2018

Forecast To Grow: Aviation Demand Forecasting and Peer-Group Learning In The Era Of demand Uncertainty And Optimism Bias

Daniel Young-Sang Suh

University of Pennsylvania, dysuh03@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Transportation Commons, and the Urban Studies and Planning Commons

Recommended Citation

Suh, Daniel Young-Sang, "Forecast To Grow: Aviation Demand Forecasting and Peer-Group Learning In The Era Of demand Uncertainty And Optimism Bias" (2018). Publicly Accessible Penn Dissertations. 2977.

https://repository.upenn.edu/edissertations/2977

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2977
For more information, please contact repository@pobox.upenn.edu.
Forecast To Grow: Aviation Demand Forecasting and Peer-Group Learning In The Era Of demand Uncertainty And Optimism Bias

Abstract
Airport sponsors, typically municipal governments in the US, along with the Federal Aviation Administration (FAA) engage in a number of planning activities to determine the long-term development needs of airport infrastructure. One of the primary tasks of these airport planning activities is to estimate future use of the airport. Airport planners use two broad categories of methods to estimate future use of the airport; 1) peer group learning (as in, considering the experiences of “peer” airports) and 2) aviation demand forecasting. Airport master planning, a federally mandated planning process for airports for infrastructure planning such as building a new runway, for instance, relies on these techniques to be effective. Yet, there are numerous challenges to how airport planners can use these techniques effectively. These challenges can be largely categorized as the problem of demand uncertainty and optimism bias; demand uncertainty stemming from the dynamic socioeconomic and aviation industry trends and optimism bias from the economic development narrative surrounding airports and the federal funding incentives for airport infrastructure projects. Demand uncertainty and optimism bias create large forecast errors and have led airport planners to make unwise infrastructure investment decisions. In this dissertation, I use publicly available aviation and census data to develop and test new methodologies that enable airport planners to 1) identify true airport peers that share similar socioeconomic trends, 2) predict the probability of a severe contraction in passenger volumes in the next 10 years, and 3) improve forecast accuracy by incorporating past forecast errors systematically into the current forecast and “ground” optimistic forecasts. I show that the methodologies can have much more immediate and robust impact on airport planning than traditional methods to curtail demand uncertainty and optimism bias.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
City & Regional Planning

First Advisor
Megan S. Ryerson

Keywords
Airport planning, Demand uncertainty, Forecasting, Infrastructure planning, Optimism bias, Peer-group learning

Subject Categories
Transportation | Urban Studies and Planning

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2977
FORECAST TO GROW: AVIATION DEMAND FORECASTING

AND PEER-GROUP LEARNING IN THE ERA OF

DEMAND UNCERTAINTY AND OPTIMISM BIAS

Daniel Y. Suh

A DISSERTATION

in

City and Regional Planning

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation
Signature ___________________________
Megan S. Ryerson, Assistant Professor of City and Regional Planning

Graduate Group Chairperson
Signature ___________________________
Eugenie L. Birch, Lawrence C. Nussdorf Professor of Urban Research and Education

Dissertation Committee
Megan S. Ryerson, Assistant Professor of City and Regional Planning, University of Pennsylvania

Shane T. Jensen, Associate Professor of Statistics, University of Pennsylvania

Erick Guerra, Assistant Professor of City and Regional Planning, University of Pennsylvania
To my family.
ACKNOWLEDGMENTS

What I write here in these acknowledgments falls far short of the immense gratitude I have for my family, friends, and mentors in my life. It would take more than these few pages to properly express my feelings of gratitude, appreciation, and incredulity for guiding, supporting, and sustaining me during my graduate studies. I plan to do my best to continually share these feelings in many years to come by cherishing these precious relationships, but for now, let me try to say a few words in order to make sure that any future reader of this dissertation will know that this work is made possible by an army of people in my life.

First, if anyone has any positive things to say about me as a scholar or about my work, please thank Dr. Megan Ryerson. I met Dr. Ryerson in my last year of my master’s program and my fascination with her work in aviation motivated me to continue on as a PhD student and actively seek out Dr. Ryerson as my advisor. Little did I know that the knowledge I accumulated and contributed to in the field of airport planning during my PhD studies is just an icing on the cake compared to the immeasurable ways I have grown both as a person and as a scholar because of my relationship with Dr. Ryerson. Her passion for mentorship and her openness to share the joys and the challenges of being an academic propelled me and sustained me beyond what I thought I could do. She celebrated my successes as if they were her own and treated my failures as nothing more than the necessary growing pains. As a first-generation college student in my family, I had many doubts and challenges but Dr. Ryerson always encouraged me through her
frank talks about the reality of the industry and the academia. I hope to continue to cultivate our relationship as mentor-mentee and colleagues. (Sorry Dr. Ryerson, you are not done with me yet)

I owe an immeasurable debt of gratitude to my family. My parents, Kwang-Bin Suh and In-Sook Chang, as many of the folks of their generation did in Korea, had to sacrifice their dreams of higher education and provided for the family. I am forever grateful that they brought me and my sister Kristine to the United States and supported our education even through financial and physical difficulties. I am so proud of this dissertation because of what it means to them. I am equally blessed to have my wife, Soo Kyung, and her parents in my life without whom none of this could have happened in a very literal sense. My wife has been at my side from the beginning of this journey – she first started at Penn as a graduate student and encouraged me to begin my graduate studies. She has been a source of support and encouragement beyond what she realizes. (I love you!) My in-laws, Dr. Sung-Hyuk Kim and Dr. Hye-Ryung Kim, have also been an invaluable part of this journey. They have not only given me timely advice on this journey, which they themselves have gone on before, but they have also given me the greatest gift of time. They have sacrificed so much to take care of our son, Nathan, which allowed me to go on this quest for higher education. Literally, this dissertation and all my accomplishments as a scholar could not have happened if it were not for them.

I am deeply grateful to my colleagues and committee members at the University of Pennsylvania. Dr. Erick Guerra has always given me insightful and detailed feedback
and advice on my work and I owe him so much for the rigor and clarity of my work. Dr. Shane Jensen has been invaluable in my development as a quantitative researcher and scholar. Dr. Jensen’s guidance and mentorship has opened new doors for me in terms of my research and ignited and solidified my passion for data-driven research approaches. I also thank Dr. Eugenie Birch for leading our doctoral program and fostering the collegiate relationships with faculty members and colleagues through doctoral seminars. My friends and colleagues have also played an integral part in my development as a scholar through their friendship and guidance. Your presence made this journey a little less lonely and more enjoyable and I owe you all a drink or two at New Deck.

Finally, I thank God for these friends, family, and mentors in my life. I do not believe in chance meetings in life. I believe everything happens for a reason because of God’s love and goodness. So, I thank God for sustaining me during this arduous and often very disorienting journey by blessing me with the abundance of his people. In turn, I give him all the glory and pray that he blesses these people who mean so much to me.
ABSTRACT

FORECAST TO GROW: AVIATION DEMAND FORECASTING AND PEER-GROUP LEARNING IN THE ERA OF DEMAND UNCERTAINTY AND OPTIMISM BIAS

Daniel Y. Suh

Dr. Megan S. Ryerson

Airport sponsors, typically municipal governments in the US, along with the Federal Aviation Administration (FAA) engage in a number of planning activities to determine the long-term development needs of airport infrastructure. One of the primary tasks of these airport planning activities is to estimate future use of the airport. Airport planners use two broad categories of methods to estimate future use of the airport; 1) peer group learning (as in, considering the experiences of “peer” airports) and 2) aviation demand forecasting. Airport master planning, a federally mandated planning process for airports for infrastructure planning such as building a new runway, for instance, relies on these techniques to be effective. Yet, there are numerous challenges to how airport planners can use these techniques effectively. These challenges can be largely categorized as the problem of demand uncertainty and optimism bias; demand uncertainty stemming from the dynamic socioeconomic and aviation industry trends and optimism bias from the economic development narrative surrounding airports and the federal funding incentives for airport infrastructure projects. Demand uncertainty and optimism bias create large forecast errors
and have led airport planners to make unwise infrastructure investment decisions. In this dissertation, I use publicly available aviation and census data to develop and test new methodologies that enable airport planners to 1) identify true airport peers that share similar socioeconomic trends, 2) predict the probability of a severe contraction in passenger volumes in the next 10 years, and 3) improve forecast accuracy by incorporating past forecast errors systematically into the current forecast and “ground” optimistic forecasts. I show that the methodologies can have much more immediate and robust impact on airport planning than traditional methods to curtail demand uncertainty and optimism bias.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................... iii

ABSTRACT ............................................................................................................................ vi

LIST OF TABLES .................................................................................................................... x

LIST OF FIGURES ................................................................................................................ xi

PREFACE (optional) .............................................................................................................. xi

CHAPTER 1. INTRODUCTION ................................................................................................. 1
  1.1 Current State of Airport Planning ................................................................................ 1
  1.2 Demand Uncertainty and Optimism Bias in Airport Planning ............................... 4
  1.3 Dissertation Outline ...................................................................................................... 8
  1.4 Chapter Bibliography .................................................................................................. 12

CHAPTER 2. CHALLENGES AND OPPORTUNITIES FOR AVIATION DEMAND FORECASTING IN AIRPORT MASTER PLANNING ........................................ 15
  2.1 Introduction ................................................................................................................ 15
  2.2 History of Airport Expansions .................................................................................... 16
  2.3 Airport Master Planning and Aviation Demand Forecasting .................................. 20
      2.3.1 Aviation Demand Forecasting Methodologies .................................................. 21
  2.4 Challenges to Forecasting .......................................................................................... 24
      2.4.1 Dynamic Changes in the 21st Century ................................................................. 27
      2.4.2 Systematic Optimism ............................................................................................ 28
      2.4.3 Evaluation of Forecast Accuracy in the Era of Demand Uncertainty and Optimism Bias ........................................................................................... 31
  2.5 Alternative Airport Planning Approaches ................................................................. 34
  2.6 Methodological Improvements to Aviation Demand Forecasting ........................... 39
  2.7 Summary ..................................................................................................................... 42
  2.8 Chapter Bibliography ................................................................................................. 43

CHAPTER 3. IDENTIFYING TRUE AIRPORT PEERS USING CLUSTER ANALYSIS .................................................................................................................. 49
  3.1 Introduction ................................................................................................................ 49
  3.2 Peer Group Learning .................................................................................................. 53
  3.3 Methodology ................................................................................................................ 56
      3.3.1 Study Airports ....................................................................................................... 57
      3.3.2 Variable Selection ................................................................................................. 57
      3.3.3 K-means Clustering Algorithm ............................................................................ 63
LIST OF TABLES

Table 2-1 High Growth Forecast vs. Actual Passenger Volumes for AUS (Base Year 2000) .................................................................................................................. 31

Table 2-2 Low Growth Forecast vs. Actual Passenger Volumes for AUS (Base Year 2000) .................................................................................................................. 32

Table 3-1 Sources and Types of Uncertainty and the Corresponding Metrics .......... 58

Table 3-2 Summary of Selected Data for 64 Airports in the Top 50 MSAs .......... 63

Table 3-3 Summary Table for 11 Airport Clusters .................................................. 66

Table 4-1 Summary Statistics for 10-Year Percent Changes in Passenger Volumes ..... 90

Table 4-2 Summary Statistics for 10-Year Change Clusters .................................. 92

Table 4-3 Overview of Variables ........................................................................... 93

Table 4-4 Summary Table for the Binary Logistic Regression Model ................. 101

Table 5-1 Summary Statistics for MFE-Adjusted Forecast Errors and ............... 123

Table 5-2 Paired Wilcoxon Signed Rank Test for MFE-adjusted Forecast Errors ... 123

Table 5-3 Summary Statistics for MGBFE-Adjusted Forecast Errors and .......... 127

Table 5-4 Paired Wilcoxon Signed Rank Test for MGBFE-adjusted Forecast Errors .. 128

Table 5-5 List of Selected Variables for Peer Identification Clustering ............... 130

Table 5-6 Summary Statistics for MPBFE-Adjusted Forecast Errors and .......... 132

Table 5-7 Paired Wilcoxon Signed Rank Test for MPBFE-adjusted Forecast Errors ... 133

Table 5-8 Summary Statistics for EMPBFE-Adjusted Forecast Errors and .......... 136

Table 5-9 Paired Wilcoxon Signed Rank Test for EMPBFE-adjusted Forecast Errors. 136

Table 5-10 Summary Statistics for MFE, MGBFE, MPBFE, and EMPBFE .......... 139
LIST OF FIGURES

Figure 2-1 10-Year TAF Forecast Errors for Top 64 Busiest Airports (1995 – 2005).... 34

Figure 3-1 Annual Passenger Volumes for Top 64 Busiest Airports in the US (1990 – 2015).................................................................................................................................................. 51

Figure 3-2 Cluster Visualization Using Principal Components (k = 11)......................... 65

Figure 3-3 Bar Plots of Cluster Means (k = 11).............................................................................. 68

Figure 4-1 10-Year Forecasted Growths and Forecast Errors of .............................. 78

