it} = 1)$	-0.07	0.03	-2.26	0.0246
	α_3	FS_{it}	0.33	0.04	7.29	<.0001
	α_4	$FUEL_t * MILES_i$	0.17	0.02	7.68	<.0001
	ρ_1	Autoregressive Parameter	0.65	0.04	-15.88	<.0001
	Model fit statistics	$Est.Var(\epsilon_t)$	0.007			
		Regress R-Square	0.998			
		Total R-Square (computed from the autoregressive model residuals)	1.000			
	Durbin-Watson Test		2.030			
Second Stage	**Parameter Notation**	**Variable**	**Coefficient**	**Std. Error**	**t-value**	**Pr>t**
	$\delta_{JAX-PHL}$	$\mathrm{I}(JAX - PHL = 1)$	-32.70	5.07	-6.45	<.0001
	$\delta_{TUS-SEA}$	$\mathrm{I}(TUS - SEA = 1)$	-32.32	4.96	-6.51	<.0001
	$\delta_{GRR-TPA}$	$\mathrm{I}(GRR - TPA = 1)$	-32.70	4.89	-6.69	<.0001

	$\delta_{CAE-LGA}$	$\mathrm{I}(CAE - LGA = 1)$	-32.23	4.85	-6.65	<.0001
	$\delta_{PWM-CLT}$	$\mathrm{I}(PWM - CLT = 1)$	-30.57	4.75	-6.44	<.0001
	$\delta_{BDL-TPA}$	$\mathrm{I}(BDL - TPA = 1)$	-32.24	5.03	-6.41	<.0001
	$\delta_{CMH-TPA}$	$\mathrm{I}(CMH - TPA = 1)$	-33.32	5.17	-6.45	<.0001
	$\delta_{CHS-LGA}$	$\mathrm{I}(CHS - LGA = 1)$	-31.48	4.82	-6.53	<.0001
	$\delta_{CHA-DCA}$	$\mathrm{I}(CHA - DCA = 1)$	-31.55	4.72	-6.68	<.0001
	$\delta_{HSV-DCA}$	$\mathrm{I}(HSV - DCA = 1)$	-29.77	4.62	-6.45	<.0001
	β_1	$FARE_{it}$	-1.20	0.19	-6.37	<.0001
	β_2	LEG_{it}	-3.11	0.26	-12.1	<.0001
	β_3	SEA_{it}	0.91	0.13	6.89	<.0001
	β_4	POP_{it}	2.21	0.36	6.11	<.0001
	β_5	FS_{it}	0.39	0.14	2.85	0.0047
	β_6	$SIZE_{it}$	1.32	0.36	3.62	0.0003
	β_7	$SIZE_{it} * \ln(FS_{it})$	-0.12	0.03	-3.63	0.0003
	β_8	DIS_i	0.00	.	.	.
	β_9	ENP_{it}	0.25	0.10	2.4	0.0168
	ρ_2	Autoregressive Parameter	0.53	0.05	-10.33	<.0001
	Model fit statistics	$Est.Var(v_t)$	0.050			
		Regress R-Square	0.998			
		Total R-Square (computed from the autoregressive model residuals)	1.000			
	Durbin-Watson Test		1.982			

All variables in Model 1-b are still significant at 95% confidence level. The signs of the variables do not change from Model 1-a except TUS-SEA route indicator variable and DIS_i. The coefficients of the route indicator variables are smaller than Model 1-a, meaning that the impact of route indicator variables on passengers is smaller if we account for time serial correlation. Positive autoregressive parameters in first-stage and second-stage models verify the detection of positive autocorrelation in Model 1-a. After considering first-order autocorrelation, the impact of LCC indicator variable on airfare reduces as shown by the absolute value of LCC parameter. Normally, LCC serves an airport in consecutive quarters, so the effect of LCC indicator will be captured by autoregressive error model. The effect of the fuel cost variable can also be captured by autoregressive error model. In the second-stage model of Model 1-b, the impact of the airfare at the local airport on the demand decreases slightly while the impact of the airfare at the substitute airport increases slightly. A further discussion of the estimated parameters in Model 1-a and Model 1-b will be presented in the section of discussion of results.

For both the first-stage model and second-stage model in Model 1-b comparing with Model 1-a, goodness of fit improves because total R-square values are close to one. The d values from Durbin-Watson tests indicate that there is no autocorrelation in the first-stage model and second-stage model of Model 1-b.

3.2.3. Model 2: Feasible Generalized Least Squares (FGLS) Model with an Additional Enplanement Variable

A new model is built to test the hypothesis that other variables excluded from Model 1 (1-a and 1-b) have impact on the local airport's demand. As autocorrelation of the data has already been detected, we will use FGLS estimation. Other variables, such as the flight legs at the substitute airport, income, and the interaction effect of the driving distance and airfare at substituent airport, have been tested. The result indicates that the addition of the enplanement variable at the substitute airport into the second-stage model is appropriate. This feasible generalized least squares model with an additional enplanement variable is called Model 2 in this thesis. The variables will not be explained again if they have been used in Model 1-a and Model 1-b. Also these variables are contained in Appendix C.

Second-stage Model

The first-stage model is the same as Model 1-b. The predicted airfare from the first-stage model is used in the second-stage demand model. The second-stage model of Model 2 is below.

$$\begin{aligned}\ln(PASS_{it}) = \sum_{i} \delta_i \cdot \mathrm{I}(i=1) + \beta_1 \ln(\widehat{FARE}_{it}) + \beta_2 \ln(LEG_{it}) + \ \beta_3 \ln(SEA_{it}) \\ + \beta_4 \ln(POP_{it}) + \beta_5 \ln(FS_{it}) + \beta_6 \ln(SIZE_{it}) + \beta_7 * SIZE_{it} \\ * \ln(FS_{it}) + \beta_8 DIS_i + \beta_9 \ln(ENP_{it}) + \beta_{10} \ln(ES_{it}) + \varepsilon_t \end{aligned} \quad \textbf{(3-5)}$$

$$\varepsilon_t = \rho_2 \varepsilon_{t-1} + v_t$$

Where ES_{it} is the total passenger enplanement from the corresponding substitute airport for the local OD pair i to all the U.S. destination airports at quarter t.

Because the first-stage model of Model 2 is the same as Model 1-b, the estimation result of the first-stage model will not be shown again. Parameter estimation results and goodness of fit for the second-stage model in Model 2 is in Table 3.8.

Table 3.8 Estimation Result of the Second-Stage Model in Model 2

Parameter Notation	Variable	Parameter Estimate	Std. Error	t-value	Pr>t
$\delta_{JAX-PHL}$	$\mathrm{I}(JAX-PHL=1)$	-31.32	5.24	-5.98	<.0001
$\delta_{TUS-SEA}$	$\mathrm{I}(TUS-SEA=1)$	-30.76	5.14	-5.98	<.0001
$\delta_{GRR-TPA}$	$\mathrm{I}(GRR-TPA=1)$	-31.13	5.07	-6.14	<.0001
$\delta_{CAE-LGA}$	$\mathrm{I}(CAE-LGA=1)$	-30.50	5.04	-6.05	<.0001
$\delta_{PWM-CLT}$	$\mathrm{I}(PWM-CLT=1)$	-29.00	4.93	-5.89	<.0001
$\delta_{BDL-TPA}$	$\mathrm{I}(BDL-TPA=1)$	-30.95	5.19	-5.97	<.0001
$\delta_{CMH-TPA}$	$\mathrm{I}(CMH-TPA=1)$	-32.02	5.33	-6.01	<.0001
$\delta_{CHS-LGA}$	$\mathrm{I}(CHS-LGA=1)$	-29.82	5.00	-5.96	<.0001
$\delta_{CHA-DCA}$	$\mathrm{I}(CHA-DCA=1)$	-29.27	4.98	-5.88	<.0001
$\delta_{HSV-DCA}$	$\mathrm{I}(HSV-DCA=1)$	-27.55	4.87	-5.66	<.0001
β_1	$FARE_{it}$	-1.11	0.20	-5.67	<.0001
β_2	LEG_{it}	-3.07	0.26	-11.99	<.0001
β_3	SEA_{it}	0.96	0.13	7.2	<.0001
β_4	POP_{it}	2.38	0.37	6.41	<.0001
β_5	FS_{it}	0.29	0.15	1.91	0.0567
β_6	$SIZE_{it}$	1.17	0.37	3.16	0.0017
β_7	$SIZE_{it} * \ln(FS_{it})$	-0.11	0.03	-3.16	0.0017
β_8	DIS_i	0.00	.	.	.
β_9	ENP_{it}	0.40	0.13	3.2	0.0015
β_{10}	ES_{it}	-0.43	0.21	-2.1	0.0368
ρ_2	Autoregressive Parameter	0.56	0.05	-11.13	<.0001
Model fit statistics	$Est.Var(v_t)$	0.049			
	Regress R-Square	0.998			
	Total R-Square	1.000			
Durbin-Watson Test		2.008			

In Model 2, the variable ES_{it} has a negative impact on demand for the local OD pair. This means that the substitute airport with higher passenger traffic will attract more passengers from the local airport. Interestingly, the absolute value of its coefficient is close to that of ENP_{it}. In other words, the higher traffic at the local airport, the more passengers it will retain; the higher traffic at the substitute airport, the more passengers from the local airport will "leak" to the substitute airport. This verifies the existence of positive feedback effects at the local airport and the substitute airport, and the sensitivities of the demand on the local OD pair are similar to the total passenger enplanements at the local airport and the substitute airport. Comparing with Model 1-b, Model 2 has a higher coefficient of ENP_{it}. It indicates that the positive feedback effect has been underestimated in Model 1-b. Other coefficients have experienced small changes between Models 1-b and 2. Goodness of fit (i.e., R-squares value) for Model 2 is also similar to Model 1-b, and the result of the Durbin-Watson test also shows no autocorrelation exists in Model 2.

3.2.4. Discussion of Results

Model 1-a is a two-stage least squares (2SLS) model to capture the endogeneity between airfare and demand. Model 1-b has been improved by correcting first-order autocorrelation and heteroskedasticity in Model 1-a. Model 2 used the same estimation method of feasible generalized least squares (FGLS) as Model 1-b but introduced one more variable in the second-stage model. In the three models, all the signs of the estimated parameters change slightly.

All variables are significant at the 95% confidence level. Variables that impact airfare include the route indicator variables, passengers, LCC indicator variable, airfare at the substitute airport, and the product of unit fuel cost and non-stop miles. Passengers impact airfare negatively when higher passenger traffic intensifies airline competition which eventually leads to lower airfare. Low-cost carrier (LCC) availability decreases airfare. The airfare at the local airport will decrease if the airfare at the substitute airport decreases. The positive sign of the product of unit fuel cost and non-stop miles suggests that airfare will increase if unit fuel cost increase or non-stop miles is longer. By introducing a positive autoregressive parameter in Model 1-b and 2, the

R-square value of airfare model largely improves. To have a better understanding of airfare models in Model 1-b, airfare against passengers has been plotted in the figure below. The LCC indicator variable is set as zero. The lagged residual is set as zero because we are interested in impact of passengers on airfare in the same time period. Values of other variables are the mean values in the dataset for the Route JAX- PHL. Based on the results shown in Figure 3.7, airfare ranges from $164 to $178 when the number of passengers ranges from 2,000 to 8,000.

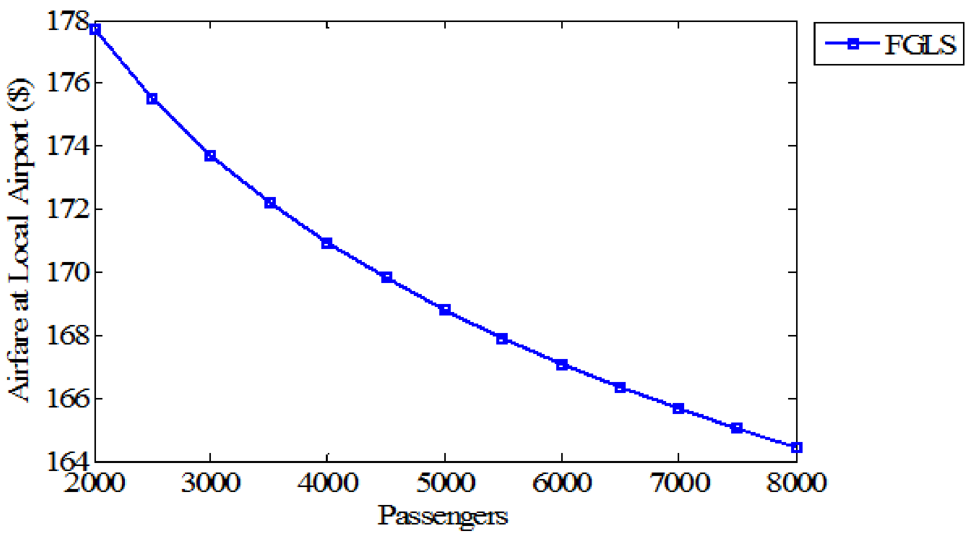

Figure 3.7 Relationship between airfare and passengers for local OD pair based on airfare model (Model 2)

Based on the result of Model 2, the local airport demand variables include the route indicator variables, airfare at the local airport, flight leg at the local airport, seasonality, population, driving distance, total enplanement of the local airport, total enplanement of the substitute airport, airfare at the substitute airport, group size, and the interaction variable of group size and airfare at the substitute airport. The estimated coefficients of the two enplanement variables reflect that the positive feedback effects exist at both the local airport and the substitute airport. The positive feedback effect means that an airport with higher passenger traffic will attract more passengers (Hansen, 1995). However, Model 1-b does not reflect the positive feedback effect at the substitute airport, and underestimates the positive feedback effect at the local airport. On the other hand, the results of Model 2 suggest that passengers may be attracted to the substitute airport not only because of lower airfare but also because more people are using the substitute airport. The attractiveness of lower airfare at the substitute airport will be eliminated if passengers travel in a group with more than three people. To show how autocorrelation would impact the estimation result, the relationship between the passengers and the airfare at the local airport as in two-stage least squares model (Model 1-a) and feasible generalized least squares model (Model 1-b) are presented and compared in Figure 3.8. The lagged residual in Model 1-b is set as zero because we are interested in the same time period. Other variables are the mean values of the 41 observations for the Route JAX- PHL. In Figure 3.8, the number of passengers based on Model 1-a is at a scale of 10^{13} which is not realistic. The number of passengers based on Model 1-b is much smaller at a scale of 10^4. The difference of Model 1-a and Model 1-b suggests that estimation result is misleading if autocorrelation is not considered.

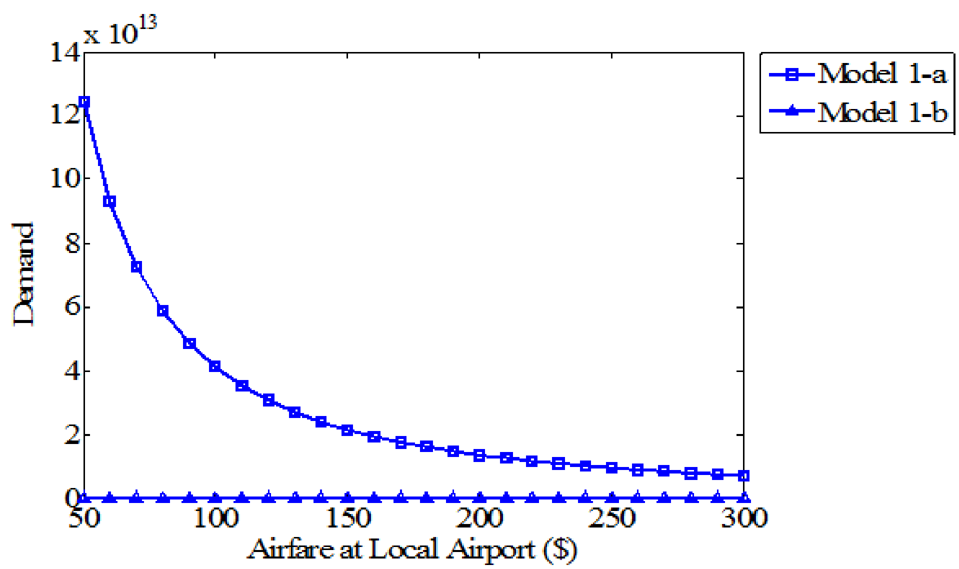

Figure 3.8 Relationship between passengers and airfare for local OD pair based on demand model (Model 1-a and Model 1-b)

A similar plot for Model 1-b and Model 2 are shown in Figure 3.9.

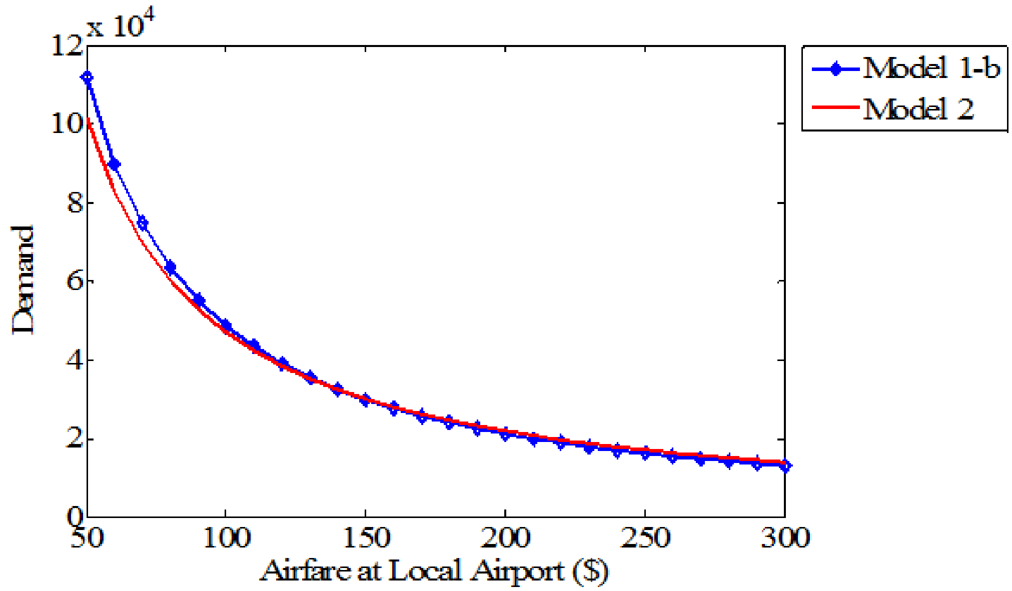

Figure 3.9 Relationship between passengers and airfare for local OD pair based on demand model (Model 1-b and Model 2)

Adding a variable in Model 2 only slightly changes the demand value comparing with Model 1-b. The number of passengers in Model 2 ranges from 1.7×10^4 to 1×10^5 when the airfare at the local airport ranges from $50 to $300. When the enplanement at the substitute airport is considered, the demand for local OD pair is less sensitive to the airfare at the local airport.

