January 2012

Estimation of the Impact of Single Airport and Multi-Airport System Delay on the National Airspace System using Multivariate Simultaneous Models

Nagesh Nayak
University of South Florida, nagesh23nayak@gmail.com

Estimation of the Impact of Single Airport and Multi-Airport System Delay on the National Airspace System using Multivariate Simultaneous Models

by

Nagesh Nayak

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Major Professor: Yu Zhang, Ph.D.
Sisinnio Concas, Ph.D.
Tony Diana, Ph.D.
Jian J. Lu, Ph.D.
Abdul Pinjari, Ph.D.

Date of Approval:
March 27, 2012

Keywords: Delay Propagation, 3SLS, Macroscopic Tool, Econometrics, System Effect

DEDICATION

I dedicate this dissertation to my grandmother, Saraswati R. Nayak, whom I lost during my studies in the United States. I would also dedicate this work to my family, dad and mom (Prabhakar and Shobha Nayak) and Deepika, who have endlessly supported me since the beginning of my studies. Apart from them I appreciate the support of my grandfather and the extended family back in India. I would also like thank Professor Prabhat Srivastava for giving me initial knowledge of transportation engineering in India.

Most importantly, I am deeply indebted to Professor Yu Zhang for her continuous guidance, motivation, and support through my graduate studies at the University of South Florida. I would also like to thank my supervisor, Dr. Sisinnio Concas for giving me wonderful opportunity to work in CUTR and being supportive while nearing graduation.

I would also like to thank my friends from India, Tejas, Mahesh, Vivek, Chirag, Mehul, Vinit and Rohan for their support throughout the journey. Graduate studies in the U.S. wouldn’t have been easier without my friends here at the USF, Arjun, Aaditya, Sudeep, Makarand, Ben, Satish, Vishal, Saniya, Himanshu, Priyanka, Priyanka, Supriya, Swati, Tejsingh, Meeta, Jinendra, Prashant, Anshul, Payal, Rahul, Dhir and Nirav. Finally, I would like to thank Sneha for having to bear with me while I was busy working. It would have been difficult for me to finish this study without her continuous love and long-lasting support.

ACKNOWLEDGEMENTS

This research was conducted under the guidance of Dr. Yu Zhang, and I sincerely thank her for keeping faith in me and giving me this wonderful opportunity to work with her. She is the first person to guide me into the field of air transportation and gave me constant support to carry out this research. She has single handedly helped me during this research by dedicating her time and expertise for me. It is a privilege for me to have had the opportunity to work with her and it’s been a great learning experience. I would like to thank her for her trust in me throughout the years.

I would also like to thank my other committee members, Dr. Sisinnio Concas and Dr. Tony Diana for helping me in the research by giving valuable pointers and ideas. The resources that they provided really helped me during this work. I would also like to acknowledge my professor s Dr. Abdul Pinjari and Dr. Jian J. Lu for teaching us various transportation courses during graduate studies. All the minute details we studied in the courses really helped me while conducting this research.

Finally I would like to thank Mr. Lawrence D. Goldstein from the Airport Cooperative Research Program (ACRP), Mr. Robert Samis from the Federal Aviation Administration and Mr. Richard Golaszewski from the GRA, Inc for their valuable comments and suggestions while conducting this research. I would like to thank the ACRP Graduate Research Award Program for supporting me in this endeavor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Airline delays lead to a tremendous loss of time and resources and cost billions of dollars every year in the United States (U.S.). At certain times, individual airports become bottlenecks within the National Airspace System (NAS). To explore solutions for reducing the delay, it is essential to understand factors causing flight delay and its impact on airports in the NAS. Major causal factors of flight delay at airports include over-scheduling, en-route convective weather, reduced ceiling and visibility around airports, and upstream delay propagation. Delay at one airport can be passed on to other airports in the NAS, in another word, operational improvement at one airport will have network effect and benefit to other airports as well. Moreover delay at different airports in a region might agglomerate to cause delay at different regions in the NAS. Hence, to optimally allocate NAS resources, e.g. capital investment for airport capacity expansion, the impact of single airport delay to the NAS and vice versa need to be investigated and quantified.

For air transportation planning and policy purposes, this study concentrates on providing answers from a macroscopic point of view without being distracted by volatile operational details. In the first part, we estimate the interaction between flight delay at one single airport and delay at the rest of the NAS (RNAS) using case study for LaGuardia (LGA) and Chicago O'Hare (ORD) airports. In the second part, this research applies multivariate simultaneous regression models to quantify airport delay spillover

effects across 34 of the 35 Operational Evolution Plan (OEP) airports and the RNAS. Observing the interactions between these two models, they are regressed with an econometric technique; three stage least square (3SLS). Thus, the regression results help us to determine the delay interactions between different airports and the RNAS and compare these airports based on delay propagation characteristics. Another significant contribution of this research is that, the estimated coefficients can be used for determining the marginal effects of all the delay causal factors presented in the model.

Also, regional airport system development has been a hot topic of research in the air transportation community in recent years. Many metropolitan regions are served with more than one airport making their operations synchronized and interdependent and are known as regional airport system. This paper studies nine different prospective regions with multi-airport systems in the U.S. and identifies various key factors affecting the delay in these regions. Econometrics models and three stage least square (3SLS) estimation method are used to explore interdependency of delay at the multi-airport system and the RNAS. Along with it, different factors affecting delay at the system and the RNAS is being identified from the research. The outcomes from this research will help aviation planners understand the spillover effects of delays from multi-airport systems and provide decision support for future NAS improvement.

CHAPTER I

INTRODUCTION

1.1 Background

Air transportation industry is considered to be one of the most important components of our economy. A report by the U.S. Department of Transportation and the FAA indicated that aviation accounts for over $1.3 trillion in economic activity, roughly 5.2 percent of U.S. Gross Domestic Product (GDP) in 2009 [1]. The report clearly states that US economy success greatly depends upon the economic success of our aviation system. Considering all these circumstances and huge economic ramifications it is imperative for us to produce an efficient air transportation system for generations. This forms the base of our inspiration to understand the air transportation system, find out the caveats like delay and capacity constraints and then finally suggest possible solutions to improve the system overall.

Airport congestion and delay have been the focus of intense research during the last few decades. The U.S. air transportation demand is constantly increasing throughout the years. Figure 1 shows the trend of domestic airline passengers and domestic flights departed from 1996 to 2011. It is seen that in recent years growth of passenger demand surpasses that of increase in number of flights.

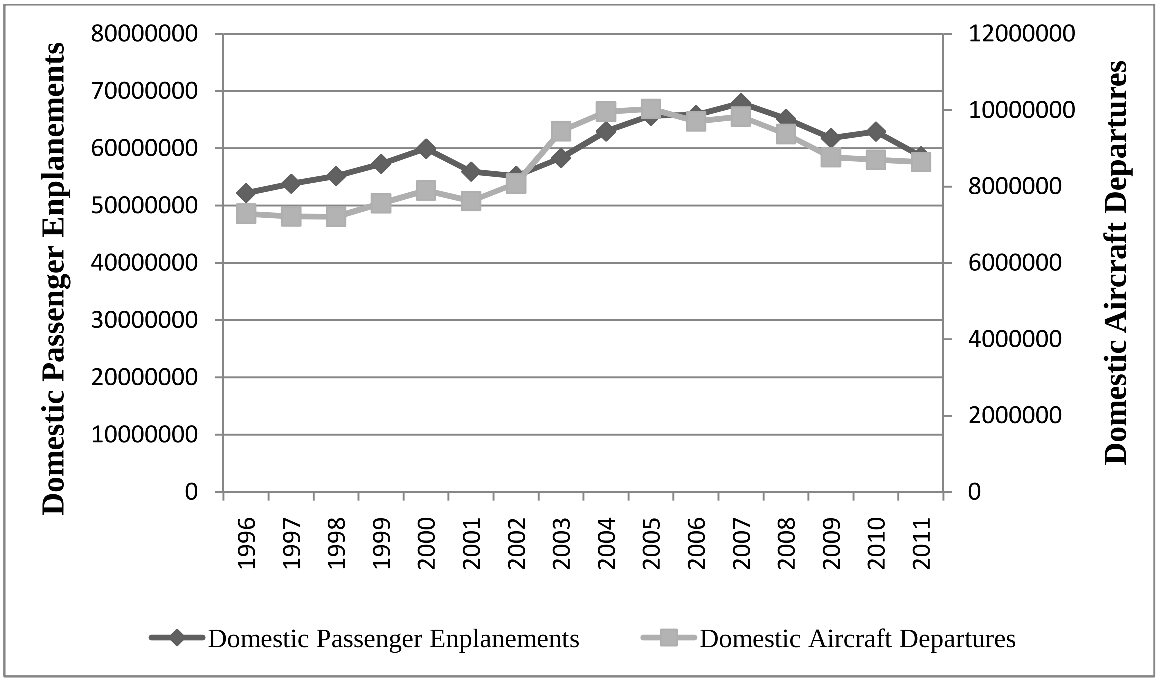


Figure 1 US Domestic Airline Passenger and Flight Trend from 1996 to 2011

Many major airports in the U.S. have significant delay problems due to this increased traffic demand and capacity imbalance. According to the Bureau of Transportation Statistics (BTS) in the U.S. Department of Transportation, less than 80 percent of arriving flights were on time for the period from January 2011to December 2011 [2]. The causes of flight delays include air carrier caused delays, late arrival of aircrafts, National Airspace System (NAS) delays, security delays, extreme weather, and delays due to cancelled or diverted flights. Among these causes, the delay due to late arrival of aircrafts accounts for more than 25 percent of total flight delays. As a result of the network structure of the NAS, delay at one airport is bound to affect delay at other airports.

Then we have a perennial question of who pays for all these flight delays? Ultimately it affects all the components of air transportation system that includes airlines, passengers, airports, etc. A recent study by Ball et al [3], estimated that in

2007, the total flight delay led to a loss of $32.9 billion to the U.S. economy. The airline passengers were the most affected group with a loss of $16.7 billion due to the loss of passenger time, flight cancellations, and additional expenses of food and so on. The second most affected groups were airlines raking up a loss of $8.3 billion. In a recessive economic period such an enormous loss is highly unacceptable and we need to take certain measures to curtail the delay.

The Next Generation Air Transportation System (NextGen) plans for a highly-efficient NAS by 2018, when the total flight delay will be reduced by 35 percent producing a benefit of $23 billion to aviation industry and saving of about 1.4 billion gallons of aviation fuel [4]. The NextGen also estimates that the total flight operations will increase by 19 percent at the 35 major U.S. airports between 2009 and 2018 [4]. Considering such enormous growth of air traffic on already-constrained resources, an appropriate action plan is needed to make this growth smooth and manageable on the airports. The addition or extension of runways at airports and the development of innovative technologies and procedures are some of the methods that need to be explored and implemented to achieve the NextGen's goal. Nevertheless, such an extensive change to the current NAS will require huge capital backing from the government and ultimately by the tax payers. According to one of the five-year plans that regulates the NAS modernization projects, popularly known as the Federal Aviation Administration (FAA) Capital Investment Plan (CIP), the FAA intends to invest about $20 billion during the years 2011 to 2015 for projects that modernize the existing system, increase airspace capacity, and introduce new technologies to achieve the planned NextGen capabilities [5]. Nevertheless, the impact that an increase in resources and therefore efficiency in a single airport will have on the efficiency of other airports remains undetermined. This delay propagation has become one of the major problems of the air transportation industry. With increasing cost of operations and the current economic crisis there is an urgent requirement of better technique to determine the factors causing delay and means to mitigate it. From an air transportation planning and policy point of view, sufficient tools are needed to test the system-wide effect of such investment activities and help further strategic planning. The research proposed in this paper will help this process by quantifying the interactions among airports in the U.S.

This research shows a collective comparison among airports and regions across the U.S. and the delay causal variables at each airport and predicts which interactions among airports are likely to create the highest or most regular delays. A case study in this research also helps to determine the benefits of capacity expansion at different airports and how it will affect the system overall.

1.2 Research Contribution

This research proposes the path of aggregate analysis conducted by the authors explained in detail in further chapters and intends not only to investigate the impact of single airport delay on other airports in the NAS (denoted as RNAS hereafter, i.e. the rest of the NAS excluding the reference airport(s) or multi-airport system) but also to explore how the delay spillover is widely dispersed across the RNAS. Causal factors of the average daily arrival delays are explored, and multivariate equations are developed for all airports under consideration along with the RNAS. The average daily arrival delay is the dependent variable in the equation for each airport and the RNAS, while simultaneously being considered as an independent variable in the equation of other airports and the RNAS. The estimated coefficients can be used to compute marginal effect of delay increase of that airport to the other airports or the RNAS. This type of model is widely used in economics and business management research studies.

Our model tries to establish the correlation between various delay causal factors at the airports and their effects on the entire system. Most previous studies estimate the delay propagated through an individual flight from an airport to the system. In our research we have tried to estimate and compare flight delay propagation from each individual airport to another in the US and vice versa. We have studied different factors causing delay and the extent of delay propagation amongst 34 Operational Evolution Partnership (OEP) airports except Honolulu International Airport (HNL) and RNAS containing the remaining of 74 ASPM airports together. This research illustrated the effectiveness of applying multivariate simultaneous equation model to study delay propagation from a single airport to other airports and to the rest of the system, and vice versa. The model estimates the effect of each of these factors using the three-staged least square (3SLS) method. This method is generally used to deal with the bidirectional relationship that exists between dependent and independent variables and suitable for the equations with correlated error terms. The estimated results help quantify the interdependency between flight delays at different airports and the NAS.

Going a step further, a collective comparison among airports for different regions in the U.S. is explored. The research includes identifying the delay causal variables at each such region and the interactions among regions that are likely to create the highest or most regular delays for the RNAS. The regional airport system is defined as a system with set of airports that serve airline traffic of a metropolitan area [6]. Previously all the individual airports served only their catchment areas. However with the increase in population, city's geographical growth, better ground transportation modes and sometimes political factors, there has been steady increase in number of airports within a region or a metropolitan area. Most of the major cities in the US are served by more than one airport. Many of these airports have coordinated operations in terms of sharing regional airspace, some act as a reliever airport in case of over shooting of capacity at the major airport(s) and also help reduce environmental effects like noise and air pollution in one specific area. Hence, it was worthy of research effort to explore the impact of these groups of airports in a region on other airports.

There is also a case of major airports situated very close to each other. Three of the world's busiest airports, namely LaGuardia (LGA), John F Kennedy (JFK) and Newark (EWR) are situated not very far from each other and have coordinated operations both in air and ground [7]. The New York airspace being one the most congested in the world with both domestic and international air traffic, the FAA has felt the need to increase the capacity of airports in the New York region. However we know that runway expansion requires enormous capital investment, project delays, public outcry and environmental concerns. Hence it is important to identify the potential for alternative airports to meet regional capacity needs and understand the potential of airport operation that can make more efficient use of existing resources and better use of limited funds for airport development. However in some cases the airports might be competing against each other for air service demand as in the case of Boston Logan (BOS), Manchester (MHT) and Providence (PCD) airports in the New England region of the U.S. [8]. The

BOS airport is operated by legacy airlines while the MHT and the PVD airports have large number of operations offered by low cost carriers (LCC). Both the airport operations completely differ from each other in terms of their management. Hence, it would be interesting to learn the impact of operations at these airports in comparison to other airports in the U.S.

In today's world, delay propagation and airport capacity constraints have become some of the major problems of the air transportation industry. Various researchers have tried to understand the microscopic perspective of delay propagation, i.e., delay propagation from an individual flight to another flight or the system (Beatty et al. [9], Schaefer and Millner [10], Wang et al. [11] Ahmad Beygi et al. [12]). However, their studies capture the details of only a few components of the NAS, such as specific airports, sectors, or individual flights, but fail to reflect the system overall. Our research takes the first step in considering all the airports in the U.S. together and estimates their effects on the NAS. It tries to determine the relationship between various delay causal factors at the airports and their effects on the entire system. It also initiates a step to determine the advantages and disadvantages of a regional airport system wherein two or more airports operate in a synchronized fashion. Total eleven regional airport systems in the US were studied in this research depending upon regional traffic share and proximity [13]. However due to the difficulty in terms of data availability the final analysis was limited to nine regions with the exclusion of the Orlando and the Tampa region. The research involved steps to determine the percentage share of air traffic demand in all airports in these regions and determine their delay at the regional level. The aggregate delay was then used to determine the combined impact on airports in other regions. The results obtained were very interesting and will be explained in detail further.

The remainder of this proposal is organized as follows. Chapter II summarizes the existing literature on delay propagation, factors affecting delay and the regional airport systems. Our approach related to this research is explained in detail in Chapter III. Chapter IV presents our earlier work related to the case study of two airports Chicago O'Hare (ORD) and LaGuardia (LGA), methodology and the outcomes. Chapter V specifies multivariate simultaneous equations and delay propagation for 34 OEP airports. Chapter VI presents the extension of the methodology to the multi-airport system. Chapter VII concludes the study and provides recommendations for further research.

CHAPTER II

LITERATURE REVIEW

The NAS can be defined as a complex agglomeration of different aviation components like airports, airspace, aircrafts, different facilities, etc working together for the safe and efficient airline operations. Since there are so many components involved and most of them are inter-connected, the delay at one component gets easily propagated to others. This research tries to understand different factors affecting delay and their immediate impacts. Different studies have already been conducted on delay in the NAS and its propagation. The following section gives an insight of all the studies conducted before.

2.1 Microscopic Methods

Beatty et al. [9] developed the concept of a delay multiplier for understanding the effect of initial flight delay on an airline’s operating schedule. They assumed that various airline resources such as crew members, aircraft, passengers, and gate space affect flight delay. The delay multiplier was used to determine all potential downstream flight delays connected to that initial flight. Their research concludes that the existence of a delay multiplier is due to the branching nature of crew and aircraft sequences. The research estimated the delay propagation from one airport to the other based on the connectivity of airline’s operating resources and its schedule.

Delay propagation has also been studied by Schaefer and Millner [10] using the detailed policy assessment tool. They modeled the propagation of delay throughout airports and airspace sectors given inputs such as air traffic demand and airport capacities. They synthesized aircraft assignment given the air traffic data from Official Airline Guide (OAG) and then used the information to simulate delay propagation according to departure and arrival queues between origin and destination airports. Three airports were analyzed using several combinations of Visual Meteorological Conditions (VMC) and Instrument Meteorological Conditions (IMC) when capacities reduced due to inclement weather. The results show that the delay augments with prolonged duration of IMC at the airports. They also concluded that although the propagation effect for the first leg was significant, it diminished along each subsequent leg.

Further research by Wang et al. [11] developed an analytical model to separate controllable factors that influence delays and their propagation in the NAS from other factors that are random variables in a given scenario. The controllable factors are scheduled and minimum airport turnaround time, slack for airport turnaround time, scheduled and minimum flight time between airports, and fixed flight time allowance, while the variable factors considered in the research were variable airport turnaround time and variable airport flight time. The model analyzed the interaction between fixed and variable delay components at each airport under both VMC and IMC conditions and emphasized the importance of schedule parameters on delay propagation in the NAS. Their study shows that airports with less slack time between flights had more delay.

A recent research by AhmadBeygi et al [12] explores a similar observation in terms of slack time between two flights. Their study indicates that the delay of one flight can propagate to disrupt one or many subsequent downstream flights that await the aircraft and crew from the delayed flight. In such case, the presence of well-planned slack between flights is critical for absorbing the disruption. All of these studies discussed above attempt to show how common resources and weighted airline schedules can be major causes of delay propagation and are microscopic in perspective. These research studies are clear indicators that the issue of delay propagation at airports is prevalent.

2.2 Macroscopic Methods

The studies discussed above attempt to show how common resources and weighted airline schedules can be major causes of delay propagation. These research studies are clear indicators that the issue of delay propagation at airports is prevalent.

A macroscopic research by Diana [14] proposed a methodology to compute delay propagation from airports based on the Discrete Fourier Transform (DFT). The airports sampled in his study vary in terms of location and traffic throughput. The research assumed that the delay propagation is similar as wave propagation where the delays represent signals and the NAS acts as the medium. Airlines anticipate delays and build precautionary buffer in their schedule to absorb the propagation effects. In his study, he applied the delay concept in airline on-time performance, i.e. only arrival flights with more than fifteen minutes delay past schedule are considered as delayed flights. Diana tried to investigate whether market concentrated airports (i.e. with higher traffic throughput) have more delay propagation effects than less concentrated airports. The outcomes show that, when delay propagation is considered as a signal through the system, it is not dependent on the degree of market concentration.

A recent study done by Laskey et al. [15] takes into consideration the dynamic aspects of flight delay, such as weather effects, wind speed, flight cancellations, and others, to estimate delay propagation in the NAS. They used Bayesian Networks (BN) to quantitatively analyze major factors affecting each delay component and the relationship among the delay components. The model studied weather effects and flight cancellations as two variables that might have an effect on flight delays. This research tried to demonstrate the system level impact due to delay at individual airports under different weather conditions. In their study, flight arrival delay was decomposed into Gate-In Delay, Turn Around Delay, Gate-Out Delay, Taxi-Out Delay, Airborne Delay, and Taxi-In Delay, each of which was considered as a dependent variable for that phase of the flight, with delays from previous phases as independent variables. The principal objective of this research was to estimate the impact of changes in tactical decisions and policies with respect to the ground delay program (GDP), rescheduling, and cancelled flights on delay in the system. Nevertheless, only three months of data were used to identify the critical phase of the flights from Chicago O' Hare International Airport (ORD) and Hatrsfield-Jackson Atlanta International Airport (ATL).

A similar study by Liu and Ma [16] used Bayesian Network to study flight delay and its propagation for airports in China. They established a direct relationship between arrival and departure delay at the airport studied. Primarily the delay was divided into normal, light, medium and heavy categories depending upon different times, ranging from less than 20 minutes for normal to more than 60 minutes for heavy. It was seen that the delay propagation is highest during medium and heavy delay period. It was also observed that flight cancellation is one technique that could be utilized to reduce flight delays. In both the Bayesian network studies discussed above, it is seen that few continuous variables needed to be discretized and this could produce erroneous results.

Hansen and Zhang [17] devised a macroscopic technique to study the delay propagation in the NAS. They studied the operational performance at LGA under different demand management regimes using multivariate simultaneous-equation regression model. The outcome of that research shows that, according to historical data from 2000 to 2004, the increase in one minute average-daily-arrival delay at the LaGuardia when compared to airline schedule causes an increase in the average-daily-arrival delay at non-LGA airports by 1.7 minute. The research indentified various factors causing arrival delay at LGA and non-LGA airports and estimated the impact of each of these factors on the total delay.

Morisset and Odoni [18] compared the capacity, schedule and reliability at major airports in Europe and the US. After studying 34 busiest airports in both US and Europe it was found that the European airports follow a conservative approach of operating at IFR rules for all weather conditions. On the contrary all the US airports operate with higher capacities using VFR rules for most of time. Due to this the delays at the US airports are very volatile and vary a lot due to weather, higher demand and constrained scheduling making it less reliable than the airports in the Europe. Our research tries to identify this effect of adverse weather and regional airport systems on the delay in the system.

2.3 Demand Management Regimes

In 1968, due to the increase in the number of air traffic operations, the airline slot management strategy called, High Density Rule (HDR) was applied at five major airports in the U.S. namely ORD, LGA, John F Kennedy International (JFK), Ronal Reagan Washington National (DCA) and Newark Liberty International (EWR) airports (Berardino [19]). Eventually it was exempted at EWR airport at very early stages. In 2000's, this slot control were gradually removed from ORD, LGA and JFK airports, however it still remained at the DCA airport. The demand management strategies at LGA, JFK and ORD have always been parallel, as shown in Figure 2 [19]. During this period, numerous demand management strategies were employed at these airports.

The HDR period at LGA was characterized by limiting the hourly slots to 68 between 6:00 am and 12:00 midnight. The slots were initially regulated by a scheduling committee composed of representatives from different airlines. Later in 1986, the scheduling committee was replaced by "use-it-or-lose-it" and "buy-sell" rules (Donohue [20]). However, with no airline willing to sell its slots, FAA granted 42 slot exemptions for various air services to LGA, especially for ones that were new entrant airlines or essential air services. As a result, by 1997, 30 new entrant exemptions were approved for LGA [20]. In April 2000, a demand management strategy called AIR-21 was introduced to eliminate slot control. During AIR-21, delay increased dramatically due to an increasing number of requests for slot exemptions. To overcome such delay, the FAA quashed the AIR-21 slot exemptions it had already granted and redistributed some of these exemptions by lottery. It also capped the number of operations per hour for commercial flights to 75 from the initial 100 under AIR-21. The terrorist attacks on

September 11, 2001, affected airport operations in many ways. Beginning in 2002, air traffic increased each following year, leading to a period of over-scheduling, and HDR completely expired by 2007 [20].

The JFK airport also had similar demand management regimes operating at the airport. The HDR strategy that was applied in 1968 expired only in January 2007. However, the operations were also affected by 9/11 incident wherein the total airport operations reduced a lot. In year 2004, there was an increase in airport capacity and subsequently increase in operations by Delta and Jet Blue airlines [19].