Figure 4-2 Annual passenger demand and 10-year TAF forecast (base year 1995, target year 2005) for growing airports .................................................................................................................. 83

Figure 4-3 Annual passenger demand and 10-year TAF forecast (base year 1995, target year 2005) for contracting airports .......................................................................................................................... 84

Figure 4-4 64 study airports in the top 50 MSAs................................................................. 89

Figure 4-5 Histograms of 10-Year Percent Changes in Passenger Volumes............. 91

Figure 4-6 Correlation Matrix................................................................................................. 98

Figure 5-1 Histogram of the Forecast Errors for 10-Year Aviation Demand Forecasts 112

Figure 5-2 The MFE-adjusted Forecast Errors and the Actual Forecast Errors (n=64) 124

Figure 5-3 Forecasted 10-year growth vs. forecast errors ............................................. 125

Figure 5-4 The MGBFE-adjusted Forecast Errors and the Actual Forecast Errors (n=52) .................................................................................................................................................. 129

Figure 5-5 Bivariate Cluster Plot (k = 7, n = 64) ................................................................. 131

Figure 5-6 The MPBFE-adjusted Forecast Errors and the Actual Forecast Errors (n=64) .................................................................................................................................................. 134

Figure 5-7 The EMPBFE-adjusted Forecast Errors and the Actual Forecast Errors (n=64) .................................................................................................................................................. 137
CHAPTER 1. INTRODUCTION

1.1 Current State of Airport Planning

Airport planning is an umbrella term that encompasses any planning activity in support of providing adequate level of airport infrastructure to serve the local, regional, and national aviation demand. Airport sponsors, typically municipal governments in the US, along with the Federal Aviation Administration (FAA) engage in a number of planning activities to determine the long-term airport infrastructure investment needs. One of the primary tasks of these airport planning activities is to estimate future use of the airport from airlines, passengers, and cargo shippers. What will the future passenger demand be? How can we use the airport most effectively to serve the demand? Airport planners use two broad categories of methods to estimate future use of the airport; 1) peer group learning and 2) aviation demand forecasting.

In peer group learning, airport planners compare their airports with airports with similar characteristics and engage in peer-to-peer exchange of information about lessons learned from past experiences and technical and planning guidance (ACRP Synthesis 46, 2013). Various airport planning documents contain information about airport peers and the associated lessons and benchmarked performance outcomes. Aviation demand forecasting, on the other hand, is an activity of estimating the future aviation demand (e.g., how many passengers will use the airport in the future) and is an integral part of determining infrastructure needs (US DOT, 2009). The estimate of how many passengers will use the airport helps airport planners assess their current level of capacity and ensure they are
growing their capacity in step with the future demand (De Neufville, Odoni, Belobaba, & Reynolds, 2013).

Appropriate use of these airport planning techniques (peer group learning and aviation demand forecasting) is critical because any decision to expand or build new infrastructure is based on these techniques. In a federally mandated airport planning process known as airport master planning, airport sponsors are required to produce aviation demand forecasts to evaluate whether their current capacity is projected to meet the future demand (De Neufville et al., 2013). In preparing airport master plans, airport planners also consult with their peer airport sponsors to learn from the peer airports’ past master planning experiences (ACRP Synthesis 46, 2013). The decision to build a new runway, which typically costs hundreds of millions of dollars and years of planning and construction (about 10 years on average), for example, follows from the airport master planning process involving these techniques. If the forecasted demand exceeds available infrastructure, congestion and delays will disrupt airport operations and incur economic and environmental costs (Ball et al., 2010). On the other hand, if airport planners oversupply infrastructure, e.g., the future demand is below expectations, airport infrastructure will be underutilized, and investments in airport infrastructure will be wasted (Redondi, 2013).

---

1 The term “demand” is used throughout the dissertation interchangeably with passenger volumes or the number of passengers boarding aircraft at an airport. In the aviation industry, aviation demand denotes a number of different meanings with respect to airports. It could mean passenger volumes at an airport, the number of aircraft takeoffs and landings at an airport (operations), or cargo volumes. I limit the use of the term to denote passenger volumes because the focus on passengers enables exploration of the research questions and discussion of the results in a clear way that does not suffer information loss due to the omission of the other meanings of demand.
Malighetti, & Paleari, 2012) and airport will be unnecessarily expensive to use (Forsyth, 
1998; Ryerson, 2016).

This mismatch between airport capacity and aviation demand, unfortunately, 
characterizes many of the major airports in the US. While frequent delays and congestions 
are the status quo for some airports, other airports are instead struggling to incentivize 
airlines and passengers to use their excess capacity (Ryerson, 2016). For instance, as I will 
uncover, the major airport in St. Louis, MO is sitting on a rarely used $1.1 billion runway 
built based on a forecast that turned out to have more than a +100% forecast error (i.e., the 
forecasted demand was almost double the actual demand). The unusually large forecast 
error for St. Louis as well as similarly large forecast errors for the major airports in 
Cincinnati and Pittsburgh is primarily due to the demand uncertainty induced by the event 
of de-hubbing, in which the dominant airline with a significant level of air service at an 
airport makes a decision to drastically reduce or remove their service entirely from the 
airport (Redondi et al., 2012). Consequences from the oversupply of infrastructure persist 
to this day for these airports as the airport sponsors are now spending even more money to 
subsidize airlines to launch new services in an effort to recapture the lost demand and fill 
the excess capacity (Ryerson, 2016).

As the primary airport planning techniques in the airport master planning process 
that guide such significant infrastructure investment decisions, peer group learning and 
aviation demand forecasting are difficult to implement effectively due to what scholars 
have identified as the problems of demand uncertainty and optimism bias. These terms
have specific meanings in the context of airport planning and in the next section, I expound on these terms in more detail.

### 1.2 Demand Uncertainty and Optimism Bias in Airport Planning

Demand uncertainty arises when aviation demand follows an unpredictable pattern or a sudden disruption is introduced to a demand pattern. Demand uncertainty in the airport planning environment became an issue for airport planning after the deregulation of the airline industry in 1978 and has intensified especially in the post-2000 era of airline mergers and consolidations. Since the deregulation of the airline industry, aviation demand at an airport was no longer controlled and calibrated by the federal regulation but instead was determined by the airlines that now had the freedom to choose their service level at any airport and on any route. This meant that the forces of market economy began to have more direct impacts on aviation demand at airports. Particularly in the beginning of the 21st century, economic downturns and fluctuating fuel prices affected airline business significantly resulting in a number of airline mergers and consolidations. Airlines that acquired the struggling airlines through mergers consolidated the existing services, often reducing or pulling services from markets/airports in order to improve efficiency in their overall portfolio of markets. These dynamics brought disproportionate changes to airports; airports that were once considered as peers may not share similar experiences anymore and conversely, airports that did not share similar trends may now be going through similar experiences. These changes also produced disproportionate levels and magnitudes of
forecast errors in aviation demand forecasts; forecasts for some airports overestimate by significant margins because of a severe contraction in passenger volumes caused by these changes (Redondi et al., 2012).

At the same time, the growth-oriented attitude of airport sponsors and the aviation industry as well as the federal funding mechanism for airport infrastructure could provide incentives for airport planners to inflate aviation demand forecasts either intentionally or not. In one sense, airport sponsors possess a generally optimistic view of the future of aviation and assume optimistically linear growths in the demand forecasts. They also view their airports as the engines of regional economic development (J. K. Brueckner, 2003). On the other hand, Wachs (1989) and Flyvbjerg (2005b) show that forecasters may also face political and organizational pressures to inflate the forecasts strategically to gain project approval. In the airport planning literature, there is no existing research that confirms the presence of such strategic behavior but the federal funding structure for airport infrastructure that is based on the forecast of future demand and airport sponsors’ aspirations to expand their airports may incentivize such behaviors.

For the purpose of this dissertation, however, I use the term optimism bias as an overarching term that refers to the act of overestimating benefits and underestimating costs regardless of whether it was done intentionally or not, in the same way the Airport Cooperative Research Program (ACRP), an industry driven applied research program, uses the term in their report (ACRP Report 76, 2012). The majority of aviation demand forecasts, as will be shown in this dissertation, are systematically optimistic (i.e.,
overestimate) and I use the term optimism bias to describe this systematic phenomenon instead of using the term to distinguish the causes behind such optimistic forecasts. The distinction between optimistic forecasts due to self-deception (i.e., assume linear growths) and intentional deception (i.e., towards gaining project approval) matters more in the context of broader policy discussions and less in the context of technical improvements to airport planning techniques. Flyvbjerg (2008) also concedes that the distinction is not useful if the primary objective is to improve forecast accuracy methodologically.

In response to the problem of demand uncertainty and optimism bias, a group of scholars have suggested new airport planning frameworks in place of the current airport master planning process (Burghouwt, 2007; Jan H. Kwakkel, Walker, & Marchau, 2010; Neufville, 2000). These alternative frameworks are Dynamic Strategic Planning, Flexible Strategic Planning, Adaptive Policy-Making, and Adaptive Airport Strategic Planning. These frameworks install various decision-points in the planning process at which planners would monitor the unfolding situations and reevaluate the feasibility of the plan and formulate a new one if needed. While the core idea behind the new planning frameworks is sound, there is little empirical evidence that these frameworks would reduce the overall planning costs (Jan H Kwakkel, Walker, & Marchau, 2012) and there is no meaningful push towards implementing them in the actual planning process.

Instead, I argue that there is ample room for methodological improvements in how airport planners can use these airport planning techniques, peer-group learning and aviation demand forecasting, which could provide more immediate and significant contributions to
improving the robustness and effectiveness of these techniques in the airport master planning process. In this dissertation, I pose and answer the following research questions:

1) What are the dynamics of socioeconomic and operational changes affecting the airports in the post-2000 era of airline mergers and consolidations?
2) What are the operational and socioeconomic characteristics of an airport on the verge of experiencing a severe contraction in passenger volumes?
3) Does reference class forecasting (a method of calibrating a forecast based on past forecast errors) produce statistically significant reductions in forecast errors for aviation demand forecasts?
   a. What is the relevant and effective definition of a reference class of the forecast errors?

Towards answering these questions, I propose new methodologies that enable airport planners to 1) identify true airport peers that share similar operational and socioeconomic trends, 2) predict the probability of experiencing a severe contraction in passenger volumes in the next 10 years, and 3) improve forecast accuracy by incorporating past forecast errors systematically into the current forecast and “ground” optimistic forecasts. These methodologies are designed to help airport planners employ publicly available data in such a way they help them make better informed infrastructure investment/planning decisions in the era of demand uncertainty and optimism bias.
My results indicate that the dynamic operational and socioeconomic trends have disrupted the traditional definitions of airport peers and more nuanced approaches to peer identification can benefit the practice of peer group learning in measurably significant ways. For instance, the de-hubbed airports are better positioned to successfully learn relevant planning lessons from one another than the traditionally defined peers whose planning needs may be drastically different. My research also indicates that airport planners can adopt more cautious approaches to forecasting and decision making by considering the disproportional impacts of operational and socioeconomic trends as well as learning from the past mistakes and errors in the forecasts, and reduce the chance of making unwise infrastructure investment decisions.

1.3 Dissertation Outline

In this dissertation, I propose and answer the research questions in order to understand the dynamics of operational and socioeconomic trends affecting airports in order to improve airport planners’ ability to implement peer-group learning and aviation demand forecasting more effectively and help airport planners make more informed infrastructure investment decisions.

In **CHAPTER 2**, I provide a history of airport planning and runway expansions and discuss the problem of demand uncertainty and optimism bias in the context of runway expansions. Using literature review and exploratory data analysis, I show that the problem of demand uncertainty and optimism bias results in systematic overestimations and I argue
that improving aviation demand forecast accuracy and addressing demand uncertainty in the planning environment is a critical issue in airport planning. I then discuss and contextualize my proposed methodologies in subsequent chapters.

In **CHAPTER 3**, I discuss how the dynamic changes in the 21st century have brought disproportionate impacts to airports. I show that this requires a new way of identifying airport peer groups, i.e., airports with similar characteristics, and develop a new airport peer identification methodology to both understand these dynamics and help airport planners identify their true peers. While the practice of peer-group learning is well-recognized and popular, there is opacity to how airport sponsors choose their peer airports. Most widely used criteria is the Federal Aviation Administration (FAA)’s airport categories based on the number of passenger enplanements (i.e., boardings) (FAA, 2017). However, I show that the simple static metric of passenger enplanements does not capture the dynamic changes and may harm the practice of peer-group learning. I propose a new methodology to identify true airport peers using a variety of dynamic operational and socioeconomic variables and show that this method results in peer groups that are reflective of the trends in the planning environment.

In **CHAPTER 4**, I develop a predictive model to estimate the probability of an airport experiencing a severe contraction in passenger volumes in the next 10 years. Airport sponsors typically forecast the aviation passenger demand 10 or more years into the future in their airport master planning documents. This is in accordance with the fact that the average length of time from planning to completion of a new runway is about 10 years. I
show that airports with a diverse mix of passenger traffic and a balanced distribution of market shares among airlines located in a region that is attracting population are more likely to have stable passenger demand and less likely to experience a sudden disruption in their passenger volumes. This insight carries a significant relevance to airport planners especially in their airport master planning and aviation demand forecasting processes for runway expansions, as it could help them assess the health of their airports and make more informed infrastructure investment decisions.