CHAPTER 4. SUPPLY-DEMAND EQUILIBRIUM MODEL

A binary logit model is used to determine the market shares of two airports for a population that is expected to be served by a local airport but may leak to a "substitute" airport outside the region under certain conditions. The airfare model in Chapter 3 is used to determine the airfare variable in the binary logit model. A numerical analysis is performed to explore the impact of variables and coefficients on equilibrium solutions.

4.1. Model Specification

Discrete choice models have been used extensively in describing passengers' airport choice behaviors, and airports' market share (de Luca, 2012; S. Hess, 2005; Hsiao, 2008; Warburg, Bhat, & Adler, 2006). The underlying objective of a discrete choice model is utility maximization, when there is a set of alternatives to choose from. A binary logit model was built to understand leakage from a local airport to a substitute airport. Also, the impact of airfare on demand has been considered.

4.1.1. Binary Logit Model Structure

Passenger's utility of choosing an airport is (Train, 2009) :

$$U_i = V_i + \varepsilon_i \tag{4-1}$$

Where

i indexes the airport, of which there are two such that $i = 1$ or 2.

U_i is the utility of choosing Airport i.

V_i is the deterministic utility of Airport i.

ε_i is the unknown part of utility that is not captured by V_i.

Passengers will choose Airport 1 when the utility of patronizing Airport 1 is higher than the utility of patronizing Airport 2. So the probability of choosing Airport 1 is $P_1 = \text{Prob}(U_1 \geq U_2)$, which can be further written as follows:

$$\begin{aligned} P_1 = \text{Prob}(U_1 \geq U_2) = \text{Prob}(\text{V}_1 + \varepsilon_1 \geq \text{V}_2 + \varepsilon_2) = \text{Prob}(\text{V}_1 + \varepsilon_1 \geq \text{V}_2 + \varepsilon_2) \\ = \text{Prob}(\varepsilon_2 - \varepsilon_1 \leq \text{V}_1 - \text{V}_2) \end{aligned} \quad \textbf{(4-2)}$$

By assuming that ε_1 and ε_2 follow the standard Gumbel distribution, their difference $\varepsilon_2 - \varepsilon_1$ follows the logistic distribution (Train, 2009). Derivation of this can be checked in Train (2009). Ultimately, the final expression for the probability of choosing Airport 1 can be written in closed form as:

$$P_1 = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)} \quad \textbf{(4-3)}$$

The probability of choosing Airport 2 (P_2) can be written in a similar fashion. We know that $P_1 + P_2 = 1$, because Airports 1 and 2 are the only airports in the choice set.

4.1.2. Supply-Demand Equilibrium Model Specification

A binary logit model has been developed to estimate airport leakage from a local airport (Airport 1) to a substitute airport (Airport 2). To make the substitute airport attractive to the "leakage" passengers, the substitute airport is supposed to provide services that are superior to those offered by the local airport. The airfare variable in the airport market share model is based on the airfare model from Chapter 3. This variable is designed to explain how demand affects airfare in the passenger-airfare relationships. Some assumptions have been made when specifying airport market share model and airfare model.

4.1.2.1. Airport Market Share Model

In this model, Airport 1, as mentioned in Equation 4-3, is treated as the local airport. The population for which this choice model applies is from the metropolitan area that is expected to be served by that local airport. But the other airport (Airport 2), located in a different metropolitan area, is also available for the passengers. Assuming that every individual in the population follows an identical airport choice pattern, the aggregated market share of the local airport (Airport 1) will be equal to the disaggregate probability of choosing the local airport. Thus, the market share of Airport 1 to a specific destination airport can be calculated in the following expression.

$$MS_1 = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)} \tag{4-4}$$

V_1 is the deterministic utility of Airport 1, and V_2 is the deterministic utility of Airport 2. The attributes in utility functions are different in disaggregate airport choice and aggregate airport market share. Generally, attributes that are specific to individuals, such as a passenger's experience, cannot apply to the airport market share model. In addition, for the important attributes such as airfare and ground access time, which vary with respect to individuals, the average values are normally used in aggregate airport market share models. In this model, three attributes that were found to be significant in previous studies for explaining airport choice are chosen, including airfare, flight frequency and ground access distance (S. Hess, 2005; S. Hess & Polak, 2010). Ground access distance is the distance from ground access origin (such as home) to the airport. The utility of departing from Airport i to the destination is

$$V_i = \alpha F_i + \beta \log(f_i) + \gamma \log(g_i) \tag{4-5}$$

Where

F_i is the average airfare from Airport i to the destination airport, i=1 or 2.

f_i is the flight frequency from Airport i to the destination airport, i=1 or 2.

g_i is the average ground access distance to Airport i for the population, i=1 or 2.

α, β and γ are parameters. They can be interpreted as weights of corresponding attributes in the utility function.

4.1.2.2. Airfare Model

The airfare term of Airport 1 (F_1) in Equation 4-5 is assumed to be a function of the number of passengers at Airport 1 on the subject origin-destination (OD) pair, as shown in Equation 4-6.

$$F_1 = f(PASS_1) \tag{4-6}$$

The airfare model is based on the results of the feasible generalized least squares model (Model 2) in the previous chapter, which indicates Equation 3-3. The lagged residual in autoregressive model is set as zero because we only consider the impact of the number of passengers on airfare in the same period. Route indicator variables and non-stop miles variable need to be specified with respect to a specific origin-destination (OD) pair. As a result, the OD pair from Jacksonville International Airport (JAX) to Philadelphia International Airport (PHL) is randomly chosen from the 10 local OD pairs in Table 3.1. Although we use the empirical model as defined for the route from JAX to PHL, the entire modeling exercise itself is based on a hypothetical situation. The airfare model for Airport 1 is below.

$$\begin{aligned} F_1 = \exp(2.75 - 0.06\ln(T \cdot MS_1) - 0.07 \cdot \mathrm{I}(LCC \\ = 1) + 0.33\ln(F_2) + 0.17\ln(FUEL * MILES)) \end{aligned} \tag{4-7}$$

Where

T is the total air passenger demand in the metropolitan region of Airport 1.

MS_1 is the market share of Airport 1. $T \cdot MS_1 = PASS_1$.

$PASS_1$ is the number of passengers departing from Airport 1 to the destination airport.

$\mathrm{I}(LCC = 1)$ is the low-cost carrier (LCC) indicator variable. $\mathrm{I} = 1$ if LCC are available at Airport 1; $\mathrm{I} = 0$ otherwise.

F_2 is average airfare from Airport 2 (i.e., the substitute airport) to the destination airport, in dollars.

$FUEL$ is unit fuel cost per gallon.

$MILES$ is non-stop miles of the flight from Airport 1 to the destination airport.

4.2. Numerical Analysis

If all the variables except MS_1 in Equation 4-7 are known, we are able to obtain a range of F_1 values by inputting MS_1, which is in the range of $0 \leq MS_1 \leq 1$. Then, by using the F_1 values in Equation 4-4, we will obtain the new values of MS_1, which are the output market shares of Airport 1. We say that an equilibrium condition exists when the output market share equals the input market share, because it is a closed loop feedback. In this section, the numerical analysis focuses on equilibrium market share given the values of variables and parameters. The following descriptions explain how the base values of parameters and coefficients are set.

1. Because we do not have empirical survey data to populate the model (Equation 4-5), the values of coefficients α, β and γ are taken from the literature (Brooke, Caves, & Pitfield, 1994; Caves, Ndoh, & Pitfield, 1991; Ndoh, Pitfield, & Caves, 1990; Pels, Nijkamp, & Rietveld, 2000).

2. Different from airfare term of Airport 1, the airfare term of Airport 2 (i.e., the substitute airport) is not in a function of the market share at Airport 2.This is because the demand "leaking" to the substitute airport is only a small part of total demand at the substitute

airport. The total demand will impact airfare at that airport. Thus, the average airfare from Airport 2 to the assumed destination is set as \$200 per passenger.

3. The non-stop miles variable ($MILES$) is set as 742 miles, which is the flight distance from JAX to PHL. The same value has been used to estimate the coefficients in Equation 4-7. Thus, the estimated airfare may be more accurate than others when $MILES = 742$ miles.

4. The flight frequency at Airport 1 is 100 flights per quarter, while the frequency at Airport 2 is 200 flights per quarter. This means that Airport 1 provides one flight per day to the assumed destination airport, while Airport 2 provides two flights per day to the destination.

5. Average ground access distance to Airport 1 is set as 30 miles. The Orlando International Airport (MCO) is the substitute airport for JAX. Based on the driving distance between JAX and MCO, which is 141 miles, the average ground access distance to Airport 2 is 171 miles.

6. The total air passenger demand from the metropolitan area of Airport 1 to the destination is assumed to be 20,000 passengers in a quarter. By assuming that one aircraft provides 200 seats on average, Airport 1 has the capacity to satisfy the total demand.

7. The local airport is not a low-cost carrier (LCC) hub, and $\mathrm{I} = 0$.

8. The unit aviation fuel cost is based on the record for 2013, which is around \$3/gallon (Bureau of Transportation Statistics, U.S. Department of Transportation, 2014c).

The parameters, descriptions, and their base values are listed in the table below. The variables and coefficients will be set to different values to assess how sensitive the market share equilibrium is to these parameters. Without specifying the exact values, the inputs of other parameters are the base values in Table 4.1.

Table 4.1 Explanation and Base Values of Parameters in Utility Function and Airfare Function

Utility Function		
Notation	*Description*	*Base Value*
α	Coefficient of Airfare	-0.04
β	Coefficient of Frequency	1.15
γ	Coefficient of Ground Access Distance	-0.04
f_1	Flight Frequency at Airport 1 per Quarter	100
f_2	Flight Frequency at Airport 2 per Quarter	200
g_1	Ground Access Distance to Airport 1 (miles)	30
g_2	Ground Access Distance to Airport 2 (miles)	171
Airfare Function		
Notation	*Explanation*	*Base Value*
T	Total Passenger Demand	20,000
$\mathrm{I}(LCC = 1)$	LCC Indicator Variable, for Airport 1	0
F_2	Airfare at Airport 2 ($)	200
$FUEL$	Unit Fuel Cost ($/gallon)	3
$MILES$	Non-stop Miles of Flight form Airport 1 to the Destination (miles)	742

According to Equation 4-7, the value of F_1 ranges from $190 to $230 when MS_1 ranges from zero to one. Meanwhile, other variables are set the base values in the airfare function. When $MS_1 = 0.46$, F_1equals to F_2, which is $200.

4.2.1. Impact of Airfare Coefficient (α) in Utility Function

In Figure 4.1, the solid line with no markers is the 45^o reference line. An equilibrium exists when the output market share (i.e., y-axis) equals to the input market share (i.e., x-axis), because, by definition, they are the same value. Therefore, the equilibrium exists where each curve intersects with the 45^o reference line. Only the equilibrium points reflect real situations, so we are only interested in how equilibrium points change when the value of a specific coefficient or variable changes. There are two types of equilibrium points: stable equilibrium and unstable equilibrium. "If at the intersection point the curve cuts the 45^o line from above as MS_1 increases, the equilibrium is stable" (Hansen, 1995). The intuition for this is that after any derivation from the stable equilibrium, the market share will return to the stable equilibrium (Sharov, 1996). On the contrary, after any derivation from the unstable equilibrium, the market share will never return to the stable equilibrium (Sharov, 1996).

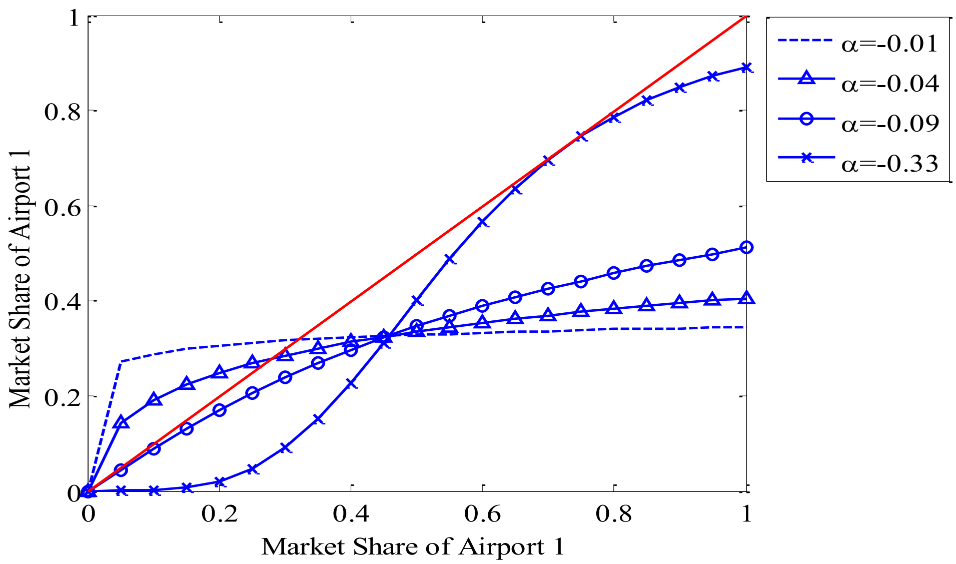

Figure 4.1 Equilibria under alternative airfare coefficients in utility function

The coefficient α is given four different values of -0.01, -0.04, -0.09, and -0.33, as shown in Figure 4.1. Only negative values of α are applied, because the sensitivity of utility to airfare is normally negative. Such a negative relationship can be interpreted to mean that an airport with higher airfare reduces the probability of passengers choosing this airport. $\alpha = -0.01$ is set to investigate equilibrium market share when α is larger than -0.04. $\alpha = -0.04$ is the base value from Table 4.1.There are two possible equilibrium points when α is larger than (approximately) -0.09. When α is smaller than -0.09, the number of equilibrium points drops to one. However, if α decreases to (approximately) -0.33, there will be two possible equilibrium points again. When α is smaller than -0.33, the number of equilibrium points increases to three.

Two equilibrium points exist when $\alpha = -0.01$. One is a stable equilibrium at (approximately) $MS_1 = 0.32$, and the other is an unstable equilibrium at $MS_1 = 0$. When $\alpha = -0.04$, the unstable equilibrium stays at $MS_1 = 0$, but the market share of Airport 1 at stable equilibrium reduces. In consideration that the absolute value of α represents how much consideration a passenger gives to airfare when choosing an airport (i.e., the weight of the airfare in utility), the market share of Airport 1 reduces when the weight of airfare increases (meanwhile $\alpha > -0.09$). In addition, based on the values of stable equilibrium when $\alpha > -0.09$, the airfare at Airport 1 (F_1) is higher than the airfare at Airport 2 ($200), according to Equation 4-7. It can be interpreted to mean that more passengers will leak to the substitute airport when airfare is increasingly important to passengers (under the circumstances that the substitute airport provides lower airfare than the local airport). When $-0.33 < \alpha < -0.09$, only one stable equilibrium exists at $MS_1 = 0$. This means that all passengers will leak to the substitute airport when the weight of airfare is in a specific range (under the circumstances that airfare at the substitute airport is $200 while the airfare at the local airport is $230). When $\alpha < -0.33$, there are three equilibrium points. Two stable equilibrium points exist at $MS_1 = 0$ and $MS_1 > 0.7$ respectively, while one unstable equilibrium exists at $0 < MS_1 < 0.7$. The stable equilibrium at $MS_1 > 0.7$ will increase when α decreases. In addition, based on the values of stable equilibrium when $\alpha < -0.33$, the airfare at Airport 1 (F_1) is lower than the airfare at Airport 2 ($200). This means that more passengers will use the local airport when airfare is increasingly important to passengers (under the circumstances that the local airport provides lower airfare than the substitute airport).

In conclusion, the airport with lower airfare always has an advantage in airport market share. This advantage will be magnified when passengers consider airfare to be more important when choosing an airport. Normally, airfare is more important to leisure passengers than business passengers, which is evident in the fact that leisure passengers are more likely to leak to the substitute airport when it provides lower airfare. However, if the local airport provides lower airfare, more leisure passengers will be retained at the local airport.

4.2.2. Impact of Frequency Coefficient (β) in Utility Function

In the utility function, coefficient β is given three values of 0.01, 1.15, and 2.90. The positive sign of β is fixed, assuming that frequency contributes positively to utility in Equation 4-5.

$\beta = 0.01$ is to show the equilibrium results when β is much lower than the base value. $\beta = 1.15$ is the base value from Table 4.1. There are two possible equilibrium points when β is smaller than (approximately) 2.90. When β is larger than 2.90, the number of equilibrium points drops to one. The equilibrium results are shown in the figure below.