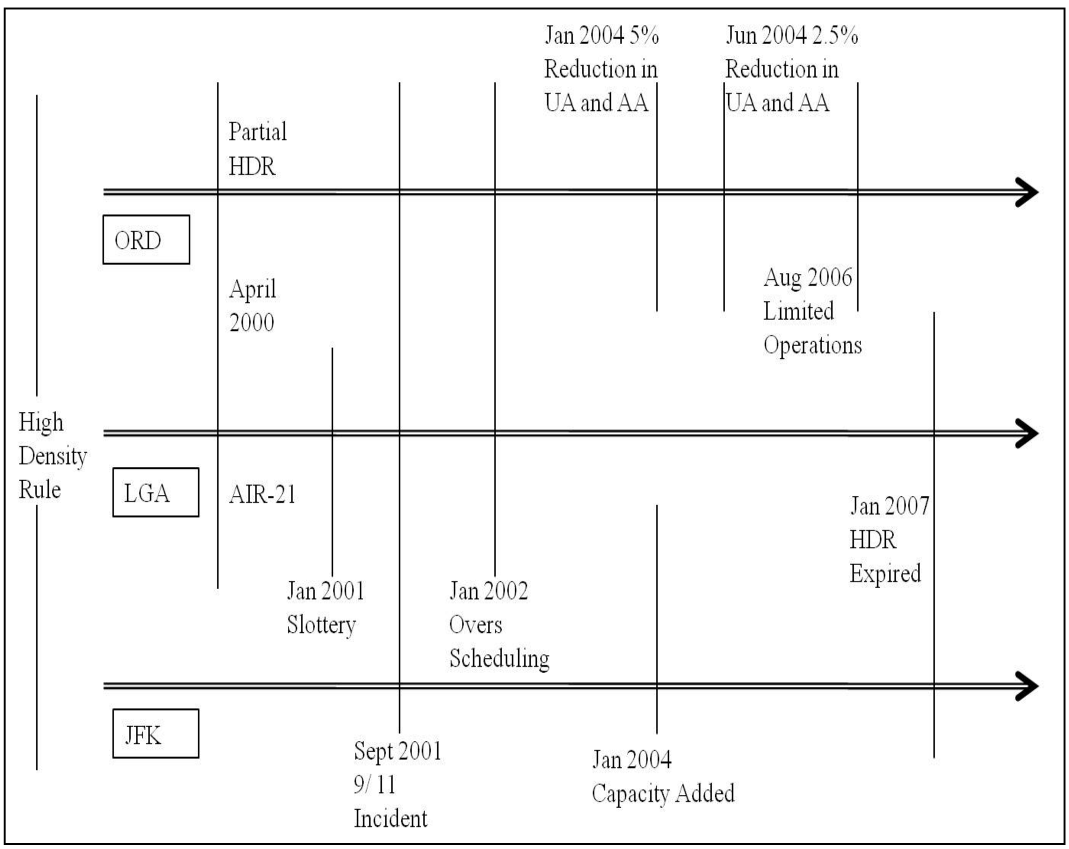


Figure 2 Demand Management Regimes at ORD, LGA and JFK Airports

ORD, similarly, has its own demand management regimes affecting air traffic operations in and around the airport. As mentioned earlier, the HDR strategy was also applied at ORD in 1968, one that resulted in the slot control phenomenon by major airlines. In the 1990s, 53 new slot exemptions were created at ORD [20]. Gradually, the HDR strategy was reduced at ORD, and its complete elimination took place by 2002. The operations at ORD reduced greatly after 9/11; however, since 2002, there has been a general increase in air traffic, creating a period of over-scheduling, with more than 100 daily operations at ORD. This period of increased operations made delay one of the major problems at ORD, resulting in the FAA negotiating a 5% reduction in American Airlines (AA) and United Airlines (UA) flights in January 2004. However, these vacated slots were quickly taken up by Northwest Airlines and Independence Air, resulting in a further reduction of AA and UA flights in June 2004 by 2.5% to reduce delays [20]. In August 2004, from a meeting between Federal officials and individual airlines, the scheduled arrivals of AA and UA flights were further reduced by 5 % during peak hours. Other airlines also agreed to some flight re-timings and limiting the number of scheduled arrivals. Finally, in August 2006, FAA stated a rule limiting the flight operations until the completion of first phase of ORD expansion in 2008 [20].

Various researchers have tried to understand the regional airport system in the US and all over the world. There also have been researches conducted on delay propagation from individual airports and causal factors of delay. However, these studies capture the details of only a few components of the NAS, such as specific airports, sectors, or individual flights, but fail to reflect the system overall. Our previous research has tried to capture the delay propagation phenomena from the system point of view. In the following section we will look at all these studies related to the present work.

2.4 Multi-Airport System

The FAA Fact 2 report has identified 14 airports in 10 major metropolitan regions in the US to be capacity constrained by 2015 and even more in 2025 [7] [21]. While FAA expects individual airports to improve their capacity, it also expects them to investigate the possibility of the Regional Airport System Plans (RASP) involving development of regional transportation system. In order to take the correct decision an airport planner needs to look at different alternatives like capital costs, aviation safety, airspace utilization, requirements, environmental impacts, delay and other operational costs, consistency with local area comprehensive and transportation plans, and land-use availability and compatibility [7]. Table 1, mentions the names of the airports to be capacity constrained after planned improvements; however, the number is expected to increase to 27 by 2025, if no improvements occur during this period [21] . Some of these metropolitan regions have been studied in this research and will be explained in detail.

Considering all these difficulties experienced by the existing system and even accomplishing planned improvements, developing a RASP for a metropolitan region might reduce regional congestion, develop airport benefits like lesser delays and more revenue generation and also produce political benefits like regional infrastructure development and positive environmental impacts. A Citigroup study in 2005 [22] also recommended decentralization of passengers and air cargo services from congested urban airports to nearby suburban airports for balanced capacity utilization. The Table 2 shows names of all the nine regions studied in this paper with the airports.

Table 1 Airports Needing Capacity Enhancement by 2015 and 2025 according to the FAA Fact 2 report

Year	Airports	Metropolitan Region
2015 and 2025, even after planned improvements	Newark Liberty International (EWR) LaGuardia (LGA) Long Beach (LGB) Oakland International (OAK) Philadelphia International (PHL) John Wayne (SNA)	New York New York Los Angeles San Francisco Bay Area Philadelphia Los Angeles
2025, even after planned improvements	Hartsfield-Jackson Atlanta International (ATL) Fort Lauderdale-Hollywood International (FLL) John F Kennedy International (JFK) McCarran International (LAS) Chicago Midway International (MDW) Phoenix Sky Harbor International (PHX) San Diego International (SAN) San Francisco International (SFO)	Atlanta Miami-South Florida New York Las Vegas Chicago Phoenix San Diego San Francisco Bay Area

All the airports in these regions, except New York and Houston are multi-jurisdictional with different organizations handling their operations and management [22]. Some of them are owned by different cities, different counties, municipalities, etc. Hence a coordinated operation between different airports in a specific region becomes a challenging and an intriguing task.

Neufville [6, 23-28] is a pioneer in conducting an extensive research on necessity and planning of multi-airport systems in the US and around the world. In his research, a multi-airport system is defined as set of airports that serve the airline traffic of a metropolitan area [6]. Early research found that a multi-airport system will only work when the level of originating traffic is high for the metropolitan region. In some cases it is also affected by the limitations experienced by the primary airport or some political circumstances. There are several other factors that affect multi-airport systems such as market forces, geographic location, airline traffic activity, government interferences,

Table 2 Metropolitan Regions and Airports Studied

Metropolitan Regions	Airports
Bay Area	San Francisco International (SFO) Oakland International (OAK) San Jose International (SJC)
Chicago Region	Chicago O'Hare International (ORD) Chicago Midway (MDW)
Dallas Region	Dallas-Fort Worth International (DFW) Dallas Love Field (DAL)
Houston Region	George Bush Intercontinental (IAH) Houston Hobby (HOU)
Los Angeles Region	Los Angeles International (LAX) Long Beach (LGB) Ontario International (ONT) John Wayne (SNA) Bob Hope Burbank (BUR)
New England Region	Boston Logan International (BOS) Manchester Boston Regional (MHT) T.F. Green Providence (PVD)
New York Region	John F. Kennedy International (JFK) Newark Liberty International (EWR) LaGuardia (LGA)
South Florida Region	Miami International (MIA) Fort Lauderdale-Hollywood International (FLL)
Washington-Baltimore Region	Washington Reagan National (DCA) Washington Dulles International (IAD) Baltimore/Washington International (BWI)

regional economic development, etc. The research also indicated that traffic at secondary airports is generally volatile since their concentration is less as compared to primary airport or them being depended on specific airlines [23].

A futuristic study by Neufville [25] explored the regional airport system development process in the 21st century. The research was based on three key elements namely; expected levels of traffic, development of airport systems and airport system management. This century has seen lot of changes in airline operations like airline mergers, global partnerships and introduction of new routes. Furthermore, due to city expansion airports those were only concerned with their regions have started competing for market shares of other airports. Thus, airport traffic which previously depended on region, population and economic activity is now also depended on airline and airport management [25]. This was studied in depth by Neufville [28], in the recent research on no-frill airlines and growth of secondary airports in the metropolitan regions. As contrary to previous airport operations, no-frill airlines like Southwest, Air Train, Jet Blue, Spirit and other low cost airlines (LCC) have developed a parallel airport network system [28]. The possible consequences of such development is a shift of passenger traffic from congested airports to low-cost competition airports, growth in sub-urban regions having low cost airports, decrease in growth of major airports, etc.

More recently, Bonnefoy and Hansman [13, 29] studied in detail the emergence of secondary airports and the regional airport system in the US. The research states that the emergence of secondary airports in the U.S. were due to factors like congestion at the core airport (LGA, SFO, ORD, IAH, etc), entry of new or low cost carriers in the secondary airport (MDW, FLL, PVD, MHT, HOU, etc) and change in dynamics at the airport level. An important observation made in the study was that most of the secondary airports developed were around airports having large proportion of originating traffic as compared to transfer passengers. These airports relieved major airports of increasing traffic and reduced congestion in the system. However it was also seen that closely located airports in the multi-airport system like New York region faced severe operational constraints at regional airspace level. The research highlighted the need for reducing air traffic interactions to increase the capacity of the system. Bonnefoy et al [30] also studied the evolution of multi-airport system from a worldwide perspective. It was seen that in the US and Europe development of multi-airport regions is due to emergence of secondary airports and growth of low-cost carriers. While in Asia it is mainly due to insufficiency of available airports and greater need of high capacity airports. The study suggests the need for protecting existing underutilized airports in the US and Europe with an eye for multi-airport regional development in the future. Whereas in Asia, there is need to reserve land and other resources to develop this system.

Brueckner et al [31] in their research tried to define the market for the airline industry between different metropolitan regions. Since all the metropolitan regions contain more than one airport that compete for passengers, the research tried to identify if these multiple airports need to be viewed as same or separate destination. In terms of airline market the competition will be higher when it is viewed as city-pairs as compared to airport-pairs. The grouping of airports was done using regression results, with separate analysis for each region with average nonstop fare as the dependent variable. All the regions were tested for effects of arrival and departure competing airports, year and quarter, routes and carrier, etc. It was found that all regions except Boston and Detroit can be grouped as city airports. For Boston, it is due to effect of LCC at secondary airports causing fare reduction at the primary airport.

Hess [32] used a mixed multinomial logit (MMNL) model to study the passenger airport choice in a multi-airport region of San Francisco Bay. The research tested different attributes fare, frequency, access-journey cost, flight time, size of the aircraft etc that affects airport choice in the bay region. It was found that fare, frequency and access-journey cost had significant impact on the airport choice. An important observation was passenger’s willingness to accept higher fares for the reduction in the access time to the airport. It was also seen that different types of passengers like residents, business and leisure have different requirements and react differently while choosing the airport.

Similarly, an earlier research by Hansen and Du [33] used a calibrated logit model to determine airport choice in the multi-airport region of the San Francisco Bay area. It was found that accessibility to the airports is a major factor affecting market shares at the airports. The airport market share depends largely upon the location distribution of trip origins as compared to other factors. The research clearly states that transportation planning could be used to improve airport accessibility and obtain consistent airport market share distribution.

We can see that, apart from the traditional approaches to increase the capacity like new runways, new commercial service airports, congestion management, etc. One of the steps we need to look at is ‘Regional Solutions” to study air travel behavior in different multi-airport regions in the US.

CHAPTER III

RESEARCH METHODOLOGY

This chapter presents the methodology to estimate the delay propagation from individual airports and the multi-airport systems through the rest of the RNAS. The NAS comprises of all the airports in the US and the massive network amongst them. It is important to understand the causal factors of delay at various airports and the interactions between them. To achieve these objectives, we apply regression methods to analyze the causal relationship between factors and delays and to capture interactions between airports. We also study interactions between different metropolitan regions in the U.S. having more than one airport. Previous studies (Bhargava et al [34], Cervero and Hansen [35] and Himes and Donnell [36]) used simultaneous equation regression models to study such interactions is different transportation studies. The research approach and methodology are explained in the following section.

3.1 Simultaneous Equation Regression Model

The multivariate simultaneous equation regression model is a statistical model widely used in economics and business management research studies. It has a set of multivariate equations, where the dependent variable in one equation could be the independent variable in other equations. In addition, the error terms in the equations can be correlated. This research applies multivariate simultaneous regression models to determine the delay spillover effects from individual airports or the regional airport system across the RNAS.

Bhargava et al [34] in their research used three stage least square (3SLS) regression to analyze the time and cost overruns in a highway construction project in Indiana, US. The authors identify that time and cost overruns are interdependent and their independent variables are not exogenous. Endogeneity spurs from the correlation between independent variables and the error terms and leads to biased and inconsistent estimates.

A study conducted by Cervero and Hansen [35], investigated the inter-relation between induced travel demand and induced road investment using a demand and supply simultaneous equation analysis of California covering the period 1976 and 1997. The authors used 3SLS to control for inter-dependability and cross-equation correlation of error terms. The study concludes that there is a strong interaction and simultaneity between both of them with causal factors like income, price, demographic and government policy being significant. Similarly, Himes and Donnell [36] developed a speed prediction model for multi-lane highway in North Carolina and Pennsylvania, US using system of equations. Due to the presence of endogenous variables, Himes and Donnell used 3SLS to find consistent estimates for lane speeds.

Due to the inter-dependability between delays at different airports, regions and the RNAS, we consider using 3SLS regression. The use of 3SLS also allows studying the physical interactions between airports in the NAS.

3.2 Definitions

We subdivided the U.S. airports into different groups depending upon the level of air traffic operations as explained below:

3.2.1 Operational Evolution Partnership (OEP) 35

The 35 OEP airports are commercial U.S. airports with significant activity [37]. These airports serve major metropolitan areas and also serve as hubs for airline operations. The names of all OEP airports are reported in Appendix I. Honolulu International Airport (HNL) is excluded from the list because it has somehow different characteristics due its distant location from the U.S. Continent.

3.2.2 National Airspace System (NAS)

The NAS consists of a complex collection of facilities, systems, equipment, procedures, and airports operated by thousands of people to provide a safe and efficient flying environment [38]. It includes more than 750 air traffic control (ATC) facilities, more than 18,000 airports, approximately 4,500 air navigation facilities and about 48,000 FAA employees [38]. In this study, 74 Aviation System Performance Metrics System (ASPM) airports are selected to represent the NAS, except HNL, Sacramento International Airport (SMF) and Palm Springs International Airport (PSP) because of their geographical location and data unavailability.

3.3 Data Source

We obtained the data for our research from government-maintained database, such as those maintained by the Federal Aviation Administration (FAA), the U.S. Department of Transportation and the U.S. Department of Commerce. The following sections describe these data sources.

3.3.1 Aviation System Performance Metrics System (ASPM)

We use quarter-hourly interval data from the ASPM database, maintained by FAA's Aviation Policy and Plans Office for the period 2000 to 2010. ASPM is an integrated database of air traffic operations, airline schedules, operations and delays, weather information, runway information and related statistics. The data are available starting January 2000 for 55 airports and for additional 20 airports starting October 2004 and for 2 airports from January 2007. ASPM records are created using data from a variety of sources with varying update cycles. Enhanced Traffic Management System (ETMS) and Aeronautical Radio, Incorporated (ARINC) supply next-day operational data, and Innovata provides flight schedule data, while US Department of Transportation's Aviation's Airline Service Quality Survey (ASQP) provides finalized schedule data, Out-Off-On-In (OOOI) data, and delay causes as reported by the carriers after the close of each month. ASPM is also further enhanced with inclusion of weather data and airport specific information [39]. The database is used for reporting and analysis of operating performance of the NAS.

3.3.2 National Oceanic and Atmospheric Administration (NOAA)

We obtained weather pattern data from the Surface Summary of Day database maintained by the NOAA [40]. NOAA is maintained by the U.S. Department of Commerce and provides daily weather forecasts, severe storm warnings and climate monitoring to scientific agencies, fisheries management, coastal restoration and supporting marine commerce. It provides reliable information regarding oceans and atmospheric conditions and was used in this research to assess weather conditions in the NAS. NOAA has its stations in every state in the U.S. and supplies information related to the environmental patterns.

3.3.3 U.S. Bureau of Transportation Statistics (BTS)

We used the BTS database to obtain passenger load factor data, flight schedule, historical trends and so on [41]. BTS, as a part of the U.S. Department of Transportation, compiles, analyzes, and makes information accessible on the nation's transportation systems. It improves the quality and effectiveness of DOT's statistical programs through research, development of guidelines, and promotion of improvements in data acquisition and use. BTS is a part of the Research and Innovative Technology Administration (RITA). The Air Carrier Statistics database, also known as the T-100 data bank, contains domestic and international airline market and segment data. All the certificated U.S. air carriers report monthly air carrier traffic information using Form T-100. The data are collected by RITA Office of Airline Information, Bureau of Transportation Statistics. All the air carrier data are available online from 1990 to the current year.

3.4 Dependent and Independent Variables

The following section gives the description of all dependent and independent variables used in the research.

3.4.1 Dependent Variable: Daily Average Arrival Delay

We define daily average arrival delay as the dependent variable in our model. In our previous study, the arrival delay of a flight was defined as difference between actual arrival time and the Official Airline Guide (OAG) scheduled arrival time. [42] This definition could not reflect the evolution of schedule padding introduced by the airlines in different time periods. It was observed that with limited airport capacity and increased air traffic demand, airlines intended to increase scheduled flight block timings (i.e., imbedding more padding in their flight schedules). It is a way for airlines to improve their on-time performance, which is defined as the percentage of flights arrive no later than 15 minutes after their scheduled arrival time. [43] [44]. Thus, the schedule-based analysis does not give us accurate enough results. Figure 3 shows us the gate-to-gate timings for flights between Atlanta (ATL) and Orlando (MCO) airports obtained from the BTS database for years 1995 to 2011. It is clearly seen that average gate-to-gate flight timings are continuously increasing throughout the years. A U.S. government report on economic analysis of flight delay clearly mentions that schedule padding in flights increased before and after 9/11 incident to compensate for flight delays [45].

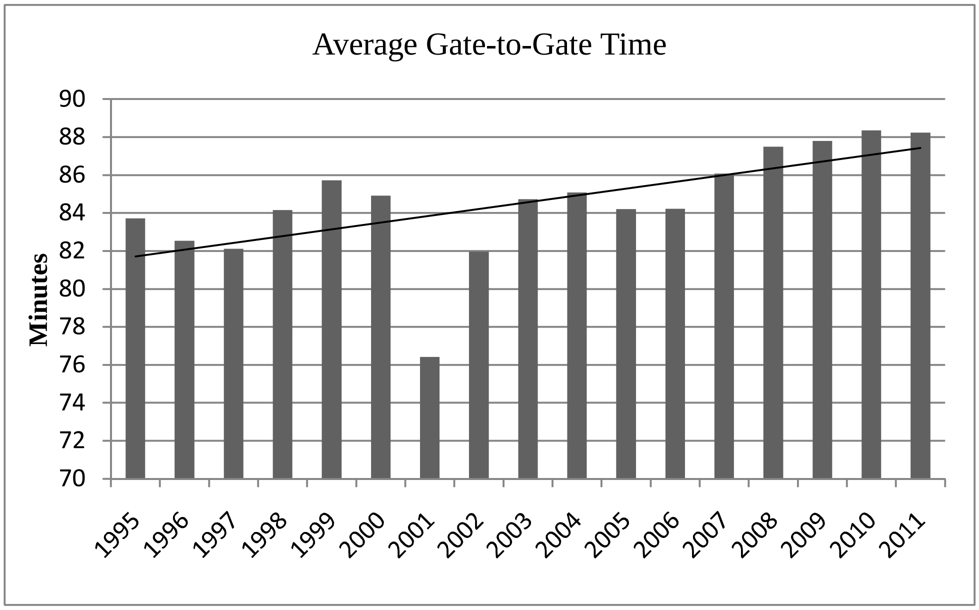


Figure 3 Increase in Schedule Time for Flights between ATL and MCO Airports

Therefore, in this research we use flight-plan-based arrival delay, which is equal to the difference between actual arrival time of a flight and predicted arrival time according to the flight plan. Then the daily average of each airport is calculated by dividing the total delay with the number of total arrivals. Note that if one flight arrived earlier than the flight-plan arrival time, the delay is considered as zero.

3.4.2 Independent Variables

In this research we investigated the effects of different factors like queuing delay, adverse weather, airline scheduling, demand management regimes, etc on airport arrival delay. The data downloaded from the ASPM database is used to compute a set of independent variables used in the model. Appendix II reports the data dictionary used to calculate these variables and describes the input variables used in the analysis.

3.4.2.1 Deterministic Queuing Delay

Deterministic queuing delay indicates the operational demand and supply relationship at each airport. The arrival count is the actual number of arrivals at the airports in 15 minutes, which is restricted by the number of flights that need to land, and by Airport supplied Arrival Rate (AAR) during the same time period. In other words, if the number of flights waiting to land is larger than the AAR rate, then the arrival count is the AAR rate; otherwise, the arrival count is the number of flights that need to land. The cumulative flight demand in one quarter-hourly interval is the remaining scheduled arrival demand until the end of the quarter-hourly interval.

Figure 4 shows the Newell Curve of cumulative number of arrivals, where the actual arrival counts are always less than arrival demand since arrival counts are either restricted by arrival demand or the capacity of the airport. The daily average queuing delay at an airport is calculated by dividing the area between the curves, which is known as total queuing delay, by the total number of arrivals at the airport for that day. The same definition applies to the RNAS as well, where the daily average arrival delay is the total arrival queuing delay at the RNAS airports divided by the total number of arrivals at those airports. The hypothesis that we would like to test is the increase of queuing delay leading to more observed flight delay.

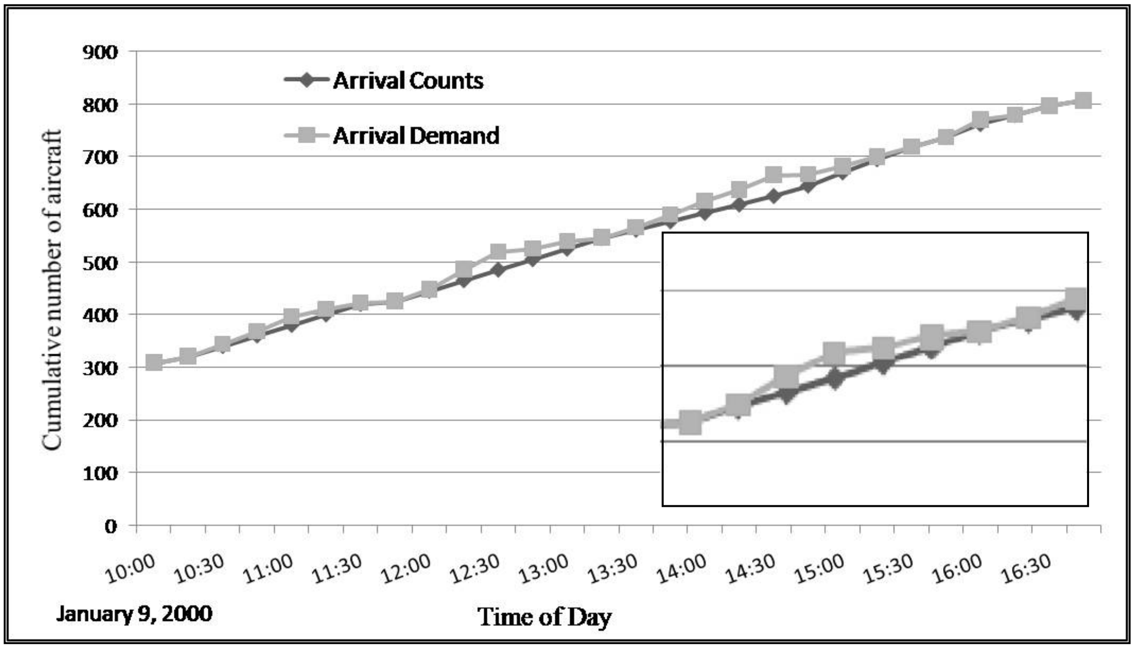


Figure 4 Queuing Diagram of Arrivals at ORD

3.4.2.2 Adverse Weather Indicator

Adverse weather is one of the most important factors causing delay. Our research introduces adverse weather into the regression model by means of two indicators. One indicator is used to capture the convective weather on the route. To measure convective weather, the U.S. is divided into 16 regions of 10 degrees latitude by 10 degrees longitude, as shown in Figure 5. For each region, the proportion of weather stations reporting thunderstorms is computed from the Surface Summary of Day database maintained by the National Oceanographic and Atmospheric Administration (NOAA). Using thunderstorm data, the thunderstorm ratio is calculated as the ratio of the number of stations reporting thunderstorms by the total number of stations. The effect of regional convective weather on airport delay is complicated. Considering flights from different origins to the reference airport, the convective weather in a particular region may hold some flights that alleviate the congestion at the reference airport. However, if the flights held are released later in a batch, then the concentrated cumulative arrive will deteriorate the operational condition at the reference airport. For this variable, we wait to see what the data tells us once we control for all other variables.