In CHAPTER 5, I develop methodologies based on the theories of reference class forecasting to “ground” the optimistic forecasts by incorporating past forecast errors into the current forecast. Behavioral economists Kahneman and Tversky (1977) originally developed the concept of reference class forecasting in which forecasters use the distribution of past forecast errors of similar projects to “de-bias” the current forecast. This method of “grounding” the forecasts has been shown to be effective in some domains, yet it has not been applied to aviation demand forecasting (Lovallo & Kahneman, 2003; Flyvbjerg, 2005a). The key task of reference class forecasting is defining a relevant reference class, a group of projects that are similar in nature. I develop four reference class identification methods for aviation demand forecasts and statistically test their performances in terms of forecast accuracy improvement. This constitutes a first serious attempt at applying reference class forecasting to aviation demand forecasting. My results indicate that identifying a reference class of airports that share similar socioeconomic and airport trends improves the forecast accuracy the best among the proposed methods.
Lastly in **CHAPTER 6**, I summarize the findings of my dissertation and discuss policy implications and suggest recommendations for implementing these methodologies in the current framework of airport master planning process to improve the decision-making process and achieve more effective planning outcomes.
1.4 Chapter Bibliography


Flyvbjerg, B. (2005b). Policy and planning for large infrastructure projects: problems, causes, cures (Vol. 3781). World Bank Publications. Retrieved from http://books.google.com/books?hl=en&lr=&id=ZyUsyXZdHUcC&oi=fnd&pg=PA1&dq=%22for+roads+20.4+percent.+The+difference+in+cost+overrun+between+the+three+project+types%22+%22Overrun+is+found+in+the+20+nations+and+5+continents+covered+by+the%22+&ots=a4JZRYlkq0&sig=PhBFRqwrPbO_OAGzeDIX4bFdAs


13
CHAPTER 2. CHALLENGES AND OPPORTUNITIES FOR AVIATION DEMAND FORECASTING IN AIRPORT MASTER PLANNING

2.1 Introduction

Airports maintain and update their infrastructure to ensure they are able to serve airlines and passengers efficiently. A growing concern among major airports in the US is the continued growth and concentration of passenger demand. When this demand grows beyond the current capacity, airports can become congested and passengers will experience delays. Airport planners meet this challenge by expanding airport capacity via runway expansions following the federally mandated planning process known as airport master planning (De Neufville & Odoni, 2003). A cornerstone of any airport master plan is the aviation demand forecast, the forecast of future airport activity. A decision to plan and expand a runway, which typically spans a period of 10 years from the planning to completion, predicates on the accuracy of the projected future passenger demand. Yet aviation demand forecasts are known to be inaccurate at best and biased at worst. Given that these forecasts guide airport infrastructure investment decisions that cost millions or even billions of dollars, there is a surprising lack of research in aviation demand forecast accuracy. In this chapter, I argue that improving aviation demand forecast accuracy is a critical issue in airport planning and lay the foundational knowledge on aviation demand forecasting for the subsequent chapters.
First, I begin with a brief history of airport expansions and a description of the integral role aviation demand forecasting plays in the airport master planning process. Then, I summarize previous research findings in literature on the types of challenges affecting the aviation demand forecasting process, specifically, demand uncertainty and optimism bias. Next, I survey the airport planning literature for proposed solutions to demand uncertainty and optimism bias. I find that the proposed alternative airport planning frameworks require an overhaul of the existing airport planning process and put in place new planning processes that are characterized as flexible and adaptable. I also find that these alternative approaches remain mostly theoretical due to a lack of proven efficacy and unrealistic assumptions about political support behind the planning processes. I instead argue that research in the systematic evaluation of aviation demand forecast accuracy as well as research in the application of reference class forecasting, a method of “grounding” forecasts by extracting information from other forecasts for similar projects/entities, can produce a much more immediate and meaningful impact on addressing the issue of demand uncertainty and optimism bias in aviation demand forecasting.

2.2 History of Airport Expansions

Airports in the U.S. are mostly legacy airports sited and built in the early and mid-20th century when the prevailing aviation technologies of the time required shorter runways than the ones used today. Runways were at most 50% shorter than those needed to accommodate the wide body aircraft of today (Bednarek, 2001). The subsequent
technological advancements and regulatory reforms brought about seismic changes to the facility requirements of airports (Altshuler & Luberoff, 2003). In the early days of aviation during the 1930s, the nascent aviation technology was changing so unpredictably fast that no meaningful group of experts on the issue of long-term airport planning existed (Barrett & Rose, 1999). Coupled with the widespread sentiment of general excitement for aviation (known as “the Winged Gospel”), the lack of true expertise meant that airports were built with enthusiasm but little regard, justifiably so for the nascent technology, for the future use. For example, the introduction of jet engines forced airports to expand the existing runways or build new ones to accommodate the transformative technical needs of the new jets. It also produced faster travel speeds for an ever-increasing number of passengers and destinations, creating dramatic growths in aviation demand.

Even more dramatic changes ensued after the regulatory shift in the aviation industry. Up until the early 1970s, the US government regulated the airline industry through the Civil Aeronautics Board (CAB) (Bednarek, 2001). This meant that the CAB controlled the market entry and exit of airlines, set airline routes (i.e., where airlines can fly), and determined air fares. With no meaningful competition allowed in the industry, this meant that passenger air travel was a luxury item for the select few who could afford the high air fares. Riding the global policy shift towards deregulation in the 1970s, however, the US government deregulated the airline industry in 1978 (Goetz & Vowles, 2009). The deregulation of airline industry allowed airlines to enter and exit the market freely, set routes, and charge air fares according to their business model, no longer subject to the
government mandate. The market mechanisms brought tremendous benefits to the industry and passengers; air fares became much more economical to a greater number of passengers. Airlines responded to the newfound freedom to optimize their operations by adopting a hub-and-spoke system in which airlines concentrate service on major hub airports and serve a greater number of routes cost effectively (Forsyth, 1998).

The net effect of the deregulation, for major hub airports, has been a general decline in the quality of airport experience due to increased congestion, delays, and longer travel times (Goetz & Vowles, 2009). The problem of congestion and flight delays is not just a matter of inconvenience; it incurs economic costs to airlines, passengers, and the US economy. According to a comprehensive study on flight delay impact by Ball et al. (2010), the total cost estimate of all US air transportation delays in 2007 was $32.9 billion. The total cost includes costs to airlines in the form of increased expenses for crew and fuel and additional costs to passengers in terms of lost time and lost demand (i.e., passengers avoid air travel because of delays). The study also estimates that air travel delays reduced the US GDP by $4 billion in 2007. The issue of airport congestion and flight delays has become the dominating airport planning and management issue since the deregulation of the airline industry.

Historically, airport planners at the federal and local levels address the issue of congestion in two ways: they can manage flight demand through policy/prices or expand their airfield capacity. Demand management involves redistributing flight demand by using pricing or incentives without much additional infrastructure investments. Although
demand management may constitute a cost-effective solution, the adoption is rare due to fears that demand management will hinder economic development (Brueckner, 2003). Yet, there is very little known about the tradeoffs between demand management and capacity (Ryerson & Woodburn, 2014). On the other hand, airport expansion typically involves building a new runway to increase capacity which requires significant investments in terms of both time and money; from planning to completion, a typical runway takes about 10 years and requires millions of dollars (sometimes above a billion dollars). Despite such high costs, airport planners have responded to the issue of congestion overwhelmingly through airport expansions.

In the case of an airport in a dire need of more capacity, expanding the existing airport infrastructure represents the path of least resistance over acquiring new land and building brand new airports (Altshuler & Luberoff, 2003). At the same time, the host cities hold onto the strong notion that airport expansion is a powerful economic development tool (J. K. Brueckner, 2003) despite the lack of conclusive empirical evidence supporting the claim (Mosbah & Ryerson, 2016). In addition, airports have incentives to make unwarranted investments in airport expansions because they are able to pass the costs to the users (i.e., airlines and passengers) (Forsyth, 2007). As a result, the overwhelming majority of Airport Improvement Program (AIP) funds, the federal grant program to airports and a good barometer of where airports concentrate their infrastructure investments, typically go to runway construction or expansion in lieu of other planning activities (Ryerson &
Woodburn, 2014). In the U.S., airport expansions have become the most prominent planning tool to meet the growth in aviation demand and congestion.

2.3 Airport Master Planning and Aviation Demand Forecasting

In this section, I contextualize and deepen the discussion on the current approach to aviation demand forecasting and the challenges that are left unaddressed. I begin by describing the formal planning process for airport expansions and the pivotal role of aviation demand forecasting in the process, both in qualitative sense and also the details of the empirical models used. I then evaluate the aviation demand forecasts of a major US airport to detail how these challenges manifest themselves in practice. I also conduct an exploratory analysis on aviation demand forecast accuracy on historic 10-year forecasts for major airports in the US and contextualize broadly the types of challenges facing the forecasting practice for aviation demand.

The FAA requires all U.S. airports receiving federal grants to produce airport master plans, a blueprint for long-term airport development including capital investments such as runways and terminals (De Neufville et al., 2013). Airport master plans are designed to be “a comprehensive study of an airport” to “meet future aviation demand” while considering “potential environmental and socioeconomic impacts” (US DOT, FAA, 2015). The FAA sets strict guidelines for an Airport Master Planning process, a linear process that typically involves the following key steps (Kwakkel, Walker, & Wijnen, 2008):
- Analyze existing conditions
- Produce an aviation demand forecast
- Determine facility requirements needed to accommodate the forecast demand
- Develop and evaluate several alternatives to meet the facility requirements
- Develop the best alternative into a detailed Master Plan

From the early years of aviation and airport development, accommodating aviation demand became the favored goal of airport planning (Barrett & Rose, 1999) and the formal airport master planning process, as outlined above, embodies this perspective through its heavy reliance on aviation demand forecasting. Airport sponsors (i.e., local governments) prepare aviation demand forecasts, the forecast of future passenger and aircraft demand in terms of passenger boardings and the number of take-offs and landings, respectively, for a variety of planning and budgeting purposes. Primarily, the overwhelming majority of aviation demand forecasts are prepared in support of airport master plans (ACRP Synthesis 2, 2007). In this sense, forecasting has become a decision-making process in its own right for airport planning because forecasting is used in systematic exploration and selection of goals and plans (Ascher, 1979).

2.3.1 Aviation Demand Forecasting Methodologies

There are largely four types of forecasting methods used for airport demand forecasting found in the literature; market share forecasting, time series model forecasting, simulation,
and econometric model forecasting (ACRP Synthesis 2, 2007). Market share forecasting measures airport traffic as a share of a larger aggregate measure and assumes the relationship to extend into the future. A time series model is a relatively simple method in which the existing data trend is extrapolated into the future. Simulation provides more disaggregate information such as how a passenger might travel through an airport terminal.

The most widely used forecasting method for aviation demand forecasts is econometric modeling (ACRP Synthesis 2, 2007). Econometric modeling involves statistical estimation of a regression model that assumes a relationship between dependent and independent variables. In its simplest form (and many airports adopt this form), the relationship between the dependent variable (e.g., aviation demand) and the independent variables (e.g., socioeconomic and airport’s operational metrics) is assumed to be linear. For any number of independent variables, the relationship is typically written as:

\[ Y = \alpha + \beta X + \epsilon \]

where:

- \( Y \) is the dependent variable
- \( X \) is a set of \( n \) independent variables (\( X = \{X_1, X_2, \ldots, X_n\} \))
- \( \alpha \) is a intercept or a constant term
- \( \beta \) is a set of the coefficients describing relationship between \( X \) and \( Y \) (\( \beta = \{\beta_1, \beta_2, \ldots, \beta_n\} \))
- \( \epsilon \) is a random error term (assumed to have a mean of zero with constant variance).
The forecaster first collects historic data of socioeconomic and airport’s operational variables ($X$) that are shown or assumed to be related to aviation demand ($Y$) and fits the model (1) to estimate the intercept ($\hat{\alpha}$) and the coefficients ($\hat{\beta}$). Using the estimates of the intercept and the coefficients, the forecaster can then use a new set of observed or simulated/forecasted data ($X'$) and let the model (1) estimate what the future aviation demand will be ($Y'$).