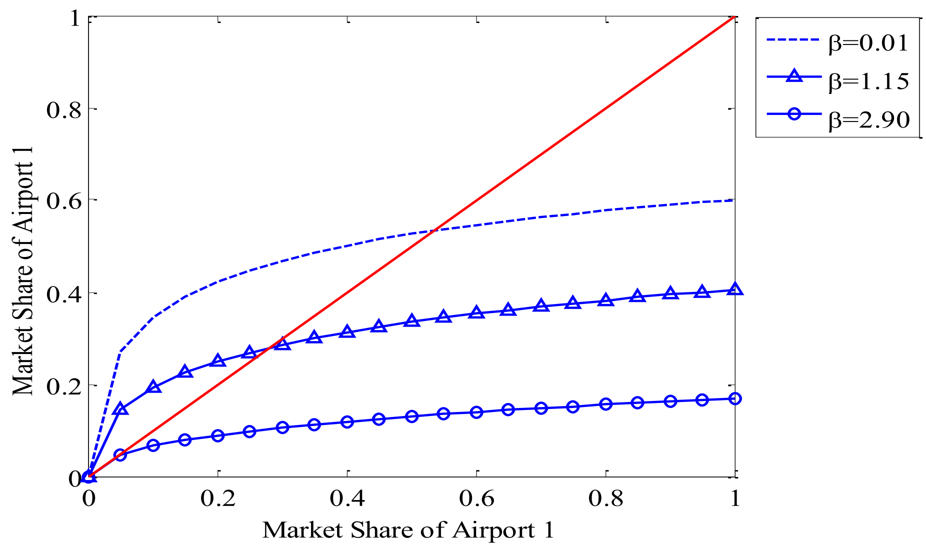

Figure 4.2 Equilibria under alternative frequency coefficients in utility function

Two equilibrium points exist when $\beta = 0.01$. One is a stable equilibrium at (approximately) $MS_1 = 0.55$, and the other is an unstable equilibrium at $MS_1 = 0$. The corresponding airfare at Airport 1 (\$190) is smaller than the airfare at Airport 2 (\$200) when $MS_1 = 0.55$. When β increases, the market share of Airport 1 at the stable equilibrium decreases. Because the absolute value of β represents the weight of frequency in utility, we can conclude that the market share of Airport 1 reduces when the weight of frequency increases (under the circumstance that frequency at Airport 1 is 100 while frequency at Airport 2 is 200). It should also be noticed that airfare at Airport 1 is also changing with respect to different equilibrium points. When $\beta > 2.90$, only one stable equilibrium exists at $MS_1 = 0$, meaning that all passengers will use the substitute airport (Airport 2) when frequency is very important to them (under the circumstance that the frequency at Airport 1 is 100, the frequency at Airport 2 is 200, the airfare at Airport 1 is \$230, and the airfare at Airport 2 is \$200).

In conclusion, this advantage of Airport 2 with higher flight frequency will be magnified when passengers consider frequency to be more important when choosing an airport. Meanwhile, in the long term, when Airport 1 has a lower market share, the airfare at Airport 1will increase, which may further reduce its market share. Normally, business passengers are more sensitive to flight frequency. Thus, it is important for the local airport to know the ratio of business passengers in its market.

4.2.3. Impact of Ground Access Distance Coefficient (γ) in Utility Function

As shown in Figure 4.3, the coefficient γ is given three values of -0.01, -0.04, and -3.50. Only negative values of γ are used, because the sensitivity of utility to ground access distance is negative. In other words, a longer ground access distance reduces the utility and reduces the probability of choosing an airport. $\gamma = -0.01$ is to show the results when γ is larger than the base value. $\gamma = -0.04$ is the base value from Table 4.1. A third value of $\gamma = -2.90$ is given to show the results when γ is smaller than the base value.

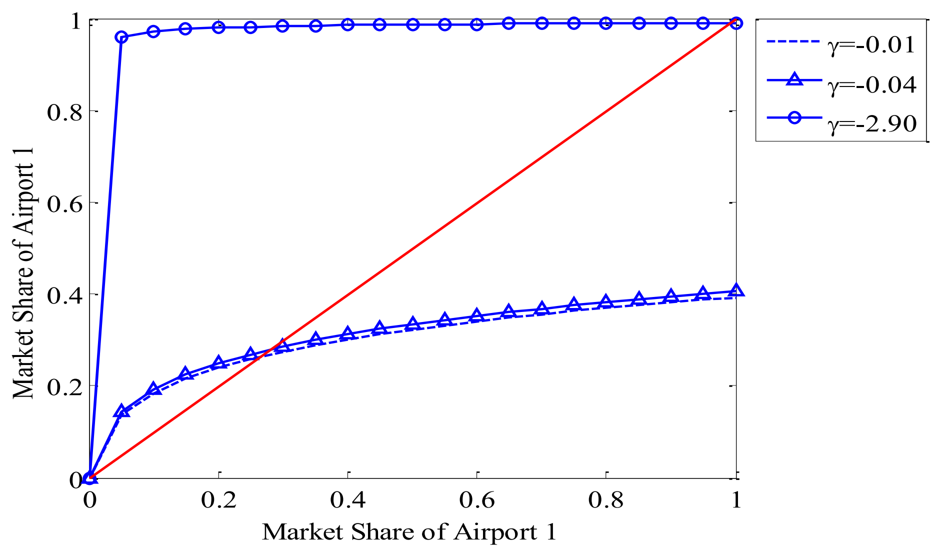

Figure 4.3 Equilibria under alternative ground access distance coefficients in utility function

There are two equilibrium points when $\gamma = -0.01$. One stable equilibrium exists at $MS_1 = 0.28$ and one unstable equilibrium exists at $MS_1 = 0$. The stable equilibrium changes very slightly when γ decreases to -0.04. At its stable equilibrium point, the airfare at Airport 1 is \$205 according to Equation 4-7. This means that around 72% of passengers in the metropolitan area of Airport 1 will leak to Airport 2 when airfare at Airport 1 is \$205 (and, as stated earlier, the airfare at Airport 2 is \$200, the frequency at Airport 1 is 100 flights per quarter, the frequency at Airport 2 is 200 flights per quarter, the ground access distance to Airport 1 is 30 miles, and the ground access distance to Airport 2 is 171 miles). When γ decreases, the market share of Airport 1 at the stable equilibrium increases. We can conclude that the market share of Airport 1 reduces when the weight of ground access distance increases (and, of course, the average ground access distance to the local airport is shorter than the distance to the substitute airport). It should also be noted that the airfare at Airport 1 also changes with respect to different equilibrium points. The market share of Airport 1 reaches one when $\gamma = -2.90$, meaning that there is no airport leakage when the weight of ground access distance is very high.

In conclusion, local airport is more likely to attract passengers that treat ground access distance as an important factor of their airport choices. If the local airport is able to increase its market share, it will provide lower airfare in the long term. Normally, business passengers are more sensitive to ground access distance. Thus, it is important for the local airport to know the ratio of business passengers in its market. Meanwhile, in the raining or snowing seasons, passengers are more likely to patronize the local airport.

4.2.4. Impact of Airfare at Substitute Airport (F_2) in Utility Function and Airfare Function

As shown in Figure 4.4, three values of \$150, \$200, and \$350 for F_2 have been provided. $F_2 = \$150$ is chosen to reflect the equilibrium when F_2 is lower than the base value. $F_2 = \$200$ is the base value from Table 4.1. $F_2 = \$300$ is chosen to reflect the equilibrium when F_2 is higher than the base value. There is only one equilibrium that is stable at $MS_1 = 0$ when $F_2 < \$150$. When F_2 exceeds \$150, the number of equilibrium points increases to two. One stable equilibrium exists at $MS_1 > 0$ and one unstable equilibrium exists at $MS_1 = 0$.

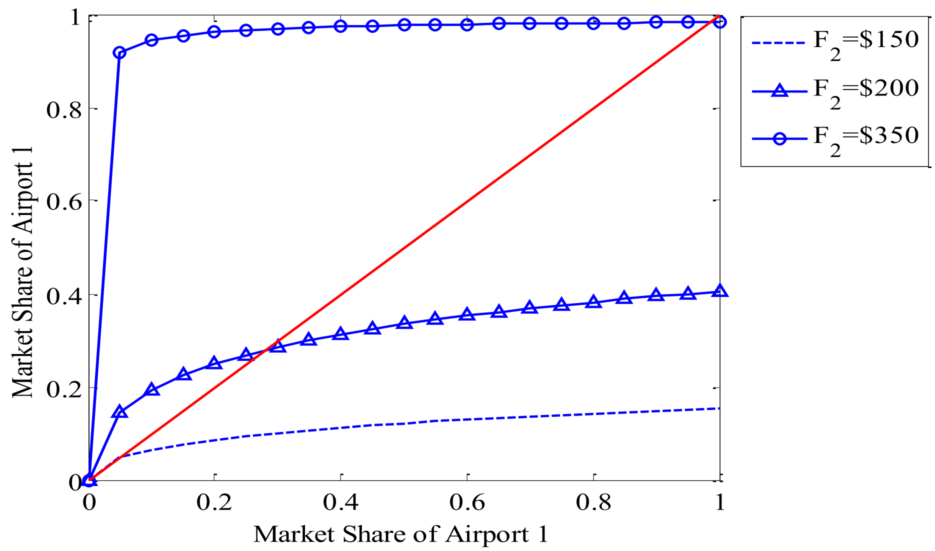

Figure 4.4 Equilibria of market share at Airport 1 under alternative airfares at Airport 2

With the increase of F_2 (and $F_2 > \$150$), the market share at stable equilibrium point also increases. Such a result not only depends on the impact of F_2 in the utility function (Equation 4-5) but also depends on the combined effect of F_2 and MS_1 on F_1 in the airfare function (Equation 4-7). Based on Equation 4-7, changes of F_2 will lead to different values of F_1. To have a better understanding of how F_2 impacts F_1 and further impacts equilibrium, the values of F_1 for different F_2 are plotted in Figure 4.5.

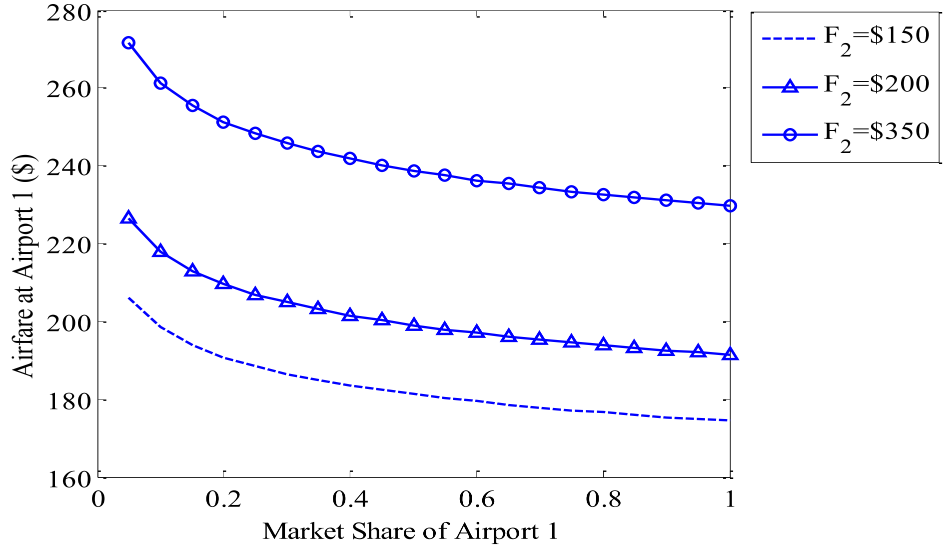

Figure 4.5 Relationship between airfare at Airport 1 (F_1) and market share of Airport 1 (MS_1) under alternative airfares at Airport 2 (F_2)

Based on Figure 4.5, when $F_2 = \$150$, F_1 is always higher than \$170. When $F_2 = \$200$, F_1ranges from \$190 to \$230. So when $F_2 = \$200$, whether F_1 is larger than F_2 depends on the value of input market share. When $F_2 = \$350$, F_1 is always lower than \$230. Comparing the three cases, the change rate of F_1 is smaller than F_2.

In Figure 4.4, when F_2 is no higher than \$150, all passengers will leak to Airport 2, under the circumstance that the airfare at Airport 1 (F_1) is higher than \$170. When F_2 increases, F_1 will also increase but at a slower rate. Thus, the difference between the airfares at the two airports is reduced, and Airport 1 will attract more passengers. If F_2 increases to \$350, F_1 will be much lower than F_2. As a result, all the passengers will use the local airport.

In conclusion, when airline competition is intense at the substitute airport, or low-cost carriers are available at the substitute airport, the average airfare at the substitute airport is likely to be lower than at the local airport. The local airports may introduce more airlines to retain its market share.

4.2.5. Impact of Frequency Variable (f_1) in Utility Function

Three values have been given for f_1. This is to reflect the impact that f_1 has on equilibrium when it is larger than, equal to and smaller than f_2. The three values are 33, 100, and 200. When f_1 is smaller than 33, there is one equilibrium point. When f_1 exceeds 33, the number of equilibrium points increases to two. $f_1 = 100$ is the base value from Table 4.1. When $f_1 = 200$, the flight frequencies at Airport 1 and Airport 2 are the same. The equilibrium results are shown in the figure below.

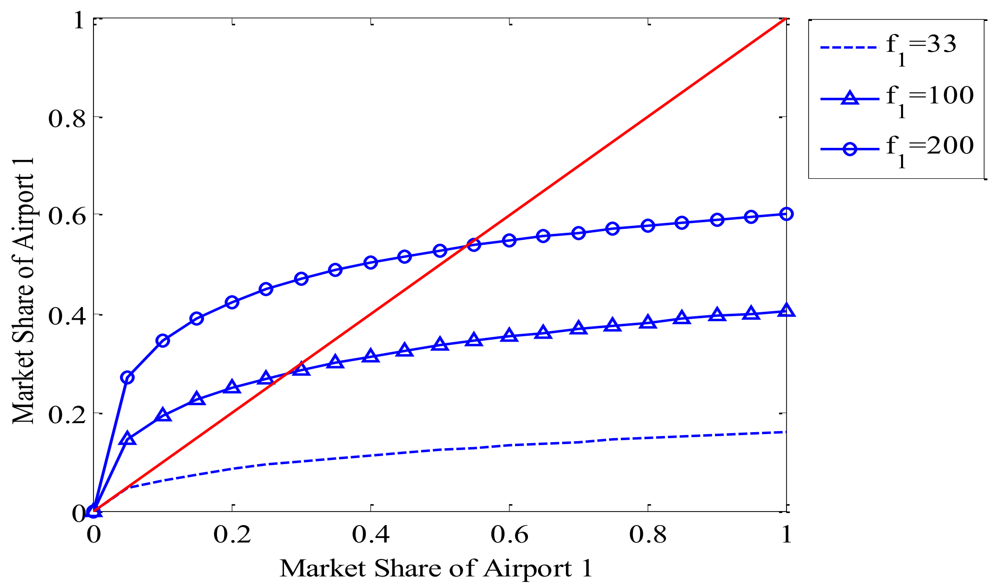

Figure 4.6 Equilibria under alternative flight frequency at Airport 1

Only one stable equilibrium exists at $MS_1 = 0$ when $f_1 < 33$ and $f_2 = 200$. This means that all passengers at the local airport will "leak" if the airport's flight frequency is much smaller than that of the substitute airport. At that point ($MS_1 = 0$), the airfare at Airport 1 is \$230 while the airfare at Airport 2 is \$200. If f_1 increases, the market share of Airport 1 at the equilibrium will increase. If the frequency is the same for the two airports when $f_1 = 200$ and $f_2 = 200$, the stable equilibrium exists at (approximately) $MS_1 = 0.55$, meaning around 45% of the market will "leak" to the substitute airport in the long term when the flight frequencies at Airport 1 and Airport 2 are 200 (and the airfare at Airport 1 is around \$200, the airfare at Airport 2 is \$200, the ground access distance to Airport 1 is 30 miles, and the ground access distance to Airport 2 is 171 miles).

4.2.6. Impact of Frequency Variable (f_2) in Utility Function

Three values of f_2 have been set as 100, 200, and 600. $f_2 = 100$ when the flight frequencies at Airport 1 and Airport 2 are the same. $f_2 = 200$ is the base value from Table 4.1. When $f_2 < 600$, two equilibrium points exist. When f_2 exceeds 600, the number of equilibrium points drops to one. The equilibrium results are shown in the figure below.

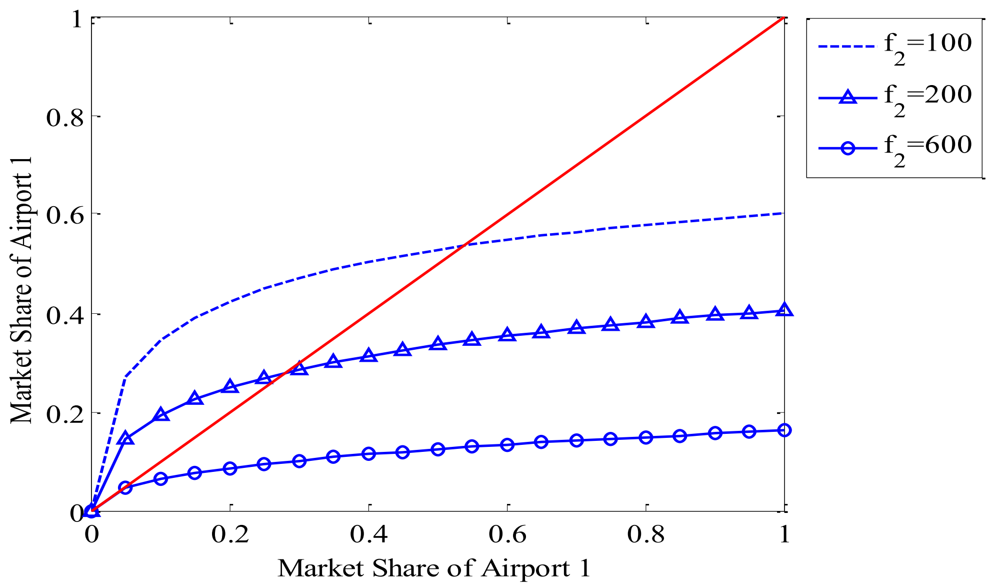

Figure 4.7 Equilibria under alternative flight frequency at Airport 2

When $f_2 < 600$, one unstable equilibrium exists at $MS_1 = 0$ and one stable equilibrium exists at $MS_1 > 0$. The market share of Airport 1 at stable equilibrium decreases when f_2 increases. This means that when Airport 2 provides higher flight frequency, it will attract all of the passengers from Airport 1. When $f_2 = f_1 = 100$, the equilibrium market share is around 0.55, which is the same equilibrium value when $f_2 = f_1 = 200$. By setting other values, it is observed that the equilibrium does not change when the frequency at Airport 1 equals the frequency at Airport 2.