The weather close to the airport directly affects the determination of airport runway configurations and utilization of runways. We propose to use the Instrument Meteorological Condition (IMC) ration to measure it. It is calculated as the proportion of the day in which the airport was under IMC conditions. It is known that an airport operating under IMC conditions has a lower capacity than that operating under VMC conditions, which causes more delays.

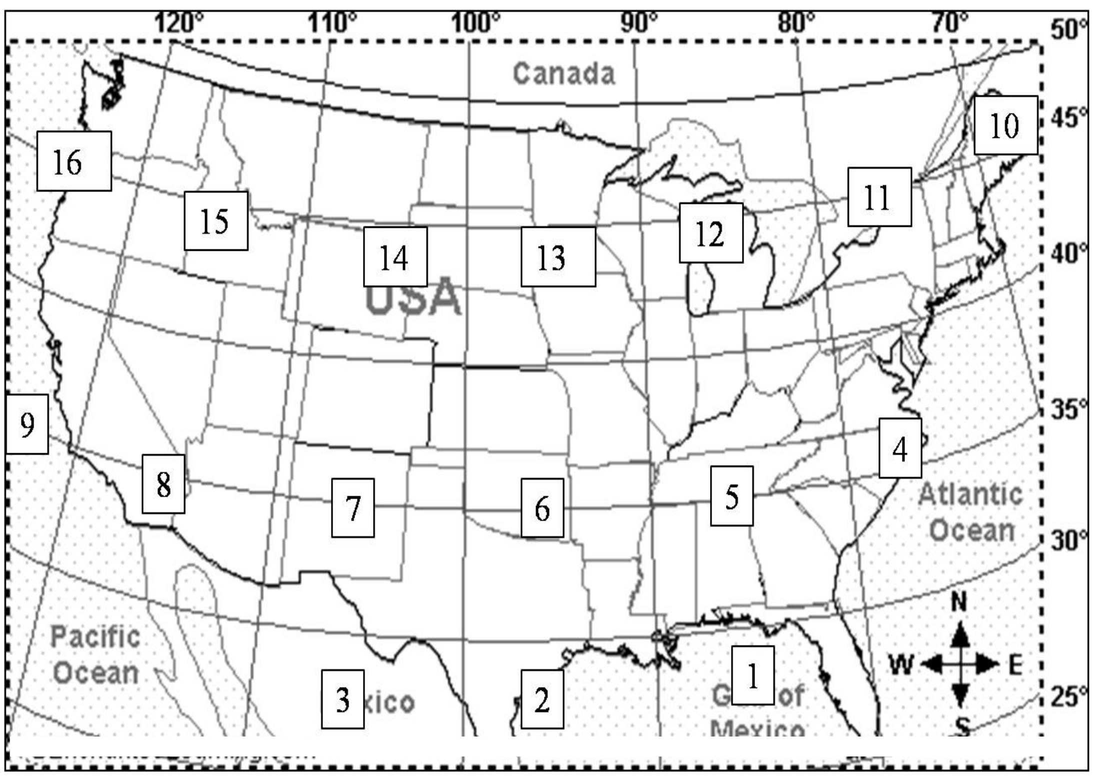


Figure 5 U.S. Weather Regions

3.4.2.3 Passenger Load Factor

The BTS database contains domestic monthly data reported by U.S. air carriers, including carrier, origin, destination, aircraft type and service class for transported passengers, freight and mail, available capacity, scheduled departures, departures performed, aircraft hours, and load factor when both origin and destination airports are located within the boundaries of the United States and its territories. In our first study for estimating the impact of individual airport, as shown in Chapter IV, passenger load factor is considered as an explanatory variable for the daily average delay. This is because higher passenger load factor, busier the airlines are more variation will be caused towards turnaround time of the flights and causes the delay at the airport. It is calculated as the monthly average ratio of the number of passengers by the number of seats available at the airport under consideration.

3.4.2.4 Aircraft Equipment Type

This variable is tested as an alternative for passenger load factor in our second study (Chapter V) for estimating the impact of individual airport. The aircraft equipment type is categorized based on International Air Transportation Association (IATA) Aircraft Size Classification Scheme as observed in Table 3. Since most airport design standards are related to aircraft size, it is necessary to understand the effect of aircraft fleet size on delay at various airports. In this research, aircrafts are classified into seven categories based on their seat capacities and categories defined by the IATA [46]. The mean weighted number of aircrafts for all seven groups for each quarter is then used as a variable for each individual airport.

Table 3 International Air Transportation Association (IATA) Aircraft Size Classification Scheme

Category	Number of Seats	Aircraft
0	< 50	Embraer 120, Saab 340
1	50 - 124	Fokker 100, Boeing 717
2	125 – 179	Boeing B727 - 200, Airbus A321
3	180 – 249	Boeing 767 – 200, Airbus A300 - 600
4	250 - 349	Airbus A340 – 300, Boeing 777 – 200
5	350 – 499	Boeing 747 – 400
6	> 500	Boeing 747 – 400 high density seating

3.4.2.5 Total Flight Operations (Air Traffic Volume)

The RNAS model also considers the total flight operations as one of the variables. It captures the effects of total air traffic volume on delay in the system. We assume that with the increase in air traffic volume, there is an increase in the airport delay.

New variables were introduced in our third study for multi-airport systems to understand the impact of different attributes causing delay propagation from the entire region to the RNAS.

3.4.2.6 LCC Airline Market Share

The term Low Cost Carriers (LCC) originated within the airline industry and refers to airlines with a lower operating cost structure than their competitors. To keep their operating costs lower, these airlines apply business models that are different from legacy airlines. For instance, they use only one type of aircraft to reduce crew and maintenance costs, and serve secondary airports to avoid congestion and high landing fees at primary airport in the same region. LLC also try to operate with cost-effective ways of handling passengers. LCC airlines operating at secondary airports are the prime reason for the development of the multi-airport region phenomena [47].

In this study, we calculate the percentage share of LLC operations at each airport in the region and include it as an explanatory variable to understand its impact on the delay in the region and in the RNAS.

3.4.2.7 Herfindahl–Hirschman Index (HHI)

The Herfindahl–Hirschman Index (HHI) is a measure of the size of the component in relation to a group and an indicator of competition among them [13]. It is seen that with the entry of LCCs, there is an increased level of competition for market share in each region. The HHI is calculated as the sum of squares of airport market shares for the time period 2000 to 2010—the higher the value, the less competitive the region. For instance, the HHI of the Los Angeles region with five airports is 0.53, whereas for the New York region with 4 airports, it is 0.33. This indicates that the New York region is much more competitive compared to the Los Angeles region, where 70 percent of operations in that region occur at LAX airport.

3.4.2.8 Demand Management Regimes

We use dummy variables to indicate various demand management regimes used at LGA, JFK and ORD airports at given time periods, as shown in Figure 1. For instance at the ORD airport, variable AIR21 takes a value of 1 from May 2000 to December 2000 and zero otherwise. This process was carried out continuously from 2000 to 2010. As shown in Table 4, there are total of 14 dummy variables indicating different operational strategies used at the three airports indicated above and HDR is used as the base for comparison.

Table 4 Summary of Demand Management Regime Applied in the Model

Period	**Demand Management Regime**
January 2000 to April 2004	High Density Rule (HDR)
May 2000 to December 2000	AIR 21
Year 2001 till September 9, 2001	Before 9/11
September 20, 2011 till December 2001	After 9/11
2002	OV2002
2003	OV2003
January 2004 till May 2004	CAP
May 2004 till December 2004	REDA
2005	REDB
January 2006 till July 2006	REDC
August 2006 till December 2006	LIM
2007	Year2007
2008	Year2008
2009	Year2009
2010	Year2010

3.4.2. 9.Seasonal Dummy Variables

We introduce dichotomous variables to indicate different seasons throughout the year. Three dummy variables introduced for different seasons namely summer, fall and winter, with spring as the base. Since the traffic demand varies and the airport operations are affected significantly for different seasons at all the airports. Assuming the seasonal weather variation has been controlled for by the weather indicators described earlier, this seasonal variable is proposed to capture airlines scheduling trends in different seasons.

3.5 Descriptive Statistics

We obtained data for all the independent variables for the period 2000 to 2010. The following tables show the sample descriptive statistics for the Hartsfield Jackson Atlanta International Airport (ATL) airport. We carried similar analyses for the other individual airports and the RNAS.

The maximum daily average arrival queuing delay of 11,805.00 minutes was observed on September 12, 2001 just after the terrorist attacks. Similarly, the least number of flights flown on a single day that is two was also on the same day. In our research we have excluded those ten days of data from September 11 to September 20, 2011 to get accurate and consistent results. Also for the thunderstorm ratios there are eleven missing values for the first day of every year from 2000 to 2010 and will be excluded from the dataset. The following Table 5 shows the descriptive statistics of the data considered.

Table 5 Descriptive Statistics

Variable	N	Mean	Standard Deviation	Minimum	Maximum
arrobdelay	4008	12.627	11.939	1.792	99.102
depobdelay	4008	18.946	11.922	4.429	206.530
arraqdelay	4008	5.384	13.011	0	347.980
arraqdelay2	4008	198.251	2714.49	0	121090.51
depaqdelay	4008	6.435	8.649	0.071	215.494
IFR_ratio	4008	0.235	0.294	0	1.000
IFR_ratio2	4008	0.141	0.249	0	1.000
Region1	4007	0.107	0.150	0	0.787
Region2	4007	0.084	0.134	0	0.727
Region3	4007	0.044	0.152	0	1.000
Region4	4007	0.054	0.093	0	0.559
Region5	4007	0.108	0.133	0	0.671
Region6	4007	0.114	0.138	0	0.807
Region7	4007	0.085	0.111	0	0.613
Region8	4007	0.026	0.048	0	0.394
Region9	4007	0.006	0.023	0	0.500
Region10	4006	0.022	0.072	0	0.733
Region11	4007	0.0490.053	0.098	0	0.789
Region12	4007	0.064	0.094	0	0.558
Region13	4007	0.079	0.115	0	0.686
Region14	4007	0.038	0.125	0	0.714
Region15	4007	0.009	0.073	0	0.528
Region16	4007	28045.89	0.026	0	0.327
EQPT1	4008	1112.09	3468.73	6875.00	35400.00
EQPT2	4008	100752.15	776.202	0	3393.00
EQPT3	4008	60385.98	33398.68	4560.00	194560.00
EQPT4	4008	62730.76	50667.85	2795.00	135020.00
EQPT5	4008	1377.33	64139.67	0	180600.00
EQPT6	4008	62.874	2842.21	0	11900.00
EQPT7	4008	0.030	259.895	0	3000.00
HDR	4008	0.061	0.171	0	1.000
AIR	4008	0.063	0.239	0	1.000
Sepb	4008	0.025	0.243	0	1.000
Sepa	4008	0.091	0.157	0	1.000
OV2002	4008	0.091	0.287	0	1.000
OV2003	4008	0.037	0.287	0	1.000
CAP	4008	0.053	0.191	0	1.000
REDA	4008	0.091	0.224	0	1.000
REDB	4008	0.052	0.287	0	1.000
REDC	4008	0.038	0.223	0	1.000
LIM	4008	0.091	0.191	0	1.000
Year2007	4008	0.091	0.287	0	1.000
Year2008	4008	0.091	0.288	0	1.000
Year2009	4008	0.091	0.287	0	1.000
Year2010	4008	0.247	0.287	0	1.000
quarter1	4008	0.249	0.431	0	1.000
quarter2	4008	0.250	0.432	0	1.000
quarter3	4008		0.433	0	1.000
quarter4	4008		0.434	0	1.000

3.6 Correlation Analysis between Independent Variables

If the independent variables used in the analysis are correlated it creates the problem of multi-collinearity. In that case, parameter estimates will become unreliable, exhibiting large p-values or confidence intervals. Hence it was necessary to check if this problem exists in our dataset. This correlation between independent variables could dealt by removing a variable, introducing variable interactions or by increasing the sample size [48]. As a part of this process we tested the correlation for independent variables for all the airports namely from January 2000 to December 2010.

Table 6 shows the relationship between average daily arrival and departure queuing delay at LGA and ORD airports. As seen in the table there is a high degree of positive correlation between both arrival and departure queuing delay. Hence in our research we have only used arrival queuing delay as our explanatory variable.

Table 6 Correlation Analysis between Arrival and Departure Queuing Delay

LGA	Arrival Queuing Delay	Departure Queuing Delay
Arrival Queuing Delay	1.000	0.819
		<0.0001
Departure Queuing Delay	0.819	1.000
	<0.0001	
ORD	Arrival Queuing Delay	Departure Queuing Delay
Arrival Queuing Delay	1.000	0.894
		<0.0001
Departure Queuing Delay	0.894	1.000
	<0.0001	

Table 7 displays the correlation between different explanatory variables used to analyze average daily arrival delay at the ORD airport. From the results it is learned that only the dummy variable ‘Over_Scheduling’ shares 33.75 percent similarity with another dummy variable 'Partial_HDR'. For all other variables there is no significant correlation between them. Table 8 includes correlation analysis results for explanatory variables at the LGA airport. No significant correlations are observed between the variables.

Table 7 Correlation Analysis for Independent Variables at the ORD Airport

ORD	Queuing	IFR ratio	Total	HDR	Partial HDR	Sep_11	Over Sche-duling	Five	Q1	Q2	Q3	Q4	Pred Variable
Queuing	1	.296	-.140	-.039	-.072	.015	.001	.123	.000	.012	-.032	.019	.091
		.000	.000	.114	.004	.538	.980	.000	.987	.627	.201	.439	.000
IFR ratio	.296	1	-.216	.054	-.005	.015	-.024	-.011	.110	-.046	-.151	.082	.177
	.000		.000	.030	.825	.531	.329	.643	.000	.063	.000	.001	.000
Total	-.140	-.216	1	-.146	-.049	-.301	.057	.354	-.084	.094	.039	-.050	-.068
	.000	.000		.000	.045	.000	.020	.000	.001	.000	.112	.042	.007
HDR	-.039	.054	-.146	1	-.183	-.080	-.252	-.100	.301	-.018	-.151	-.151	.125
	.114	.030	.000		.000	.001	.000	.000	.000	.459	.000	.000	.000
Partial HDR	-.072	-.005	-.049	-.183	1	-.184	-.581	-.229	-.132	.050	.143	-.055	.438
	.004	.825	.045	.000		.000	.000	.000	.000	.042	.000	.025	.000
Sep_11	.015	.015	-.301	-.080	-.184	1	-.253	-.100	-.174	-.175	.015	.360	-.140
	.538	.531	.000	.001	.000		.000	.000	.000	.000	.546	.000	.000
Over Sche-duling	.001	-.024	.057	-.252	-.581	-.253	1	-.316	-.057	-.055	.060	.060	-.417
	.980	.329	.020	.000	.000	.000		.000	.021	.025	.015	.015	.000
Five	.123	-.011	.354	-.100	-.229	-.100	-.316	1	.178	.176	-.190	-.190	.023
	.000	.643	.000	.000	.000	.000	.000		.000	.000	.000	.000	.353
Q1	.000	.110	-.084	.301	-.132	-.174	-.057	.178	1	-.381	-.331	-.331	.016
	.987	.000	.001	.000	.000	.000	.021	.000		.000	.000	.000	.516
Q2	.012	-.046	.094	-.018	.050	-.175	-.055	.176	-.381	1	-.332	-.332	-.002
	.627	.063	.000	.459	.042	.000	.025	.000	.000		.000	.000	.935
Q3	-.032	-.151	.039	-.151	.143	.015	.060	-.190	-.331	-.332	1	-.289	.027
	.201	.000	.112	.000	.000	.546	.015	.000	.000	.000		.000	.284
Q4	.019	.082	-.050	-.151	-.055	.360	.060	-.190	-.331	-.332	-.289	1	-.042
	.439	.001	.042	.000	.025	.000	.015	.000	.000	.000	.000		.093
Pred Variable	.091	.177	-.068	.125	.438	-.140	-.417	.023	.016	-.002	.027	-.042	1
	.000	.000	.007	.000	.000	.000	.000	.353	.516	.935	.284	.093	

Table 8 Correlation Analysis for Independent Variables at the LGA Airport

LGA	Queing	IFR	Total	HDR	AIR21	Slottry	Sep11	2002	2003	2004	Q1	Q2	Q3	Q4	Predicted Variable
Queing Delay	1	.020	.126	-.058	.155	.037	-.051	-.176	-.108	.248	-.058	.042	.003	.014	.064
		.414	.000	.019	.000	.139	.041	.000	.000	.000	.019	.089	.909	.567	.011
IFR	.020	1	-.173	.034	.014	-.006	-.083	-.062	.082	.005	.008	.086	-.052	-.049	.290
	.414		.000	.170	.565	.808	.001	.012	.001	.848	.755	.000	.035	.047	.000
Total	.126	-.173	1	-.066	.081	.166	-.297	-.105	-.003	.166	-.031	.098	-.054	-.017	.052
	.000	.000		.007	.001	.000	.000	.000	.888	.000	.212	.000	.027	.482	.039
HDR	-.058	.034	-.066	1	-.118	-.117	-.080	-.151	-.151	-.100	.301	-.018	-.151	-.151	.084
	.019	.170	.007		.000	.000	.001	.000	.000	.000	.000	.459	.000	.000	.001
AIR21	.155	.014	.081	-.118	1	-.174	-.119	-.224	-.224	-.148	-.258	-.026	.152	.152	.388
	.000	.565	.001	.000		.000	.000	.000	.000	.000	.000	.289	.000	.000	.000
Slottery	.037	-.006	.166	-.117	-.174	1	-.118	-.223	-.223	-.147	.089	.091	.031	-.224	.124
	.139	.808	.000	.000	.000		.000	.000	.000	.000	.000	.000	.207	.000	.000
Sep_11	-.051	-.083	-.297	-.080	-.119	-.118	1	-.151	-.151	-.100	-.174	-.175	.015	.360	-.132
	.041	.001	.000	.001	.000	.000		.000	.000	.000	.000	.000	.546	.000	.000
Year2002	-.176	-.062	-.105	-.151	-.224	-.223	-.151	1	-.286	-.189	-.034	-.033	.036	.036	-.248
	.000	.012	.000	.000	.000	.000	.000		.000	.000	.167	.181	.145	.145	.000
Year2003	-.108	.082	-.003	-.151	-.224	-.223	-.151	-.286	1	-.189	-.034	-.033	.036	.036	-.222
	.000	.001	.888	.000	.000	.000	.000	.000		.000	.167	.181	.145	.145	.000
Year2004	.248	.005	.166	-.100	-.148	-.147	-.100	-.189	-.189	1	.178	.176	-.190	-.190	.072
	.000	.848	.000	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000	.004
Q1	-.058	.008	-.031	.301	-.258	.089	-.174	-.034	-.034	.178	1	-.381	-.331	-.331	.023
	.019	.755	.212	.000	.000	.000	.000	.167	.167	.000		.000	.000	.000	.358
Q2	.042	.086	.098	-.018	-.026	.091	-.175	-.033	-.033	.176	-.381	1	-.332	-.332	.016
	.089	.000	.000	.459	.289	.000	.000	.181	.181	.000	.000		.000	.000	.517
Q3	.003	-.052	-.054	-.151	.152	.031	.015	.036	.036	-.190	-.331	-.332	1	-.289	.023
	.909	.035	.027	.000	.000	.207	.546	.145	.145	.000	.000	.000		.000	.364
Q4	.014	-.049	-.017	-.151	.152	-.224	.360	.036	.036	-.190	-.331	-.332	-.289	1	-.065
	.567	.047	.482	.000	.000	.000	.000	.145	.145	.000	.000	.000	.000		.010
Predicted Variable	.064	.290	.052	.084	.388	.124	-.132	-.248	-.222	.072	.023	.016	.023	-.065	1
	.011	.000	.039	.001	.000	.000	.000	.000	.000	.004	.358	.517	.364	.010	

Table 9 Correlations Analysis for Thunderstorm Ratio

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16
R1	1.000	**0.311**	0.054	**0.460**	0.514	**0.269**	**0.540**	**0.379**	-0.013	**0.296**	**0.351**	**0.274**	**0.328**	**0.498**	**0.380**	**0.121**
R2	**0.311**	1.000	**0.292**	**0.247**	**0.344**	**0.466**	**0.339**	**0.191**	-0.033	**0.163**	**0.205**	**0.163**	**0.167**	**0.208**	**0.170**	0.033
R3	0.054	**0.292**	1.000	**0.072**	**0.100**	**0.194**	**0.147**	**0.081**	-0.036	0.061	**0.070**	0.041	0.015	0.039	0.043	0.006
R4	**0.460**	**0.247**	**0.072**	1.000	**0.629**	**0.269**	**0.384**	**0.289**	-0.026	**0.254**	**0.566**	**0.328**	**0.222**	**0.357**	**0.322**	**0.141**
R5	**0.514**	**0.344**	**0.100**	**0.629**	1.000	**0.534**	**0.405**	**0.264**	-0.046	**0.248**	**0.429**	**0.488**	**0.292**	**0.371**	**0.301**	**0.134**
R6	**0.269**	**0.466**	**0.194**	**0.269**	**0.534**	1.000	**0.402**	**0.117**	-0.049	**0.155**	**0.255**	**0.354**	**0.315**	**0.261**	**0.162**	0.030
R7	**0.540**	**0.339**	**0.147**	**0.384**	**0.405**	**0.402**	1.000	**0.498**	-0.051	**0.270**	**0.343**	**0.280**	**0.345**	**0.647**	**0.492**	**0.120**
R8	**0.379**	**0.191**	**0.081**	**0.289**	**0.264**	**0.117**	**0.498**	1.000	**0.091**	**0.188**	**0.229**	**0.146**	**0.111**	**0.317**	**0.363**	**0.067**
R9	-0.013	-0.033	-0.036	-0.026	-0.046	-0.049	-0.051	**0.091**	1.000	-0.014	-0.017	-0.014	-0.025	**-0.065**	-0.009	**0.093**
R10	**0.296**	**0.163**	0.061	**0.254**	**0.248**	**0.155**	**0.270**	**0.188**	-0.014	1.000	**0.476**	**0.132**	**0.137**	**0.292**	**0.235**	**0.084**
R11	**0.351**	**0.205**	**0.070**	**0.566**	**0.429**	**0.255**	**0.343**	**0.229**	-0.017	**0.476**	1.000	**0.389**	**0.133**	**0.323**	**0.285**	**0.122**
R12	**0.274**	**0.163**	0.041	**0.328**	**0.488**	**0.354**	**0.280**	**0.146**	-0.014	**0.132**	**0.389**	1.000	0.471	**0.293**	**0.182**	**0.074**
R13	**0.328**	**0.167**	0.015	**0.222**	**0.292**	**0.315**	**0.345**	**0.111**	-0.025	**0.137**	**0.133**	**0.471**	1.000	**0.498**	**0.222**	**0.063**
R14	**0.498**	**0.208**	0.039	**0.357**	**0.371**	**0.261**	**0.647**	**0.317**	**-0.065**	**0.292**	**0.323**	**0.293**	**0.498**	1.000	**0.600**	**0.130**
R15	**0.380**	**0.170**	0.043	**0.322**	**0.301**	**0.162**	**0.492**	**0.363**	-0.009	**0.235**	**0.285**	**0.182**	**0.222**	**0.600**	1.000	**0.316**
R16	**0.121**	0.033	0.006	**0.143**	**0.134**	**0.030**	**0.120**	**0.067**	**0.093**	**0.084**	**0.122**	**0.074**	**0.063**	**0.130**	**0.316**	1.000

The correlation results for thunderstorm ratio show very interesting characteristics. Regions that are very close to each other geographically, shows certain degree of correlation. For instance Region 7 and Region 14 located adjacent to each other are 41.8 percent correlated. Similar correlations could be seen for Region 1 and 5, Region 4 and 5, Region 4 and 11, Region 12 and 5 and so on. As seen in the Table 9, numbers that are displayed in *bold* are the ones that are correlated. However the coefficient of correlation is very small for all the cases.

A correlation analysis was also conducted for all the independent variables at different airports. It was seen that there was no correlation indicating the independence of explanatory variables for different airports used in the analysis.

After conducting this preliminary analysis we learn that there is insignificant amount of correlation between independent variables for the same airport. In the following section we discuss the mathematical format of the MSERM and the regression techniques applied.

3.7 Regression Methods

Since the equations developed in this research, mentioned in later chapters include both endogenous and exogenous explanatory variables, that means that the dependent variable in one equation of interest is the independent variable in another or more equations and vice versa. This could create the problem of identification if no enough variables are excluded from each equation. Also selecting the right estimation technique to solve this complex simultaneous equation models becomes very important.