For instance, the City of Austin Aviation Department (hereafter referred to as the City of Austin) used econometric modeling for their aviation demand forecasts for Austin-Bergstrom International Airport (AUS) in the 2003 Master Plan Update (City of Austin Aviation Department, 2003). The City of Austin assumed that there is a relationship between aviation demand and population, per capital personal income, and the average cost of air travel. For this Master Plan Update, the City of Austin produced the aviation demand forecasts in 2000 (base year) for the next 5, 10, and 20 years. More formally, the model takes on a mathematical expression in the following equation:

$$Y = \alpha + \beta_1 Population + \beta_2 Income + \beta_3 Cost + \epsilon$$

Once the relationship between the dependent and independent variables is established using historic data (i.e., $\hat{\beta}$’s are estimated), the forecasters used future estimates of independent variables (5-,10-, 20-year projections for population, income, and air travel cost) to estimate the future aviation demand. This rather parsimonious model represents the typical forecasting model used in aviation demand forecasting. The use of
such parsimonious models in aviation demand forecasting can be partially attributed to the
decision-makers’ preference for easily implementable models. In Yokum and Armstrong’s
(1995) telling study on the opinions of forecasting experts about preferred features in
forecasting models, decision-makers rated “ease-of-use” criteria higher than the other
groups that included researchers, practitioners, and educators.

2.4 Challenges to Forecasting
The current approach to forecasting, i.e., econometric modeling, is a classic example of
what Ascher (1979) calls the “insider’s” approach. In the insider’s approach, the relevant
concerns are limited to “the basic scientific information and techniques at the forecaster’s
disposal” (Ascher, 1979). In other words, the primary elements of evaluation become “the
adequacy of the data, the a priori validity of the assumptions, the biases that are likely to
stem from the formal characteristics of the techniques, and the context of the trends
themselves” (Ascher, 1979).

For example, embedded in the econometric modeling process is a set of assumptions
about the relationship between the dependent and independent variables as well as the
presumed trajectories of the independent variables (i.e., what the socioeconomic and
operational metrics will be in the future). This is both the strength and the weakness of the
econometric models. You can test the underlying assumptions by fitting the model and
evaluating the model parameters and develop a more informed set of assumptions. On the
other hand, because the outcome is determined by its core assumptions (Wachs, 1990),
there is a danger in using a set of incorrect assumptions that may lead to wildly inaccurate estimation of the reality. For example, incorrect assumptions of exogenous inputs or independent variables, can result in over- or under-estimation of the dependent variable (Pickrell, 1992). Furthermore, assumptions do not hold well for long-term forecasts (Ascher, 1979); as the forecast target year gets farther into the future, there are more uncertainties and risks that play into the dynamics of the forecasts.

A typical approach to addressing the forecast uncertainty in aviation demand forecasting is through the use of scenarios. Under this approach, the forecaster uses scenarios or different assumptions of how the future will play out in order to adjust the magnitudes of the input variables and evaluate the corresponding levels of the forecast. For example, the City of Austin attempted to account for “the uncertainty associated with a twenty-year planning horizon” by producing forecasts for High, Medium, and Low Growth scenarios (City of Austin Aviation Department, 2003). They are three forecasts with different growth assumptions about the underlying socioeconomic factors. In the High Growth scenario, for instance, the assumption is that population and/or income will have relatively high growth. This is the assumption the City of Austin selected on the basis that the Greater Austin region had shown “strong local economy, continued population growth, and high per capita income” (City of Austin Aviation Department, 2003).

Yet there are largely two types of challenges that are particularly detrimental to the current approach to aviation demand forecasting. First, since the deregulation of the airline industry, changes in the industry have been very dynamic and unpredictable. Especially, in
the post-2000 era, a number of airlines merged with other airlines and consolidated their services, resulting in tenuous relationships between airlines and airports. Airports that used to be home to a major airline, for example, may have lost the airline altogether and may be vying for new air service (Ryerson, 2016a). Second, the prevalent view of airports as regional economic development engines as well as the federal funding incentives based on the forecasts of aviation demand biases the aviation demand forecasts to be optimistic. There are indeed clear and measurable economic impacts from airports (Brueckner, 2003; Green, 2007) as well as intangible benefits such as civic pride (Ryerson & Woodburn, 2014). However, the urban boosterism, i.e., promoting a city or a region, has become one of the dominating drives in airport planning. Ryerson and Woodburn (2014) show that airport EIS documents “put significant focus on growing operations to preserve their hub status”. In addition, Airport Improvement Program (AIP) funding, which provides funding for a significant portion of capital investment projects for airports, is granted to airports based on the aviation demand forecasts (FAA, 2015). Because airports are competing with each other for the limited AIP funding, it may create incentives for airports to inflate their forecasts.

Uncertainty induced by the dynamic and volatile conditions in the airline industry and optimism bias in aviation demand forecasts, left unaddressed, become the main sources of forecasting error. I now take a closer look at each of these challenges.
2.4.1 Dynamic Changes in the 21st Century

There are inherent uncertainties to forecasting that are amplified by the deregulation of airline industry in the 1970s and to a greater extent, by the airline mergers and consolidations in the post-2000 era. Following the deregulation, airlines and airports became highly susceptible to changing economic and market conditions; the number of airlines increased dramatically, routes were expanded, fares declined, and airlines adopted hub-and-spoke system in which airlines concentrated service on key hub airports (Moore, 1986). In an effort to reduce cost and increase revenue in the face of volatile fuel costs and economic downturns, airlines merged and consolidated their hub operations in the 2000s. The adoption of hub-and-spoke system and airline mergers and consolidations brought uneven changes across airports in the U.S. (Fuellhart, Ooms, Derudder, & O’Connor, 2016), propping some airports as “fortress hubs” where a single airline controls a vast majority of the market while other airports encountered reduced levels of service (Goetz & Vowles, 2009).

These dynamic changes pose immediate and existential challenges to many airports. Airline mergers, for example, directly impact the well-being of airport; Redondi, et al. (2012) found that airports that lost the hub airlines due to mergers and consolidations did not recover the lost traffic within five years. Lambert-St. Louis International Airport (STL), one of the “de-hubbed” airports, is sitting on a rarely-used $1.1 billion third runway after passenger traffic declined by more than a half between 2000 and 2004 because a major airline declared bankruptcy (ACRP Report 76, 2012). Goetz and Szyliowicz (1997) also
show that the new Denver International Airport (DIA) experienced a protracted planning and construction period and incurred excessive costs after its intended hub airlines merged and consolidated service.

Against such demand uncertainty, transportation planning and demand forecasting is inadequately based on notions of habit and stability (Marsden & Docherty, 2013). As evidenced by the City of Austin’s practice (City of Austin Aviation Department, 2003), the common practice in aviation demand forecasting with regards to addressing the issue of uncertainty is to produce alternative “high”, “medium”, and “low” point estimates. There are however a number of issues with this approach. First, this approach ignores any model-related statistical uncertainty (ACRP Synthesis 2, 2007). In other words, there may be underlying structural assumptions about the model that are inaccurate and simply changing the levels of the inputs via the alternate scenarios and assumptions merely perpetuates the structural flaws in each scenario (Draper, 1995). Second, the typical use of the point estimates as inputs in econometric modeling compounds uncertainty because point estimates can be highly biased (Zhao & Kockelman, 2002). If the input variables such as population and income projections are highly biased, the resulting aviation demand forecasting model will only propagate the problem of uncertainty in the forecasts.

2.4.2 Systematic Optimism

Another challenge facing aviation demand forecasting is optimism bias. The prevailing mindset in the aviation industry is that the growth of the aviation industry is indisputable
around the world under the underlying assumptions of economic growth and global consumer capitalism (May & Hill, 2006). At the same time, local governments tend to overgeneralize the indisputable benefits of air service to the local economy (e.g., direct and indirect creation of jobs and facilitation of business travel) (Cidell, 2015) and believe that providing higher quantity of air service via airport expansions will increase the economic benefits (Mosbah & Ryerson, 2016).

The urban boosterism and unwarranted investments into airports has a long history since the inception of airport development in the US. As the fledgling air transportation system was beginning to gain prominence as a serious form of transportation in the 1920s and 1930s, cities began to pitch themselves as the center of the national airport transportation network (Bednarek, 2000). For example, in their 1930 regional airport plan, the City of Philadelphia envisioned an air transportation system in which Philadelphia played a central role along with its outlying system of airports feeding and distributing air traffic into and out of Philadelphia (Bednarek, 2000). Other cities and metropolitan areas followed suit and vied to be the centers of the air transportation system. Airport construction frenzy reached its peak in the postwar era of the 1940s underscored by the “build-it-and-they-will-come” mentality (Barrett & Rose, 1999). Consequently, cities built airports out of all proportion to the volume of estimated traffic in the hopes that they can bring air traffic and prominence to their airports and cities. This mentality has largely persisted to this day (Ryerson & Woodburn, 2014).
At the same time, the federal funding mechanism may provide perverse incentives for airports to inflate their forecasts. The federal government is strongly committed to the maintenance, development, and expansion of the National Airspace System (NAS) and, more so than any other modes of transportation, the federal government plays an outsized role in regulation of aviation (US DOT, 2009). As an entity in charge of the NAS, Federal Aviation Administration (FAA) funds airport capital projects through their Airport Improvement Program (AIP) designed to improve, upgrade, and expand infrastructure in support of the NAS. The funding allocation is based on the estimated future activity at airport (i.e., aviation demand forecasts) (FAA, 2017b). Yet local governments have the most detailed knowledge of their infrastructure needs and this information asymmetry can incentivize airport sponsors to “overstate future activity demand” (ACRP Synthesis 2, 2007).

Towards alleviating the symptoms of optimism bias, FAA also produces the official forecasts of aviation demand called the Terminal Area Forecast (TAF) for airports in the US and requires that the forecasts prepared by airport sponsors remain within certain ranges of the TAFs. Specifically, FAA requires 5-year forecasts to be within 10% range of the TAF and 10-year forecasts within 15% (FAA, 2008). However, FAA gives airport sponsors leeway to negotiate (i.e., work with the FAA to update the TAF) in the case their forecasts fall outside the ranges of the TAF (FAA, 2008).

The type of inflationary incentives in the federal airport funding structure is by no means unique to airport planning. Pickrell (1992) found that local officials in eight US
cities showed bias towards high-capital transit investments due to the structure of the federal transit grant programs that levies “financial risk of forecasting errors” to the federal treasury rather than local government. In other words, because the federal government provides a majority of the funding for transit projects and local governments prepare projections of the ridership to justify the investments, local governments have incentives to inflate their forecasts and the federal government ends up taking on the risk.

2.4.3 Evaluation of Forecast Accuracy in the Era of Demand Uncertainty and Optimism Bias

In the face of demand uncertainty and optimism bias, the aviation demand forecasts tend to perform very poorly. Table 2-1 shows the passenger demand forecasts for Austin-Bergstrom International Airport (AUS) for the 5- and 10-year forecast horizons under the preferred scenario of High Growth for AUS as well as the actual passenger volumes in the respective target years.

**Table 2-1 High Growth Forecast vs. Actual Passenger Volumes**

<table>
<thead>
<tr>
<th>Year</th>
<th>Forecast (High Growth)</th>
<th>Actual*</th>
<th>Forecast Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>5,497,000</td>
<td>3,645,956</td>
<td>51%</td>
</tr>
<tr>
<td>2010</td>
<td>6,623,000</td>
<td>4,201,136</td>
<td>58%</td>
</tr>
</tbody>
</table>

* Source: (FAA, 2015b)
The likely interaction between demand uncertainty in the planning environment and optimism bias provides two interesting observations. First, the length of forecast horizon seemed to have had very little impact on the forecast accuracy. Generally, the longer the forecast horizon, the less accurate forecast becomes (Button, 2014). Under this framework, one would expect a substantially smaller forecast error in the 2005 forecast than that of the 2010 forecast. Second, the magnitudes of the forecast errors for both years are substantially large. Both forecasts overestimated by more than half of the actual traffic for respective years.

Because the forecasters assumed the High Growth scenario was the most likely scenario (City of Austin Aviation Department, 2003), there is a possibility that the observed errors in Table 2-1 could be mainly due to the assumptions behind the scenario. However, the evaluation of the forecast errors under the Low-Growth scenario (Table 2-2) indicates that the problem extends beyond the scenario selection.

Table 2-2 Low Growth Forecast vs. Actual Passenger Volumes for AUS (Base Year 2000)

<table>
<thead>
<tr>
<th>Year</th>
<th>Forecast (Low Growth)</th>
<th>Actual*</th>
<th>Forecast Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>4,855,000</td>
<td>3,645,956</td>
<td>33%</td>
</tr>
<tr>
<td>2010</td>
<td>5,535,000</td>
<td>4,201,136</td>
<td>32%</td>
</tr>
</tbody>
</table>

* Source: (FAA, 2015b)
The Low Growth forecasts, which are the most conservative forecasts that were deemed by forecasters as unrealistically low, still overestimated by more than 30% for both forecast years. Short and long forecast horizons also did not impact the forecast accuracy. Borrowing Kain (1990)’s more revealing terminologies, “optimistic, very optimistic, and incredibly optimistic” seem to be more appropriate terms for the Low, Medium, and High Growth forecasts.