In conclusion, passengers at the local airport will "leak" if the substitute airport provides flights with higher frequency. With fewer passengers using the local airport, the average airfare at that airport will increase in the long term, which will further reduce its market share. In order to retain the market share, the local airport needs to be sensitive to the frequency changes at the substitute airport, and to make sure that airlines at the local airport provide sufficient flights for different destinations.

4.2.7. Impact of Ground Access Distance Variable (g_1 and g_2) in Utility Function

Five sets of values for g_1 *and* g_2 have been provided to show market share equilibria of five locations around Airport 1. Their locations are shown in Figure 4.8.

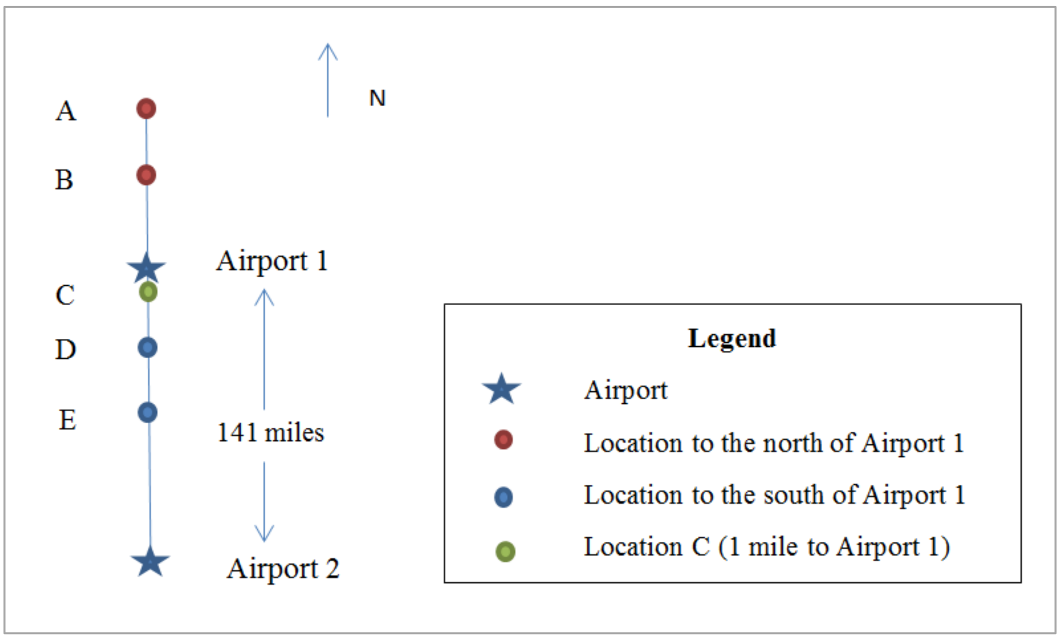

Figure 4.8 Five locations used to show impact of ground access distance on equilibria

The star on the north is Airport 1 and the other star is Airport 2. The five locations are the five circles in the figure, labeled A to E. Location A and Location E are 30 miles from Airport 1 while Location B and Location D are 21 miles from Airport 1. Location C is 1 mile to the south side of Airport 1. Because the distance between Airport 1 and Airport 2 are assumed to be 141 miles, the ground access distance from Location A to Airport 2, minus the ground access distance from Location A to Airport 1, should be equal to 141 miles. Thus, the ground access distances to Airport 1 (g_1) and Airport 2 (g_2) from Location A are 30 miles and 171 miles respectively. For Location E, the sum of its ground access distances to Airport 1 and Airport 2 is equal to 141 miles. Thus, the ground access distances to Airport 1 (g_1) and Airport 2 (g_2) from Point E are 30 miles and 111 miles respectively. The equilibrium results are shown in Figure 4.9.

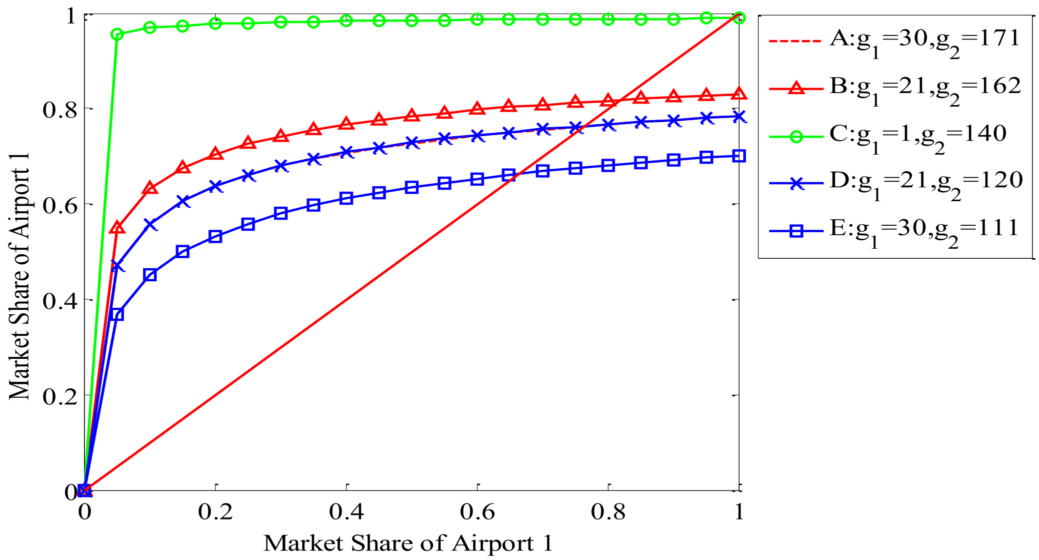

Figure 4.9 Equilibria under alternative locations and ground access distances

Two equilibrium points exist for all five locations. One unstable equilibrium exists at $MS_1 = 0$ and one stable equilibrium exists at $MS_1 > 0.65$. The market share of Airport 1 at the stable equilibrium is approximately one for Location C ($g_1 = 1$ and $g_2 = 140$). Location C is very close to Airport 1 and can be treated as the center of the catchment area of Airport 1. When the ground access distance to Airport 1 (g_1) increases, market share of Airport 1 at stable equilibrium decreases. It is consistent with findings in previous studies that as the radius around an airport spreads, market share reduces (Fuellhart, 2007; Lieshout, 2012). However, the reduction rate of market share is asymmetric on the two sides of the airport. As shown in Figure 4.9, the curves of Location A and Location D overlap, and the two locations have the same equilibrium point. However, compared with Location A, Location D is closer to Airport 1. We can conclude that the reduction rate of such market share on the south side of Airport 1 is higher than on the north side. This is because for locations on the south side of Airport 1, when the ground access distance to the airport (g_1) increases, the ground access distance to Airport 2 (g_2) decreases.

In conclusion, passengers who live between the local and substitute airport are more likely to use the substitute airport than passengers in other locations. This research has shown that an existing local airport should provide lower airfare and higher flight frequency to attract passengers. Municipalities planning to build a medium-size airport should consider major hub airports in other areas that are reachable to the market.

4.3. Discussion

A binary logit model has been specified to determine the market shares of two airports. The airfare model in Chapter 3 is used to determine the airfare variable in the binary logit model. Log-forms of the frequency variable and the ground access distance variable are applied in the utility functions while the airfare variable is in linear form. The log-form indicates that the impact of a variable changes very slowly (to a point where it does not change at all) when its value exceeds a critical value. Many studies have verified that frequency has such a relationship with utility (de Luca & Di Pace, 2012; Harvey, 1987; S. Hess & Polak, 2005a; S. Hess, Adler, & Polak, 2007). However, the form of ground access distance (or ground access time) variable varies (Lian & Rønnevik, 2011; Phillips et al., 2005; Suzuki et al., 2003), and whether log-form or linear-form is more appropriate for this variable needs more work.

The numerical analysis is based on some data from previous studies (Brooke et al., 1994; Caves et al., 1991; Ndoh et al., 1990; Pels et al., 2000), some artificially constructed values, and distances based on the case of the Jacksonville International Airport (JAX) in Florida as discussed in Chapter 3. The input values are chosen to reflect the advantage of airfare and flight frequency at Airport 2, which may attract passengers from Airport 1. Parameters in the utility functions are given different values to show the sensitivity of equilibrium market share with respect to coefficients, airfare, frequency and ground access distances.

The coefficients in the utility function (Equation 4-5) can be treated as the weight of the corresponding variable. It was found that the advantage of lower average airfare for either a local airport or substitute airport will be magnified, when passengers consider airfare to be more important when choosing an airport (i.e., the weight of airfare increases). If the weight of flight frequency increases, the advantage of higher flight frequency at Airport 2 will be magnified. Meanwhile, in the long term, when Airport 1 has a lower market share, its airfare will increase, which may further reduce its market share. If the weight of ground access distance increases, the advantage of a local airport (Airport 1) is magnified. In this case, when Airport 1 has a higher market share, its airfare will decrease in the long term, which may further increase its market share. The increase of airfare at Airport 2 leads to a higher market share for Airport 1. Because of the positive frequency coefficient in the utility function, when Airport 1 provides higher flight frequency, its market share increases. When the frequencies at Airport 1 and Airport 2 are the same, Airport 1's market share at equilibrium remains at (approximately) $MS_1 = 0.55$. Five locations on the north and south sides of Airport 1 have been chosen to show how the ground access distance impacts equilibrium. When the radius around Airport 1 spreads, the market share reduces. However, the reduction rate is asymmetric on the north and south sides because the ratio of ground access distances to Airport 1 and Airport 2 are asymmetric in the two directions.

In future work, this type of model may be populated by real data collected through a survey of air passengers. Firstly, the forms and coefficients of variables in the utility functions need to be verified using survey data. Secondly, the airfare model is obtained for 10 U.S. airports as shown in Chapter 3. If airport leakage is identified through a survey for another airport, the airfare model should be rebuilt. Thirdly, when the equilibrium model is applied to a specific airport, the boundary of the area served by that airport can be explored further.

CHAPTER 5. CONCLUSIONS AND DISCUSSIONS

This chapter provides an overview of this research, and summarizes the major findings and contributions of the previous two chapters. The limitations of this research are discussed, along with suggestions for future work in data collection and model building.

5.1. Research Overview

The overall objective of this research is to investigate what factors affect airport leakage and how they affect airport leakage, in the context of models that consider the two-way interactions between air transportation demand and supply. These included empirical instrumented models that were estimated using two-stage least squares (2SLS) and feasible generalized least squares (FGLS) methods, as well as a theoretically-derived equilibrium model based on a binary logit specification. The focus of the empirical models is to find variables that are significant by replacing the endogenous airfare variable with an instrumental variable in an airport demand model. The focus of the supply-demand equilibrium model is to find equilibrium solutions when considering airfare (supply) and airport market share (demand) endogeneity.

In the empirical model in Chapter 3, 10 medium-size airports were identified as “local” airports in the airport leakage problem. This is based on a number of criteria. These criteria include: 1) the substitute airport should belong to Operational Evolution Partnership (OEP) 35; 2) the average airfare at the substitute airport should be lower than at the local airport; 3) the average flight leg at the substitute airport should be lower than at the local airport; 4) 50 percent more passengers should be using the substitute airport than the local airport to the destination; 5) the air trip needs to be over 500 miles; 6) the destination should not be a low-cost carrier (LCC) hub. There are three additional selection criteria: that there are sufficient observations in the dataset, that every local airport has only one destination, and that every local airport has only one substitute airport. The two-stage least squares (2SLS) and feasible generalized least squares (FGLS) models first estimated airfare by passengers and other attributes of the air trip, and then input the predicted airfare (instrumental variable) into the demand model. The 2SLS model is based on a previous study (Suzuki & Audino, 2003). The FGLS models are able to correct the autocorrelation and heteroskedasticity in the 2SLS model. Significant variables that impact passengers include airfare at the local airport, airfare at the substitute airport, and the driving distance between the local and substitute airports. Other variables include route indicator variables, the flight leg at the local airport, seasonality, population, group size, total passenger enplanement at the local airport, total passenger enplanement at the substitute airport, and the interaction variable of group size with airfare at the substitute airport.

The supply-demand equilibrium model in Chapter 4 applied the airfare function from the FGLS model and combined it with a binary logit model. The binary logit model is able to estimate market share for each of the two airports, assuming that all passengers choose their airport to maximize utility. Airfare, flight frequency, and ground access distance were considered as the three variables in the deterministic utility function. The total market in this model is a population that is expected to use the local airport (Airport 1). Because the input in the model is the market share of Airport 1 and the output is also the market share of Airport 1, equilibrium exists when the output equals the input. Unstable and stable equilibria were obtained when coefficients and variables in the binary logit model were set to different values. This shows the sensitivity of the equilibrium market share with respect to airfare, flight frequency and ground access distance coefficients, airfare and frequency at the substitute airport, frequency at the local airport, and different combinations of ground access distances. The coefficient in the utility function represents the weight of the corresponding variable in the utility function. Five sets of ground access distances were chosen to represent locations on the north and south sides of Airport 1. The equilibrium results greatly depend on the values that are assumed in numerical analysis.

5.2. Research Findings

Both models from Chapters 3 and 4 show that major hub (or substitute) airports will impact the demand at medium-size local airports. This finding further supports the hypothesis of this research that airport leakage exists, when there are major hub (or substitute) airports near metropolitan regions served by medium-size airports, and these hub airports provide better air services. In addition, a variety of factors were found to affect demand at local airports, such as airfare, ground access distance, enplanement, and so on. All factors that impact airport leakage are listed in Table 5.1. This table shows how the demand or market share at the local airport will change, when each factor changes.

Table 5.1 Factors Impacting Demand at the Local Airport

Feature	If feature should	Then local airport demand will
Airfare at local airport	↑	↓
Airfare at substitute airport	↑	↑
Average group size	↑	↑
Population in metropolitan area of local airport	↑	↑
Seasonal fluctuation of passenger traffic to the destination airport	↑	↑
Total passenger enplanement at local airport	↑	↑
Total passenger enplanement at substitute airport	↑	↓
Average flight leg at local airport	↑	↓
Weight of airfare	↑	↓ / - / ↑
Weight of flight frequency	↑	↓
Weight of ground access distance	↑	↑
Flight frequency at local airport	↑	↑
Flight frequency at substitute airport	↑	↓
Ground access distance to local airport	↑	↓
Ground access distance to substitute airport	↑	↑

It was found in the empirical models of Chapter 3, if a substitute airport provides lower airfare, the demand at the local airport will decrease. Alternately, if airfare at the substitute airport increases, more passengers will use the local airport. The positive impact of the travelers' group size shows when a larger group is travelling, passengers prefer to use a local airport. Dividing the parameter of airfare at the substitute airport by the parameter of the interaction variable of the group size and airfare at the substitute airport shows that the positive impact of airfare at the substitute airport will be eliminated when there are more than three people in the group. In other words, lower airfare at the substitute airport is not as attractive to passengers from the local airport when passengers travel in a group of three or more.

In addition to the airfare at a substitute airport, total enplanement at a substitute airport was found to impact the demand for a local airport. Furthermore, this impact is negative. Enplanement at an airport has a positive impact on that airport's demand. This verifies the existence of positive feedback effects at the local airport and the substitute airport. In other words, the higher the traffic at a local airport, the more passengers the airport will retain; the higher the traffic at a substitute airport, the more passengers the substitute airport will attract.

In the numerical analysis of the supply-demand equilibrium model, passengers may be attracted to the substitute airport to take advantage of lower airfare and higher flight frequency. If the substitute airport reduces its airfare, the airfare at the local airport will also reduce. As a combination effect of the two airfares, the equilibrium market share changes. Similarly, if the substitute airport provides higher flight frequency, more passengers will "leak" to the substitute airport from the local airport. In the long term, the average airfare at the local airport will increase, which will further reduce the market share at the local airport. In addition, it was found that market shares are different for locations even if their ground access distances to the local airport are identical.

5.3. Research Contribution

There are three contributions in this research.

- It has demonstrated that airport leakage exists when a major hub (or substitute) airport is located within a reasonable driving distance of the metropolitan region of a local airport, and provides lower airfare, higher flight frequency, and more direct flights. The sensitivity of airport demand (or airport leakage) with respect to a variety of factors have been tested in this research.

- The interaction between airfare (supply) and demand has been considered through a feasible generalized least squares (FGLS) model, and a supply-demand equilibrium model.

- FGLS estimation was used to understand the interaction between airfare and air passengers because autocorrelation (i.e., time serial correlation) and heteroskedasticity (i.e., the variance of error term is unequal with respect to the time variable) are present.

5.4. Limitations of the Research

The major limitations of this research are listed below.

- The empirical model focused on two competing airports in two metropolitan regions. In the origin-destination selection, simplification has been made by considering only one substitute airport. However, this cannot be applied to all the cases of airport leakage. There are local airports that are competing with two or more airports. These substitute airports may be located in one metropolitan region (i.e., a multi-airport region) or in different metropolitan regions. In either case, the airport should be studied in a different model.

- Airport leakage was identified, based on certain selection criteria, at 10 local airports in the United States. However, whether the criteria are sufficient to support the existence of airport leakage is unknown.

- All the 10 local airports were assumed to be independent in the 2SLS model and the FGLS models. However, some of these 10 airports have the same substitute airport. Whether this fact will impact the result is unknown.

- Because both the airfare and passenger models have route indicator variables, the models cannot be used for other routes.

- Due to the time period unit of data, airfare and passengers models were estimated on a quarterly basis, it may be biased to use yearly or monthly data. In addition, population and income data are only available in years, which may impact the model estimation result.