3.7.1 Problem of Identification

The problem of identification may occur in a multi-equation model where the equations have both endogenous and exogenous explanatory variables. Consider a linear system of M equations, with M > 1. According to the order condition, an equation cannot be identified from the data if less than (M− 1) variables are excluded from that equation. For instance, for a model with four equations, at least three variables from each particular equation have to be exclusive to make sure there is no identification problem. The simultaneous equation system considered in this research identifies daily average arrival delay at all the airports and RNAS as the endogenous variable. All other independent variables are exogenous since they are uncorrelated and unique for different airports. Hence, in our system of 35 equations, more than 34 exogenous variables are exclusive from each equation. So there is no identification problem in our proposed simultaneous equation regression model.

3.7.2 Regression Techniques

Regression analysis is defined as a way of estimating or predicting the mean or average value of the dependent variable on the basis of the known values of the independent or explanatory variables [49].

We define the equation as, $Y_i = \beta_0 + \beta_1 X_i + u_i$

Where, β_0 and β_1 are unknown parameters also called as regression coefficients.

And $\beta_0 + \beta_1 X_i$ are called the systematic component, while u_i is the random component.

If $\hat{\beta}_0$and $\hat{\beta}_1$ are the estimates of β_0 and β_1.

Then $\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$ is the sample regression function and $\hat{Y}$ is the predicted value of Y

The Ordinary Least Square (OLS) method is the most popular economic method for estimating the unknown parameters in a linear regression model. This is a basic method on which all other methods are dependent.

3.7.2.1 The Ordinary Least Square (OLS) Method

In econometrics, the Ordinary Least Square (OLS) method is the most popular method used to obtain estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ [49]. This method chooses the values of $\widehat{\beta_0}$ and $\hat{\beta}_1$ such that it minimizes the sum of squared residuals

$$\sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_1)^2$$

A few of the basic assumptions for the OLS method are as follows

- Error term has zero mean, $\sum e_i = 0$
- Error term is uncorrelated with regressors, $\sum e_i X_i = 0$
- Error term has constant variance $(V(u_i | \mathrm{X_i}) = \sigma^2)$ and error terms are uncorrelated with each other$(E\left(u_i u_j\right) = 0\ for\ i \neq j)$.

$(V(u_i | \mathrm{X_i}) = \sigma^2)$ indicates homoskedasticity or constant variance assumption and if $V(u|$ X) depends upon X, then it indicates that the error term exhibits heteroskedasticity. Also $(E\left(u_i u_j\right) = 0\ for\ i \neq j)$ is known as no autocorrelation assumption

We need to overcome these problems faced in the OLS method; using different approaches that are modifications of the OLS method and are discussed in brief in the next part. All these approaches are categorized into two parts. The difference between the two approaches is that, the system estimation method takes into consideration full information like parameter restrictions and correlation of the error term while the single equation system ignores it.

We have identified two econometric approaches to estimate the simultaneous linear equation models as shown below, including the single equation estimation method and the system estimations method [50]. Both these techniques are explained in brief in following sections.

3.7.2.2 Single Equation Estimation Methods

This method considers one equation at a time, estimating the structural form as does the OLS method. It uses the information as to which variables, both endogenous and exogenous, is included in the other equations of the model but excluded from the equation being estimated. In this group there are, following methods: the indirect least squares method (ILS) and the two-stage least squares method (2SLS).

3.7.2.2.1 Indirect Least Squares Method (ILS)

The ILS method uses OLS to estimate the reduced form of equations, and then converts the OLS estimates from the reduced form into the estimates of the structural form of equations. This method produces estimates that are consistent, but not unbiased

[51]. This method is used for just-identified system of equations. The 2SLS method is similar to ILS if the system is just identified.

3.7.2.2.2 Two Staged Least Square (2SLS)

A common approach when confronted with autocorrelation and heteroskedasticity problem in the linear regression context is to try to use the technique of instrumental variables (IV), also known as the two-stage least squares (2SLS) [52]. This method does not give unbiased estimates, but does give consistent estimates. The first step involves, estimating the model in by least squares to get consistent estimates of the endogenous variables, and compute the model predictions. In the second step, we estimate the model in by least squares, but replacing endogenous variable with the predictions obtained in the first stage. The key assumption needed for consistency of the IV estimator is that the instruments and error term are uncorrelated.

3.7.2.3 System Estimation Methods:

This approach estimates the entire model of the simultaneous linear equations together, using all information's available on each of the equations of the system. We consider two methods in this approach: three-stage least squares method (3SLS) and full-information maximum likelihood method (FIML).

3.7.2.3.1 Three Stage Least Square (3SLS)

The 3SLS method combines two statistical techniques; one is the two-stage least square (2SLS) and the other is the seemingly unrelated regression (SUR). The 3SLS method generalizes the two-stage least-squares method by taking account of the correlations between equations in the same way that SUR generalizes OLS. Three-stage least squares method contains three steps: first-stage regressions to get predicted values for the endogenous regressors; a two-stage least-squares step to get residuals to estimate the cross-equation correlation matrix; and the final 3SLS estimation step. The first two stages of the 3SLS method are similar to the previously discussed 2SLS method. The third stage which is SUR is an extension of a linear regression model allowing correlated errors between equations. It is a way of improving the efficiency of estimation equations jointly, as it provides consistent estimates for linear equations.

3.7.2.3.2 Full Information Maximum likelihood Method (FIML)

The FIML method obtains maximum likelihood estimates of a nonlinear simultaneous equations model. The model should have N equations for N endogenous variables. FIML is an asymptotically efficient estimator for simultaneous models with normally distributed errors. Some of the key aspects of FIML are as follows [53]:

- FIML does not require instrumental variables.
- FIML requires that the model include the full equation system, with as many equations as there are endogenous variables. With 2SLS or 3SLS you can estimate some of the equations without specifying the complete system.
- FIML assumes that the equation errors have a multivariate normal distribution. If the errors are not normally distributed, the FIML method may produce poor results. 2SLS and 3SLS do not assume a specific distribution for the errors.

CHAPTER IV

A MACROSCOPIC TOOL FOR MEASURING DELAY PERFORMANCE IN THE NATIONAL AIRSPACE SYSTEM: CASE STUDY OF ORD AND LGA AIRPORTS

We conducted the case study of delay propagation from individual airports (LGA and ORD) to the RNAS. This research follows a similar path of macroscopic analysis that was conducted and not only investigates the impact of single airport delay to the RNAS but also to explore how the delay spillover is widely dispersed across the Operational Evolution Partnership (OEP) 34 airports (see Appendix I). The remaining 40 airports in NAS, except the 34 OEP airports (excluding HNL), are grouped together and are known as the Rest of the NAS (RNAS). RNAS delay is considered aggregately using a single multivariate equation in the simultaneous equation regression model. Therefore, our model consists of 35 equations determining average daily arrival delay at 34 OEP airports and one equation for the RNAS. Causal factors of the average daily arrival delays are explored, and multivariate equations are developed for all airports under consideration along with the RNAS. The average daily arrival delay is the dependent variable in the equation for each airport and the RNAS, while simultaneously being considered as an independent variable in the equation of other airports and the RNAS.

4.1 Methodology

In our previous study, we developed a set of multivariate simultaneous equations for both individual airports and the RNAS. We regressed these models using two-staged least square (2SLS), as seen in Figure 6.a. As observed in Figure 6.b and 6.c, we used the predicted value of the average observed arrival delay at the RNAS as the independent variable for average observed arrival delay at an individual airport and vice versa. This predicted value is the dependent variable created at the end of the first stage of regression and, along with the other variables, was used in the second stage to regress arrival delays with full models along with heteroskedastic error correction. The auto-correlation, however, are insignificant in this case.

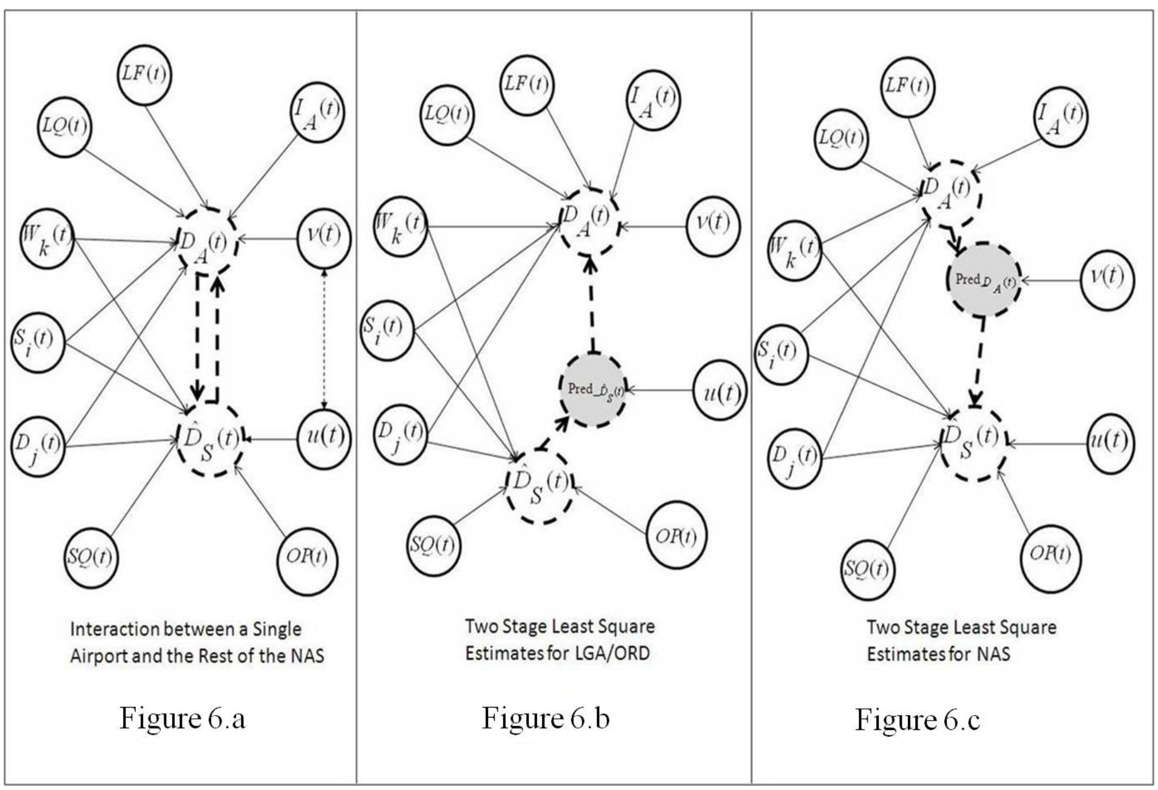


Figure 6 Two Stage Least Square Regression

The data we used were from ASPM covering the period of January 2000 to June 2004. The model for the individual airport decomposes average daily delay at LGA or ORD into components related to different delay casual factors explained earlier. The explanatory variables include average arrival deterministic queuing delay, average observed arrival delay at other airports, adverse weather, seasonal effects, demand management regimes, and other factors. Whereas, the NAS model decomposes average daily delay at airports other than the airports under consideration (LGA or ORD). The explanatory variables include observed delays at LGA or ORD, convective weather, total operations, seasonal effects, demand management regimes, and other factors.

4.1.1 Equation 1 for Individual Airport

The equation for the individual airport decomposes average daily delay at a reference airport into components related to different delay-causing factors. The explanatory variables include average arrival deterministic queuing delay, average observed arrival delay at other airports, adverse weather, and other factors.

$$D_A(t) = \beta_1 \times D_S(t) + \beta_2 \times LQ(t) + \beta_3 \times LQ^2(t) + \beta_4 \times LF(t) + \beta_5 \times I_A(t) + \beta_6 \times {I_A}^2(t) + \sum_k \lambda_{kA} W_k(t) + \sum_i \omega_{iA} S_i(t) + \sum_j \theta_{jA} D_j(t) + v(t)$$

4.1.2 Equation 2 (daily average arrival delay at RNAS)

The model for the RNAS decomposes average daily delay at airports other than the airports under consideration (LGA or ORD). The explanatory variables include observed delays at LGA or ORD, convective weather, total operations, and other factors.

$$D_S(t) = \gamma_1 \times OP(t) + \gamma_2 \times D_A(t) + \gamma_3 \times SQ(t) \quad + \sum_k \lambda_{kS} W_k(t) +$$

$$\sum_i \omega_{iS} S_i(t) + \sum_j \theta_{jS} D_j(t) + u(t)$$

$D_A(t)$ average observed arrival delay against schedule at individual airport on day t;

$\widehat{D}_S(t)$ = average observed arrival delay at airports other than LGA or ORD on day t;

Pred_$\widehat{D}_S(t)$ = predicted average observed delay at airports other than LGA or ORD on day t (not shown in the above listed models, obtained from the first stage of 2SLS and used in the second stage);

$LQ(t)$ = average arrival deterministic queuing delay at individual airport on day t;

$LF(t)$ = passenger load factor in the aircraft at the airport on day t;

$I_A(t)$ = daily IMC ratio recorded at individual airport on day t;

$D_S(t)$ = average observed arrival delay against schedule at other airports on day t;

$\widehat{D}_A(t)$ = average observed arrival delay at individual airport (LGA or ORD)on day t;

$Pred\_\widehat{D}_A(t)$= predicted average observed delay at individual airport (LGA or ORD) on day t; (not shown in the above listed models, obtained from the first stage of 2SLS and used in the second stage);

$OP(t)$ = total operations (arrivals)of system on day t;

$SQ(t)$ = weighted average arrival deterministic queuing delay of system on day t;

$W_k(t) =$ weather index of different region k on day t;

$S_i(t)$ seasonal dummy variable, set to 1 if daily arrival delay is observed in quarter i and 0 otherwise;

$D_j(t)$demand management regime dummy variable, set to 1 if daily arrival delay is observed in time period j and 0 otherwise;

$v(t), u(t) =$ stochastic error terms; and

$\lambda, \omega, \theta,$ and γ are coefficents

4.2 Research Results

Table 10 and Table 11 show the regression results. We assume that the mean of delay is zero if all the independent variables are zero. The R-square values from Table **10**; clearly indicate that the model captured about 77.4 percent and 82.4 percent of the variation in the average daily arrival delay at LGA and ORD, respectively. The estimated coefficient for average queuing delay is 0.235 for LGA and 1.270 for ORD, while for the quadratic term of average queuing delay, the coefficients are negative. Nevertheless, the combined effect of linear and quadratic terms of average queuing delay is positive. It is also found that a one-minute delay at other airports in NAS may cause increases of 0.946 minute and 0.553 minute delays at LGA and ORD, respectively. Adverse weather, as measured by the IMC ratio, is the principal factor of delay at both LGA and ORD. For the thunderstorm ratio, however, only specific regions show significant contributions.

Region 11, comprising the northeastern part of the U.S., is a major delay contributor to LGA. Regions 12 and 13, which include the upper-middle regions of the U.S., are delay contributors to ORD. The estimates for the seasonal effect, however, show smaller magnitude while compared to other factors. Interestingly, for both airports, the summer seasonal effect shows the least amount of delay when compared to other seasons. Significant factors affecting delay are demand management regimes (time-period fixed effects). HDR was considered as the base in the regression. These estimates provide a better perspective of different demand management regimes applied for different time periods (see Figure 1) and the success of their application in terms of operations and delay reduction.

We then graphically decompose the delays according to the causal factors, as shown in Figure 7 and Figure 8. For LGA, the delay increased by more than 12 minutes during the AIR-21 period in comparison to HDR and gradually reduced during the slottery period. The lowest delay was reached post-9/11 when there were fewer air traffic operations and it slowly increased through 2004. For ORD, the general phenomenon was the same, with high delays during partial HDR periods, touching low levels post-9/11, and sharply shooting up in 2004 to more than 2 minutes. As shown in Figure 7.a, average delay of other airports in the NAS and passenger load factors are the major factors affecting average arrival delay at LGA. Average arrival queuing delay and delay in the system are the major contributing factors for the average arrival delay at ORD (Figure 7.b).

Table 10 Estimation Results of Arrival Delay at Individual Airport (LGA/ORD)

	Variable	LGA			ORD		
		Estimate	SE	P-Value	Estimate	SE	P-Value
LQ(t)	Average Queuing Delay	0.235	0.02	<0.0001	1.270	0.05	<0.0001
$LQ^2(t)$	Quadratic Average Queuing Delay at Airport	-0.001	0.00	<0.0001	-0.007	0.00	<0.0001
$D_S(t)$	Predicted arrival delay at NAS	0.946	0.08	<0.0001	0.553	0.11	<0.0001
$I_A(t)$	IMC Ratio	24.900	2.68	<0.0001	21.717	3.41	<0.0001
$I_A^2(t)$	Square of IMC Ratio	-9.568	2.82	0.0007	-9.414	3.73	0.0115
LF(t)	Passenger Load Factor	0.075	0.02	0.0013	0.020	0.03	0.4731
$W_K(t)$	Thunderstorm Ratio						
	Region 11	45.280	3.64	<0.0001			
	Region 12				44.144	3.64	<0.0001
	Region 13				11.775	2.79	<0.0001
$S_i(t)$	Seasonal Dummy Variables						
	Quarter 1	-3.832	0.79	<0.0001	-1.539	1.35	0.2537
	Quarter 2	-8.567	0.96	<0.0001	-4.622	1.51	0.0022
	Quarter 3	-6.489	0.96	<0.0001	-3.353	1.50	0.0252
$D_j(t)$	Demand Management Regimes						
	AIR-21	5.122	1.12	<0.0001			
	Slottery	-1.227	1.17	0.2942			
	Partial HDR				0.231	2.04	0.9271
	Post 9/11 Period	-10.050	2.09	<0.0001	-7.160	2.53	0.0047
	Year 2002	-3.033	1.14	0.0079			
	Year 2003	-3.480	1.024	0.0007			
	Year 2004	-6.101	1.216	<0.0001			
	Overscheduling				-3.891	1.87	0.0374
	5% Reduction in UA & AA				2.264	1.93	<0.2405
R^2	R-Square	0.7741			0.8254		

Table 11 Estimation Results of Arrival Delay for RNAS

	Variable	LGA			ORD		
		Estimate	SE	P-Value	Estimate	SE	P-Value
$SQ(t)$	Average Queuing Delay	1.176	0.04	<0.0001	0.963	0.05	<0.0001
$D_A(t)$	Predicted Arrival Delay at LGA/ORD	0.082	0.01	<0.0001	0.052	0.00	<0.0001
$OP(t)$	Total Operations (arrivals) in the System	0.001	0.00	<0.0001	0.001	0.00	<0.0001
$W_k(t)$	Thunderstorm Ratio						
	Region 04	4.010	0.95	<0.0001	6.511	0.94	<0.0001
	Region 05	4.863	0.79	<0.0001	5.345	0.78	<0.0001
	Region 06	5.056	0.61	<0.0001	3.623	0.58	<0.0001
	Region 11	2.495	1.27	0.0493	11.682	1.10	<0.0001
	Region 12	11.572	0.92	<0.0001	5.625	1.12	<0.0001
$S_i(t)$	Seasonal Dummy Variables						
	Quarter 1	0.666	0.47	0.1577	0.242	0.48	0.6123
	Quarter 2	-2.657	0.49	<0,0001	-3.275	0.52	<0.0001
	Quarter 3	-3.163	0.51	<0.0001	-3.802	0.53	<0.0001
$D_i(t)$	Dummy Variable for Demand Management Regimes						
	AIR-21	2.086	0.61	0.0007			
	Slottery	0.865	0.61	0.1578			
	Partial HDR				1.447	0.57	0.0111
	Post 9/11 Period	-0.176	0.78	0.8207	-0.795	0.83	0.3396
	Year 2002	-0.651	0.56	0.2416			
	Year 2003	-0.768	0.56	0.1701			
	Year 2004	0.263	0.63	0.6773			
	Overscheduling				-1.095	0.53	0.0383
	5% Reduction in UA & AA				-1.376	0.67	0.0411
R^2	R-Square	0.944			0.941		

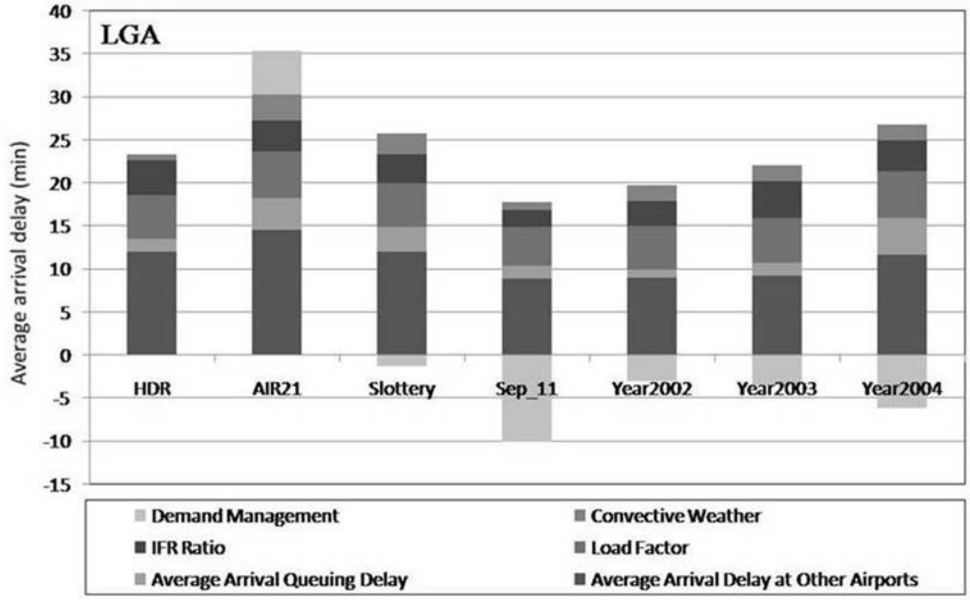


FIGURE 7.a LGA

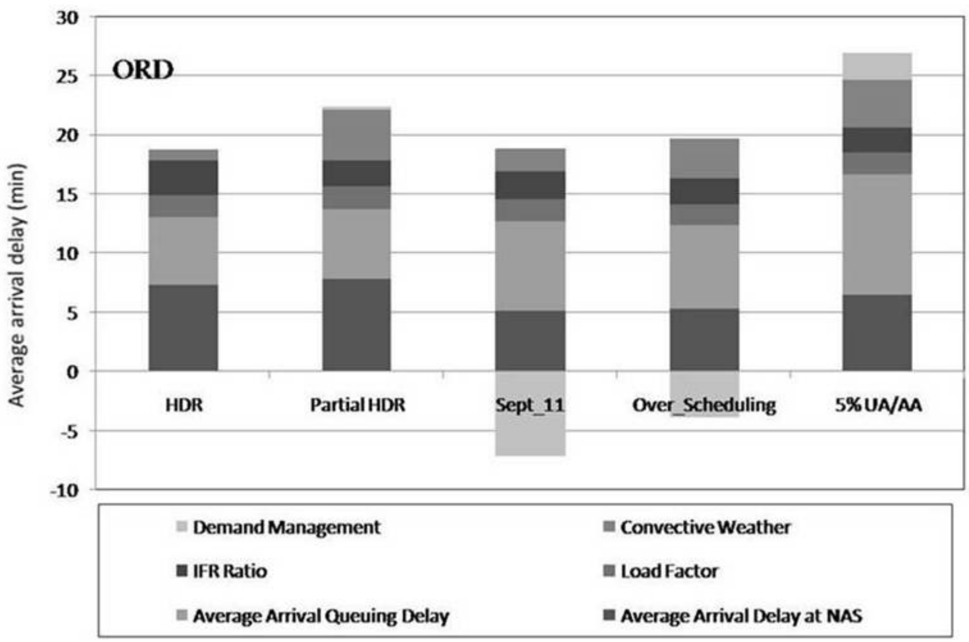


Figure 7.b ORD

Figure 7 Decomposition of LGA and ORD Average Arrival Delay

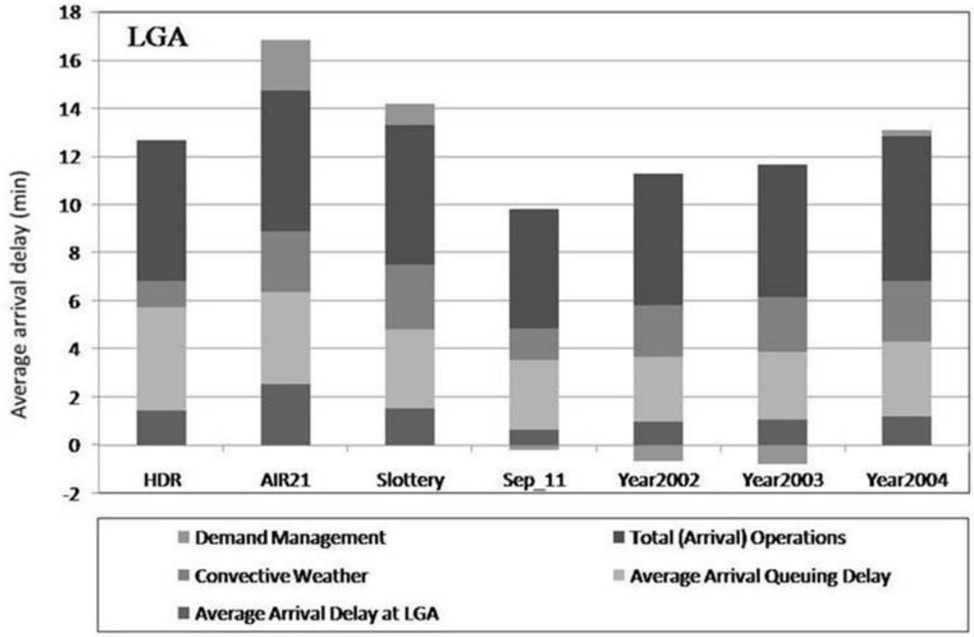


FIGURE 8.a LGA

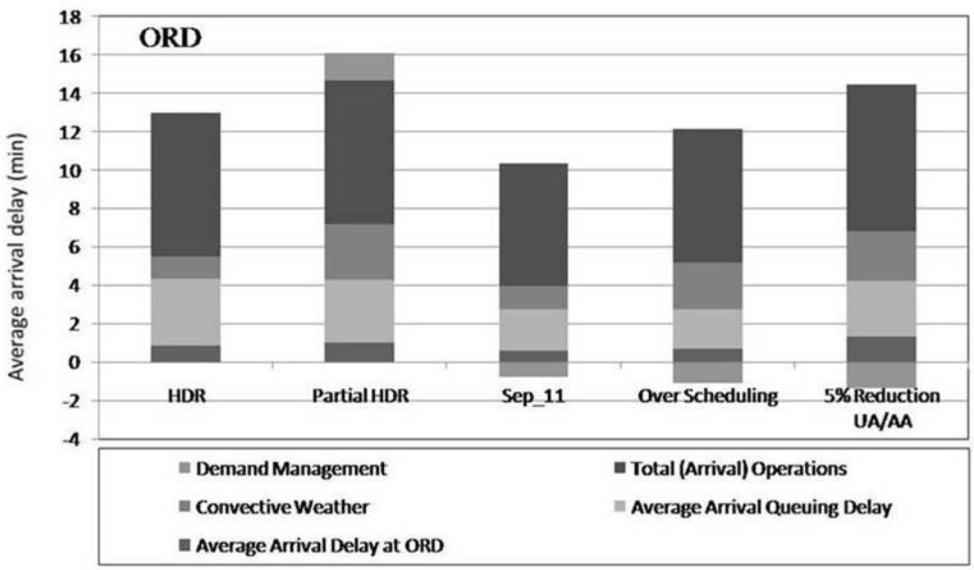


FIGURE 8.b ORD

Figure 8 Decomposition of RNAS Average Arrival Delay Considering LGA and ORD

The estimates for the RNAS model are shown in Table 11. These are the regression estimates for average arrival delay for flights to 31 benchmark airports other than LGA or ORD. The RNAS model for LGA explains a 94.35 percent variation in average arrival delay, whereas the model for ORD shows a 94.06 percent variation. The queuing delay, total operations, and thunderstorm ratio are all significant factors affecting arrival delay in the NAS. It is also seen that a one-minute increase of delay at LGA causes a 0.082-minute increase in delay in the NAS, while a one-minute delay at ORD causes a 0.052-minute delay in the NAS. Thus, if we consider the ratio of non-LGA to LGA arrivals of about 34 to 1, the effect of a one-minute delay at LGA on non-LGA airports is 34 * 0.082 = 2.788 minutes. Similarly, considering the ratio of non-ORD to ORD arrivals as 34 to 1, the effect on other airports of a one-minute delay at ORD is 34 * 0.052 = 1.768minutes. The decomposition of the RNAS at LGA (Figure 8.a) and ORD (Figure 8.b) produced results similar to those of individual airports. This is an indication that different demand management strategies applied at an individual airport have a definite impact on the whole system. The delay in the NAS due to LGA was more during the AIR-21 period, and the delay due to ORD was more influential during the partial HDR period before sharply increasing in 2004 due to over-scheduling.