The likely culprits behind these inaccurate forecasts, i.e., demand uncertainty and optimism bias, are not limited to a few cases similar to AUS but are systematically pervasive and pose challenges to the aviation demand forecasting and airport planning processes. Since Pickrell (1992)’s work on optimism bias in ridership forecasts in transit projects, there seems to be some evidence that ridership forecast accuracy has improved over time (Button, Doh, Hardy, Yuan, & Zhou, 2010). However, there is very little evidence this is the case for aviation demand forecasts. Figure 2-1 shows annual boxplots of the 10-year TAF forecast errors from the base year 1995 to the base year 2005 for the top 64 busiest airports. These boxplots indicate that the aviation demand forecasts generally overestimated (i.e., positive forecast errors) and the overestimations in fact became worse over the years until about 2001 when the economy was reeling from a recession caused by the “dot-com bubble”. This indicates that the optimistic assumptions about the future seemed to be only temporarily tempered by the macro economic trends.
2.5 Alternative Airport Planning Approaches

Against the backdrop of the systematic optimism and demand uncertainty in the aviation demand forecasts, the planning community has increasingly challenged the proposition of unquestioningly accommodating the projected demand growth and called for more comprehensive evaluations of the benefits of aviation growth contrasted with sustainable solutions to the demand for mobility (May & Hill, 2006; Freestone & Baker, 2011; Ryerson & Woodburn, 2014; Ryerson, 2016b). In reality, there is very little to suggest that airport planners are responding to these call-to-action challenges in any meaningful way in airport expansions. Ryerson and Woodburn (2014) found that among the 17 airports with runway projects between 2000 and 2013, only one considered broader factors beyond
accommodating growth through expansion in their Environmental Impact Statements (EIS), the federally required document evaluating the environmental impacts of the proposed actions (e.g., expanding a runway). May and Hill (2006) also argue that no rational argument about the environment and the welfare of residents seemed to have carried any real weight in decision-making about airport expansions.

On the other hand, a group of airport planning scholars argue that the problem of demand uncertainty in aviation demand forecasting and airport master planning can be alleviated by using a different planning/decision-making process altogether. They challenge the “predict and provide” framework used in the airport master planning process and offer more flexible and dynamic planning frameworks. Specifically, these “alternative” airport planning frameworks include Dynamic Strategic Planning (DSP), Flexible Strategic Planning (FSP), Adaptive-Policy Making (APM), and Adaptive Airport Strategic Planning (AASP). Kwakkel et al. (2010) provides detailed summaries of each of these alternative planning frameworks; I summarize them here to the extent that it is useful to my discussion of demand uncertainty.

**Dynamic Strategic Planning**

De Neufville (2000) introduced the concept of Dynamic Strategic Planning (DSP) in the context of developing rational policies for large-scale engineering projects. De Neufville first used the case of low-emission automobiles in the U.S. and later applied the concept directly to airport planning (De Neufville & Odoni, 2003). DSP takes more incremental,
“wait-and-see”, approach in which planners make commitment to a first stage of the plan and then reevaluate the development plans in the subsequent stages (Neufville, 2000). DSP deals with the issue of demand uncertainty primarily through flexibility created by real options (Neufville, 2000). A real option is a right to take a future action when opportunities arise; an example would be reserving land use for future development if there rises a need for airport expansion (Kwakkel et al. 2010). If the actual demand exceeds the forecasted demand, for example, airport planners can exercise the option of expanding their terminals or runways. While there are reported cases of real options applications to airports (ACRP Report 76, 2012b), Dynamic Strategic Planning (DSP) as a practical planning framework lacks a prescribed and concrete planning process and remains largely theoretical.

**Flexible Strategic Planning**

Building on the concepts of Dynamic Strategic Planning (DSP), mainly its use of real options, Burghouwt (2007) developed Flexible Strategic Planning (FSP) as an alternative to airport master planning. FSP is designed to complement traditional airport master planning by adding two extra dimensions of pro-activity and flexibility (Burghouwt, 2007). Burghouwt (2007) suggests airport planners should shape the future proactively through the use of scenario planning, contingency planning, monitoring, and experimentation. Burghouwt also adopts de Neufville’s real options in injecting flexibility into airport planning. However, as Kwakkel et al. (2010) point out, Burghouwt does not detail how FSP could be applied in practice. By far, Chakraborty, Kaza, Knaap, and Deal (2011)
represent the best attempt at operationalizing the concepts of FSP in non-airport planning settings; their approach, named Robust and Contingent Plans, is applied to hypothetical models for Baltimore-Washington metropolitan region. Although Chakraborty et al. (2011) show that their model yielded more preferable results than traditional scenario planning model, they concede that lack of data and complexity of stakeholder interactions introduce challenges to practical application.

**Adaptive Policy-Making**

Walker, Rahman, and Cave (2001) first introduced the concept of Adaptive Policy-Making (APM) as a general adaptive approach to policy-making in the face of uncertainty. Kwakkel et al. (2010) later suggested the use of APM in airport planning in conjunction with other alternative planning approaches. APM makes adaptation explicit by designating signposts throughout planning and implementation phases, monitoring the changing conditions at each signpost, and adapting policies accordingly (Walker et al., 2001). The main concepts of continual monitoring and adaptation have found their way into other areas of planning as well; for example, Quay (2010) suggests a new planning/governance framework called Anticipatory Governance in response to uncertainty created by climate change using the main concepts from APM. However, as van der Pas, Walker, Marchau, van Wee, and Kwakkel (2013) discovered through APM workshops with stakeholders on operationalizing APM, it “lacks proof of its efficacy” and there is “only very limited insight into how to operationalize the concept”.

37
Adaptive Airport Strategic Planning

By recognizing that the above alternative planning approaches (DSP, FSP, and APM) share similarities and remain theoretical, Kwakkel et al. (2010) attempted to synthesize them into a concrete planning approach for airport planning. Named Adaptive Airport Strategic Planning (AASP), this approach borrows key ideas from the three alternative approaches; specifically, real options, contingency planning, and signposts and monitoring (Kwakkel et al. 2010). The specifics of each stage under AASP are too complex and numerous to list in this chapter. However, AASP approach is detailed enough that Kwakkel, Walker, and Marchau (2012) were able to conduct computational simulation to compare the performance of AASP to traditional airport master planning. The results indicate that AASP performs better than traditional approach in general; yet the authors concede ironically that unforeseen uncertainties can change the results significantly (Kwakkel et al. 2012). Additionally, AASP does not address the very realistic possibility and inevitability of changes in political actors and stakeholders who can introduce discontinuity in the implementation of AASP (Kwakkel et al. 2010).

The alternative airport planning approaches (DSP, FSP, APM, and AASP) are attempts at addressing the issue of demand uncertainty in the planning/decision-making process. Although specific details in each approach are different, they share some similarities such as continual monitoring of external conditions and reevaluating the next steps at various
signposts based on the feedback loop (e.g., go or no-go points). They are designed to respond to dynamic changes and uncertainties in the long planning process adaptively. Yet they mostly remain theoretical mainly because these are untested theoretical frameworks with no empirical evidence supporting the efficiency and efficacy. The adaptive nature of these approaches also underestimates the amount of political will needed to plan and push through a large infrastructure project by naively assuming the political and financial support will be flexible enough to adapt to the mercurial visions of these adaptive approaches (Kwakkel et al., 2010).

2.6 Methodological Improvements to Aviation Demand Forecasting

The alternative airport planning approaches require extensive changes in the current airport planning process and a systematic overhaul of how airport infrastructure is planned and built. Setting aside the fact that there is a lack of proven efficacy (i.e., Do these approaches actually help airport planners achieve better outcomes? What tradeoffs need to be made to implement them and are they worth it?), their adoption (if ever) will likely be a slow and prolonged process. Hanging the collective hope for a better-informed airport planning process on these alternative approaches seems risky and premature. Instead, there seem to be at least two areas in aviation demand forecasting where immediate and meaningful contributions can be made; namely, 1) incorporating demand uncertainty explicitly in the forecasts and 2) improving forecast accuracy by leveraging the information of the systematic optimism (ACRP Synthesis 2, 2007).
First, an area ripe for exploration in the demand uncertainty research is a better understanding of how various socioeconomic and airport’s operational metrics inform aviation demand. While most airport planners and forecasters assume strong relationships between socioeconomic conditions and airport health and use such metrics in their models, they limit the usefulness of the metrics as inputs in a deterministic model that produces point estimates of the future airport activity. The type of demand uncertainty stemming from airline strategy (e.g., airline de-hubbing), for example, can be better addressed by extracting probabilistic insights from these operational and socioeconomic metrics because airline’s strategy and decision-making at airport level is highly dependent on the markets (i.e., cities and airports) they serve (J. K. Brueckner, 2003).

Forecast accuracy on the other hand seems to be a forgotten topic in aviation demand forecasting. Although the issue of optimism bias in demand forecasting has been an actively researched area in transportation planning (Wachs, 1989; Kain, 1990; Pickrell, 1992; Flyvbjerg, Skamris Holm, & Buhl, 2005b; Button et al., 2010), the literature on aviation demand forecast accuracy is very limited (ACRP Synthesis 2, 2007). Maldonado (1990) represents by far the most comprehensive analysis on aviation demand forecasts; his research evaluated aviation demand forecast accuracy for 22 master plans in the FAA New England region and found that forecasts in general perform poorly. However, Maldonado’s research limits the discussion of forecast inaccuracy to the issue of uncertainty and does not (and probably was not designed to) address the issue of optimism bias in the aviation demand forecasts. Instead, Maldonado follows the concepts of Dynamic
Strategic Planning (Neufville, 2000) to address the problem of uncertainty (after all, he was one of de Neufville’s students).

Ascher (1979) provides one of the earliest expressions of bewilderment on the persistence of the “insider’s” approach to forecasting as well as a suggestion for a step towards improvement in forecast accuracy. Ascher (1979) wonders why forecasters adhere to the insider’s approach when “the record of the past forecasts reveals some avoidable biases”. Instead, he suggests an “outsider’s” approach to forecasting in which “one can appraise and adjust current forecasts in light of the known behavioral biases of forecasters” (Ascher, 1979). This mode of the outsider’s approach to forecasting has been seriously considered only recently through the application of reference class forecasting, a method of “grounding” forecasts by extracting information from other forecasts for similar projects/entities, in demand and cost forecasts for transportation projects (Flyvbjerg, 2008). Yet there have been no known research efforts to apply reference class forecasting in aviation demand forecasting. There is a pressing need for a systematic evaluation of forecast accuracy for aviation demand forecasts. A research in this area can provide the foundational knowledge on the pervasiveness and nuances of forecast inaccuracy and help formulate approaches to improving forecast accuracy in aviation demand forecasts. For example, a systematic understanding of forecast accuracy can enable the use of reference class forecasting or any similar “outsider’s” approach to forecasting in aviation demand forecasting.
2.7 Summary

As an integral decision-making process in airport master planning, aviation demand forecasting requires more concerted research efforts to address the problems of demand uncertainty and optimism bias. Short of overhauling the current airport master planning process, a renewed focus on the methodological improvements and contributions to aviation demand forecasting can have much more immediate and meaningful impact on airport planning. In a data-rich industry that is aviation, this line of research can leverage various statistical and forecasting tools and produce robust results. In the following chapters, I conduct statistical analyses to contextualize the dynamic operational and socioeconomic trends and build methodologies that help incorporate demand uncertainty and optimism bias into aviation demand forecasting.
2.8 Chapter Bibliography


43


45


CHAPTER 3. IDENTIFYING TRUE AIRPORT PEERS USING CLUSTER ANALYSIS

3.1 Introduction

One of the key activities of airport planning is peer-group learning, in which airport planners engage in peer-to-peer exchange of information on planning lessons (ACRP Synthesis 46, 2013). In various planning documents, airport planners mention their peer airports to benchmark their own performances against their peers in order to develop plans that would keep the airports competitive against their peers. For example, some of the planning documents for the major airports in Dallas (DFW), Philadelphia (PHL), and San Francisco (SFO) make statements such as “our goal is to rank first in our peer group of large hub airports in the Americas” (DFW, 2016), “Philadelphia’s peers include Miami and Charlotte” (Fitch Ratings, 2017), and “this report benchmarks key performance metrics between SFO and its peers” (City & County of San Francisco, 2015).

While the practice of peer group learning is widespread in the aviation industry, the peer group selection criteria are often either unclear or simplistic. The planning documents that mention peer airports often do not specify how the peer airports were selected. Other times, they use the Federal Aviation Administration (FAA)’s hub designations based on enplanements (or passenger boardings) for peer categories, such as when DFW considers the large hub airports (airports that handled 1% or more of the total US passenger boardings) as their peers. Although San Francisco International Airport (SFO) provides
relatively detailed descriptions of their peer selection methodology in its Airport Services Benchmarking document (City & County of San Francisco, 2015), it still relied on fairly simple metrics of enplanement, percentage of O-D (Origin-Destination) passengers, and number of airlines as the selection criteria.