- The supply-demand equilibrium model was analyzed numerically with assumptions. Although some assumptions are based on previous studies, the findings from the equilibrium model may vary with different values of coefficients and variables.

5.5. Future Work and Recommendations

Future work can be conducted, including an air passenger survey. Through this survey, we will obtain information about passengers' ground access origins, to identify whether airport leakage exists for the subject airport. Then, an airfare and airport passenger model can be built specifically for airports where airport leakage has been observed. Furthermore, models can be built differently for business travelers and leisure travelers if the trip purpose is investigated in the survey. Furthermore, survey data are helpful to estimate coefficients in a binary logit model or other discrete choice models. In this research, more attributes can be considered in the utility function in addition to airfare, flight leg and ground access distance. Meanwhile, real values will improve equilibrium results.

More research opportunities will be created by different combinations of approaches. In the study of the supply-demand equilibrium model, we are able to obtain the market share equilibria for five locations around the local airport. In combination with geographical software, the supply-demand equilibrium model will be able to show the distribution of market share in the entire metropolitan area of a local airport. Based on the literature review, a spatial competition model assumes even distribution of the market in the airport catchment area. Spatial competition model may be improved in combination with geographical approaches. Also, a spatial competition model can combine with a discrete choice model to account for characteristics of the market when highlighting the impact of airport accessibility on airport demand. In consideration of passengers' spreading out to airport competitors when airlines also compete, the geographical approach is able to show market distribution. Meanwhile, game theory with mathematical optimization and a discrete choice model is able to show the supply-and-demand interaction.

REFERENCES

Adler, N. (2001). Competition in a deregulated air transportation market. *European Journal of Operational Research, 129*(2), 337-345.

Adler, N., Pels, E., & Nash, C. (2010). High-speed rail and air transport competition: Game engineering as tool for cost-benefit analysis. *Transportation Research Part B: Methodological, 44*(7), 812-833.

Ayyangar, L. (2007). Skewness, multicollinearity, heteroskedasticity - you name it, cost data have it! Solutions to violations of assumptions of ordinary least squares regression Models Using SAS. Retrieved Dec. 06, 2014, from http://www2.sas.com/proceedings/forum2007/131-2007.pdf

Barbot, C. (2009). Airport and airlines competition: Incentives for vertical collusion. *Transportation Research Part B: Methodological, 43*(10), 952-965.

Basso, L. J., & Zhang, A. (2007). Congestible facility rivalry in vertical structures. *Journal of Urban Economics, 61*(2), 218-237.

Brooke, A., Caves, R., & Pitfield, D. (1994). Methodology for predicting european short-haul air transport demand from regional airports: An application to east midlands international airport. *Journal of Air Transport Management, 1*(1), 37-46.

Brueckner, J. K. (2003). Airline traffic and urban economic development. Urban Studies, 40(8), 1455-1469.

Brueckner, J., & Flores-Fillol, R. (2006). Airline schedule competition: Product-quality choice in a duopoly model.

Bureau of Transportation Statistics, U.S. Department of Transportation. (2014a). Air carrier: T-100 domestic market (U.S. carriers). Retrieved Nov. 24, 2014, from http://www.transtats.bts.gov/Fields.asp?Table_ID=258

Bureau of Transportation Statistics, U.S. Department of Transportation. (2014b). Air carriers: T-100 domestic segments (U.S. carriers). Retrieved Nov. 24, 2014, from http://www.transtats.bts.gov/Fields.asp?Table_ID=259

Bureau of Transportation Statistics, U.S. Department of Transportation. (2014c). Airline fuel cost and consumption (U.S. carriers - scheduled). Retrieved Nov. 24, 2014, from http://www.transtats.bts.gov/fuel.asp?pn=0&display=data1

Bureau of Transportation Statistics, U.S. Department of Transportation. (2014d). Airline origin and destination survey (DB1B). Retrieved Nov. 24, 2014, from http://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=125&DB_Name=Airline%20Origin%20and%20Destination%20Survey%20%28DB1B%29

Caves, R., Ndoh, N., & Pitfield, D. (1991). Route choice modeling applied to the choice between mature airports and emergent airports in their shadow. Paper presented at the *31st RSA European Congress, Lisbon, Portugal,* 27-30.

Census Bureau, U.S. Department of Commerce. (2014a). American fact finder. Retrieved Nov. 24, 2014, from http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml

Census Bureau, U.S. Department of Commerce. (2014b). Annual community survey. Retrieved Dec. 04, 2014, from http://factfinder2.census.gov/faces/nav/jsf/pages/programs.xhtml?program=acs

Cherchi, E., & de Dios Ortúzar, J. (2002). Mixed RP/SP models incorporating interaction effects. *Transportation, 29*(4), 371-395.

Cohas, F. J., Belobaba, P. P., & Simpson, R. W. (1995). Competitive fare and frequency effects in airport market share modeling. *Journal of Air Transport Management, 2*(1), 33-45.

de Luca, S. (2012). Modelling airport choice behaviour for direct flights, connecting flights and different travel plans. *Journal of Transport Geography, 22*, 148-163.

de Luca, S., & Di Pace, R. (2012). Modelling passenger departure airport choice: Implicit vs. explicit approaches. *Procedia-Social and Behavioral Sciences, 54*, 875-885.

Dmitry, P. (2012). Airport benchmarking and spatial competition: A critical review. *Transport and Telecommunication, 13*(2), 123-137.

Dougherty, C. (2011). *Introduction to econometrics* Oxford University Press.

Elwakil, O. S., Windle, R. J., & Dresner, M. E. (2013). Transborder demand leakage and the US–Canadian air passenger market. *Transportation Research Part E: Logistics and Transportation Review.*

Evans, W. N., & Kessides, I. N. (1993). Localized market power in the US airline industry. *The Review of Economics and Statistics,* 66-75.

Federal Aviation Administration. (2009). OEP 35. Retrieved, 2014, from http://aspmhelp.faa.gov/index.php/OEP_35

Federal Aviation Administration. (2014a). Airport categories. Retrieved Dec. 04, 2014, from http://www.faa.gov/airports/planning_capacity/passenger_allcargo_stats/categories/

Federal Aviation Administration. (2014b). Airport operations and ranking Reports Using the air traffic activity data system. Retrieved Dec. 05, 2014, from http://www.faa.gov/news/media_resources/atadsguide/

Fröhlich, K., & Niemeier, H. (2011). The importance of spatial economics for assessing airport competition. *Journal of Air Transport Management, 17*(1), 44-48.

Fuellhart, K. (2007). Airport catchment and leakage in a multi-airport region: The case of harrisburg international. *Journal of Transport Geography, 15*(4), 231-244.

Gillen, D. W., Oum, T. H., & Tretheway, M. W. (1990). Airline cost structure and policy implications: A multi-product approach for canadian airlines. *Journal of Transport Economics and Policy,* 9-34.

Gillen, D., & Morrison, W. (2003). Bundling, integration and the delivered price of air travel: Are low cost carriers full service competitors? *Journal of Air Transport Management, 9*(1), 15-23.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the Econometric Society,* 1293-1301.

Grubesic, T. H., & Matisziw, T. C. (2011). A spatial analysis of air transport access and the essential air service program in the united states. *Journal of Transport Geography, 19*(1), 93-105.

Gupta, S., Vovsha, P., & Donnelly, R. (2008). Air passenger preferences for choice of airport and ground access mode in the new york city metropolitan region. *Transportation Research Record: Journal of the Transportation Research Board, 2042*(1), 3-11.

Hansen, M. (1990). Airline competition in a hub-dominated environment: An application of noncooperative game theory. *Transportation Research Part B: Methodological, 24*(1), 27-43.

Hansen, M. (1995). Positive feedback model of multiple-airport systems. *Journal of Transportation Engineering, 121*(6), 453-460.

Hansen, M., & Kanafani, A. (1989). Hubbing and airline costs. *Journal of Transportation Engineering, 115*(6), 581-590.

Harvey, G. (1987). Airport choice in a multiple airport region. *Transportation Research Part A: General, 21*(6), 439-449.

Hess, S. (2004). A model for the joint analysis of airport, airline, and access-mode choice for passengers departing from the san francisco bay area. Paper presented at the *European Transport Conference, Strasbourg, France.*

Hess, S. (2005). Analysing air-travel choice behaviour in the greater london area. Paper presented at the *45th Congress of the European Regional Science Association.*

Hess, S., & Polak, J. W. (2005a). Accounting for random taste heterogeneity in airport choice modeling. *Transportation Research Record: Journal of the Transportation Research Board, 1915*(1), 36-43.

Hess, S., & Polak, J. W. (2005b). Mixed logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management, 11*(2), 59-68.

Hess, S., & Polak, W. (2010). Airport choice behaviour: Findings from three separate studies. *Competition in European Airports, Ashgate, Aldershot Forthcoming.*

Hess, S., Adler, T., & Polak, J. W. (2007). Modelling airport and airline choice behaviour with the use of stated preference survey data. *Transportation Research Part E: Logistics and Transportation Review, 43*(3), 221-233. doi:10.1016/j.tre.2006.10.002

Hsiao, C. (2008). *Passenger demand for air transportation in a hub-and-spoke network* ProQuest.

Hsu, C., & Wen, Y. (2003). Determining flight frequencies on an airline network with demand–supply interactions. *Transportation Research Part E: Logistics and Transportation Review, 39*(6), 417-441.

Hsu, C., & Wu, Y. (1997). The market size of a city-pair route at an airport. *The Annals of Regional Science, 31*(4), 391-409.

Hutchinson, B. G. (1993). Analyses of canadian air travel demands. *Journal of Transportation Engineering, 119*(2), 301-316.

Iowa Department of Transportation, & Iowa Department of Economic Development. (2001). Committee on air Service Report to the General assembly. Retrieved Dec. 04, 2014, from http://www.iowadot.gov/pdf_files/asc_report.pdf

Ippolito, R. A. (1981). Estimating airline demand with quality of service variables. *Journal of Transport Economics and Policy,* 7-15.

Ishutkina, M. A. (2009). *Analysis of the interaction between air transportation and economic activity: A worldwide perspective.* (No. Report No. ICAT-2009-2).Massachusetts Institute of Technology Cambridge, MA 02139 USA.

Lian, J. I., & Rønnevik, J. (2011). Airport competition–Regional airports losing ground to main airports. *Journal of Transport Geography, 19*(1), 85-92.

Lieshout, R. (2012). Measuring the size of an airport's catchment area. *Journal of Transport Geography, 25*, 27-34.

Loo, B. P. Y. (2008). Passengers' airport choice within multi-airport regions (MARs): Some insights from a stated preference survey at hong kong international airport. *Journal of Transport Geography, 16*(2), 117-125. doi:10.1016/j.jtrangeo.2007.05.003

Ndoh, N. N., Pitfield, D. E., & Caves, R. E. (1990). Air transportation passenger route choice: A nested multinomial logit analysis. *Spatial Choices and Processes,* 349-365.

Nussle, J. (2008). Update of statistical area definitions and guidance on their uses. Retrieved Dec. 03, 2014, from http://www.whitehouse.gov/sites/default/files/omb/assets/omb/bulletins/fy2009/09-01.pdf

O'Connor, W. E. (2001). *An introduction to airline economics* Greenwood Publishing Group.

Park, H. M. (2006). Autocorrelation and AR(1) process. Retrieved Dec. 06, 2014, from http://mse415.free.fr/autocorrelation.pdf

Pels, E., Nijkamp, P., & Rietveld, P. (1997). Substitution and complementarity in aviation: Airports vs airlines. *Transportation Research Part E: Logistics and Transportation Review, 33*(4), 275-286.

Pels, E., Nijkamp, P., & Rietveld, P. (1998). Equilibrium airfares, frequencies and airport taxes in a multiple airport region: An application of the nested logit demand model.

Pels, E., Nijkamp, P., & Rietveld, P. (2000). Airport and airline competition for passengers departing from a large metropolitan area. *Journal of Urban Economics, 48*(1), 29-45.

Pels, E., Nijkamp, P., & Rietveld, P. (2001). Airport and airline choice in a multiple airport region: An empirical analysis for the san francisco bay area. *Regional Studies, 35*(1), 1-9.

Pels, E., Nijkamp, P., & Rietveld, P. (2003). Access to and competition between airports: A case study for the san francisco bay area. *Transportation Research Part A: Policy and Practice, 37*(1), 71-83.

Phillips, O. R., Weatherford, L. R., Mason, C. F., & Kunce, M. (2005). Passenger leaks and the fate of small community air service. *Economic Inquiry, 43*(4), 785-794.

Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric models and economic forecasts* Irwin/McGraw-Hill Boston.

Ryerson, M. S., & Kim, H. (2013). Integrating airline operational practices into passenger airline hub definition. *Journal of Transport Geography, 31*, 84-93.

SAS. (2014). Testing for autocorrelation. Retrieved Dec. 06, 2014, from http://support.sas.com/documentation/cdl/en/etsug/63348/HTML/default/viewer.htm#etsug_model_sect044.htm

Sharov, A. (1996). Equilibrium: Stable or unstable? Retrieved Dec. 07, 2014, from https://home.comcast.net/~sharov/PopEcol/lec9/equilib.html

Suzuki, Y. (2007). Modeling and testing the "two-step" decision process of travelers in airport and airline choices. *Transportation Research Part E: Logistics and Transportation Review, 43*(1), 1-20.

Suzuki, Y., & Audino, M. J. (2003). The effect of airfares on airport leakage in single-airport regions. *Transportation Journal,* 31-41.

Suzuki, Y., Crum, M. R., & Audino, M. J. (2003). Airport choice, leakage, and experience in single-airport regions. *Journal of Transportation Engineering, 129*(2), 212-218.

Suzuki, Y., Crum, M. R., & Audino, M. J. (2004). Airport leakage and airline pricing strategy in single-airport regions. *Transportation Research Part E: Logistics and Transportation Review, 40*(1), 19-37.

Swan, W. M., & Adler, N. (2006). Aircraft trip cost parameters: A function of stage length and seat capacity. *Transportation Research Part E: Logistics and Transportation Review, 42*(2), 105-115.

Train, K. (2009). *Discrete choice methods with simulation* Cambridge university press.

Travelmath. (2014). Driving distance calculator. Retrieved Nov. 20, 2014, from http://www.travelmath.com/drive-distance/

Warburg, V., Bhat, C., & Adler, T. (2006). Modeling demographic and unobserved heterogeneity in air passengers' sensitivity to service attributes in itinerary choice. *Transportation Research Record: Journal of the Transportation Research Board, 1951*(1), 7-16.

Wei, W., & Hansen, M. (2003). Cost economics of aircraft size. *Journal of Transport Economics and Policy,* 279-296.

Wei, W., & Hansen, M. (2005). Impact of aircraft size and seat availability on airlines' demand and market share in duopoly markets. *Transportation Research Part E: Logistics and Transportation Review, 41*(4), 315-327.

Wei, W., & Hansen, M. (2007). Airlines' competition in aircraft size and service frequency in duopoly markets. *Transportation Research Part E: Logistics and Transportation Review, 43*(4), 409-424.

Wei, W. (2006). Impact of landing fees on airlines' choice of aircraft size and service frequency in duopoly markets. *Journal of Air Transport Management, 12*(6), 288-292. doi:http://dx.doi.org/10.1016/j.jairtraman.2006.07.008

Wiltshire, J. (2013). Airport competition. Retrieved Nov. 27, 2014, from http://www.iata.org/whatwedo/Documents/economics/airport-competition.pdf

Windle, R., & Dresner, M. (1995). Airport choice in multiple-airport regions. *Journal of Transportation Engineering, 121*(4), 332-337.

Windle, R., & Dresner, M. (1999). Competitive responses to low cost carrier entry. *Transportation Research Part E: Logistics and Transportation Review, 35*(1), 59-75.

Wooldridge, J. (2012). Serial correlation and heteroskedasticity in time serial regressions. (pp. 412) Cengage Learning.

Zellner, A., & Theil, H. (1962). Three-stage least squares: Simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society,* 54-78.

Zhang, A., & Czerny, A. I. (2012). Airports and airlines economics and policy: An interpretive review of recent research. *Economics of Transportation, 1*(1), 15-34.

Zhang, A., Fu, X., & Yang, H. G. (2010). Revenue sharing with multiple airlines and airports. *Transportation Research Part B: Methodological, 44*(8), 944-959.

Zhang, S., Derudder, B., & Witlox, F. (2013). The impact of hub hierarchy and market competition on airfare pricing in US hub-to-hub markets. *Journal of Air Transport Management, 32*, 65-70.

Zou, B., & Hansen, M. (2012a). Flight delays, capacity investment and social welfare under air transport supply-demand equilibrium. *Transportation Research Part A: Policy and Practice, 46*(6), 965-980.

Zou, B., & Hansen, M. (2012b). Impact of operational performance on air carrier cost structure: Evidence from US airlines. *Transportation Research Part E: Logistics and Transportation Review, 48*(5), 1032-1048.