4.3 System-Wide Benefit of Capacity Expansion of Individual Airport

It is interesting to know the NAS-wide delay reduction as a result of expansion of a single airport. Given the estimation results of 2SLS equations of LGA and the RNAS or ORD and the RNAS, scenario analysis can be conducted to predict the delay reduction, assuming certain percentages of capacity enhancement at each individual airport. The entire process was done in two steps, as shown in Figure 9. The first step produces output in the form of predicted arrival delay for a single airport. This value is compared with baseline observed delay to determine the percentage change of arrival delay at that airport. This predicted delay from the first step along with other variables is then used in the second step to determine the predicted arrival delay in the rest of the NAS. The predicted value can then be compared with baseline delay to determine system-wide improvement. For LGA and ORD, we assume there are 10%, 20%, and 30% capacity increases.

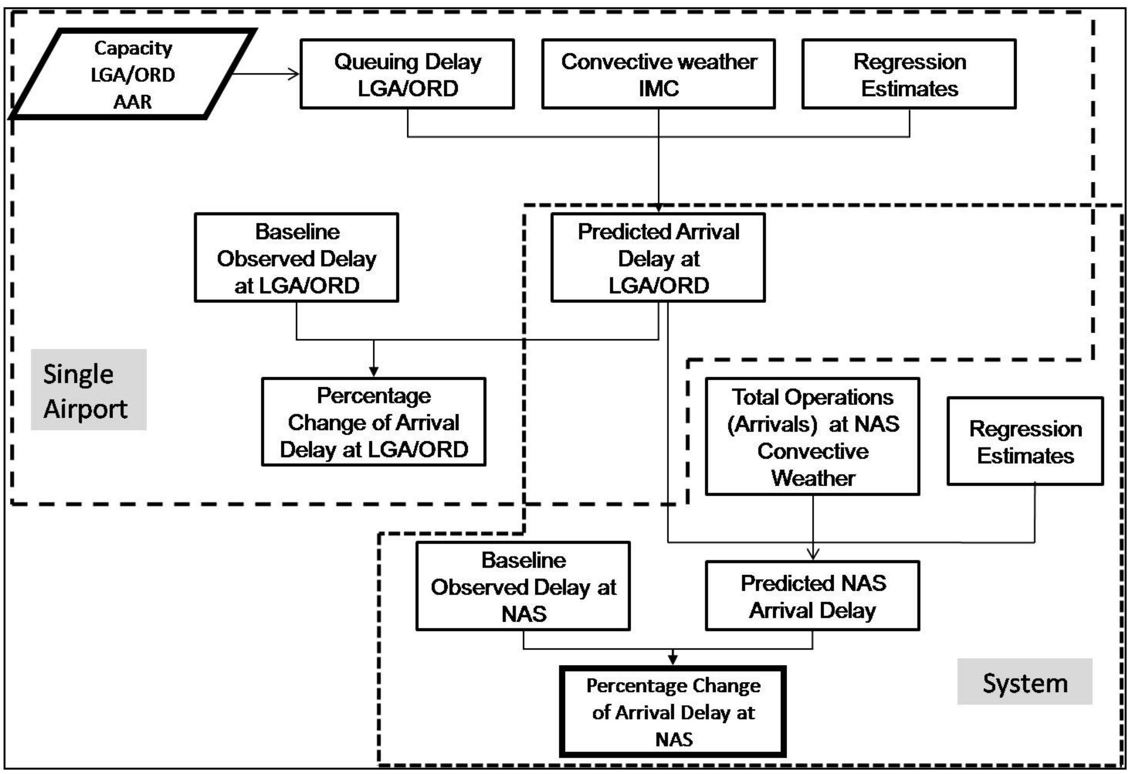


Figure 9 Scenario Analyses

The outcomes of this scenario analysis for LGA and ORD are shown in Table 12. The results are noteworthy indicators of the effects of capacity increments on delay reduction. The comparative results show that capacity increase at ORD can yield better outcomes compared to LGA in terms of percentage delay reduction. This event can be due to a high congestion rate at ORD, as it was ranked first in terms of the number of total operations till 2004, and was later overtaken by ATL [54].

Table 12 Comparison of Scenario Analysis of LGA and ORD Airports

Capacity	LGA				ORD			
	Baseline	10% Increase	20% Increase	30% Increase	Baseline	10% Increase	20% Increase	30% Increase
Airport Delay (minutes)	53.18	52.21	50.56	48.98	18.64	11.48	8.84	7.77
% Delay Reduction at Airport	Base	1.83%	4.93%	7.90%	Base	38.48%	52.60%	58.39%
NAS Delay (minutes)	6.44	6.36	6.21	6.06	8.39	8.02	7.89	7.83
% Delay Redn NAS	Base	1.36%	2.34%	2.29%	Base	4.40%	6.02%	6.67%

4.4 Research Outcomes

Airport delay has always been a major problem for the aviation industry. Most previous studies estimate the delay propagated through an individual flight from an airport to the system. This research illustrated the utility of multivariate simultaneous equations to study delay propagation from a single airport to the system, and vice versa. The model developed for LGA and ORD takes into account all the delay causal factors mentioned earlier and also has the scope to include more in the future. The estimated results clearly point toward the existing interdependency between flight delay at an individual airport and the NAS. The delay at LGA and ORD significantly depends on delay at other airports and, similarly, LGA and ORD are major contributors to delay in the system.

The decomposition of delays for different demand management regimes from the year 2000 to June 2004 explains the variation in delay throughout the period. The decomposition tries to establish the correlation between various delay causal factors at the airports and their effects on the entire system. For LGA, it shows that maximum delay occurred during the AIR-21 period with slot exemptions. The delay gradually reduced during the Slottery regime and reached the lowest point during the post-9/11 period. However, the results up to 2004 show that the delay slowly increased to the pre-9/11 Slottery period level. ORD shows a slightly different variation for delay, with the peak of its delay during 2004. The FAA had to curtail the operations of UA and AA; however, these emptied slots were taken over by other airlines, thus nullifying the efforts of the FAA to reduce delay. The decomposition for the NAS showed results similar to that of individual airports, with total operations in the system being one of the major factors affecting delay.

The research also predicts the system-wide impact of capacity enhancement or improvement in demand management strategies on delay in the NAS. The results indicate that with an increase in capacity there is a proportionate reduction in delay at the airport and the NAS. However, this phenomenon is more predominant at ORD than at LGA. Through further observation, it can be seen that the major contributing factor for delay at ORD is queuing delay, while adverse weather is a major problem at LGA. This analysis helps to determine the effectiveness of capacity improvements and can be used as a decision making tool for airport improvement projects that require massive capital investments in the future.

Furthermore, we estimate the impact of single airport delay on other OEP 34 airports and the rest of NAS using multivariate simultaneous models [55]. The variables used in the model were similar to those described in this chapter. Nevertheless, instead of defining average daily arrival delay as the actual arrival times minus scheduled arrival times (if the results are positive), we identify arrival delay by comparing actual arrival times and arrival times based on flight plans. In this way, we eliminate the noise caused by schedule buffer variations from the airlines. The research approach, methodology and the results produced from the study are presented in the following chapters.

CHAPTER V

A COMPREHENSIVE MULTI-EQUATION SIMULTANEOUS MODEL FOR ESTIMATING DELAY INTERACTIONS BETWEEN AIRPORTS AND NATIONAL AIRSPACE SYSTEM

Previously we investigated the delay propagation from one individual airport to the RNAS and vice versa, using LGA and ORD as our case studies. This study follows a similar path of macroscopic analysis not only investigating the impact of single airport delay to the RNAS but also to explore how the delay spillover is widely dispersed across the Operational Evolution Partnership (OEP) 34 airports (see Appendix I). Causal factors of the average daily arrival delays are explored, and a comprehensive multi-equation simultaneous model is developed for all airports under consideration along with the RNAS. The average arrival delay of each OEP34 airports is expressed with a multivariate equation. According to the definition, the RNAS in this chapter represents the airports in ASPM75 excluding the OEP 34 airports. In total, there are 35 equations in this model.

5.1 Multivariate Simultaneous-Equation Regression Model (MSERM)

5.1.1 Specification of Multivariate Simultaneous-Equation Regression Model (MSERM)

In this study, multivariate simultaneous equations are generated for 34 OEP airports and RNAS. The causal factors for individual airport and the RNAS are slightly different, according to the experiments of specification. For individual airport, each of the equations contains causal factors including supply-demand imbalance indicator, delays occurred at other airports and the RNAS, weather factor, and others. Analogously, the delay of the RNAS is affected by factors, such as the total operations in the RNAS, delays from 34 OEP airports, weather factor, and others. Figure 10 sketches the simultaneous characteristic of the system.

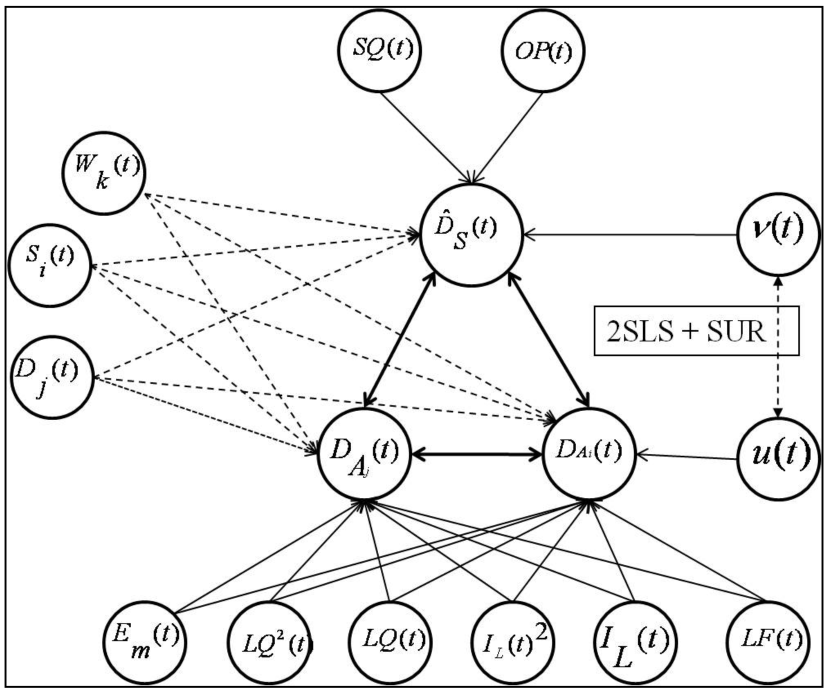


Figure 10 Interactions between a Single Airport and the Rest of the NAS

5.1.2 Model Variables

Airport data were collected from the ASPM database for the period of 2000 to 2010. As compared to the previous study, the causal factors for the delay at the individual airports include the additional explanatory variable ‘aircraft equipment type’ to study the impact of aircraft fleet size on the delay at airports. Table 13 lists the factors affecting average daily arrival delay at individual airports and the RNAS. Table 4 displays different demand management regimes operational at JFK, LGA and ORD airports and applied in the model.

Table 13 Causal Factors of Delay at Individual Airport and the RNAS

Individual Airport	**Rest of NAS (RNAS)**
Dependent Variable: Average Daily Arrival Delay	
Independent Variables:	
Average Arrival Deterministic Queuing Delay	
Arrival Delay at Other individual OEP Airport and RNAS	Average Delay at Individual OEP Airport
Adverse Weather Indicators	
Aircraft Equipment Type	Total Flights
Seasonal and Demand Management Dummy Variables	

5.2 Model Specification

The linear regression technique is one of the methods used for explaining the relationship between the variables. The flexibility of this technique derives from the possibility of being able to replace the variables in the regression equations with functions of the original variables. Applying polynomials, multiplying or dividing variables by each other, applying logarithms and exponentials, and taking reciprocals are just a few of the variable transformations available to generate nonlinear fits. In our previous research we have applied quadratic variable transformations to study average queuing delay and the IMC ratio as defined before. Even though variables may be transformed so that the equation is nonlinear in the original units of the variables, as long as the equation remains in the form of an intercept plus a slope multiplying transformed or untransformed variables, it remains a linear regression.

5.2.1 Equation 1-34 (for Individual Airport)

The model decomposes average daily delay into components related to different delay casual factors. The explanatory variables include average arrival deterministic queuing delay, average observed arrival delay at other airports, average observed arrival delay in the RNAS, adverse weather, seasonal effects, demand management regimes at JFK,LGA and ORD airports, aircraft equipment type, and others. The demand management dummy variable though used only at three airports, their effects would be studied for all the airports with each dummy variable equal to shortest period of demand management at any airport. For e.g. AIR-21 management was used at LGA from April 2000 to December 2001, hence this would be applied to all the 34 airports plus RNAS.

$$D_{Aj}(t) = \alpha + \beta_0 . D_S(t) + \sum_{i\epsilon\{i\neq j\}} \beta_i D_{Ai}(t) + \rho_1 . LQ(t) + \rho_2 . LQ^2(t) + \rho_3 . I_L(t) + \rho_4 . I_L^2(t) + \rho_5 . E_m(t) + \sum_K \lambda_{KL} W_K(t) + \sum_i \omega_{iL} S_i(t) + \sum_j \theta_{mL} D_m(t) + v(t)$$

5.2.2 Equation 35 (for RNAS)

The model for the RNAS decomposes daily average delay at the remainder of the airports that excludes the 34 OEP airports. The explanatory variables include variable delays at individual airports, convective weather, total operations, seasonal effects, yearly dummy variables, and other factors.

$D_S(t) = \alpha + \gamma_0 . OP(t) + \sum_i \gamma_i D_{Ai}(t) + \eta . SQ(t) + \sum_k \lambda_{kS} W_k(t) +$

$\sum_i \omega_{iS} S_i(t) + \sum_m \theta_{mS} D_m(t) + u(t)$

The notations in the above two models are described as follows:

$D_{Aj}(t) =$ Average observed arrival delay against flight plan at individual airport on day t;

$D_{Ai}(t) =$ Average observed arrival delay against flight plan at other individual airport (i) on day t;

$D_S(t) =$ Average observed arrival delay at airports other than individual airport on day t;

$LQ\ (t) =$ Average arrival deterministic queuing delay at individual airport on day t;

$I_L(t) =$ Daily IMC ration recorded at individual airport on day t;

$Pred\_D_A(t) =$ Predicted average observed delay at individual airport on day *t*; (not shown in the above-listed models, obtained from the first stage of 3SLS, and used in the second stage);

$OP(t) =$ Total operations (arrivals) of the system on day t;

$E_m(t) =$ Aircraft type operating at individual airport on day t;

$SQ\ (t) =$ Weighted average arrival deterministic queuing delay of the system on day t;

$W_K(t) =$ Weather index of region k on day t;

$S_i(t) =$ Seasonal dummy variable, set to 1 if daily arrival delay is observed in quarter i and 0 otherwise;

$D_m(t) =$ Demand Management Dummy Variable, set to 1 if daily arrival delay is observed in time period j and 0 otherwise;

$v(t), u(t) =$ Stochastic error terms; and

$\beta, \rho, \lambda, \omega, \theta$ are coefficients.

5.3 Research Results

Table 14 shows a part of the results, outcomes for equation ATL and the RNAS, from regression using 3SLS regression method. The table shows that the average daily arrival delay at ATL increases by 0.815 minutes if there is a corresponding increase of average queuing delay at the airport. This is due to capacity constraints and increased air traffic operations at ATL in last few years [56]. The next few rows in Table 14 show the interactions between ATL and other airports, as well as with the RNAS. For instance, the delay at ATL is significantly affected by the RNAS, as represented by the parameter in front of Ds(t). For adverse weather effects, it can be seen that Region 5 has the significant impact on arrival delay at ATL, more thunderstorms in this region leading to more delay at ATL. In contrast, more thunderstorms in Region 1 lead to less delay at ATL. If we recall Figure 5, we can see that Region 5 is where ATL is located. It is intuitively right that convective weather in this region will affect the airspace could be used, so as to lead to more delay at ATL. Region 1 covers Mexico Bay and Florida. If there are more thunderstorm, more flights from MCO, MIA, TPA will be held on the ground and waiting for clearance. Under this circumstance, arrival demand at ATL will be lower and the arrival delay will be less. The equipment type is insignificant in the case of the ATL airport. This might be possible due to availability of enough gates at the ATK airport. Similarly, the table shows results for seasonal and demand management dummy variable.

While going through the regression results from other equations (which are not listed in this dissertation due to the limitation of space), the estimated coefficients for average queuing delay for most of the airports except BWI, DCA, PDX, SAN, TPA and RNAS indicate that supply and demand imbalance is likely to be a major contributing factor to average daily arrival delays. However, the negative coefficient for the quadratic term of average queuing delay shows that this factor reduces as average queuing delay increases. This study explores the delay propagation from other airports and the RNAS to an individual airport and vice versa. The estimation results show that the other airports around the same geographical region or the other airports operating as a hub for the same carrier contribute significantly on the delay at the reference airport. For instance, the airports significantly affect the arrival delay at ATL are BWI, MCO, MEM, PDX and RNAS which are mostly located in the eastern part of the country. Similar regional phenomena can be observed and are summarized in Table 15.

Counter-intuitively, several airports have negative delay propagation effects on some other airports. For example, the delay increase at DFW will reduce the delay at ATL, BOS, CLT, CVG, DTW, LAX and PHX. The IMC ratio is likely to impact the delay at almost all the airports except BWI, FLL and PDX. Most of the airports are affected significantly by the convective weather index in the same region where they are located except CVG, LAS, PDX, SLC and SAN. It is also observed that a few airports like DEN, BWI and MEM are affected by thunderstorms occurring at destinations. In addition, convective weather at region 2, 10 and 13 which represent congested states contribute considerably to delay at the rest of the NAS airports.

As long as the weather pattern is captured by the convective weather index and IMC ratios, seasonal dummy variables in the model only reflect the seasonal difference of airline scheduling. The estimates for the seasonal effect show that their impact on delay is very small in comparison to other factors. Interestingly, for most of the airports, the winter seasonal effect shows highest amount of delay as compared to other seasons. However for the airports in the southern parts of the country like MCO, ATL, TPA, DFW and LAS, delays are higher during spring. The demand management regimes, even though implemented at only some airports, dummy variables were generated and applied for all the 34 airports and the RNAS. The dummy variable parameters show a large impact on average daily arrival delay. The estimated coefficients for the dummy variables provide a better perspective on how delays vary in comparison to different time periods. According to the FAA, 34 OEP airports are categorized into different regions (different from the convective weather regions that we have defined earlier) [57]. The trends of average arrival delay for all the airports along with the NAS are shown in Figure 11 to Figure 18.