The problem stemming from using such simple and static metrics as peer identification criteria becomes evident when one considers the disproportionate patterns of passenger growths over the years. **Figure 3-1** shows annual passenger volumes for the top 64 airports in the US from 1990 to 2015 with some of the distinct patterns highlighted. Although the FAA’s hub designation would indicate that the large hub airports such as the airports in Atlanta (ATL), Chicago (ORD), Los Angeles (LAX), and Dallas (DFW) are in the same category, ATL has seen a distinctively unique pattern of tremendous growths in passenger volumes. Likewise, airports that were once considered as peers based on the passenger volumes in the 1990s such as airports in New York (JFK), Charlotte (CLT), St. Louis (STL), and Pittsburgh (PIT), have experienced divergent patterns of growths in passenger volumes particularly after 2000 and may no longer share the same planning lessons among themselves.
The current method of identifying peers, therefore, can either result in the airports learning the wrong lessons from their peers or be used to support aspirational goals such as when the airport planners for Denver International Airport looked to Atlanta (ATL)’s economic and passenger growths in order to support their infrastructure project approval (Wallis, 1992). In the current landscape of airports where disproportionate changes have happened in terms of passenger volumes and in potentially more domains, airport planners need to have more nuanced approaches to peer identification.
In this chapter, I aim to understand the types of dynamic changes that characterize groups of airports together. Specifically, I pose the following research question that can help identify the types of trends that are relevant to peer identification:

- What are the dynamics of socioeconomic and operational changes affecting the airports in the post-2000 era of economic downturns and airline mergers and consolidations?

Towards answering this question, I apply a clustering technique to a set of operational and socioeconomic variables that reflect the changing dynamics in the economy and the airport industry according to literature. In order to capture the changes over time, I build an expanded list of dynamic (i.e., change over time) metrics to group similar airports on these metrics together, which is a significant departure from the current peer-group learning practice that relies heavily on a limited number of static metrics only.

First, I survey the literature on how peer-group learning is done and why the current way is deficient and I motivate the need for a new framework and methodology for peer identification. Second, I lay out my methodology and describe the variables I used for my analysis. Third, I present the findings and provide brief descriptions of how airport planners can benefit from this methodology.

The key contribution of this research is in setting the framework for airport planners to identify true peer airports that have experienced similar patterns of socioeconomic and
operational changes. I find that the use of dynamic variables, i.e., tracking change over time such as percent change in passenger volumes, in addition to static variables (e.g., passenger moved for a given time period) is critical in bringing out more subtle nuances and telling a more complete story about the dynamic changes airports have gone through in the 2000s.

3.2 Peer Group Learning

The FAA approaches airport planning from the national scale and provides funding and guidance to airports; funding is allocated based on the FAA-assigned peer group (which noted above is a function of the number of passengers at that airport in the past year) (FAA, 2015a). While municipal airport planners are more focused on local dynamics, they do engage with the national system when planning in a way that is influenced by these FAA peer group designations. Airport planners will consider the successes, failures, and general actions of airports within a peer group that they themselves define. These peer groups tend to have smaller, more focused membership than the large categories defined by the FAA, and are a function of more criteria: for example, peer groups may be defined by both passengers moved and airports with similar operational qualities (airline hub vs. spoke airport, urban airport vs. suburban airport, etc.). The benefits of peer-group learning are largely two-fold; first, airport planners look at the planned and implemented projects at peer airports for insightful guidance and support for their own projects and plans; second, airports benchmark their performance against that of airports with similar characteristics.
“peers”), against an industry standard, or against an industry “best practice” (ACRP Report 19A, 2011). The benefits of airport performance evaluation and improvement based on peer benchmarking are wide-ranging from airport efficiency improvements to better planning decisions (Sarkis & Talluri, 2004).

Peer-group learning is also one of the most powerful tools for project approval. During the airport expansion planning process, airports study and cite the environmental review documents prepared by other airports during their environmental review processes. Wallis (1992), in discussing the background and details leading to the approval and construction of the new Denver International Airport, quoted the former head of the Denver’s new airport office, as stating “We looked at Dallas and Atlanta and began to realize what a big issue economic development generated by the airport was.” The Director goes on to cite a “dollar-for-dollar” multiplier generated in off-site construction with Atlanta’s expansion; that the airport is the state’s largest economic generator; and that 800 international firms established offices in Atlanta attracted by the international airport. These comparisons were critical, as the author goes on to explain that funding and land for the Denver airport was ultimately approved by residents in a neighboring county that wanted the direct economic development benefits in the form of jobs. Airport Cooperative Research Program (ACRP) details how airports use their peer groups but does not identify how they are identified beyond the traditional FAA-based enplanements criteria and loosely defined similar operational qualities (ACRP Synthesis 46, 2013).
Traditionally, the FAA classifies airports based on the number of passengers carried or enplaned. In short, the categories include “large airports” which carry at least 1% of all annual passenger boardings at U.S. airports; “medium airports” which carry at least 0.25% but less than 1% of annual passenger boardings; “small airports” which carry at least 0.05% but less than 0.25% of annual passenger boardings; and “non-hub primary airports” which carry at least 10,000 passengers but less than 0.05% of annual passenger boardings (FAA, 2017a). Large airports (including airports like those in Atlanta, Chicago, New York, and Los Angeles) typically have large infrastructure and many serve as hubs for major airline, which are airports with concentrated passenger traffic and flight operations as transfer points. Medium airports (such as those in Austin and San Antonio) have many commercial operations and may serve as a focus city for a smaller, domestically-focused airline, but often do not serve as hubs for a major airline. Smaller and non-primary airports are minor airports with few commercial operations. Airport planners, those who are in charge of guiding the development of airport facilities, tend to identify peers in their airport categories.

In a regulated and stable environment of pre-deregulation in the 1970s or even until the pre-2000 era when airline consolidations were less common, peer identification on one or a few operational metrics may have been straightforward and useful because the capacity and operations of the airlines were relatively stable and predictable. Since the deregulation in late 1970s, airlines adopted the hub-and-spoke system in which airlines concentrate passenger traffic and flight operations at hub airports that serve as transfer points to more
destinations (spokes). In the first decade of the 2000s, when airline mergers and high fuel prices led airlines to consolidate their hubbing operations on a few airports, major airports which are situated in the most major cities saw their air service strengthen, while airports in the smaller of the major metropolitan areas have lost significant service (Ryerson & Kim, 2013; Kim, 2016). In the post-2000 era of airline mergers and consolidations, it is possible that airports traditionally considered peers when considering only a single metric such as passengers moved may have been disproportionately affected by changing macroeconomic and industry trends. Airports not typically considered as peers may have experienced similar changes in planning environments and may benefit from benchmarking against one another.

3.3 Methodology

In this descriptive research, I aim to uncover the operational and socioeconomic trends affecting major US airports in the post-2000 era. This endeavor requires a method that can recognize the patterns in the operational and socioeconomic trends over the years. To carry out this task, I choose to conduct a cluster analysis, an exploratory data mining technique that groups a set of objects in such a way that the objects in the same cluster or group share more similar attributes (defined by the input variables) than those in other clusters. For the variables chosen for the cluster analysis (discussed below), I collect the data for 2000 and 2014 (the last available year when this research was conducted) and calculate their percent
changes from 2000 to 2014. I call these variables dynamic variables to indicate that they
are measuring the changes during this period.

3.3.1 Study Airports

I scope the sample to the 64 large, medium, and small airports as defined by the FAA, that
are located within the top 50 metropolitan statistical areas (MSAs) by population. A MSA
is a geographic region that shares close economic ties to the core city with relatively high
population density. In my scoping choice, I reflect the intrinsic ties between airports and
regional economy. These 64 airports also serve a large share of commercial service
activities in the U.S. (more than 85% of U.S. commercial airport enplanements handled in
2016 (FAA, 2016)).

3.3.2 Variable Selection

For these 64 airports, I collect the data that will be used as the input into the cluster analysis.
The selection criteria for the variables are two-fold. First, the variables should be indicators
of dynamic changes that affect airport operations and development. I referenced ACRP
Report 48 (“Impact of Jet Fuel Price Uncertainty on Airport Planning and Development”) and
ACRP Report 76 (“Addressing Uncertainty about Future Airport Activity Levels in
Airport Decision Making”) because of their extensive discussion of the sources of demand
uncertainty in airport planning. Second, data for the variables should be readily available
for all airports included in the analysis. Sparsity of data is an issue especially for the cluster
analysis since the algorithm treats missing data by dropping the entire observation or
airport (i.e. list-wise deletion) from the analysis.
I defined a list of 15 variables and narrowed down the list based on data availability and collinearity among variables. Table 3-1 shows a full list of variables. The strikethroughs indicate variables I dropped from analysis due to data availability.

Table 3-1 Sources and Types of Uncertainty and the Corresponding Metrics

*strikethroughs indicate data unavailability

<table>
<thead>
<tr>
<th>Uncertainty Category</th>
<th>Static Variables</th>
<th>Dynamic Variables</th>
<th>Measure</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic conditions</td>
<td>Population (Pop)</td>
<td>Proportional change in population (pch_Pop)</td>
<td>Population in the MSA where airport is located</td>
<td>Census</td>
</tr>
<tr>
<td></td>
<td>Income (Inc)</td>
<td>Prop. change in income (pch_Inc)</td>
<td>Per capita real income in the MSA where airport is located</td>
<td>Bureau of Economic Analysis (BEA)</td>
</tr>
<tr>
<td></td>
<td>Employment (Emp)</td>
<td>Prop. change in employment (pch_Emp)</td>
<td>Employment in service sectors</td>
<td>Census</td>
</tr>
<tr>
<td></td>
<td>Local GDP (GDP)</td>
<td>Prop. change in local GDP (pch_GDP)</td>
<td>Local GDP in the MSA where airport is located</td>
<td>BEA</td>
</tr>
<tr>
<td>Headquarters</td>
<td></td>
<td>Prop. change in number of headquarters</td>
<td>Number of fortune 500 headquarters in each MSA</td>
<td></td>
</tr>
<tr>
<td>Overall volume of traffic</td>
<td>Total passengers (Pax)</td>
<td>Prop. change in total passengers (pch_Pax)</td>
<td>Total number of passengers who boarded aircraft at airport</td>
<td>FAA Passenger Boarding and All-Cargo Data</td>
</tr>
<tr>
<td></td>
<td>Total aircraft operations (Ops)</td>
<td>Prop. change in total aircraft operations (pch_Ops)</td>
<td>Total number of takeoffs and landings at airport</td>
<td>FAA Air Traffic Activity System (ATADS)</td>
</tr>
<tr>
<td>Air cargo volume</td>
<td></td>
<td>Prop. change in air cargo volume</td>
<td>Volume of air cargo served at airport</td>
<td></td>
</tr>
<tr>
<td>Mix of traffic</td>
<td>Domestic vs International (Dom)</td>
<td>Prop. change in percentage of domestic departures (pch_Dom)</td>
<td>Proportion of departures performed that is domestic</td>
<td>Bureau of Transportation Statistics (BTS) Domestic Segment Data (T-100)</td>
</tr>
<tr>
<td></td>
<td>OD vs connecting (OD)</td>
<td>Prop. change in percentage of OD flights (pch_OD)</td>
<td>Proportion of passengers who are making OD trips</td>
<td>BTS Origin and Desintation Survey (DB1B)</td>
</tr>
<tr>
<td>Airport capacity</td>
<td>Arrival Acceptance Rate (AAR)</td>
<td>Prop. change in AAR</td>
<td>Number of aircraft able to land at airport in a period</td>
<td></td>
</tr>
<tr>
<td>Airline strategy</td>
<td>Airline Concentration (HHI)</td>
<td>Prop. change in airline concentration (pch_HHI)</td>
<td>HHI (sum of squared market shares) at airport calculated from OAG seats or calculated from BTS data</td>
<td>BTS On-Time Performance Data</td>
</tr>
<tr>
<td></td>
<td>Percentage of seats flown by low-cost airlines</td>
<td>Prop. change in percentage of seats flown by low-cost airlines</td>
<td>Percentage of seats flown by low-cost airlines</td>
<td></td>
</tr>
<tr>
<td>Technology change</td>
<td>Changes in Technology and Fleet Mix ( ssize)</td>
<td>Prop. change in average seat size (pch ssize)</td>
<td>Average seat size at airport (larger aircrafts have lower costs per seat)</td>
<td>BTS T-100</td>
</tr>
</tbody>
</table>
Variable Descriptions

Economic Conditions

Air traffic has historically been correlated with economic conditions (J. K. Brueckner, 2003; Green, 2007; Blonigen & Cristea, 2015). The study period between 2000 and 2014 captures one of the biggest economic downturns in the U.S. history. I therefore consider economic conditions in each of the top 50 MSAs as key variables for defining airport peer groups. The economic variables are population, income, employment in service sectors, local GDP, and number of corporate headquarters. I use service sector employment as opposed to total employment because the literature indicates a stronger systematic relationship between employment in service sector and air passenger demand (Alkaabi & Debbage, 2007; Bilotkach, 2015). Likewise, literature indicates that the location of corporate headquarters has a strong relationship with the availability of direct non-stop flights (Bel & Fageda, 2008).