APPENDIX A. DESCRIPTIVE STATISTICS OF DATASET

Table A. 1 Descriptive Statistics of Leakage and Substitute Origin-Destination (OD) Pairs (Quarter 1, 2004 - Quarter 1, 2014)

Local Airport	Destination Airport	Passengers for Local OD Pair	Average Airfare for Local OD Pair	Flight Legs for Local OD Pair	Non-stop Flight Distance for Local OD Pair	Substitute Airport	Passengers for Substitute OD Pair	Average Airfare for Substitute OD Pair	Flight legs for Substitute OD Pair	Driving Distance to Substitute Airport
JAX	PHL	112,265	142.63	1.24	742	MCO	612,363	125.71	1.07	144
TUS	SEA	65,489	173.11	1.42	1,216	TPA	395,242	156.18	1.11	181
GRR	TPA	29,561	158.64	1.72	1,041	DTW	237,723	131.90	1.16	120
CAE	LGA	26,531	166.61	1.43	617	CLT	266,991	159.78	1.07	88
PWM	CLT	17,653	182.75	1.68	812	BOS	169,340	160.65	1.21	96
BDL	TPA	140,991	142.15	1.23	1,111	JFK	287,696	134.20	1.03	106
CMH	TPA	102,344	138.50	1.30	829	DTW	237,723	131.90	1.16	155
CHS	LGA	62,314	171.01	1.18	641	CLT	266,991	159.78	1.07	148
CHA	DCA	11,309	185.57	1.49	523	ATL	385,501	164.42	1.04	106
HSV	DCA	42,162	277.50	1.11	613	ATL	385,501	164.42	1.04	151

APPENDIX B. BOX PLOT OF EACH VARIABLE IN DATASET

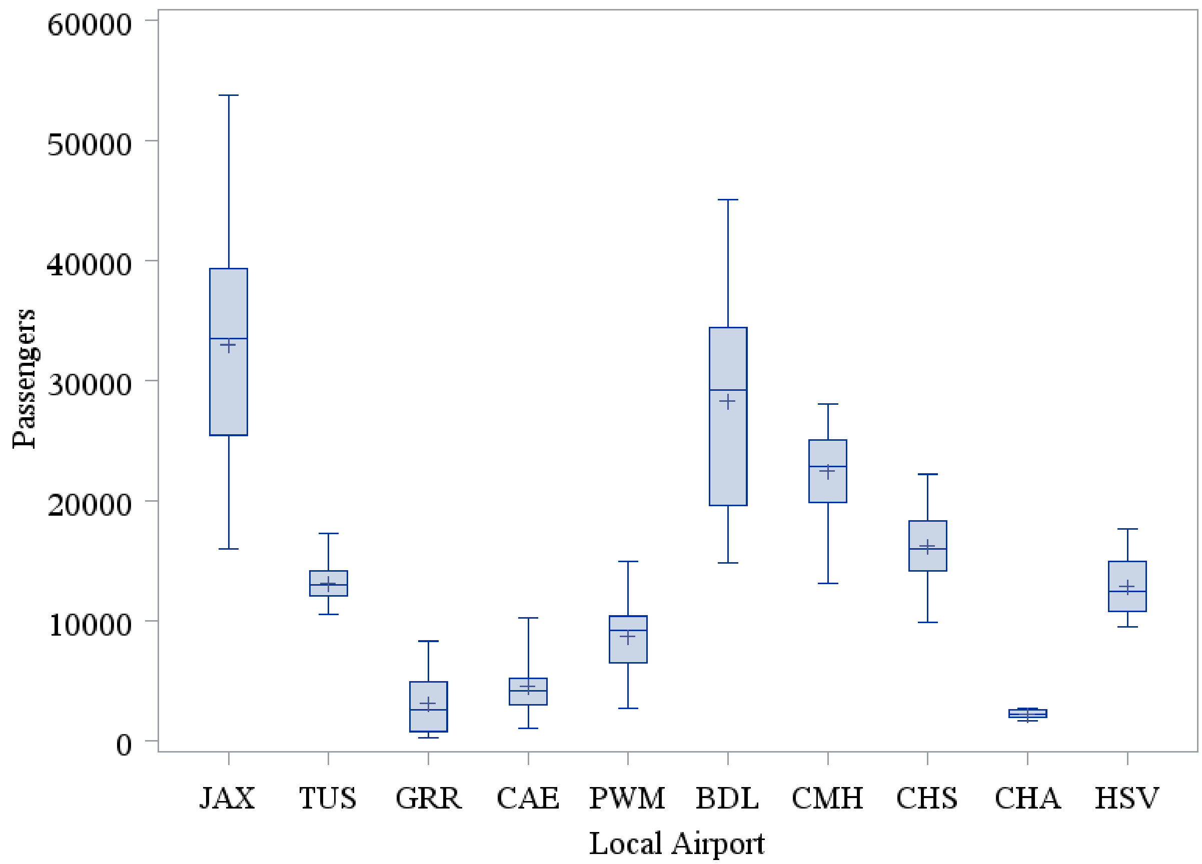

Figure B. 1 Box plot of passengers per quarter for the local OD pair with respect to each local airport

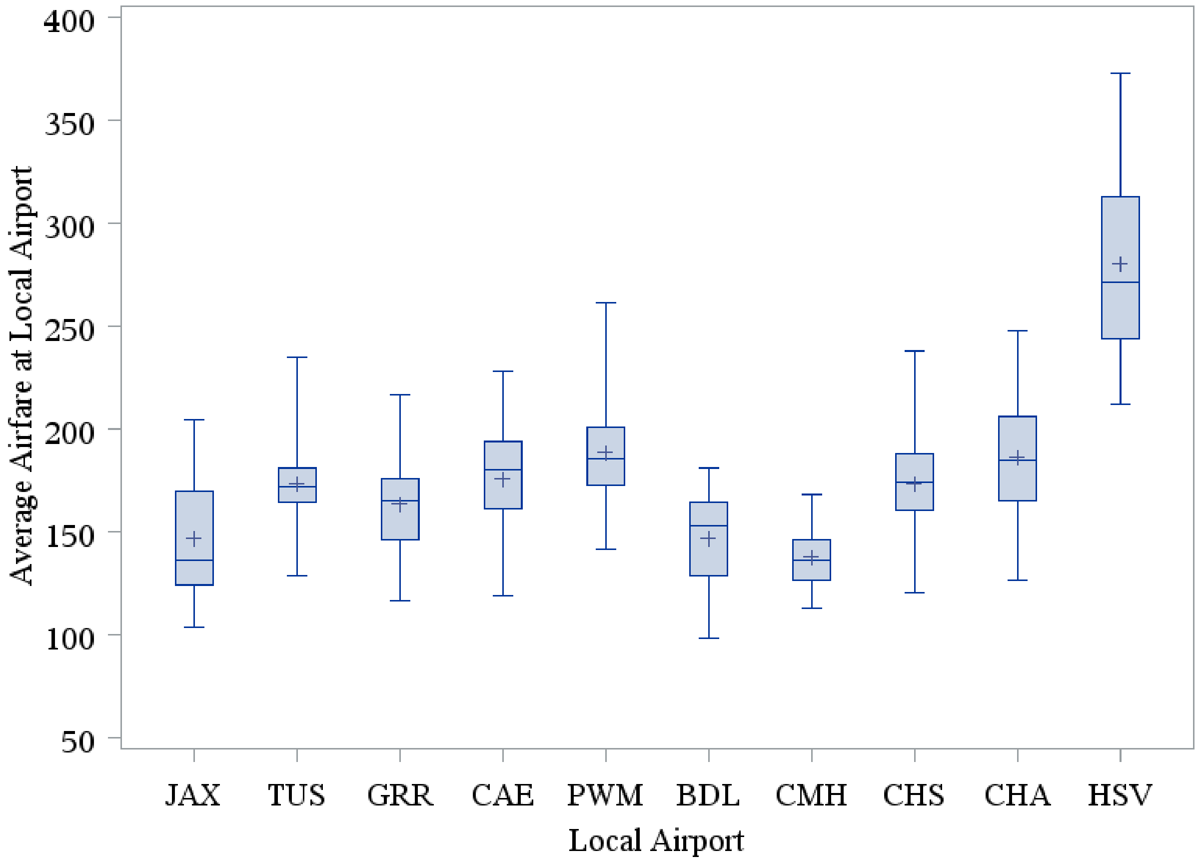

Figure B. 2 Box plot of average airfare per passenger per quarter for the local OD pair with respect to each local airport

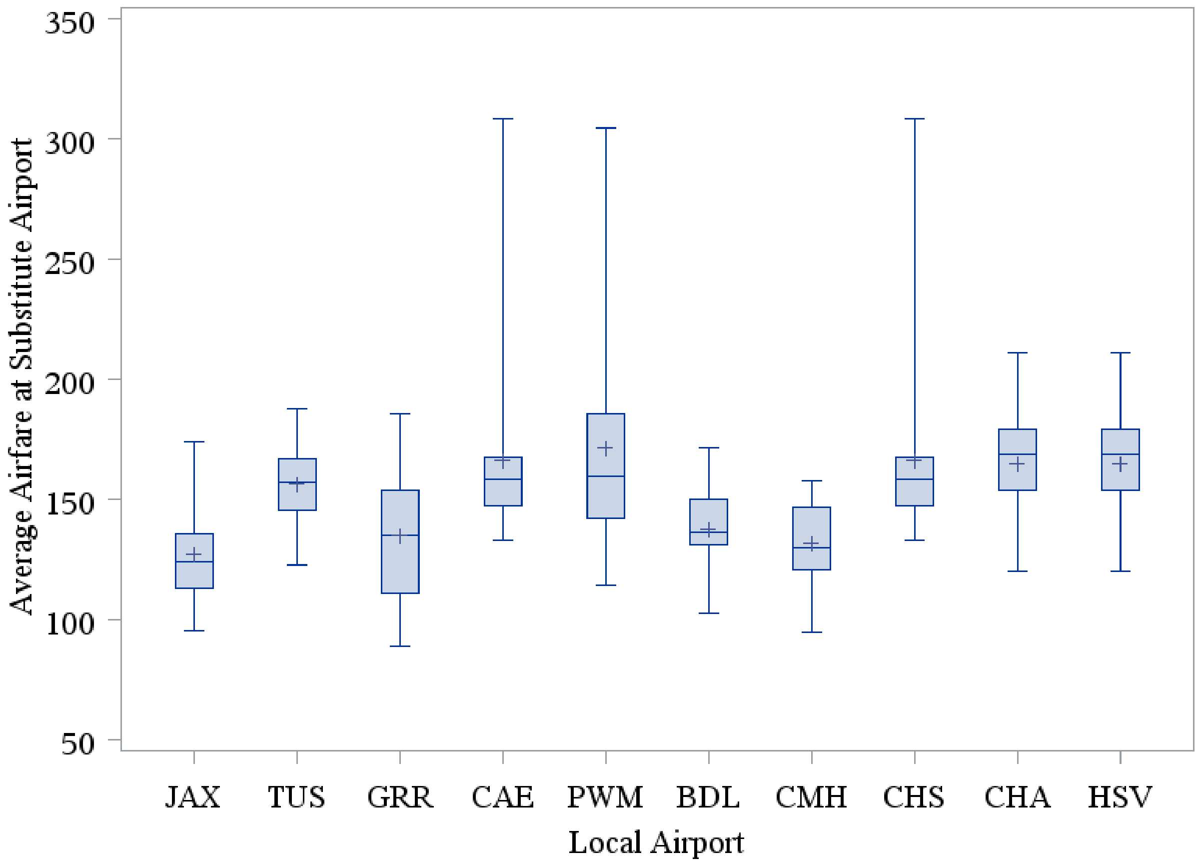

Figure B. 3 Box plot of average airfare per passenger per quarter for the (corresponding) substitute OD pair with respect to each local airport

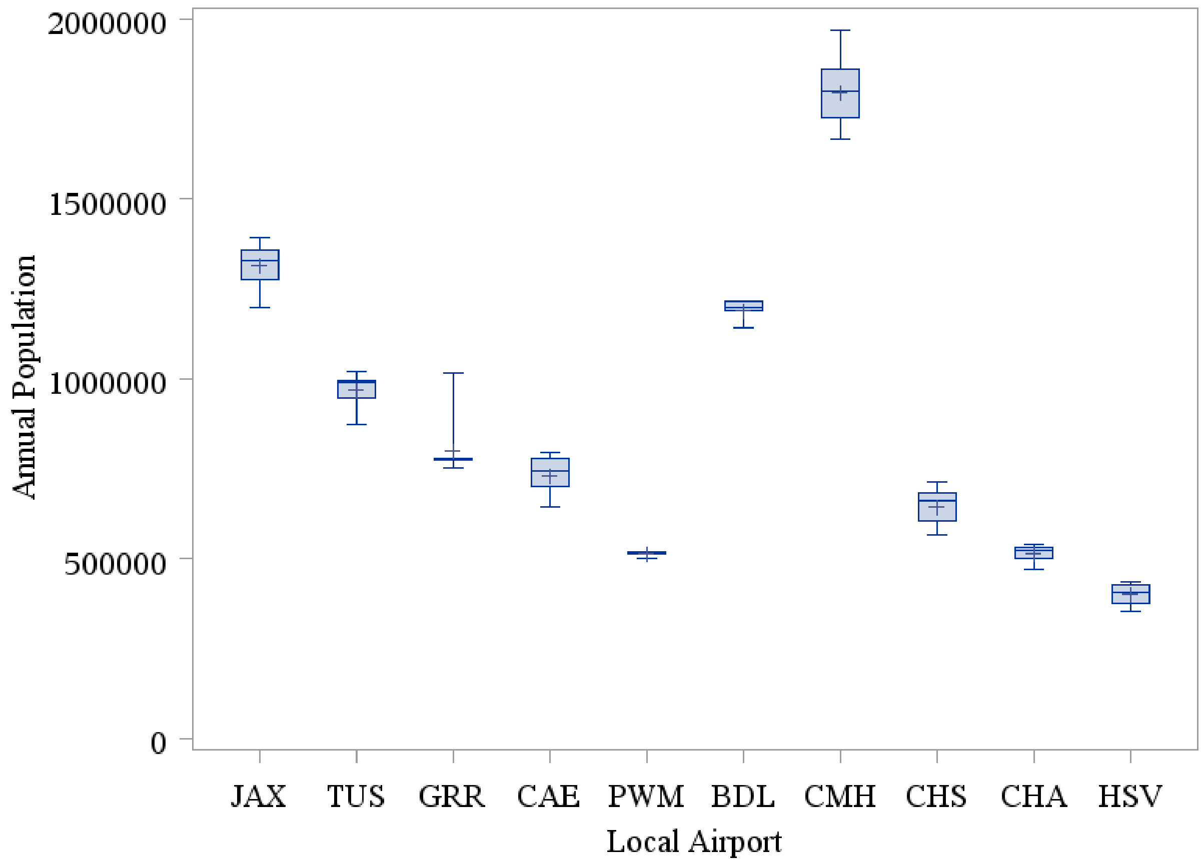

Figure B. 4 Box plot of population in the metropolitan area with respect to each local airport

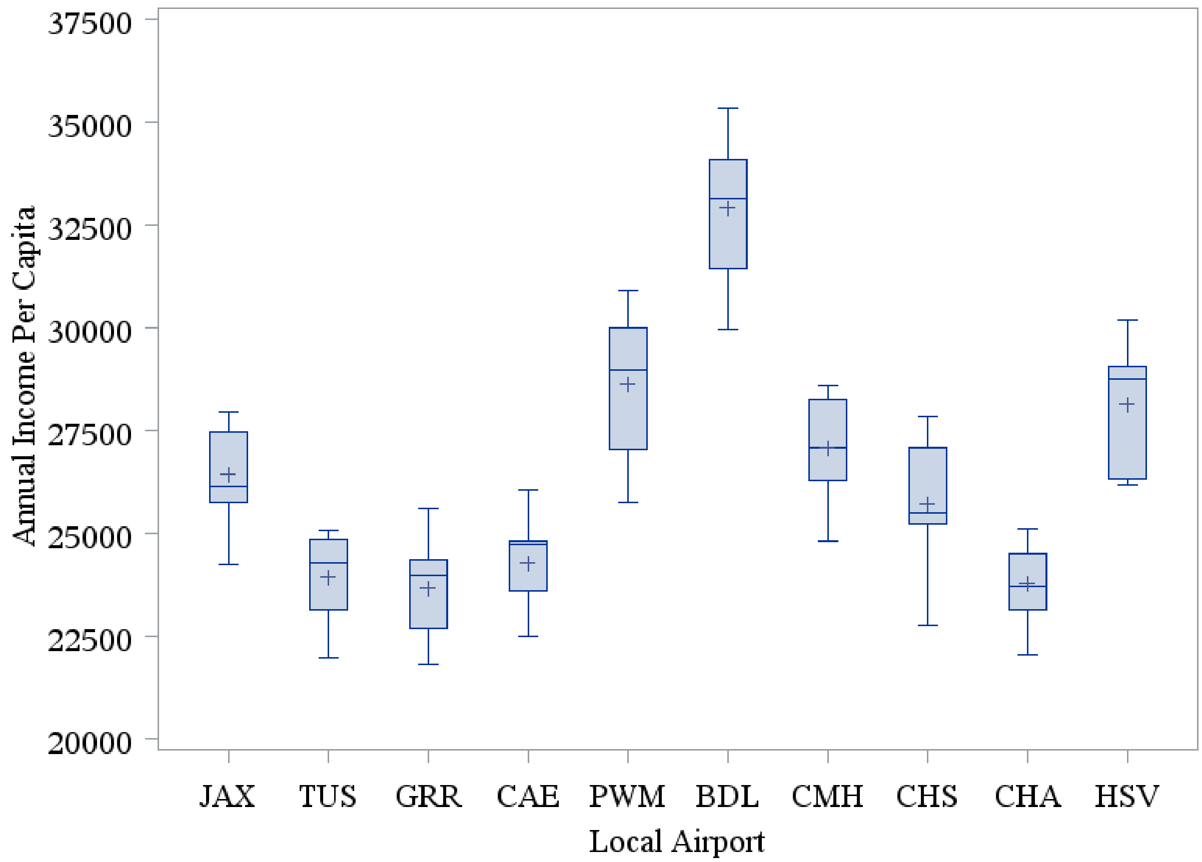

Figure B. 5 Box plot of annual per capita income in the metropolitan area with respect to each local airport