Table 14 Estimation Results of Arrival Delays at an Individual Airport (ATL) and the RNAS

	Variable	Atlanta (ATL)			System		
		Estimate	SE	P-Value	Estimate	SE	P-Value
	Intercept	-6.08850	1.612001	0.0002	-0.09578	0.250255	0.7019
$LQ(t)$	Average Queuing Delay	0.815279	0.016347	<.0001	0.008169	0.020269	0.6869
$LQ^2(t)$	Quadratic Average Queuing Delay at Airport	-0.00258	0.000068	<.0001			
$D_S(t)$	Predicted arrival delay at						
	ATL				0.010759	0.004269	0.0118
	BOS	0.003150	0.018070	0.8616	0.015715	0.005199	0.0025
	BWI	0.253522	0.088139	0.0040	0.120480	0.024952	<.0001
	CLE	0.042338	0.054490	0.4372	0.019515	0.015686	0.2135
	CLT	0.017762	0.050061	0.7228	0.037763	0.014593	0.0097
	CVG	0.011691	0.056531	0.8362	0.020028	0.016063	0.2125
	DCA	0.050516	0.073527	0.4921	0.106542	0.020753	<.0001
	DEN	-0.08094	0.036804	0.0279	0.035417	0.010392	0.0007
	DFW	-0.13410	0.033219	<.0001	0.153919	0.008041	<.0001
	DTW	-0.22495	0.046100	<.0001	0.045573	0.013238	0.0006
	EWR	0.032303	0.024109	0.1804	0.032462	0.006858	<.0001
	FLL	-0.31707	0.067745	<.0001	-0.03299	0.019346	0.0882
	IAD	-0.37491	0.060995	<.0001	-0.05680	0.017759	0.0014
	IAH	-0.01707	0.023049	0.4590	0.048768	0.007271	<.0001
	JFK	-0.02792	0.035584	0.4328	-0.04884	0.010340	<.0001
	LAS	-0.02789	0.036788	0.4485	0.009075	0.010576	0.3909
	LAX	-0.12154	0.060266	0.0438	0.118548	0.016883	<.0001
	LGA	0.012100	0.023295	0.6035	0.041000	0.006332	<.0001
	MCO	0.526641	0.131260	<.0001	0.224844	0.036287	<.0001
	MDW	-0.04617	0.049609	0.3521	0.050133	0.013709	0.0003
	MEM	0.208091	0.052253	<.0001	-0.02290	0.015378	0.1365
	MIA	-0.00966	0.068913	0.8886	0.037790	0.019356	0.0510
	MSP	0.016713	0.022930	0.4661	0.004489	0.006582	0.4953
	ORD	-0.00094	0.018501	0.9597	0.025652	0.005198	<.0001
	PDX	0.252337	0.106883	0.0183	0.170080	0.029714	<.0001
	PHL	-0.02964	0.020885	0.1559	-0.04474	0.005872	<.0001
	PHX	0.043615	0.038512	0.2575	0.043902	0.010956	<.0001
	PIT	-0.11213	0.082175	0.1725	0.058391	0.023375	0.0125
	SAN	-0.05863	0.109247	0.5915	0.098975	0.030358	0.0011
	SEA	-0.10822	0.058938	0.0664	-0.03307	0.016679	0.0475
	SFO	-0.00046	0.012309	0.9699	-0.00136	0.003543	0.7012
	SLC	-0.05233	0.039920	0.1900	0.013350	0.011579	0.2490
	STL	-0.01448	0.025239	0.5661	0.045254	0.007162	<.0001
	TPA	0.338615	0.113945	0.0030	-0.01013	0.033109	0.7597
	Total System	0.003150	0.018070	<.0001			
$T(t)$	Total Flights				0.000067	0.000022	0.0027
$I_A(t)$	IMC Ratio	3.139454	1.064630	0.0032			
$I_A^2(t)$	Square of IMC Ratio	4.279547	1.229735	0.0005			

Table 14: Continued							
$E_m(t)$	Equipment 1	2.891E-6	0.000080	0.9712			
	Equipment 2						
	Equipment 3						
	Equipment 4	0.000036	0.000017	0.0322			
	Equipment 5	0.000023	0.000012	0.0434			
	Equipment 6	-0.00021	0.000086	0.0136			
WK(t)	Thunderstorm Ratio						
	Region 01	-10.0904	1.244729	<.0001	-1.19338	0.321979	0.0002
	Region 02				0.539538	0.276836	0.0514
	Region 03				-1.13248	0.521833	0.0301
	Region 04				0.009421	0.377856	0.9801
	Region 05	24.56031	1.350248	<.0001	-0.60270	0.329344	0.0673
	Region 06				-0.41505	0.363538	0.2536
	Region 07				0.147004	0.624867	0.8140
	Region 08				0.508938	0.364675	0.1629
	Region 09				-1.14208	0.449766	0.0111
	Region 10				-0.07357	0.484183	0.8792
	Region 11				0.631972	0.300177	0.0353
	Region 12				-1.19395	1.000187	0.2327
	Region 13				-1.19338	0.321979	0.0002
	Region 14				0.539538	0.276836	0.0514
	Region 15				-1.13248	0.521833	0.0301
	Region 16				0.009421	0.377856	0.9801
$S_i(t)$	Seasonal Dummy Variables						
	Quarter 2	1.091572	0.378151	0.0039	0.313029	0.106023	0.0032
	Quarter 3	1.372655	0.474541	0.0038	0.367125	0.131297	0.0052
	Quarter 4	0.428131	0.360006	0.2344	0.108400	0.096186	0.2598
$D_j(t)$	Demand Management Regimes						
	AIR	1.257618	0.810436	0.1216	-0.16492	0.205676	0.4227
	Before 9/11	0.437741	0.838518	0.6029	0.034505	0.209368	0.8691
	After 9/11	1.582508	1.078529	0.1440	1.001999	0.275613	0.0003
	OV 2002	0.555405	0.955752	0.5645	1.188643	0.201927	<.0001
	OV 2003	0.374392	1.039075	0.7217	0.615717	0.204152	0.0026
	CAP	-0.81224	1.351016	0.5457	1.194191	0.234043	<.0001
	RED A	-1.28303	1.226865	0.2930	1.099363	0.240780	<.0001
	RED B	-0.95103	1.078013	0.3755	0.992495	0.224898	<.0001
	RED C	0.847884	1.413631	0.5507	1.125077	0.231898	<.0001
	LIM	4.785715	1.487951	0.0013	0.978558	0.268300	0.0003
	Year 2007	3.452351	1.470327	0.0191	0.883978	0.236155	0.0002
	Year 2008	3.077824	1.457181	0.0350	1.082171	0.228994	<.0001
	Year 2009	4.308317	1.387312	0.0019	0.509018	0.225555	0.0241
	Year 2010	3.449082	1.512166	0.0228	1.714824	0.215350	<.0001
	System Weighted MSE	4.5296					
	Degrees of Freedom	136284					
R^2	System Weighted R-Square	0.7335					

Figure 11 shows that the average arrival delays at all the airports in ASO region, except ATL, CLT and CVG, remained almost the same throughout 2000 to 2010. After 2005 (REDB), average daily arrival delay at ATL and CLT increased continuously from 2005 to 2010. On the contrary the average daily arrival delay at CVG reduced from 2000 to 2010. In Region AWP, as shown in Figure 12, the delay at LAX decreased drastically after 9/11 and slowly approached the level of pre 9/11 in 2006. For SFO, LAS and SAN in the same region, however, the delay increased immediately after 9/11.

Figure 13 shows the delay trends of the airports in ANM region, which comprises airports in the north-west of the country. The average arrival delay at those airports was higher in 2007, but still lower than the pre 9/11 level. However, the average daily arrival delay at DEN increased dramatically post 2005.

The north-central part of the U.S. is represented by AGL region (Figure 14), which consists of many connecting airports for east-west air traffic. The arrival delays at most of the airports reduced after year 2000 expect the ORD airport. It was also noticed that after reduction of United and American Airlines in 2004, the delay at the ORD airport reduced a bit as compared to earlier estimates. Nevertheless, the delay at MSP airport has significantly reduced from 2000 to 2010. The ASW region (Figure 15) consisting of airports from Texas state had arrival delay showing opposite trends throughout the time period. The average daily arrival delay had its peak value for DFW in 2004, while for IAH it reached its peak in 2010. The north-eastern part of the country that has a few of the world's busiest airports is represented by AEA region (Figure 16).

Table 15 Interactions between Individual Airports and the NAS

Airports	Airports Contributing to Average Arrival Delay	Airports Reducing Average Arrival Delay
ATL	BWI (0.254), MCO (0.526), MEM (0.208), PDX (0.252), NAS (0.588)	DFW (-0.134), DTW (-0.224), FLL(-0.317), IAD(-0.374)
BOS	BWI (0.453), LGA (0.127), PIT (0.410), NAS (1.089)	CLE (-0.241), DFW (-0.185), LAX (-0.343)
BWI	BOS (0.036), DCA (0.418), IAD (0.275), JFK (0.123), MDW (0.155), PHL (0.091), TPA (0.253), NAS (0.261)	EWR (-0.104), MCO (-0.344), ORD (-0.075)
CLE	BWI (0.276), DTW (0.186), PIT (0.546), NAS (0.284)	BOS (-0.036), DCA (-0.255)
CLT	CVG (0.113), DCA (0.371), EWR (0.052), MCO (0.332), PIT (0.169), NAS (0.325)	DFW (-0.089), FLL (-0.131)
CVG	CLT (0.157), DTW (0.141), LGA (0.057), MEM (0.126), ORD (0.041), PIT (0.417), STL (0.068)	DFW (-0.067), MDW (-0.107), SAN (-0.252)
DCA	BWI (0.611), CLT (0.157), IAD (0.155), PHL (0.058), NAS (0.326)	CLE (-0.143), MEM (-0.157)
DEN	MEM (0.137), MSP (0.056), PDX (0.417), SLC (0.126), NAS (0.302)	SEA (-0.189)
DFW	IAH (0.076), LGA (0.076), NAS (1.769)	EWR (-0.086), LAX (-0.343), MDW (-0.165)
DTW	CLE (0.284), MCO (0.327), MDW (0.205), PDX (0.321) and NAS (0.533)	DEN (-0.074), DFW (-0.068),
EWR	CLE (0.317), IAD (0.284), JFK (0.613), LGA (0.484), PDX (0.763), PHL (0.281), NAS (0.768)	BWI (-0.700), TPA (-0.560)
FLL	LGA (0.042), MCO (0.834), MIA (0.581)	IAD (-0.126)
IAD	BWI (0.572), DCA (0.295), DEN (0.085), EWR (0.064), LGA (0.093), ORD (0.037), PIT (0.228)	FLL (-0.182), JFK (-0.146), MDW (-0.108), MSP (-0.061)
IAH	DFW (0.191), LAX (-0.313), MEM (0.309), SAN (0.433), NAS (0.827)	MDW (-0.255)
JFK	BWI (0.734), EWR (0.307), LGA (0.068), MCO (0.956), ORD (0.095)	IAD (-0.479), MDW (-0.327)
LAS	PHX (0.072), SAN (0.847)	LAX (-0.127)
LAX	DCA (0.138), SAN (0.950), SEA (0.087), MAS (0.426)	DFW (-0.101), MDW (-0.144), PDX (-0.184)
LGA	BOS (0.111), EWR (0.577), FLL (0.515), IAD (0.411), JFK (0.212), SEA (0.395), NAS (1.022)	IAH (-0.155), MCO (-1.119), PDX (-1.071), STL (-0.176)
MCO	DTW (0.058), FLL (0.237), JFK (0.083), TPA (0.617), NAS (0.272)	BWI (-0.143), LGA (-0.057), MSP (-0.028)
MDW	BWI (0.532), DTW (0.226), ORD (0.369), PIT (0.273), TPA (0.332), NAS (0.421)	CVG (-0.195), IAD (-0.201), JFK (-0.161), LAX (-0.208)
MEM	CLE (0.175), CLT (0.128), CVG (0.129), IAH (0.083), MSP (0.047), ORD (0.046), PDX (0.309)	DCA (-0.194), JFK (-0.076), MIA (-0.158), SEA (-0.111)
MIA	FLL (0.611), ORD (0.034)	DCA (-0.167), MDW (-0.093)
MSP	BWI (0.350), DEN (0.123), EWR (0.126), PDX (0.344), TPA (0.613)	IAD (-0.269), JFK (-0.205), MIA (-0.240), PHL (-0.097)
ORD	JFK (0.200), MDW (1.827), MSP (0.167)	BWI (-0.905), DTW (-0.247)
PDX	DEN (0.043), EWR (0.044), MEM (0.067), SAN (0.229), SEA (0.456), SFO (0.026), SLC (0.118), NAS (0.262)	IAH (-0.043), LAS (-0.048), LAX (-0.116), TPA (-0.133)
PHL	BWI (1.311), DCA (0.372), EWR (0.414), LGA (0.121), MEM (0.255), PIT (0.546)	IAD (-0.746), JFK (-0.292), MSP (-0.121), NAS (-0.527)
PHX	DCA (0.115), DEN (0.094), SAN (0.447), SEA (0.138), SLC (0.069), NAS (0.497)	BWI (-0.137), DFW (-0.057), IAD (-0.124), PDX (-0.250)
PIT	CLE (0.239), CVG (0.111), DCA (0.107), MCO (0.207), MDW (0.104), NAS (0.200)	DEN (-0.053), DTW (-0.075)

Table 15: Continued		
SAN	IAH (0.035), LAS (0.184), LAX (0.374), PDX (0.162), PHX (0.051), SFO (0.025), NAS (0.172)	SEA (-0.097)
SEA	ORD (0.040), PDX (1.323)	DEN (-0.065), SFO (-0.039), SLC (-0.095)
SFO	ORD (0.071), PDX (0.513), SAN (1.042)	CVG (-0.207), LAX (-0.326), MCO (-0.605), MDW (-0.216)
SLC	DEN (0.112), MEM (0.096), PDX (0.637), SAN (0.297)	SEA (-0.185), SFO (-0.023)
STL	CVG (0.210), EWR (0.088), MDW (0.169), MEM (0.147), NAS (0.766)	LAS (-0.088), LAX (-0.148)
TPA	BWI (0.143), LGA (0.032), MCO (0.0782), MIA (0.091)	BOS (-0.024), DCA (-0.097)
RNAS (System)	BWI (0.120), DCA (0.106), DFW (0.154), EWR (0.032), IAH (0.049), LAX (0.118), LGA (0.041), MCO (0.224), ORD (0.025), PDX (0.170), PHX (0.043), STL (0.045)	JFK (-0.048), PHL (-0.044)

This region consists of the largest number of airports as compared to other regions. For all the airports, except LGA, the average arrival delay has always been positive. The average daily arrival delay at PHL, JFK, EWR and IAD significantly increased after 2005. The average arrival delay at BOS (Figure 17) had an increment, while for STL it reduced after 2004. Figure 18, shows estimates for the RNAS and it is seen that the average daily arrival delay increased constantly from 2000 to 2010.

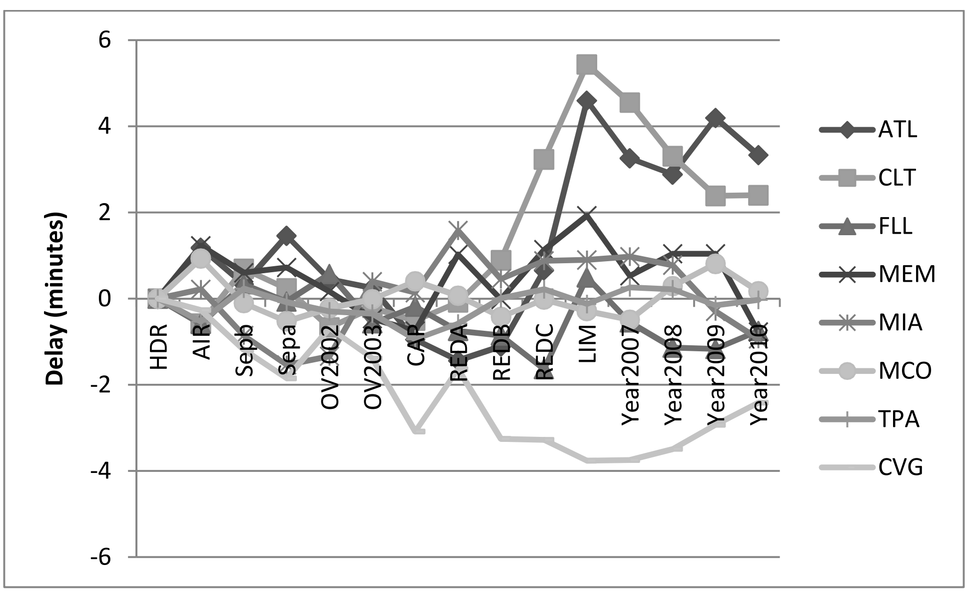


Figure 11 Airport Arrival Delay from 2000-2008 for ASO Region

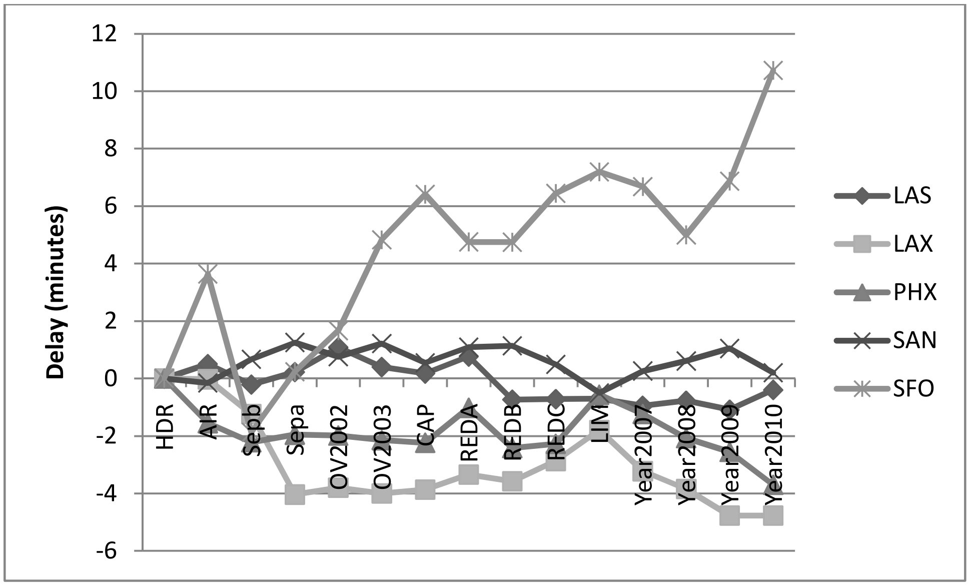


Figure 12 Airport Arrival Delay from 2000-2008 for AWP Region

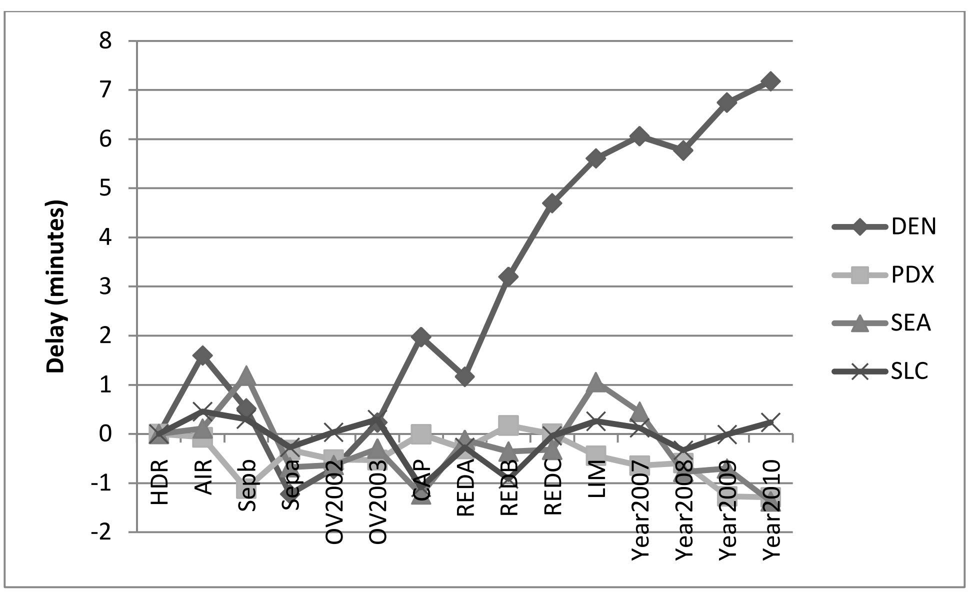


Figure 13 Airport Arrival Delay from 2000-2008 for ANM Region

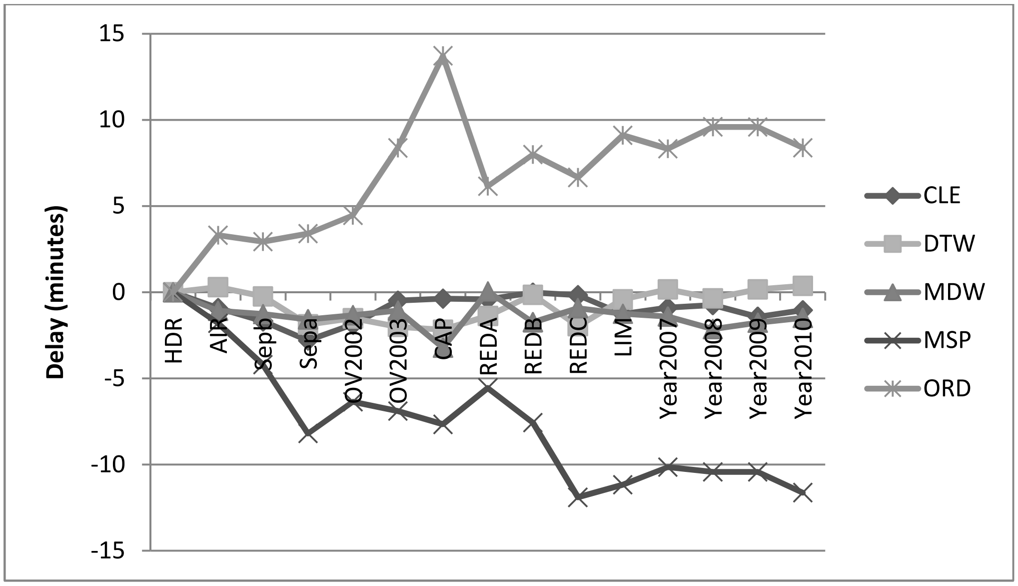


Figure 14 Airport Arrival Delay from 2000-2008 for AGL Region

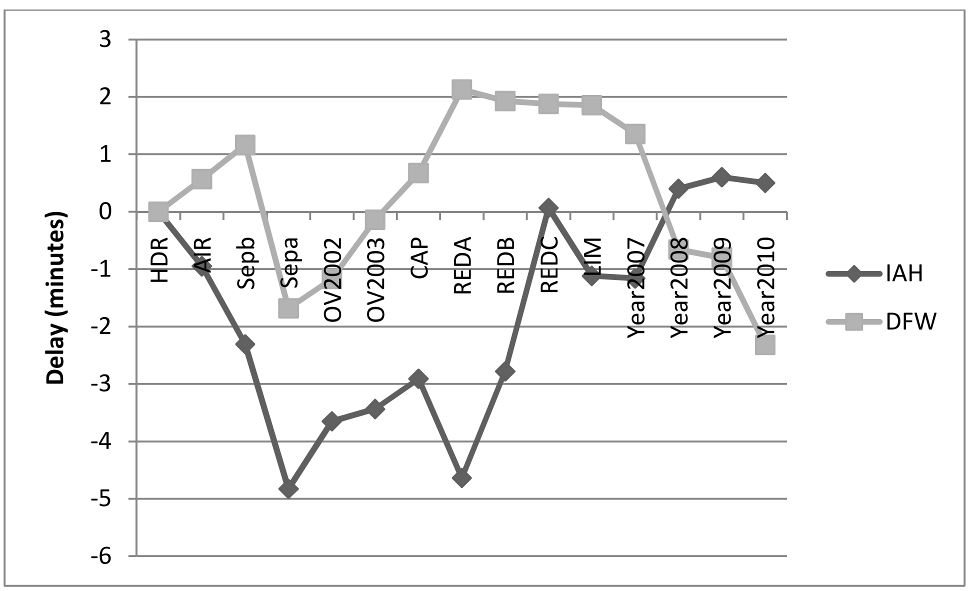


Figure 15 Airport Arrival Delay from 2000-2008 for ASW Region

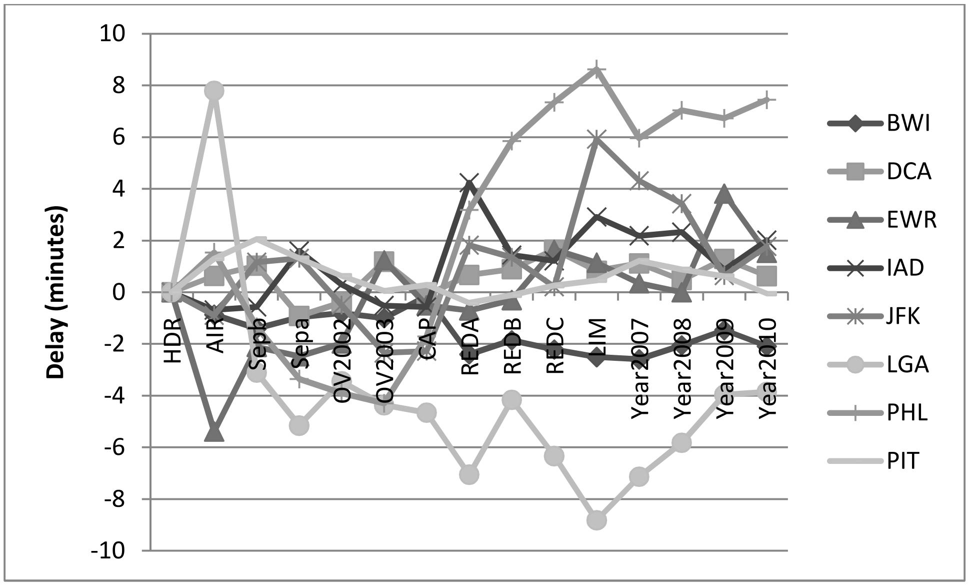


Figure 16 Airport Arrival Delay from 2000-2008 for AEA Region

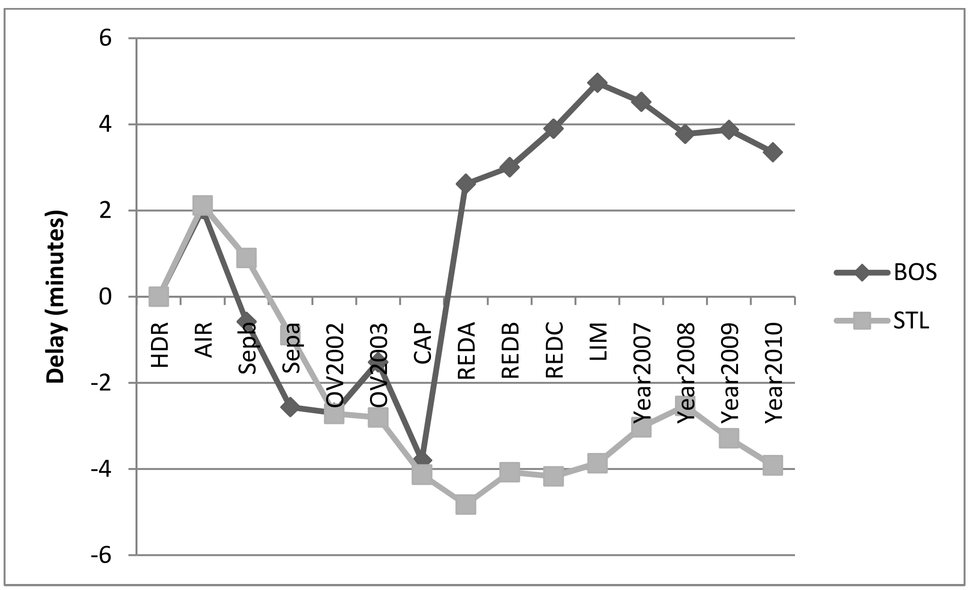


Figure 17 Airport Arrival Delay from 2000-2008 for ANE (BOS) and AAL (STL) Regions

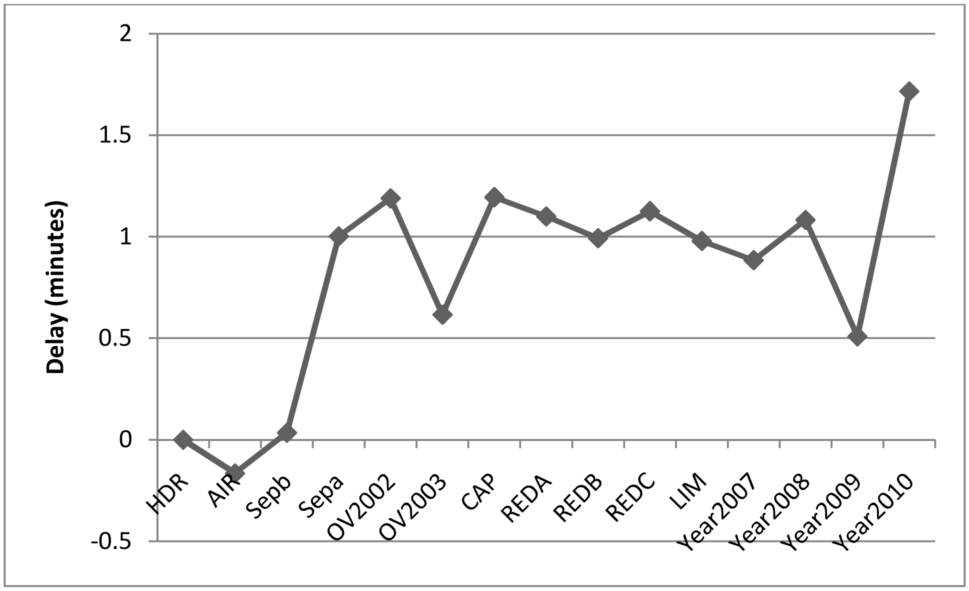


Figure 18 Airport Arrival Delay from 2000-2008 for RNAS

CHAPTER VI

ESTIMATION OF FLIGHT DELAY PROPAGATION OF THE MULTI-AIRPORT SYSTEMS IN THE US

With the increase in population, city's geographical growth, better ground transportation modes and sometimes political factors, there has been steady increase in number of airports within a region [13]. Most of the major cities in the U.S. are served by more than one airport. Many of these airports have coordinated operations in terms of sharing regional airspace, some act as a reliever airport in case of over shooting of capacity at other airport(s) and also help reduce environmental effects like noise and air pollution in one specific area. For instance, the San Francisco bay area consists of three major airports namely SFO, OAK and SJC along with many small airports. The flight routes at all the three airports are usually conflicting with each other [58]. All these airports need to take additional care to maintain air-borne safety of the flights that might result in increase of flight delay. Hence, research is warranted to explore the impact of these groups of airports in a region on other airports.