Overall Volume of Traffic

Dynamic changes in the economy and the airline industry affect both the overall volume of traffic and mix of traffic. I consider the overall volume of traffic at an airport by using total passengers and total aircraft operations. Although passenger volumes and aircraft operations traditionally are considered to be correlated, the practice of capacity discipline by airlines, i.e., decreasing the number of seats on weaker routes and increasing them on
more profitable routes, may make the correlation tenuous (Ryerson & Kim, 2013). The result has been that in some airports fewer operations serve the same or more passengers.

*Mix of Traffic*

The mix of traffic in terms of passengers reflects the types of markets and functions that an airport serves. Airports with a high proportion of connecting passengers as opposed to origin-destination passengers tend to be major hub airports (Ryerson & Kim, 2013) and may have different facility needs than those with mainly origin-destination traffic such as more gates and terminal space to accommodate connecting passengers. Similarly, airports with a high proportion of international traffic tend to be hub airports that act as gateways to international passengers and transfer them to different parts of the country. I include the proportion of departures that are domestic and proportion of passengers who are making origin-destination trips in my list of variables to indicate the mix of traffic.

*Airport Capacity*

Airports are constrained by the number and length of runways as well as terminal space to accommodate passengers and aircraft. One variable/metric for airport capacity is Arrival Acceptance Rate (AAR), a widely-accepted measure of how many aircraft can land at an airport for a certain interval.
**Airline Strategy**

Airlines’ decisions about their service level at an airport greatly influence the vitality of the airport. When a few airlines have a large market share of the airport operations, the impact of the airlines’ decisions is more powerful. A hub airline’s decision to de-hub from an airport, for example, has a long-lasting impact on the airport leaving it with excess capacity and overbuilt infrastructure (Redondi et al., 2012). I measure the level of airline concentration by using the Herfindahl-Hirschman Index (HHI), a frequently applied economic concept that measures the amount of competition among firms in an industry (Liu, Hansen, & Zou, 2013). A higher HHI indicates a higher concentration while a lower HHI means a more even distribution of the market shares among airlines. I also consider the percentage of seats flown by low-cost airlines. Low-cost airlines such as Southwest Airlines have different business models than the legacy airlines (Delta, American, United) and serve different types of passengers (more OD passengers) than legacy airlines.

**Technology Change**

Although there are numerous types of technological changes that affect airports (e.g., NextGen), I focus on aircraft technology specifically via the number of seats per aircraft. Larger aircraft (with more seats) have lower costs per seat (Ryerson & Hansen, 2010). I create a proxy for fleet mix by calculating average number of seats per aircraft at each airport.
Data Availability

In my data collection efforts, I discovered that some of the data were not readily available publicly. This is an issue because the cluster analysis algorithm ($k$-means) used for this research performs poorly with missing data. If an observation has a missing data in one variable, $k$-means drops the entire observation from analysis (list-wise deletion).

The variables that were difficult to collect without compromising the number of observations include airport capacity (Arrival Acceptance Rate), corporate headquarters concentration, air cargo volumes, share of domestic passengers, and the percentage of seats flown by low-cost carriers. Corporate headquarters location data were not readily available for both 2000 and 2014 while not every airport handles cargo. Arrival Acceptance Rate (AAR) is no longer available publicly through the FAA’s ASPM database.

Selected Variables

After narrowing down the list of variables based on data availability, I am left with 7 dynamic variables. Summary statistics for each of the selected variables are provided in Table 3-2.
This study scope takes a significant departure from the current peer-group learning practice by expanding on the metrics. Instead of using a single static variable, e.g., number of enplanements in a year, I use 7 dynamic operational and socioeconomic variables that track the changes over the 14-year period between 2000 and 2014.

### 3.3.3 K-means Clustering Algorithm

I use the $k$-means clustering algorithm to group similar airports among the 64 airports using the selected variables. $K$-means takes its name from the fact that the user pre-determines the $k$ number of clusters or groups and the algorithm assigns membership to each of the observations based on their Euclidean distance to the nearest mean or a cluster center. $K$-means is a widely-used clustering method in both transportation (Golob & Recker, 2004)
and aviation research (Ryerson & Kim, 2013; Woodburn & Ryerson, 2013). I use the gap statistics (Tibshirani, Walther, & Hastie, 2001) to determine the number of clusters \((k)\) to use for the algorithm. The gap statistics computes a goodness of clustering measure for increasing numbers of clusters and selects the number of cluster that maximizes this statistic. Because there is no guarantee that the naturally observed data would have clustering patterns, the gap statistics is used to evaluate statistically if and how many clusters exist in the data. The gap statistics for my data indicates that 11 clusters \((k=11)\) are optimal.

3.4 Discussion of Results

I ran the k-means clustering algorithm using the data of 64 airports with the 7 operational and socioeconomic variables and grouped the airports into 11 clusters. In two-dimensional data (i.e., only two variables), the clustering patterns can be visualized in scatter plots. For my data, however, each observation or airport is represented in 7 dimensions (7 variables) and cannot be visualized in such high dimensions. Instead, I use Principal Component Analysis which maps the data points into two-dimensional space in order to visually inspect the quality of the clustering (Figure 3-2). The plot indicates that the k-means algorithm resulted in good separations of the data, particularly for the clusters 2, 3, 5, and 11.
The summary statistics for the 11 clusters in Table 3-3 show the airports in each cluster with their traditional FAA hub categories as well as their cluster center means or averages for each of the variables. Cluster 2, which shows the most distinct pattern of separation from the rest of the airports (Figure 3-2), includes the medium and small hub airports in St. Louis (STL), Pittsburgh (PIT), Cincinnati (CVG), and Memphis (MEM). This makes sense because these are the airports that have experienced the event of de-hubbing during the study period and therefore, would have shown tremendous changes in the operational and socioeconomic metrics.
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Airport</th>
<th>Hub</th>
<th>Population</th>
<th>Service Sector</th>
<th>Passengers</th>
<th>Operations</th>
<th>O-D</th>
<th>Passengers</th>
<th>HHI</th>
<th>Avg. Number of Seats per Aircraft</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IND</td>
<td>M</td>
<td>28.14%</td>
<td>19.82%</td>
<td>-11.72%</td>
<td>-41.40%</td>
<td>11.43%</td>
<td>0.97%</td>
<td>-21.05%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CMI</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PBI</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BDL</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FLL</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GNT</td>
<td>M</td>
<td>-1.91%</td>
<td>6.63%</td>
<td>-16.44%</td>
<td>-55.64%</td>
<td>15.02%</td>
<td>24.16%</td>
<td>-3.02%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BUR</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PVD</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORF</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BHM</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>STL</td>
<td>M</td>
<td>8.64%</td>
<td>9.09%</td>
<td>-62.39%</td>
<td>-61.74%</td>
<td>192.82%</td>
<td>-65.67%</td>
<td>-17.00%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPT</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVG</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MEM</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>OAK</td>
<td>M</td>
<td>-4.16%</td>
<td>77.98%</td>
<td>16.47%</td>
<td>-8.30%</td>
<td>7.71%</td>
<td>-45.30%</td>
<td>-22.74%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MCI</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SJC</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CLE</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MKE</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MHT</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>ORD</td>
<td>L</td>
<td>24.99%</td>
<td>8.85%</td>
<td>49.25%</td>
<td>5.45%</td>
<td>13.03%</td>
<td>-23.04%</td>
<td>-14.09%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHL</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LGA</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLC</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>LAX</td>
<td>L</td>
<td>42.33%</td>
<td>37.34%</td>
<td>28.02%</td>
<td>-19.34%</td>
<td>23.04%</td>
<td>22.88%</td>
<td>-6.61%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JFK</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DEN</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SFO</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CLT</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAH</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>DFW</td>
<td>L</td>
<td>22.75%</td>
<td>-22.74%</td>
<td>21.69%</td>
<td>-24.35%</td>
<td>7.58%</td>
<td>-33.52%</td>
<td>-20.69%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHX</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MIA</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BNA</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>MSP</td>
<td>L</td>
<td>19.70%</td>
<td>22.48%</td>
<td>24.49%</td>
<td>-22.19%</td>
<td>1.24%</td>
<td>53.13%</td>
<td>3.39%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IAD</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DCA</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORC</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RIC</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SDF</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SEA</td>
<td>L</td>
<td>-6.59%</td>
<td>17.24%</td>
<td>11.41%</td>
<td>-27.59%</td>
<td>12.17%</td>
<td>-8.09%</td>
<td>-9.41%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BWI</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDW</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAN</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TPA</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HOU</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SNA</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SMF</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>EWR</td>
<td>L</td>
<td>42.85%</td>
<td>82.27%</td>
<td>43.01%</td>
<td>-28.26%</td>
<td>20.76%</td>
<td>-6.18%</td>
<td>-2.69%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BOS</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PDX</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHL</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RDU</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BUF</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>AUS</td>
<td>M</td>
<td>36.53%</td>
<td>7.19%</td>
<td>25.34%</td>
<td>-9.94%</td>
<td>-15.25%</td>
<td>3.34%</td>
<td>-8.70%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DAL</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAT</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PSP</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ATL</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The relative magnitudes of these changes for the de-hubbed airports in the cluster 2 become more evident when these cluster means are visually represented for each of the variables. Figure 3-3 shows 7 bar plots for each of the 7 operational and socioeconomic data. Each bar plot shows the cluster mean for each of the 11 clusters ordered from 1 to 11. Cluster 2, the second bar from the left in each of the bar plots, indeed shows distinct patterns from the rest of the clusters. On average, the host cities for the de-hubbed airports in this cluster had very modest rates of growth in population (8%) and service sector employment (9%). At the same time, on average, the airports in this cluster experienced the highest rate of loss in passenger volumes (-62%) and aircraft operations (-63%) while the proportion of the O-D passengers almost tripled, conversely meaning that the proportion of connecting passengers declined tremendously.
Figure 3-3 Bar Plots of Cluster Means ($k = 11$)
Another cluster that showed a relatively good pattern of separation was cluster 5. Cluster 5 includes large hub airports located in major metro areas in Los Angeles (LAX), New York (JFK), Denver (DEN), San Francisco (SFO), Charlotte (CLT), and Houston (IAH). The distinguishing features of this cluster are the changes in passengers and aircraft operations. On average, these airports gained about 50% in passenger volumes and, while all other clusters have seen reduced numbers of aircraft operations on average, these airports instead gained 5% in aircraft operations.

Lastly, cluster 11 contains only one airport, the airport in Atlanta (ATL). Because the k-means algorithm is sensitive to outliers, the fact that ATL became its own cluster indicates that the operational and socioeconomic changes characterizing ATL are not shared by any other airport.

In the following section, I use two case studies, Hartsfield-Jackson Atlanta International Airport (ATL) and Lambert-St. Louis International Airport (STL), to illustrate the types of insights gained from this research in terms of peer identification.

### 3.4.1 Case Studies

**Hartsfield-Jackson Atlanta International Airport (ATL)**

Hartsfield-Jackson Atlanta International Airport (ATL) is the world’s busiest airport by passenger traffic. It is one of the 30 large hub airports in the US according to the FAA’s hub classification, and is the major connecting hub for Delta Airlines, one of the three major network airlines in the U.S. By the FAA’s hub classification, ATL has quite a few options in terms of peers. In fact, many of the large hub airports (as defined by FAA) often
consider themselves as peers of ATL. Denver International Airport (DEN), as mentioned before, cites ATL as one of its peers in their various airport performance reports (Dennis J. Gallagher, 2014). Based on the FAA hub classification and ranking, the sixth ranked DEN is right to consider the top ranked ATL as their peer; they are both large hub airports with substantial amounts of passenger traffic. But such generalization also masks obvious and subtle differences between DEN and ATL.

My cluster analysis indicates that ATL may be in the league of its own. As the world’s busiest airport, ATL handles a tremendous amount of connecting and international traffic. Consequently, ATL is the only cluster that gained the connecting share of passengers (+15%) when all other clusters lost the connecting shares. In addition, while all 64 airports on average lost about 30% in aircraft operations between 2000 and 2014, ATL fared much better with the smallest decline in operations of only 5%.

Such dynamic changes that define ATL are unparalleled among the 64 airports. They show clearly that the types and mixture of traffic that ATL handles are quite different from those of DEN, for example. That is not to say that DEN can never be in a peer group with ATL in the future. It only highlights the fact that ATL and DEN have experienced different types of changes over the years and any peer group learning between them should reflect such stark differences. Unparalleled changes at ATL and in the Atlanta region indicate that ATL may be operating with a drastically different business model and/or target markets; other airports such as DEN that want to learn lessons from ATL as their peer should do so with care and the understanding that the results might not be directly transferrable.
**Lambert-St. Louis International Airport (STL)**

My cluster analysis yields beneficial results especially for airports that have gone through substantial changes over time. Lambert-St. Louis International Airport (STL) is an airport that experienced a tremendous change in passenger volumes during my study period. In the 1980s, STL experienced a substantial growth in airport activity due to strong traffic growth driven by the hub airline Trans World Airlines (TWA). In response, a 1994 airport master plan update for STL proposed that a third runway should be built to accommodate the future aviation demand. The aviation demand forecast for STL projected the enplaned passengers to grow to 20 million by 2006 from 13 million in 1994. However, financial difficulties and mergers resulted in the collapse of TWA and withdrawal of a hub airline from STL in 2001. Consequently, the enplaned passengers at STL declined by more than half between 2000 and 2004. Due to delays in the planning process as well as the already established forward momentum for the project, the construction of the proposed third runway was allowed to proceed and was completed in 2006 at a cost of $1.1 billion. However, STL never recovered the lost traffic and the third runway is heavily underused while sections of the airport are partially closed (ACRP Report 76, 2012).