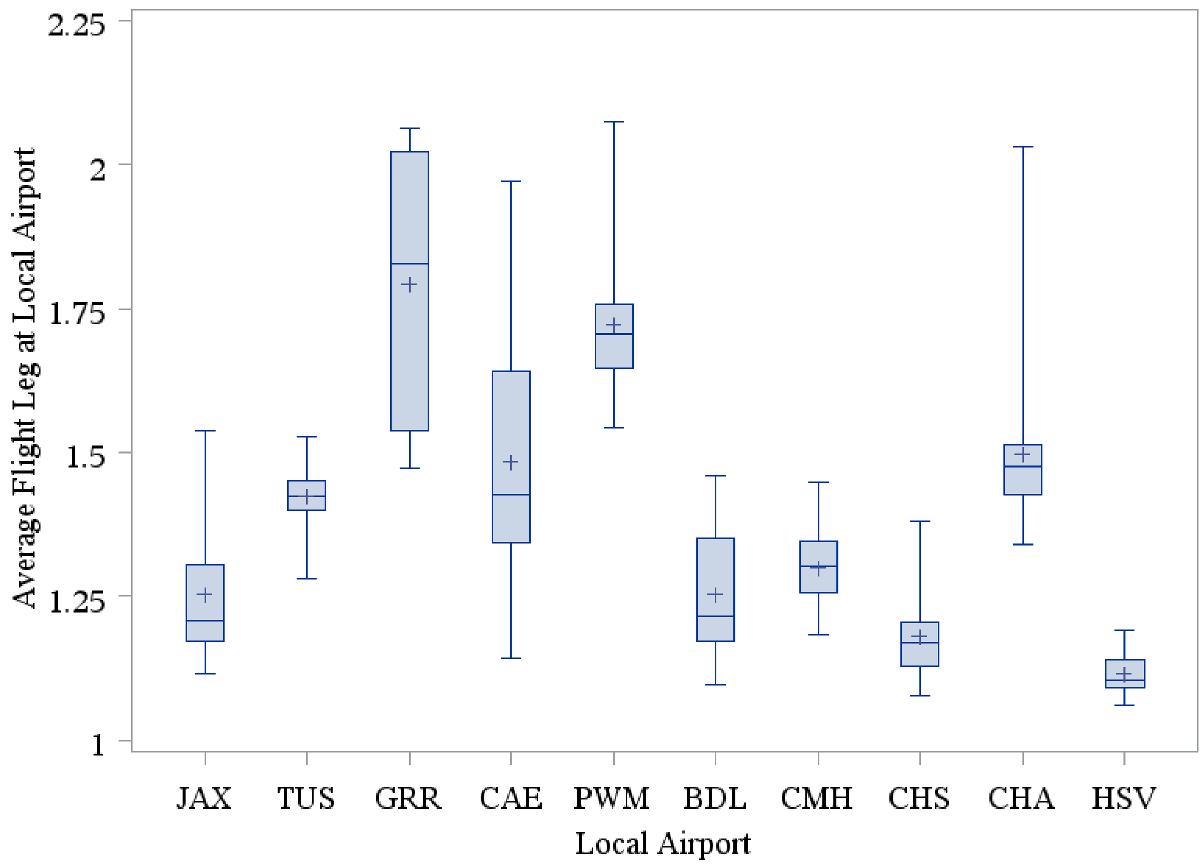

Figure B. 6 Box plot of average flight leg per passenger per quarter for the local OD pair with respect to each local airport

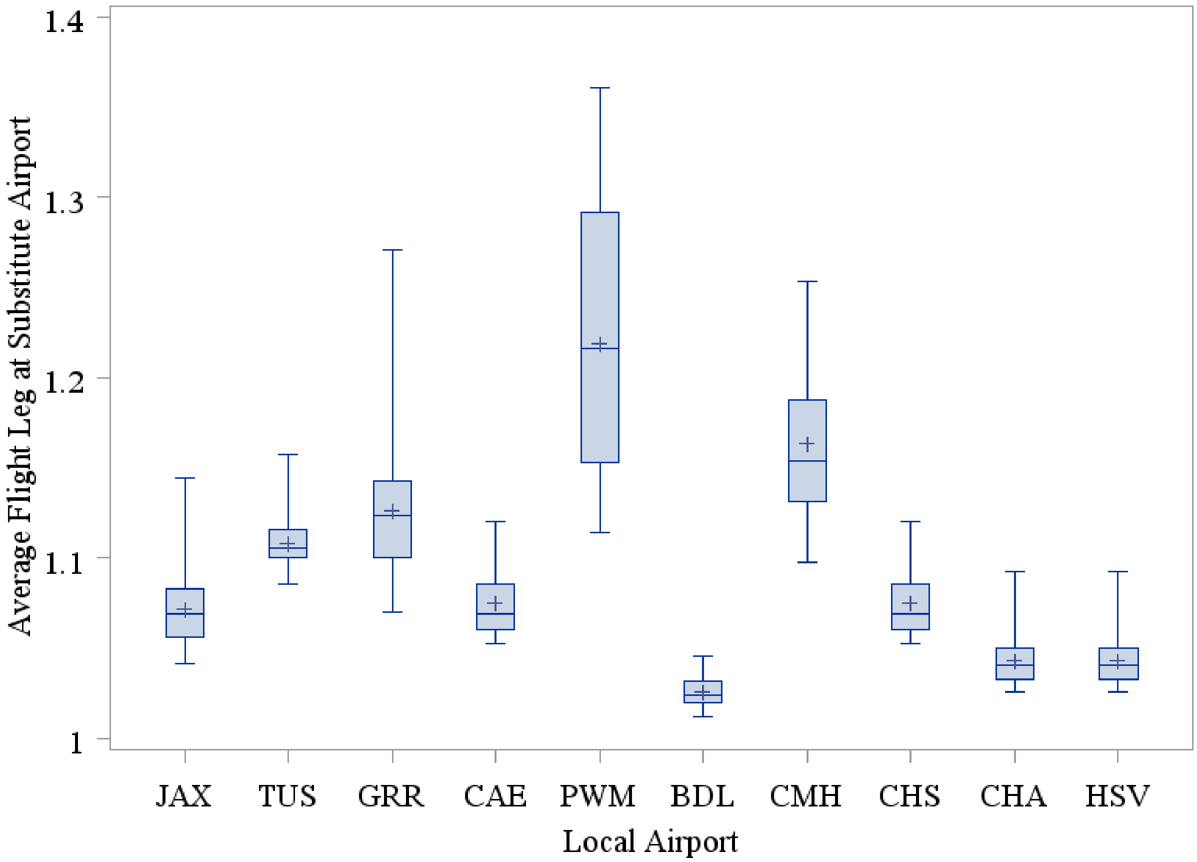

Figure B. 7 Box plot of average flight leg per passenger per quarter for the (corresponding) substitute OD pair with respect to each local airport

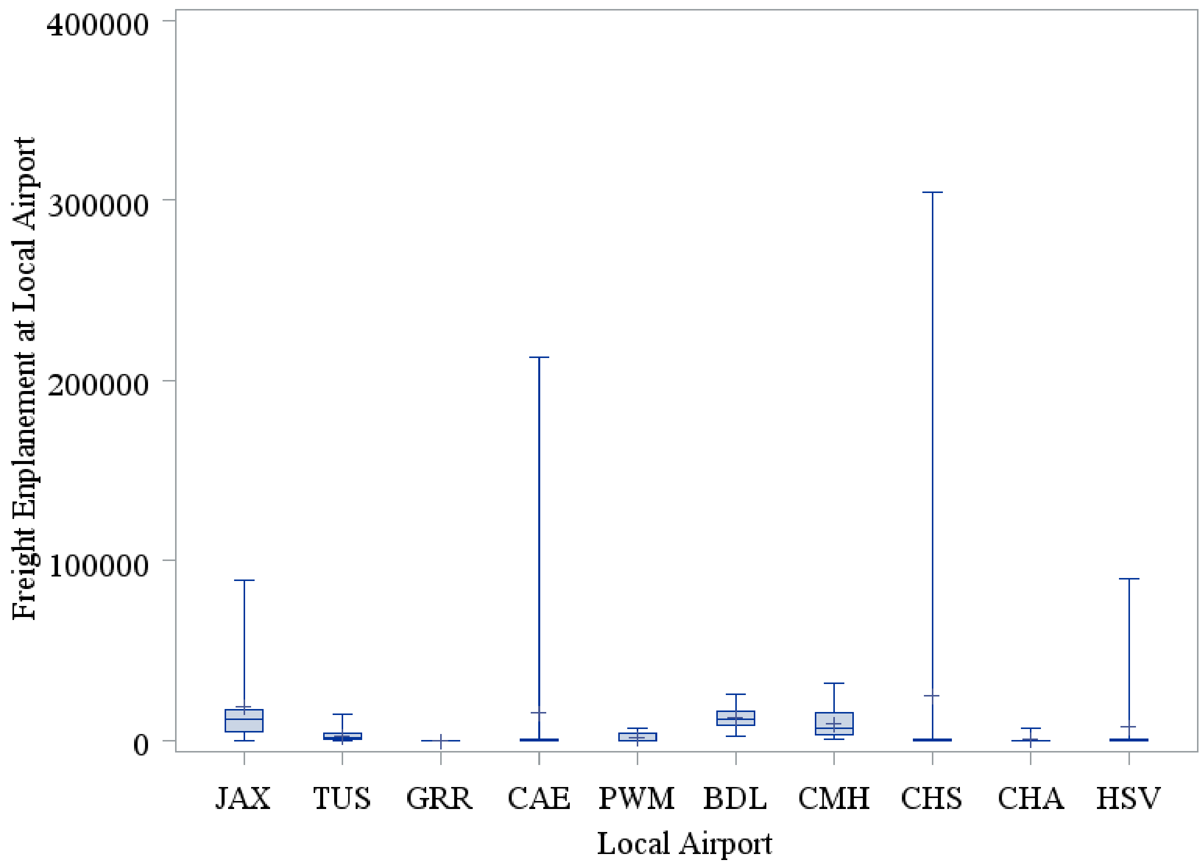

Figure B. 8 Box plot of freight enplanement per quarter for the local OD pair with respect to each local airport

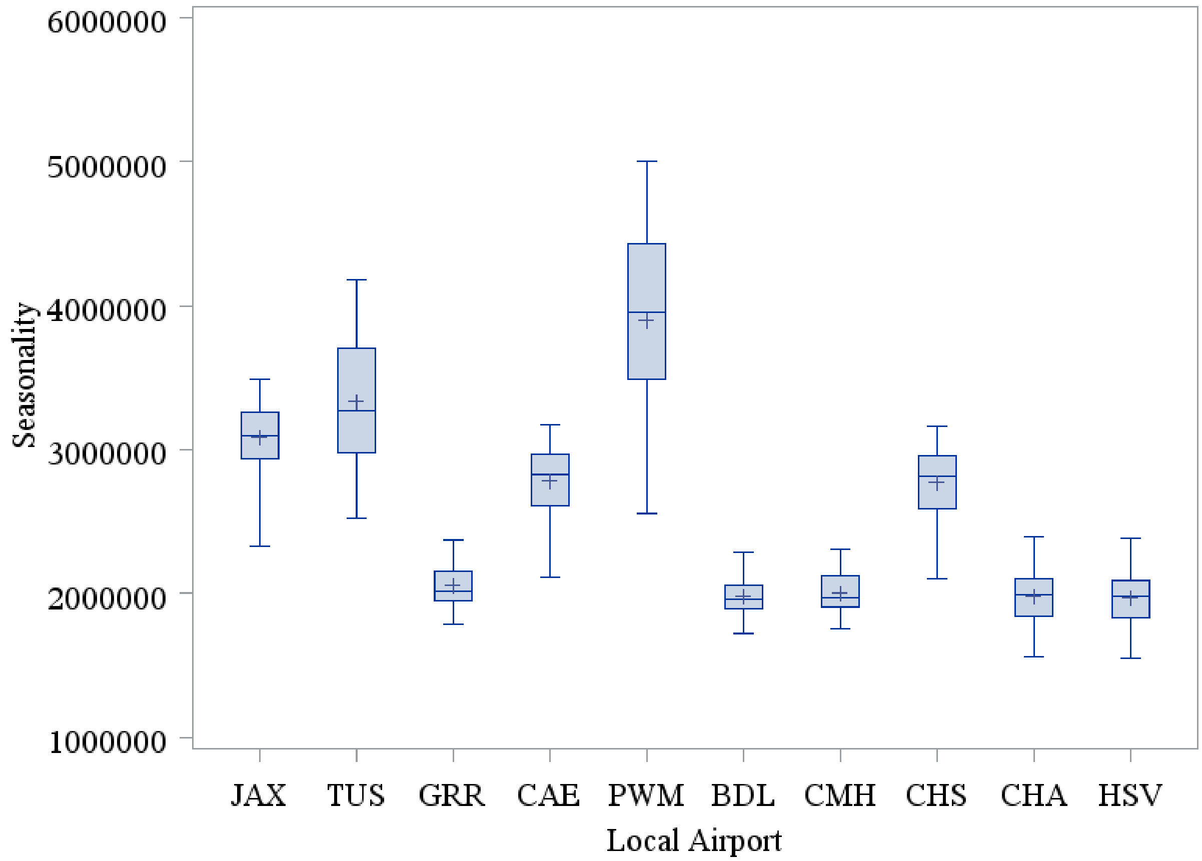

Figure B. 9 Box plot of seasonality per quarter for the local OD pair with respect to each local airport

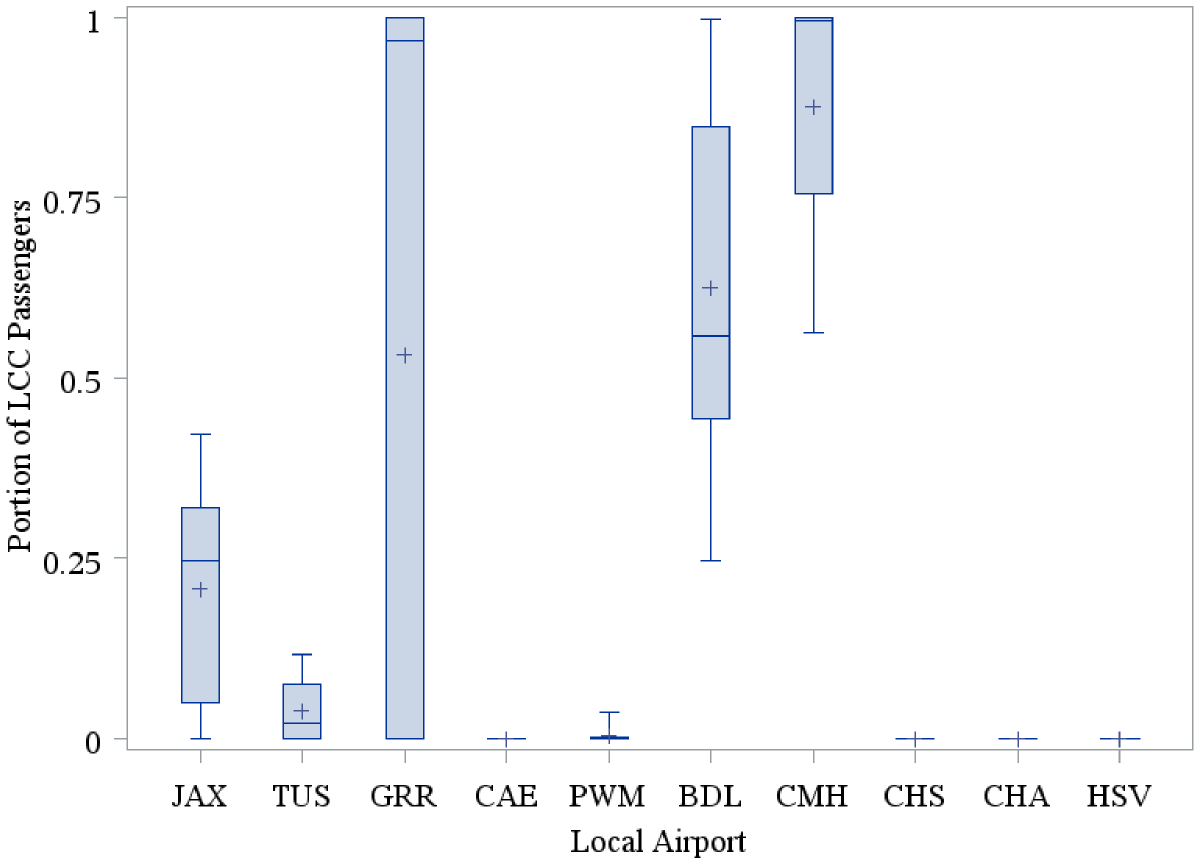

Figure B. 10 Box plot of passenger portion served by LCC per quarter for the local OD pair with respect to each local airport