Additionally, it is seen that while traffic at major airports is stable, traffic at reliever airports is volatile depending upon its demand [6]. As seen in some cases, airports might be competing against each other for air service demand due to competing airlines, close proximity, increasing demand, efficient service, etc. In the case of BOS and MHT airports as shown in Figure 19, the BOS airport is operated by legacy airlines while the MHT airport has large number of operations by low cost carriers (LCC). Both the airport operations completely differ from each other in terms of their management. Hence, it would be interesting to learn the impact of operations at these airports on other airports in the country.

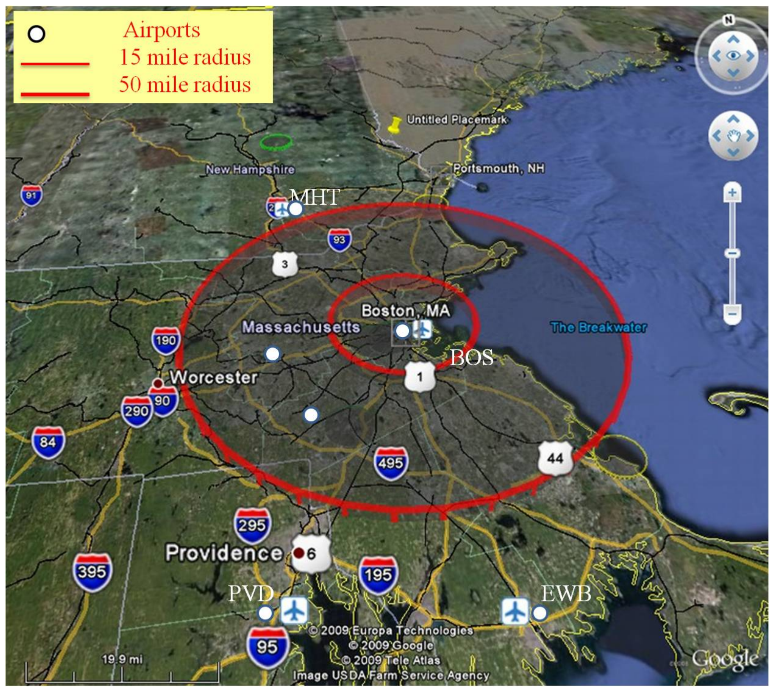


Figure 19 Air Service Area at the Greater Boston Region

Our previous studies estimated and compared flight delay propagated from one individual airport to another and vice versa, as well as the delay propagated from that airport to the RNAS and the effect of RNAS delay to that airport (Zhang and Nayak [42] [55]). The outcomes of our studies provide decision-support for future airport capacity expansion and a framework to evaluate the nation-wide effectiveness of capacity expansion or delay reduction at individual airports. In this study, we expanded our study to the multi-airport systems.

The multi-airport system is defined as a system with a set of airports that serve air traffic a metropolitan area [6]. The New York airspace is one of the most congested airspaces in the world, with both domestic and international air traffic. FAA always faces the challenge of mitigating the delays in the New York region. Increasing airport capacity by adding more runways could be a solution; however, it will lead to enormous capital investment, projected flight delays, public outcry, and environmental concerns. Atkins [59], for example, studied the interdependencies between proximate airports in the San Francisco Bay Area and found that the interdependencies resulted in reduction of airport capacity and operational efficiency. Thus, it is worthwhile to investigate the interdependency between airports for seeking solutions to improve the operational efficiency of regional airport systems. Also, given the multiple regional airport systems in the U.S. and their different characteristics, it is interesting to see how the operational performance of each regional airport system affects the RNAS (in this study, the NAS with other airports except the studied regional airport system.)

6.1 Research Approach

The objective of this study is to quantify the interdependency of airports in a multi-airport system and to investigate the delay propagation from the system to the RNAS and vice versa. Hence, the first step was to collect data for multi-airport systems. A total of 11 multi-airport systems in the U.S. were identified based on regional traffic share and proximity [13]. Orlando and Tampa regions contain airports for which data are not available in the Aviation Service Performance Management database and were not included. Table 2 shows the final list of metropolitan regions and airports. All the airports in these regions, except New York and Houston, are multi-jurisdictional, with different authorities in charge of the operations and management of the different airports in each region [22].

The second step was to define regional level performance indicators, as well as indicators reflecting the characteristics of multi-airport systems. Research on three airports in the Bay area (SFO, OAK, and SJC) [60] estimated that domestic origin-destination passenger share at SFO fell from 66 percent in 1990 to 42 percent in 2003. Some of the reasons cited were the introduction of LCC airlines or transfer of legacy airline operations to regional carrier affiliates. However, it was observed that the introduction of Virgin Atlantic (an LCC) at SFO increased its share to 51 percent. These numbers are a clear indicator of passengers responding to airline fares for airports located in the same region. Another important indicator is the LCC market share. In some regions, for instance, in the case of BOS, MHT, and PVD in the New England region, the operations at BOS are dominated by legacy airlines while at MHT and PVD, a large number of operations are offered by LCCs [8]. Correspondingly, the operations of these airports differ from each other in terms of their management.

We present elasticity estimates to demonstrate the different effects of the multi-airport system toward the RNAS in the fourth step. Meanwhile, the same analysis was conducted to identify the major factors leading to the change of flight delays at each regional airport system during different seasons of the year.

6.2 Model Variables

Figure 20 shows the interactions between multi-airport systems and the RNAS. The variables in the figure were explained earlier and are shown in the following table.

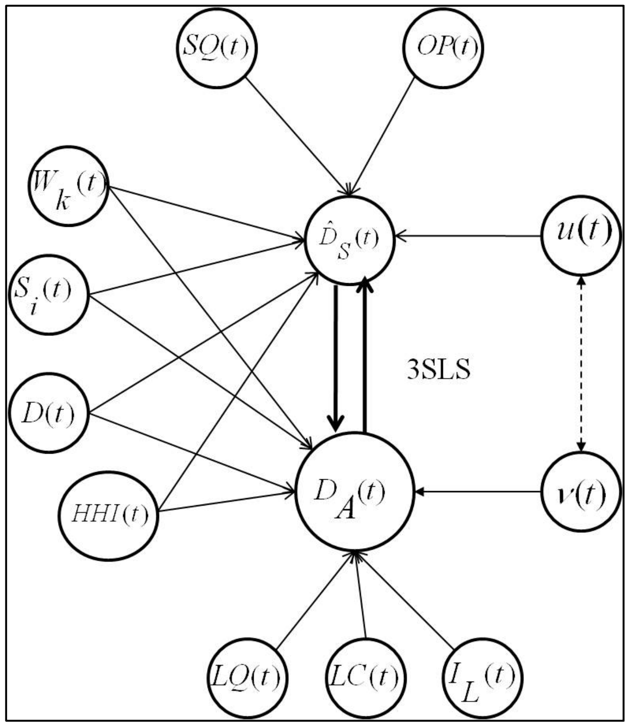


Figure 20 Interactions between Multi-Airport System and the RNAS

Average daily arrival delay is the dependent variable in our model. This average daily arrival delay is the aggregate average daily arrival delay from all airports in the region. Only arrival delay is used as the delay metric, since there is a high correlation between arrival and departure delay for both the region and the RNAS. For each multi-airport system, the average daily arrival delay is a function of average arrival delay at RNAS, deterministic queuing delay caused by the over-scheduling or supply-demand imbalance due to capacity deficiency in the system, adverse weather, HHI, and LCC Market Share, together with dummy variables indicating seasonal and yearly effects.

Similarly, the average daily arrival delay at RNAS depends upon average arrival delay in the regional airport system under consideration, deterministic queuing delay, total flights, and other explanatory variables.

Table 16 Causal Factors of Delay at Region and the RNAS

Region	**Rest of NAS (RNAS)**
Dependent Variable: Average Daily Arrival Delay	
Independent Variables:	
Average Arrival Deterministic Queuing Delay	
Arrival Delay at the RNAS	Average Delay at particular region
Adverse Weather Indicators	
HHI Index for Region	
LCC Airline Market Share	Total Flights
Seasonal and Demand Management Dummy Variables	

Table 17 Correlation Matrix for LCC Share at Four Airports in the New York Region

	EWR	ISP	JFK	LGA
EWR	1.000	0.720	0.584	0.224
ISP	0.720	1.000	0.924	0.778
JFK	0.584	0.924	1.000	0.850
LGA	0.224	0.778	0.850	1.000

We also conducted a collinear diagnosis of explanatory variables. We found that some of the LCC market shares between airports in the same region show high correlations. As shown in Table 17, the LCC market shares of EWR and ISP are highly correlated, as well as those of JFK and ISP. We also found that in most of the regions, the IMC ratios at airports in the same region are correlated.

In such cases, principal component analysis is applied to remove the multi-collinearity from the analysis. It is a statistical technique that generates a linear combination of a number of variables and uses them for further analysis [61] and is a simple variable reduction procedure. The observed variables are weighted in such a way that the resulting components account for a maximal amount of variance in the data set. The number of principal components generated is equal to the number of variables. The first few components cover up maximum variance and are used based on different criteria. For instance, when the principal component analysis is applied to the above case, it was found that Component 1 = (0.398 EWR_Share) + (0.558 ISP_Share) + (0.552 JFK_Share) + (0.475 LGA_Share). Component 1 covers 77.20 percent of the variance, so it was used in regression model.

Table 18 Principal Component Analysis for LCC Share in the New York Region

	Prin1	Prin2	Prin3	Prin4
EWR	0.398	0.792	0.163	0.434
ISP	0.558	0.072	0.318	-0.763
JFK	0.552	-.134	-0.822	0.048
LGA	0.475	-.591	0.444	0.477

Another test conducted was the test for heteroskedasticity, as known as the White test. The White test examines the null hypothesis that the variances of the residuals are homogenous. The test results showed that the p-value was high so that the null hypothesis was accepted.

6.3 Model Format

6.3.1 Equation 1 for Individual Region

The model decomposes average daily delay into components related to different delay casual factors. The explanatory variables include average arrival deterministic queuing delay aggregated for all airports in the region, average observed arrival delay in the RNAS excluding the region under consideration, proportion of LCCs at the airport, HHI index of the region, adverse weather, seasonal effects, and others. Since demand management strategies were implemented at airports in the New York and Chicago regions, to capture the effects of those strategies, the entire study time period was divided into several time windows, and dummy variables were introduced for the windows.

$$D(t) = \alpha + \beta_0 . D_S(t) + \rho_1 . LQ(t) + \rho_2 . LQ^2(t) + \rho_3 . I_C(t) + \rho_5 . LC(t) + \rho_6 . HHI(t) + \sum_K \lambda_K W_K(t) + \sum_i \omega_i S_i(t) + \sum_j \theta_m D_m(t) + v(t)$$

6.3.2 Equation 2 for RNAS

The model for the RNAS decomposes daily average delay at the remainder airports, excluding the airports in the region under consideration. The explanatory variables include delays in the region, convective weather, total operations, seasonal effects, yearly dummy variables, and other factors.

$$D_S(t) = \alpha + \gamma_0 . OP(t) + \gamma_1 . D(t) + \gamma_2 . SQ(t) + \gamma_3 . HHI(t) + \sum_k \lambda_{kS} W_k(t) + \sum_i \omega_{iS} S_i(t) + \sum_m \theta_{mS} D_m(t) + u(t)$$

The notations in the above two models are described as follows:

$\mathrm{D}\,(\mathrm{t}) =$ Weighted average observed arrival delay against flight plan in the region under consideration on day t;

$\mathrm{D_S}(\mathrm{t}) =$ Weighted average observed arrival delay in RNAS on day t;

$\mathrm{LQ}\,(\mathrm{t}) =$ Weighted average arrival deterministic queuing delay in the region under consideration on day t;

$\mathrm{SQ}\,(\mathrm{t}) =$ Weighted average arrival deterministic queuing delay in RNAS on day t;

$\mathrm{I_C} =$ Daily IMC ratio component for the region on day t;

$\mathrm{OP}(\mathrm{t}) =$ Total operations (arrivals) in RNAS on day t;

$\mathrm{HHI}\,(\mathrm{t}) =$ Herfindahl–Hirschman Index in the region on day t;

$\mathrm{W_K} =$ Weather index of region k on day t;

$\mathrm{S_i} =$ Seasonal dummy variable, set to 1 if daily arrival delay is observed in quarter i and 0 otherwise;

$\mathrm{D_m} =$ Yearly/Demand Management Dummy Variable, set to 1 if daily arrival delay is observed in time period m and 0 otherwise;

v(t) and u(t) = Stochastic error terms; and

$\alpha, \beta, \lambda, \omega, \theta,$ and γ are coefficients.

6.4 Estimation Results

We used 3SLS regression to estimate the coefficients in the 2-equation simultaneous equation models and compare the outcomes from different regions. Table 19 shows the regression estimates for the region equations for all nine regions considered in this research. Due to space limitations, we have not listed the regression results for RNAS equations, but we have summarized the effect of delay in each region on the RNAS in Table 19.

The 3SLS results shown in Table 19 indicate that for all regions, average queuing delay, delay from the RNAS, and weather are significant factors (highlighted in bold). It also shows that for all regions, the imbalance between capacity and demand is a major contributing factor to average daily arrival delay. However, the negative coefficient for the quadratic term of average queuing delay shows that this factor diminishes as average queuing delay increases. The R-square values from the table clearly indicate that the model captured considerable variation in the average daily arrival delay in all the regions.

The estimated coefficients of IMC ratio components show that delay increases due to the increase of adverse weather conditions, which is intuitively correct and easy to observe. The estimates for LCC market share components are not alike for different regions. The outcomes indicate that a drift towards the monopoly of LCC operations at a particular secondary airport results in the increase of delay for the Washington DC, Boston, New York, Los Angeles, Chicago, Houston, and Dallas regions. This might be due to induced demand of passengers who generally opted for car or train for short trips. Also it was noticed that Legacy airlines shifted their operations to new international markets to sustain themselves economically [62]. Hence LCC operations even though being beficial to airline passengers put huge pressure on airports and the NAS. However, for the South Florida region, the introduction of LCCs at primary airports increases regional delay, and the introduction of LCCs to secondary airport reduces the delay. As can be observed in Table 20 this might be due to a lower percentage of LCC operations in the South Florida region, not reaching its tipping point. San Francisco Bay area airports behave distinctly, as the increase of LCCs reduced regional delay considerably. This might be due to the presence of regional airlines operating short haul distances in and around the Bay area itself.

Table 19 Estimation Results of Arrival Delays at Different U.S. Regional Airport Systems

	Reg 1	Reg 2	Reg 3	Reg 4	Reg 5	Reg 6	Reg 7	Reg 8	Reg 9
R-square	0.591	0.583	0.592	0.482	0.631	0.492	0.594	0.508	0.486
Intercept	-8.833	6.667	0.918	3.546	32.274	9.295	50.256	16.116	8.638
Average queuing delay	0.046	0.140	1.251	0.558	0.201	0.486	0.199	0.408	0.216
Quadratic average queuing delay at airport	0.000	0.000	-0.013	-0.002	0.000	-0.009	0.000	-0.001	0.000
Predicted arrival delay at RNAS	1.014	0.951	0.356	0.394	1.413	0.694	0.415	0.385	0.564
IMCcomp1	0.235	2.086	0.489	0.437	1.627	0.945	2.262	1.141	1.378
IMCcomp2				0.690					
LCCcomp1	0.401	0.062	-3.184	-0.713	2.011	52.043	85.102	57.202	4.754
LCCcomp2	0.335			-0.518	1.292	-23.304	11.457	-2.387	-44.233
HHI_region	21.583	-3.431	-5.688	0.435	-85.808	-6.247	-77.972	-16.831	-19.357
Thunderstorm Ratio									
Region 01	0.073	1.921	0.426	-0.660	-0.561	13.367	0.011	0.760	-0.631
Region 02	-0.789	-1.459	-1.546	-1.393	-1.880	-0.386	0.407	0.689	17.853
Region 03	-0.092	-0.039	-0.667	-0.738	-0.089	0.778	-0.250	-0.227	0.732
Region 04	13.583	3.926	-0.271	-1.420	9.718	4.771	1.605	2.691	0.494
Region 05	-1.759	-5.111	-1.074	-0.316	-7.051	-2.595	-3.726	-5.621	-4.979
Region 06	-3.756	-4.119	2.041	1.531	-8.302	-4.362	0.783	22.973	14.633
Region 07	0.779	2.496	-1.220	-1.641	2.783	-0.100	-1.540	3.774	-1.938
Region 08	1.047	5.303	0.884	7.467	4.992	-0.194	-3.464	0.406	1.746
Region 09	-4.435	-8.069	19.159	25.204	-8.743	-4.582	-8.425	-7.497	-9.231

Table 19: Continued									
Region 10	-0.540	-2.152	0.336	0.739	-3.116	-0.341	4.291	0.449	1.581
Region 11	5.370	17.070	1.202	1.271	26.613	0.051	-10.043	0.764	-0.189
Region 12	-0.095	-3.740	1.867	0.300	-6.815	-6.453	37.806	-4.017	-4.976
Region 13	-2.147	-0.042	0.608	0.696	-0.670	-1.567	6.195	-2.227	-0.951
Region 14	1.081	1.294	1.338	0.842	-0.429	-1.200	1.012	-3.251	-0.891
Region 15	-3.267	-2.257	1.539	1.885	-5.081	-0.352	1.912	2.195	3.824
Region 16	0.066	-3.921	8.294	1.565	-0.118	4.086	-7.081	0.014	-7.012
Demand Mgt. Regimes									
Sepb	0.121	-0.445	-2.847	-1.761	-1.651	-0.017	-0.280	-0.187	-0.994
Sepa	2.894	-2.742	0.463	-4.671	-1.769	1.596	-5.755	-1.488	-1.566
Year2002	1.411	-2.872	0.954	-1.266	-1.945	1.013	-5.658	-0.514	-0.090
Year2003	1.181	-2.432	1.821	-1.374	-2.779	2.288	-4.840	-0.628	0.013
CapA							-0.023		
Year2004	-0.485	-3.141	3.104	0.105	-7.748	2.907	-3.533	-0.203	1.184
Year2005	-1.327	-4.168	5.419	1.510	-5.416	3.207	-2.864	-0.712	1.190
LOA							-0.419		
Year2006	-1.141	-3.887	6.972	2.225	-7.309	2.710	-1.694	0.004	3.837
Year2007	-0.825	-3.474	8.281	1.365	-7.167	4.192	-2.510	1.262	2.546
Year2008	-1.582	-4.361	10.941	2.091	-10.235	5.991	-1.918	0.443	3.105
Year2009	-1.340	-3.774	13.287	2.252	-9.621	7.428	-4.898	-0.516	1.901
Year2010	-0.824	-2.675	16.364	2.403	-10.712	8.035	-2.867	0.470	2.745
quarter2	0.828	0.460	-0.208	-1.011	1.587	0.662	-0.976	-0.623	-0.054
quarter3	0.531	1.219	-0.443	-0.699	1.922	0.104	-0.525	-0.467	-0.520
quarter4	-0.598	-0.243	1.399	0.556	0.050	0.248	0.685	0.449	0.939

Note: Figures in bold are significant for 95% level of confidence.

Highlighted cells for LCCcomp1 and LCCcomp2 indicate individual airport share and no correlation.

Highlighted cells for dummy variables indicate slot management instead of yearly dummy.

The column headings represent the following regions:

Reg 1 = Washington Metropolitan Area, comprising DCA, IAD and BWO airports
Reg 2 = New England Area, comprising BOS, MHT and PVD airports
Reg 3 = San Francisco Bay Area, comprising SFO, OAK and SJC airports
Reg 4 = Greater Los Angeles Area, comprising LAX, LGB, SNA, BUR and ONT airports
Reg 5 = New York Metropolitan Area, comprising JGK, LGA, EWR and ISP airports
Reg 6 = South Florida Metropolitan Area, comprising MIA and FLL airports
Reg 7 = Chicago Metropolitan Area, comprising ORD and MDW airports
Reg 8 = Dallas–Fort Worth Metropolitan Area, comprising DFW and DAL airports
Reg 9 = Greater Houston Area, comprising IAH and HOU airports

Table 20 Percentage of LCC Operations at Airports in Each Region

Region	**Airports (% LCC share)**				
Washington–Baltimore	DCA (6.10%)	BWI (59.14%)	IAD (23.97%)		
New England	BOS (13.85 %)	MHT (37.16%)	PVD (38.62%)		
SF Bay Region	SFO (8.91%)	OAK (61.94%)	SJC (47.73%)		
Los Angeles	LAX (19.39%)	LGB (57.00%)	ONT (41.31%)	SNA (26.83%)	BUR (58.61%)
New York	EWR (6.23 %)	ISP (75.32 %)	JFK (30.55%)	LGA (12.06 %)	-
South Florida	MIA (2.22%)	FLL (35.32%)			
Chicago	ORD (7.59%)	MDW (84.80%)			
Dallas	DFW (2.08 %)	DAL (85.33%)	-	-	-
Houston	IAH (1.17 %)	HOU (86.72 %)			

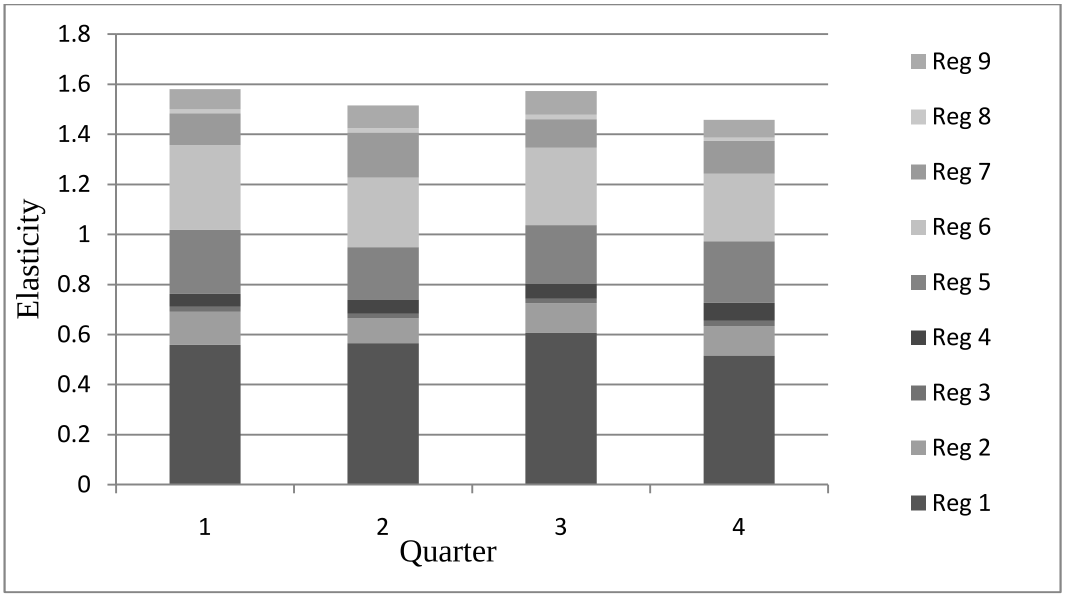
The estimates for the HHI index for each region (Table 21) indicate that for every region except the Washington-Baltimore and Los Angeles areas, an increase in HHI will lead to the reduction of delay. This is a clear indication that with an increase in competitiveness, there are more interactions and conflicts, thus increasing delay. The results for the convective weather indicate that most of the airports are affected significantly by the convective weather in the same region. The results from the yearly dummy variables or demand management regimes for some regions show the average daily arrival delay trend from 2000 to 2010. For all the regions except a few, the average daily arrival delay decreased relatively from 2000 to 2010. In the regional airport systems of the San Francisco Bay area, Los Angeles, Miami, and Houston, the arrival delay decreased drastically after 9/11 and slowly increased to the level of pre–9/11 in 2010. The increase of delay in the Miami and San Francisco Bay area regions was higher than any other regions. The arrival delay at most of the airports decreased after 9/11 and then increased gradually afterwards. The slot management techniques at airports in New York, Washington, and Chicago had a definite impact on reducing delay in the regions.