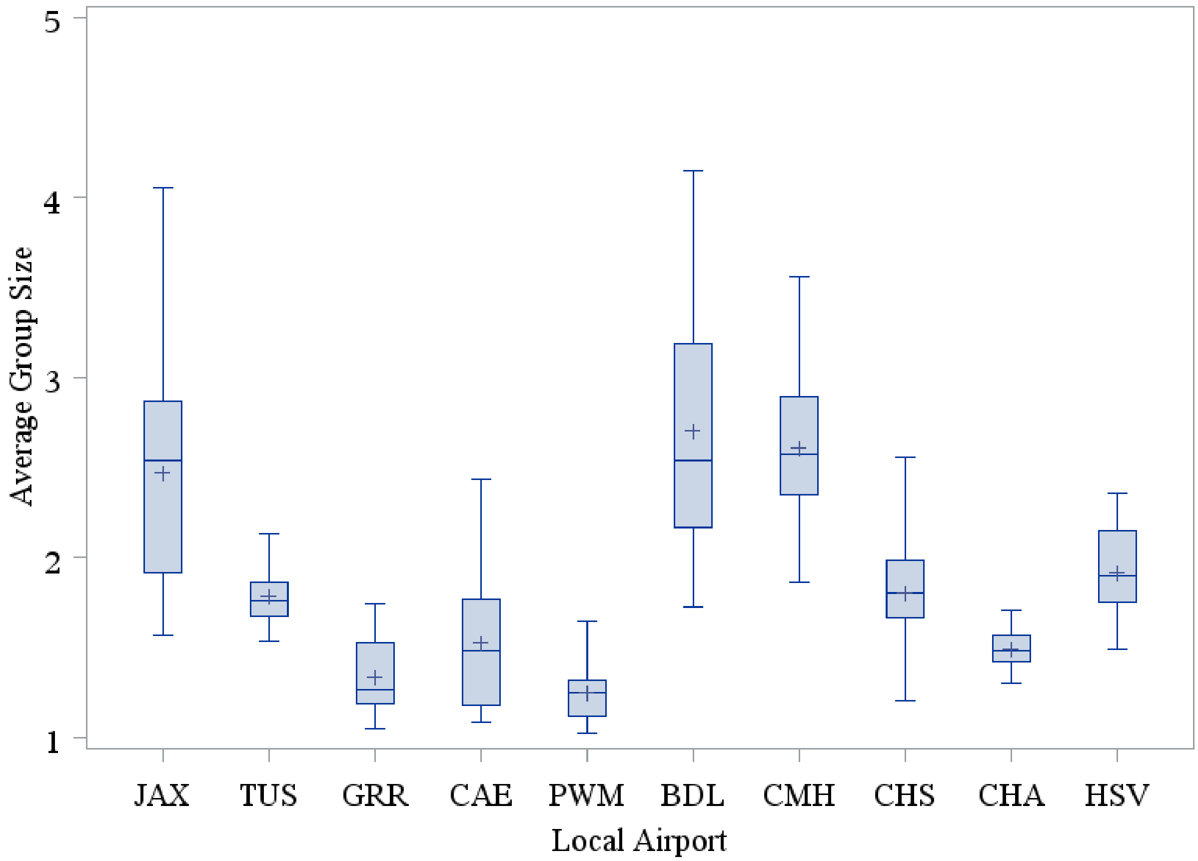

Figure B. 11 Box plot of average group size per quarter for the local OD pair with respect to each local airport

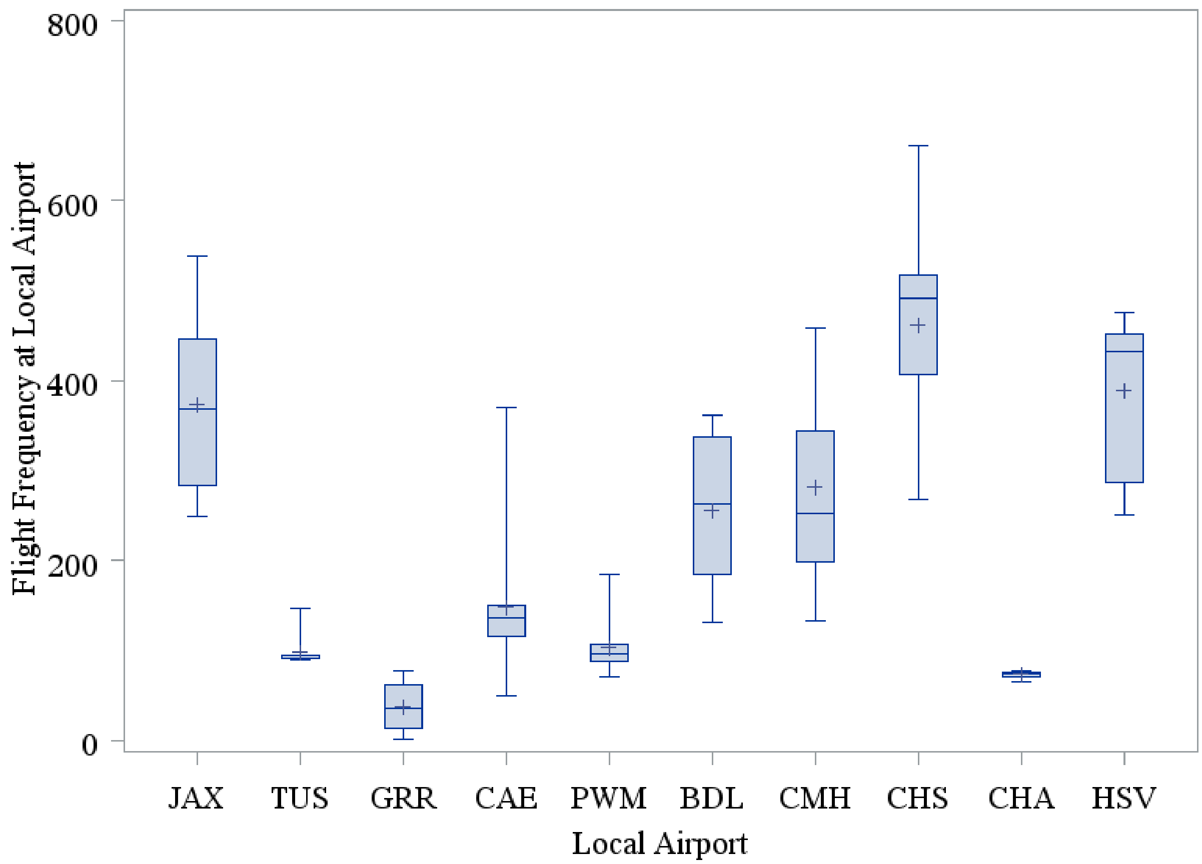

Figure B. 12 Box plot of flight frequency per quarter for the local OD pair with respect to each local airport

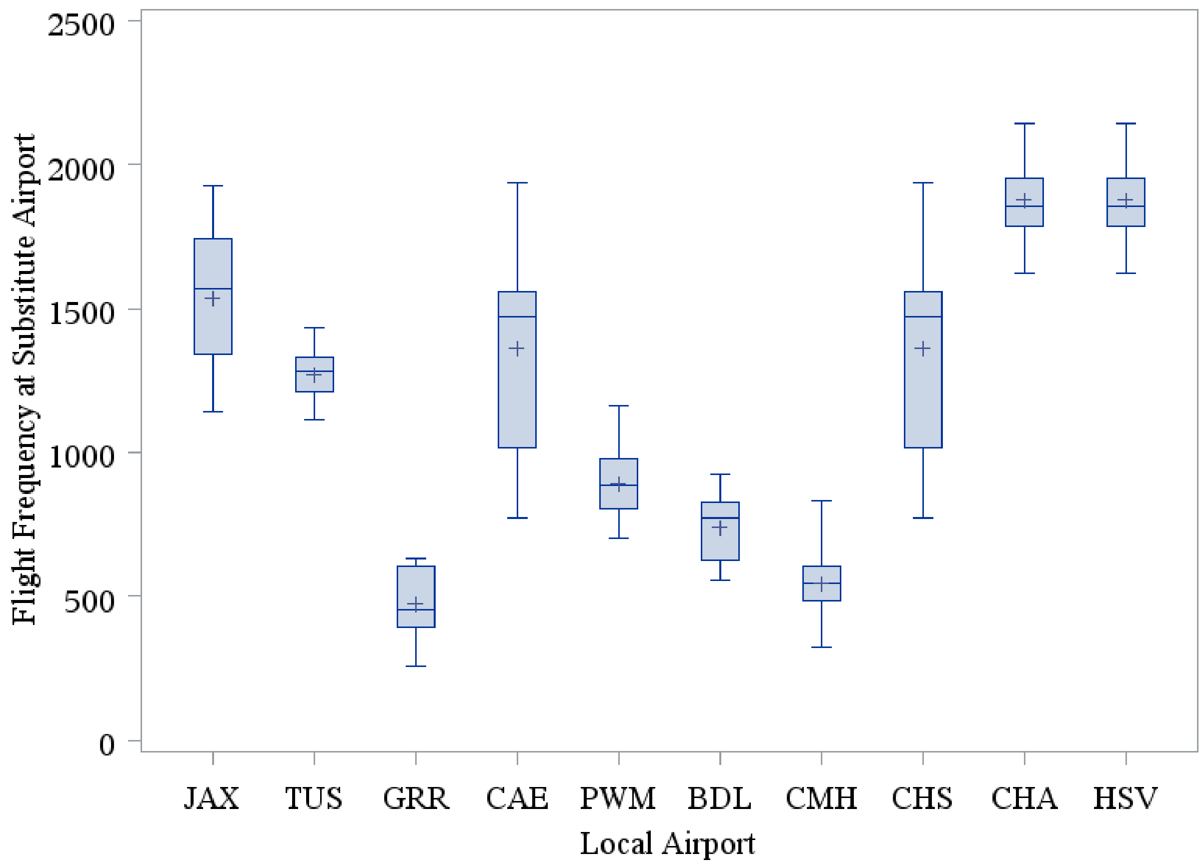

Figure B. 13 Box plot of flight frequency per quarter for the (corresponding) substitute OD pair with respect to each local airport

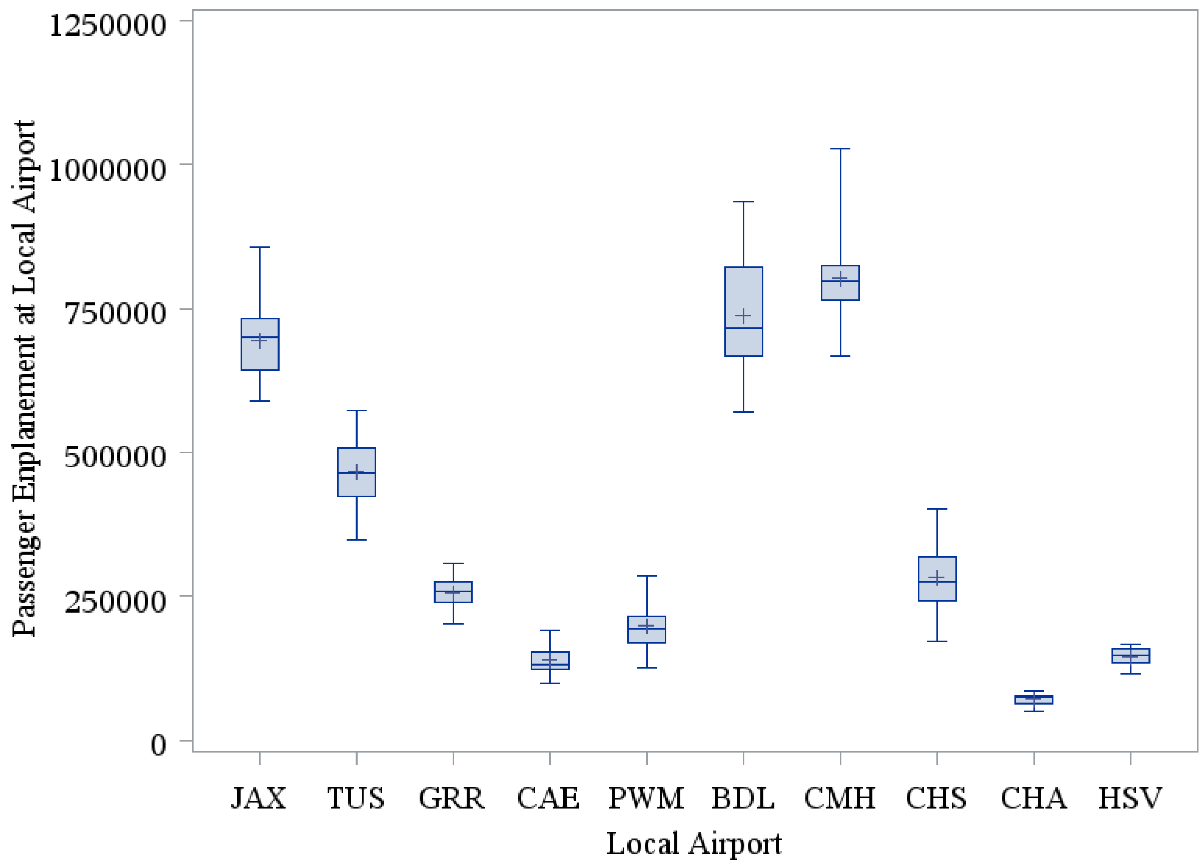

Figure B. 14 Box plot of passenger enplanement per quarter with respect to each local airport

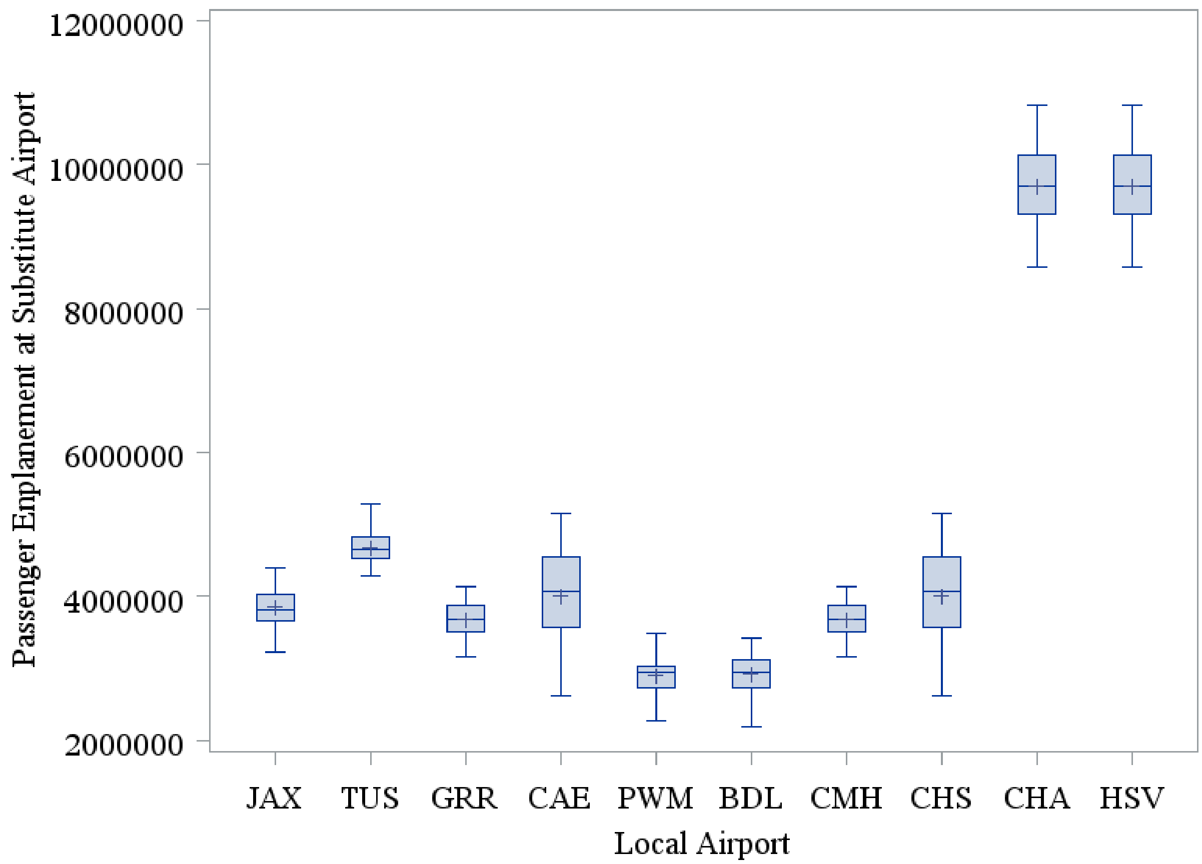

Figure B. 15 Box plot of passenger enplanement per quarter at the (corresponding) substitute airport with respect to each local airport

APPENDIX C. LIST OF VARIABLES

$FARE_{it}$ is the average airfare per passenger for the local OD pair i at quarter t.

I(i = 1) is the route indicator variable. I = 1 if the route is for the local OD pair i; and I = 0 otherwise.

$PASS_{it}$ is the number of passengers for the local OD pair i at quarter t.

I(LCC_{it} = 1) is the low-cost carrier (LCC) indicator variable. I = 1 if 25% or more passengers used low-cost carriers (LCC) for the local OD pair i at quarter t; and I = 0 otherwise.

FS_{it} is the average airfare per passenger for the substitute OD pair corresponding to the local OD pair i at quarter t.

$FUEL_t$ is the unit aviation fuel cost per gallon for U.S. domestic services provided by U.S. carriers at quarter t.

$MILES_i$ is the non-stop miles of from origin airport to destination airport for the local OD pair i.

$PASS_{it}$ is the number of passengers for the local OD pair i at quarter t.

$ln(\widehat{FARE}_{it})$ is the fitted log value of airfare per passenger for the local OD pair i at quarter t.

LEG_{it} is the average flight leg per passenger for the local OD pair i at quarter t.

SEA_{it} is the seasonality variable, represented by total number of passengers from all U.S. airports except the local airport and substitute airport to the destination airport for the local OD pair i at quarter t.

POP_{it} is the annual population in the year of quarter t in the metropolitan area served by the local airport (i.e., origin airport) of local OD pair i.

FS_{it} is the average airfare per passenger for the substitute OD pair corresponding to the local OD pair i at quarter t.

$SIZE_{it}$ is the average group size of passengers for the local OD pair i at quarter t.

DIS_i is the driving distance between local airport and the corresponding substitute airport for the local OD pair i; in miles.

ENP_{it} is the total passenger enplanement from local airport to all U.S. destination airports minus the number of passengers of the local OD pair i at quarter t.

ES_{it} is the total passenger enplanement from the corresponding substitute airport for the local OD pair i to all U.S. destination airports at quarter t.

$\alpha_1 \ldots \alpha_4$, and λ_i are parameters in the airfare model, where the subscript i denotes the local OD pair, particular to each of the 10 OD pairs represented in the dataset.

$\beta_1 \ldots \beta_9$, and δ_i are parameters in the demand model.

μ_t is the error term in the airfare model at quarter t.

ε_t is the error term in the demand model.

ρ_1 is the first-order autoregressive parameter in the airfare model.

ρ_2 is the first-order autoregressive parameter in the demand model.

ϵ_t is the error term in the autoregressive error model (of the airfare model), which is assumed to be normally and independently distributed with mean 0 and variance σ^2, $\epsilon_t \sim N(0, \sigma^2)$.

v_t is the error term in the autoregressive error model (of the demand model), which is assumed to be normally and independently distributed with mean 0 and variance σ^2, $v_t \sim N(0, \sigma^2)$.