Table 21 HHI for Each Region

Region	Washington-Baltimore	New England	SF Bay Region	Los Angeles Region	New York Region	South Florida	Chicago Region	Dallas Region	Houston Region
HHI	0.344	0.750	0.390	0.532	0.334	0.514	0.678	0.757	0.673

6.5 Elasticity Analysis

Figure 21 shows the elasticity of the regional delay to the delay of RNAS in different quarters of the year 2010. The significant effect comes from Regions 1, 5 and 6, i.e., Washington–Baltimore, New York, and Miami regions. On the contrary, changes in the delays in the California and Texas regions have very little effect on the delay of RNAS. It shows that in the first quarter, a one percent increase in delays in different regions will lead to about 1.6 percent increase in the delay in the RNAS.

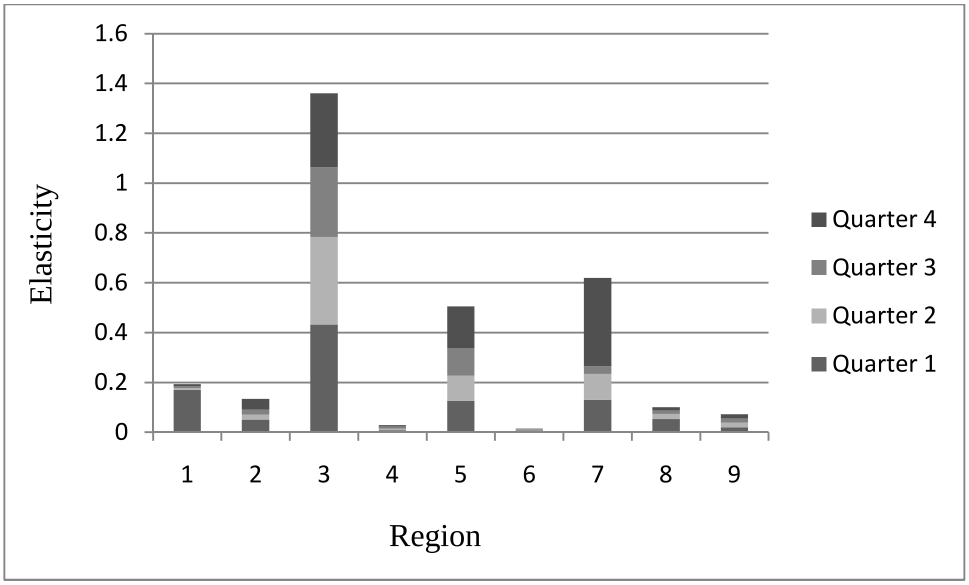


Note: The lables represent the following regions:

Reg 1 = Washington Metropolitan Area, comprising DCA, IAD and BWO airports
Reg 2 = New England Area, comprising BOS, MHT and PVD airports
Reg 3 = San Francisco Bay Area, comprising SFO, OAK and SJC airports
Reg 4 = Greater Los Angeles Area, comprising LAX, LGB, SNA, BUR and ONT airports
Reg 5 = New York Metropolitan Area, comprising JGK, LGA, EWR and ISP airports
Reg 6 = South Florida Metropolitan Area, comprising MIA and FLL airports
Reg 7 = Chicago Metropolitan Area, comprising ORD and MDW airports
Reg 8 = Dallas–Fort Worth Metropolitan Area, comprising DFW and DAL airports
Reg 9 = Greater Houston Area, comprising IAH and HOU airports

Figure 21 Effect of Delay at Each Region on RNAS for Different Quarters in 2010

Figure 22 presents the effects of average queuing delay on the average arrival delay in all regions. The elasticities were calculated for all four quarters of 2010. It can be seen from Figure 22 that the average queuing delay has a greater impact to the regions of San Francisco, New York, and Chicago, i.e., the increase of imbalance between capacity and demand will lead to a significant increase of arrival delays in these regions. This might be due to closed locations, congested airspace, and all three regions being in the top five ranked regions in domestic origin-destination passenger demand [60].



Note: Column headings represent the following regions:
Reg 1 = Washington Metropolitan Area, comprising DCA, IAD, and BWO airports
Reg 2 = Greater Boston Area, comprising BOS, MHT, and PVD airports
Reg 3 = San Francisco Bay Area, comprising SFO, OAK, and SJC airports
Reg 4 = Greater Los Angeles Area, comprising LAX, LGB, SNA, BUR, and ONT airports
Reg 5 = New York Metropolitan Area, comprising JGK, LGA, EWR, and ISP airports
Reg 6 = South Florida Metropolitan Area, comprising MIA and FLL airports
Reg 7 = Chicago Metropolitan Area, comprising ORD and MDW airports
Reg 8 = Dallas-Fort Worth Metropolitan Area, comprising DFW and DAL airports
Reg 9 = Greater Houston Area, comprising IAH and HOU airports

Figure 22 Average Queuing Delay Elasticity for Different Quarters in Year 2010

CHAPTER VII

CONCLUSION

The literature review uncovered that airport delay always has been a major problem for the aviation industry. Although several previous studies estimated the delay propagated through an individual flight from an airport to the system, a review of the literature shows that research pertaining to interactions among the entire air transportation system has not been conducted. The NextGen also identifies airport congestion and flight delay as two of the important issues for the aviation industry [4]. This research illustrates the effectiveness of applying a multivariate simultaneous equation model to study delay propagation from a single airport to other airports and to the rest of the system, and vice versa.

In the first section, the model developed for LGA and ORD takes into account most of the delay causal factors. The model estimates the effect of each of these factors using 2SLS regression. This approach is generally used to deal with the bidirectional relationship that exists between dependent variables, in this case, a single airport and the system. The estimated results clearly point toward the existing interdependency between flight delay at an individual airport and the NAS. It is seen that the delay at LGA and ORD significantly depends on delay at other airports and, similarly, LGA and ORD are major contributors to delay in the system.

The research also studies the system-wide benefit of capacity enhancement or improvement in demand management strategies on delay in the NAS. The results indicate that with an increase in capacity there is a proportionate reduction in delay at the airport and the NAS. However, this phenomenon is more predominant at ORD than at LGA. Through further observation, it can be seen that the major contributing factor for delay at ORD is queuing delay, while adverse weather is a major problem at LGA. This analysis helps to determine the effectiveness of capacity improvements and can be used as a decision making tool for airport improvement projects that require massive capital investments in the future.

In the second part, models were developed for 34 OEP airports takes into account all the delay causal factors mentioned earlier. The model estimates the effect of each of these factors using the 3SLS regression. This method is analyzing the bi-directional relationship that exists between dependent and independent variables and is suitable for equations with correlated error terms. The estimated results help to quantify the interdependency between flight delay at different airports and the NAS.

The regression results show that queuing delay and adverse weather are major delay causal factors at most of the studied airports. Aircraft equipment type is seen to be one of the important delay contributors in the case of a few airports, where large aircraft operations result in increasing average daily arrival delay. Airports located in same geographic regions had more interactions than others. Major airports such as PDX, MCO, ORD, and EWR had a higher impact on average arrival delay at other airports. From schedule-based models, it was found that a few airports had a negative impact on arrival delay at other airports. However, this scenario is subdued in the case of flight-plan-based results. The graphical representation for different time periods from the year 2000 to 2008 demonstrates a significant delay variation. Most airports had their delay reduced after 9/11 and gradually returned to pre-9/11 levels, with a peak in 2007.

In the report of NextGen Concepts of Operations, 15 metropolitan areas are identified as regional airports that have potential to provide additional capacity. The 15 metropolitan areas are Atlanta, Charlotte, Chicago, Houston, Las Vegas, Los Angeles, Minneapolis, New York, Philadelphia, Phoenix, San Diego, San Francisco, Seattle, South Florida, and Washington–Baltimore [7]. The change in air transportation structure will usher the growth of air travel and economic activity and the change in regional geography, demographic, and industrial distributions. The “Southwest Effect” seen in the New England region demonstrated the potential of regional airport system planning. Investigation of the regional airport system in Boston [8] shows a positive system development, with passenger demand shared among the airports in the region. More importantly, the benefits from the air transportation industry could be shared by a greater area and could encourage the development of the regional economy. Given this dynamically-changing background, it is important to have a tool to estimate the evolution of regional airport systems and their impact on the NAS. This study provides such a tool for decision makers and aviation planners.

Delay propagation has been studied extensively, primarily from a macroscopic perspective. To the best knowledge of the authors, this study is the first effort to investigate delay propagation considering multi-airport systems. This study illustrates the effectiveness of applying a multivariate simultaneous equation model to study delay propagation from a multi-airport system to the rest of the NAS, and vice versa. The regression results show that queuing delay and adverse weather are major delay causal factors in most of the studied regions. However, the delay elasticity caused by these two factors varies among different regions. Two variables were introduced to indicate the characteristics of multi-airport systems—HHI and LCC market share. The estimated coefficients of HHI show that more evenly-distributed operations among different airports lead to increases in regional-level arrival delay. This explains the inter-dependability among different airports, existing conflicts in airspace, and the need for proper regional level airport and airspace planning. The effects of LCC market share are not consistent. In most regions, the increase in LCC market shares at secondary airports leads to an increase in regional arrival delay. However, this is not true for the South Florida region, with FLL having the lowest LCC operations among secondary airports, as seen in Table 20.

Hence, it is necessary for airport planners to find the threshold for airline operations at individual airports. Furthermore, the outcomes of this study show that delays in the Washington–Baltimore, New York, and South Florida regions have greater impacts on delays in the RNAS.

To further this research, we need to explore other explanatory variables such as capacity ratio, runway configuration, wind speed, and demand management regimes for all the airports. To further this research, we can perform spatial analysis pertaining to individual airports and multi-airport regions. We also need to explore in depth each individual airport and regional delay trends and the impacts they have on the system. Figure 11 to Figure 18 shows us the trend of average daily arrival delay at each of the 34

OEP airports. It would be interesting to explore the factors affecting those delays and ways to reduce them.

We also need to conduct experiments on the specification of the model and the methodology used. Furthermore an important and necessary research would be to explore the economic implications of these delays on different airports and the regions in the US.

REFERENCES

Graham, T., *The Economic Impact of Civil Aviation on the U.S. Economy*. 2011, US Department of Transportation, Federal Aviation Administration: Washington, DC.

Airline On-Time Statistics and Delay Causes. 2011, Bureau of Transportation Statistics, Airline Service Quality Performance 234.

Michael Ball, et al., *Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States*. 2010, NEXTOR.

Next Generation Air Transportation System (NextGen). 2011, NextGen Integration and Implementation Office, Federal Aviation Administration: Washington, DC.

FAA's NextGen Implementation Plan 2011. 2011, NextGen Integration and Implementation Office, Federal Aviation Administration: Washington, DC.

Neufville, R.d., *Management of Multi-Airport Systems: A Development Strategy*. Journal of Air Transport Management, 1995. 2(2): p. 99-110.

Dillingham, G., *National Airspace System: Regional Airport Planning Could Help Address Congestion If Plans Were Integrated with FAA and Airport Decision Making*. 2009, Government Accountability Office.

Julie, C., *Regional cooperation and the regionalization of air travel in Central New England*. Journal of Transport Geography, 2006. 14(1): p. 23-34.

Beatty, R., et al., *Preliminary Evaluation of Flight Delay Propagation through an Airline Schedule*, in *2nd USA/Europe Air Traffic Managment R&D Seminar*. 1998: Orlando, FL.

Schaefer, L. and D. Millner, *Flight Delay Propagation Analysis with the Detailed Policy Assessment Tool*, in *2001 IEEE Systems, Man and Cybernetics Conference*. 2001: Tucson, AZ.

Wang, P.T.R., L.A. Schaefer, and L.A. Wojcik, *Flight Connections and Their Impacts on Delay Propagation*, in *Digital Avionics Systems Conference*. 2003. p. 5.B.4 - 5.1-9.

AhmadBeygi, S., A. Cohn, and Y. Gua, *Analysis of the Potential for Delay Propagation in Passenger Aviation Flight Networks.* Journal of Air Transport Management, 2008. 14(5): p. 280-286.

Bonnefoy, P.A. and R.J. Hansman, *Emergence of Secondary Airports and Dynamics of Regional Airport Systems in the United States*. 2005, MIT International Center for Air Transportation, Massachusetts Institute of Technology: Cambridge, MA.

Tony, D., *Do market-concentrated airports propagate more delays than less concentrated ones? A case study of selected U.S. airports.* Journal of Air Transport Management, 2009. 15(6): p. 280-286.

Laskey, K., N. Xu, and C.-H. Chen. *Propagation of Delays in the National Airspace System*. in *Proceedings of the Twenty-Second Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-06)*. 2006. Arlington, Virginia: AUAI Press.

Yu-Jie, L. and M. Song. *Flight Delay and Delay Propagation Analysis Based on Bayesian Network*. in *Knowledge Acquisition and Modeling, 2008. KAM '08. International Symposium on*. 2008.

Hansen, M. and Y. Zhang, *Operational Consequences of Alternative Airport Demand Management Policies: Case of LaGuardia Airport, New York.* Transportation Research Record: Journal of the Transportation Research Board, 2005. 1915(-1): p. 95-104.

Morisset, T. and A. Odoni, *Capacity, Delay, and Schedule Reliability at Major Airports in Europe and the United States.* Transportation Research Record: Journal of the Transportation Research Board, 2011. 2214(-1): p. 85-93.

Berardino, F. *History of the Slot Exchange Market in the US and Some Implications*. in *Nextor Workshop on Allocation and Exchange of Airport Access Rights*. 2007. Queenstown, MD.

Donohue, G., et al. *Air Transportation Network Load Balancing using Auction-Based Slot Allocation for Congestion Management*. in *NEXTOR Wye River Conference*. 2004. Fairfax, VA.

Capacity Needs in the National Airspace System 2007-2025: An Analysis of Airports and Metropolitan Area Demand and Operational Capacity in the Future., M.C. Blakey, Editor. 2007, The MITRE Corporation Center for Advanced Aviation System Development: Washington, DC.

Erie, S.P., et al., *Regional Airport Management Study*. 2005, Citigroup Technologies Corporation: Venice, CA.

Neufville, R.d., *Planning Multi-Airport Systems in Metropolitan Regions in the 1990's*. 1994, Massachusetts Institute of Technology: Cambridge, MA.

Neufville, R.d., *Policy Guidelines for the Option of a Development of a Multi-Airport System, the basis of a Dynamic Strategic Plan to provide the capability for flexible response to future challenges.* 1995, Massachusetts Institute of Technology: Cambridge, MA.

Neufville, R.d., *Airports in the 21st.Century, Prospects for the Future.* Transportation Research Board, 2001: p. 90-115.

Neufville, R.d., *Airport Systems Planning, Design and Management.* Technology Policy and Innovation, 2003. 8.

Neufville, R.d., *Airports of the Future: The Development of Airport Systems*, in *International Symposium and Exposition in Celebration of 100 years of powered flight*. 2003: Dayton, Ohio.

Neufville, R.d., *Multi-Airport Systems in the Era of No-Frills Airlines*, in *Transportation Research Board conference*. January 2004.

Bonnefoy, P.A. and R.J. Hansman, *Scalability Of The Air Transportation System And Development Of Multi-Airport Systems: A Worldwide Perspective*, in *Department of Aeronautics & Astronautics*. 2008, Massachusetts Institute of Technology: Cambridge, MA. p. 440.

Bonnefoy, P.A., R.d. Neufville, and R.J. Hansman, *Evolution and Development of Multi-Airport Systems: A Worldwide Perspective.* ASCE Journal of Transportation Engineering, 2010. 136(11): p. 1021-1030.

Brueckner, J.K., D. Lee, and E. Singer, *City-Pairs vs Airport-Pairs: A Market Definition Methodology for the Airline Industry*. 2010, University of California, Irvine LECG, LLC.

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

Hansen, M. and Q. Du, *Modeling Multiple Airport Systems: A Positive Feedback Approach*, in *Working Paper: University of California Transportation Center*. 1993, University of Cahfornia, Berkeley.

Bhargava, A., et al., *Three-Stage Least-Squares Analysis of Time and Cost Overruns in Construction Contracts.* Journal of Construction Engineering and Management, 2010. 1207.

Cervero, R. and M. Hansen, *Induced Travel Demand and Induced Road Investment: A Simultaneous Equation Analysis.* Journal of Transport Economics and Policy, 2002. 36(3): p. 469-490.

Himes, S.C. and E.T. Donnell, *Speed Prediction Models for Multilane Highways: Simultaneous Equations Approach.* Journal of Transportation Engineering 2010. 136(10).

OEP 35, Federal Aviation Administration. August 25, 2010]; Available from: http://aspmhelp.faa.gov/index.php/OEP_35.

IFR Operations in the NAS- Chapter I. Federal Aviation Administration (FAA) Library Manuals.

FAA ASPM System Overview Washington, DC.

National Oceanic and Atmospheric Administration. January 15, 2011; Available from: http://www.noaa.gov/about-noaa.html

Research and Innovative Technology Administration, Bureau of Transportation Statistics

Zhang, Y. and N. Nayak, *Macroscopic Tool for Measuring Delay Performance in National Airspace System.* Transportation Research Record: Journal of the Transportation Research Board, 2010. 10-2796(2177): p. 88-97.

Ryerson, M.S., M. Hansen, and J. Bonn. *Fuel Consumption and Operational Performance*. in *Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)*. 2011. Berlin, Germany.

Mccartney, S., *Why a Six-Hour Flight Now Takes Seven*, in *The Wall Street Journal*. 2010: New York.

Schumer, C.E. and C.B. Maloney, *Your Flight has Been Delayed Again: Flight Delays Cost Passenger, Airlines, and the U.S. Economy Billions*, J.E. Committee, Editor. 2008: Washington, DC.

Trani, A.A., *Airport Planning and Design: Aircraft Classifications*. 2010, Virginia Polytechnic Institute and State University: Falls Church, Virginia.

Neufville, R.d., *Low-Cost Airports for Low-Cost Airlines: FLexible Design to Manage the Risk.* Journal of Transportation Planning and Technology, 2007.

Motulsky, H. *Multicollinearity in multiple regression*. 2002 March 23, 2011]; Available from: http://www.graphpad.com/articles/Multicollinearity.htm.

Wooldridge, J., *Introductory Econometrics: A Modern Approach*. Fourth Edition ed. 2008.

Simultaneous Equation Models, College of Engineering, Purdue University, Indiana

Savin, N.E., *Two-stage least squares and the k-class estimator*, in *The New Palgrave Dictionary of Economics*, S.N. Durlauf and L.E. Blume, Editors. 2008

Gibbons, S. *OLS, GLS, 2SLS/IV, Maximum Likelihood*.

University, O.S. *Estimation Methods*.

Air Traffic Activity System (ATADS), FAA Performance and Operations Data. July 13, 2009]; Available from: http://aspm.faa.gov/opsnet/sys/Main.asp.

Nayak, N. and Y. Zhang, *Estimation and Comparison of Impact of Single Airport Delay on National Airspace System with Multivariate Simultaneous Models.* Transportation Research Record: Journal of the Transportation Research Board, 2011(2206 / 2011): p. 52-60.

Technology, A. *Hartsfield-Jackson Atlanta International Airport, United States of America*. August 13, 2010]; Available from: http://www.airport-technology.com/projects/hartsfield_jackson/

ASPM. *FAA Regions*. March 28, 2010]; Available from: http://aspmhelp.faa.gov/index.php/FAA_Region

Regional Airport System Plan Update - Baseline Runway Capacity and Delays Report. 2010, Flight Transportation Associates: Orlando, FL.

Atkins, S. *Observation and measurement of metroplex phenomena*. in *Digital Avionics Systems Conference, 2008. DASC 2008. IEEE/AIAA 27th*. 2008.

Regional Airport System Planning Analysis. 2011, SH&E an ICF international Company: California.

Inc, S.I., *SAS/STAT® 9.2 User's Guide: The PRINCOMP Procedure*, in *SAS/STAT® 9.2 User's Guide*. 2008, SAS Institute Inc: Cary, NC.

Gorman, L. *What Happened to the Airline Industry?* The National Bureau of Economic Research 2012; Available from: http://www.nber.org/digest/mar09/w14503.html.

APPENDICES

Appendix I- OEP Airports

ATL	Atlanta Hartsfield International
BOS	Boston Logan International
BWI	Baltimore-Washington International
CLE	Cleveland-Hopkins International
CLT	Charlotte/Douglas International
CVG	Cincinnati-Northern Kentucky
DCA	Ronald Reagan National
DEN	Denver International
DFW	Dallas-Fort Worth International
DTW	Detroit Metro Wayne County
EWR	Newark International
FLL	Fort Lauderdale-Hollywood International
IAD	Washington Dulles International
IAH	George Bush Intercontinental
JFK	New York John F. Kennedy International
HNL	Honolulu International
STL	Lambert St. Louis International
LAS	Las Vegas McCarran International
LAX	Los Angeles International
LGA	New York LaGuardia
MCO	Orlando International
MDW	Chicago Midway
MEM	Memphis International
MIA	Miami International
MSP	Minneapolis-St Paul International
ORD	Chicago O'Hare International
PDX	Portland International
PHL	Philadelphia International
PHX	Phoenix Sky Harbor International
PIT	Greater Pittsburgh International
SAN	San Diego International Lindbergh
SEA	Seattle -Tacoma International
SFO	San Francisco International
SLC	Salt Lake City International
TPA	Tampa International

Appendix II- Data Dictionary

- LOCID: Every Airport has specific airport ID and it is the fir column of the dataset

- YYYYMM : Year and Month

- DAY: Day

- HOUR: Local Hour (0 to 23)

- QTR: Quarter Hour (1 to 4)

    1 = 00 – 14 minutes

    2 = 15 – 29 minutes

    3 = 30 – 44 minutes

    4 = 45 – 59 minutes

- DlaSchOffA: Average OAG-Based Departure Delay Minutes

- DlaSchArrA: Average OAG-Based Arrival Delay Minutes

- DlaFPOffA: Average Flight Plan Based Departure Delay Minutes

- DlaFPArrA: Average Flight Plan Based Arrival Delay Minutes

- MetricDep: Count of ASPM Departures

- MetricArr: Count of ASPM Arrivals

- SchDep: Count of Scheduled Departures

- SchArr: Count of Scheduled Arrivals

- MC: Meteorological Conditions Flag (I-instrument, V-Visual)

- ADR: Airport supplied Departure Rate

- AAR: Airport supplied Arrival Rate

- OBPAX: Average observed number of passengers at the airport in a month

- Seats: Average number of seats in aircraft at the airport in a month

- ETMS_EQPT: IATA Aircraft Equipment Code from Enhanced Traffic Management Systems (ETMS)

- THUN: Value ‘1’ if the station reports thunderstorm, ‘0’ otherwise

ABOUT THE AUTHOR

Mr. Nagesh Nayak earned his Ph.D and M.S. in Civil and Environmental Engineering from the USF. He joined USF in August 2008 after he obtained his B.E. in Civil Engineering from University of Mumbai in India. His research interests include airport management, air traffic flow management, transportation planning and transportation network modeling. He was a research assistant for Professor Yu Zhang from Fall 2008 to Spring 2010. Since 2011, he worked at CUTR as a Research Assistant with Dr. Sissinio Concas on the TRIMMS model. Nagesh has got three research papers published, and the fourth has been accepted recently. Nagesh, together with his major advisor, Professor Yu Zhang, received the Fred Burggraf Award presented by Transportation Research Board of the National Academies, which recognizes excellence in transportation research by researchers 35 years of age or younger. Nagesh is also a recipient of the Graduate Research Award from Airport Cooperative Research Program (ACRP) for annual year 2009-2010, won ITS Florida - Anne Brewer Shanklin Scholarship Award in 2009 and 2011 and CUTR Georgia Brosch Award in 2011. Besides exceptional academic performance, Nagesh also participates actively in student organizations. He was the President of ITE Student Chapter at USF in 2009 and the President of Students of India Association in 2010. He is also a part of USF sports team at American College Cricket tournaments.