The plight of STL is not a shared experience for most of airports in its FAA hub category of medium hub airports except for those airports that have similarly lost their hub airlines (i.e, PIT, CVG, and MEM). Airports that are near STL (ranked 31) in the FAA rankings of medium hub airports such as Houston Hobby Airport (HOU, ranked 32) and
Austin-Bergstrom International Airport (AUS, ranked 34), for example, both instead gained tremendous amounts of passenger traffic between 2000 and 2014 (+40% and +52%, respectively). Single static point-in-time metric of passenger traffic used in the FAA hub classification does not tease out this essential bit of information about STL that has significant implications on how STL needs to plan. One of the practical implications for STL is that they can look to their peers identified in the cluster analysis (PIT, CVG, and MEM) and study how their peers have addressed the issue of the lost passenger volumes and excess capacity. They have the shared goal of luring air service in order to recapture the lost traffic and may be able to learn from one another’s experience more so than with any other airports in their hub category.

3.5 Summary

Airport planners traditionally have identified their airport peers based on the FAA’s hub designations. However, I demonstrated that such peer groups may not reflect the dynamic changes affecting airports. With the percent changes in operational and socioeconomic variables as the clustering variables, a different and more nuanced picture emerges. This highlights an important lesson that identifying and understanding relevant trends matter in how airports identify their peers and benchmark their performance and help improve their plans. Static measures such as number of operations will group airports with similar numbers of operations. Dynamic measures such as proportional change in number of operations between two points in time, on the other hand, cluster airports based on their
upward or downward trajectories along this particular dimension. I showed that the traditional hub designations matter less when one considers airports using dynamic measures.

The clustering scheme I employed in this analysis is by no means the only way to consider airport peers. It does, however, challenge the conventional notion that airports of similar sizes are necessarily comparable. Dynamic measures and socioeconomic characteristics provide more informative metrics to airport planners, especially in the post-2000 era of airline mergers and consolidations. Peer benchmarking becomes more fruitful if an airport can learn from its peers that have experienced similar changes and can offer practical lessons on how to deal with such changes. A challenge and opportunity for airport planners is to explore a variety of factors that impact airport planning and operations and incorporate them into peer identification and benchmarking. Planners need to remain flexible and adaptable to volatile changes through identifying and monitoring metrics.

Flexibility and adaptability are increasingly becoming critical elements in airport planning. Uncertainties ranging from volatile macroeconomic conditions to man-made disasters and climate change leave less and less room for static planning. With vast sums of money and time invested in airports, airport planners wield great power in planning and building infrastructures that have lasting economic and environmental impacts. Airport planners need to embrace flexibility and adaptability in order to direct and use this power most effectively in service of airports, air transportation system, and larger communities that they serve.
3.6 Chapter Bibliography


FAA. (2016). Passenger Boarding (Enplanement) and All-Cargo Data for U.S. Airports - Previous Years. Retrieved August 22, 2016, from


CHAPTER 4. PREDICTING THE PROBABILITY OF A SEVERE CONTRACTION IN PASSENGER VOLUMES

4.1 Introduction

Airport planners and aviation industry experts operate under the predominant assumption that economic growth and global consumer capitalism will continue to propel the aviation industry and increase the passenger volumes around the world (May & Hill, 2006). This assumption underlies the practice of aviation demand forecasting; Federal Aviation Administration (FAA) prepares the official aviation demand forecasts (Terminal Area Forecast or TAF) for the commercial airports in the US “independent of the ability of the airport and the air traffic control system to furnish the capacity required to meet demand” (FAA, 2017). In other words, the TAFs are the demand-driven forecasts of passenger volumes without regard for supply constraints.

The dominant assumption of growth in airport planning is problematic especially when significant infrastructure investment decisions are made based on this tenuous assumption. The 10-year forecasted growths and forecast errors of the TAFs in Figure 4-1 show that the growth assumption may not best reflect the reality of demand growths and could instead jeopardize the airport planning process. As expected, most of these 10-year forecasts of passenger volumes for the 64 study airports from 1995 to 2005 projected overwhelmingly positive growths in the 10 years, indicated by a high concentration of points to the right of the vertical line at 0. This observation by itself is not a cause for alarm;
however, the fact that the majority of these forecasts produced substantially large and positive forecast errors (i.e., the actual passenger volumes were lower, sometimes significantly lower, than the forecasts), is a reason for concern because the long-term airport plans are based on these forecasts.

**Figure 4-1** 10-Year Forecasted Growths and Forecast Errors of Terminal Area Forecasts (64 Airports, 1995-2005)

This systematic pattern of high forecasted growth and high forecast error has very real and wildly varying impacts on airport infrastructure investment decisions and their outcomes. For example, cities such as St. Louis and Pittsburgh experienced a severe contraction in passenger volumes at their major airports, never recovering the same levels of passenger demand subsequently, in the midst of expanding them based on the optimistic forecasts.
Now their newly built infrastructures are left rarely used and these airports are spending even more money to attract and recapture the lost demand (Ryerson, 2016a). On the other hand, there are airports such as ones in Miami, FL (MIA) and San Francisco, CA (SFO) whose passenger demand, although optimistically forecasted, eventually caught up to the level forecasted. The distinction between these demand patterns is critical to airport planners; airports that have a steady, stable growth and eventually meet the forecasted demand will be able to justify (*post facto*) heavy investments in expansion while those with contracting demand may see their investments (typically in the hundreds of millions of dollars) wasted.

Airport planners and forecasters recognize this problem and attempt to address this demand uncertainty in their forecasts by employing the High, Medium, and Low assumptions about the underlying conditions of the predictors. Yet, as the evaluation of Austin’s aviation demand forecasts in their 2003 master plan update shows (City of Austin Aviation Department, 2003), this approach perpetuates the statistical/model errors into each scenario. Furthermore, forecasters and decision-makers inevitably must choose a scenario they prefer (in the case of Austin, they went with the High scenario) and make decisions based on this forecast, essentially providing only cursory treatment to the demand uncertainty.
I argue that the current practice of aviation demand forecasting is inadequate against the uncertainty of a severe contraction in passenger volumes. Because the baseline assumption for airport planning is growth in passenger volumes, detecting the signs of a severe future passenger contraction is an important insight into airport planning that may prevent wasteful investments. In this chapter, I pose the following question:

- What are the operational and socioeconomic characteristics of an airport on the verge of experiencing a severe contraction in passenger volumes?

To answer this question, I build a binary logistic regression model to predict the probability of an airport experiencing a severe contraction in passenger volumes in the next 10 years. This model allows me to identify the predictors (i.e., the operational and socioeconomic variables) that are highly informative in assessing this probability. This insight carries a significant relevance to airport planners especially in their airport master planning and aviation demand forecasting processes for runway expansions, as it could help them reconsider unwise investment decisions.

I begin first by conducting an exploratory data analysis on aviation demand forecast accuracy and demand growth in order to contextualize the problem. Next, I briefly summarize the existing data-driven procedures for incorporating the demand uncertainty
in aviation demand forecasting and their limitations. Then I introduce my methodology and expound on the procedure. Finally, I summarize and discuss the results.

4.2 Beyond Forecast Accuracy

In this section, I show through an exploratory data analysis that aviation demand forecasts tend to overestimate and that there are different patterns of demand over time which result in disproportionate types of impact from forecast errors on decision-making.

In the literature, the limited number of available research on aviation demand forecast accuracy shows aviation demand forecasts are overwhelmingly inaccurate. For instance, the Airport Cooperative Research Program (ACRP), an industry-driven research program, produced a report (ACRP Report 76, 2012) that evaluated a number of forecasts and concluded they were wildly inaccurate. Maldonado (1990) provides a more detailed and nuanced evaluation of aviation demand forecast accuracies for the forecasts used in the master plans for airports in the FAA New England region and likewise concludes that these forecasts were highly inaccurate. At the same time, Ryerson and Kim (2013) suggest that airports have experienced drastically different and disproportionate changes in the 21st century due to changes in the economy, fluctuations in the fuel price, and airline mergers. Beyond the well-established notion that the aviation demand forecasts are inaccurate, these dynamic changes and demand uncertainty require more nuanced understanding of forecast accuracy.
Towards generating this foundational knowledge about the relationship between forecast accuracy and demand patterns, I use the publicly available aviation demand forecast data and actual annual demand and evaluate their performance. I use the official aviation demand forecasts of Federal Aviation Administration (FAA) known as the Terminal Area Forecast (TAF). As the federal entity in charge of regulating US air transportation, FAA produces TAFs annually for all US airports to help federal, state, and local authorities plan in regards to airport and air traffic operations (FAA, 2017). Because the TAFs are updated in coordination with local sponsors engaged in the airport master planning process and are readily available online, they serve as good barometers for the forecasts used in airport master plans for airport expansions and other functions.

Figure 4-2 shows the annual passenger demand (boardings) at airports in San Francisco (SFO) and Miami (MIA) from 1995 to 2016 along with the 10-year TAF forecasts with a base year 1995 and a target year 2005 (shown in red). As expected, the 10-year forecasts for both MIA and SFO (red bars) overestimated by a substantial margin. This falls in line with the general pattern of optimistic forecasts uncovered in Figure 4-1. At the same time, the annual demand for both MIA and SFO (grey bars) show a pattern of growth subsequently to the forecast target year (2005).
Based strictly on the measure of forecast accuracy, heavy infrastructure investments at these airports may not be advisable. However, the general growth patterns in demand in the subsequent years may act as *post facto* justification for the investments. In other words, the use of a new runway, for instance, at these airports may be justified by the eventual growth in demand.

On the other hand, there are airports that may not be able to justify infrastructure investments either way. Figure 4-3 also shows the 10-year TAF forecasts (red bars) and annual demand (grey bars) for airports in St. Louis (STL) and Pittsburgh (PIT). Right away, the substantially large margin of error characterizes both forecasts. The forecasts overestimated by more than twice the actual demand in 2005. Additionally, the annual demand for both STL and PIT show a pattern of a severe contraction; the annual demand in 2015 was almost half of the demand in 1995 both airports.
These airports for St. Louis and Pittsburgh, formerly prosperous industrial cities, have lost tremendous amounts of passengers along with population since their peak from decades ago. Airlines’ strategic decisions contributed to the major contractions in passengers at these airports because the airports’ major hub airlines experienced financial difficulties and declared bankruptcies during this period (Redondi et al., 2012). Consider that while the contraction was happening in St. Louis, STL was in the middle of constructing a new runway at a cost of $1.1 billion. The new runway is now largely sitting unused.

The difference between airports like SFO and MIA and those similar to STL and PIT is drastic in terms of the future growth and planning needs. Investment in new runways at the former types of airports may be contested but will likely be necessary and justified given the continual growth in passenger volumes. On the other hand, the same type of investments at airports such as STL and PIT are clearly unwise and wasteful considering
that the passenger trend will never justify the significant investments in runways that last for a long time regardless of whether they are used or not.

4.3 Current Forecasting Methods to Incorporate Demand Uncertainty

In this section, I summarize the currently available data-driven procedures to incorporate the demand uncertainty in forecasts and present their limitations in order to both motivate the use and highlight the novelty of my proposed method. These methods do not treat the problem of a severe contraction in passenger volumes explicitly but they are instead designed to address the inherent demand uncertainty, hereto defined as the stochastic variations in the passenger volumes. In short, they do not consider the possibility of a severe contraction in passenger volumes in any meaning way. Yet, I find the following discussion helpful because it motivates the need for a predictive methodology that can provide additional information to the forecasts.

First, in time-series modeling, prediction intervals are used to recognize the uncertainty associated with model specification. Time-series models typically identify historic patterns in the variable of interest (in our case, aviation demand) and extrapolates them into the future (i.e., forecast target year/period). The output of such time-series models is a point estimate of the value of the variable of interest in the future. Inherently, there is uncertainty associated with this point estimate and a prediction interval provides interval estimates of probability around a range of point estimates (Chatfield, 2001). For example, the prediction interval might indicate that there is 95% a probability that the
passenger demand will be in the range of 15-20 million in the target year of 2025. However, the prediction interval only estimates the uncertainty in the model specification (how the relationship between the dependent and independent variables are functionally described). It does not incorporate the uncertainty due to the dynamic changes in the economy, airline strategy, etc. and most master plans with forecasts using time-series modeling also do not disclose the prediction intervals in their documents (ACRP Report 76, 2012).

Second, Ascher (1979) and Flyvbjerg (2008) argue that incorporating the past empirical errors into the current forecast can help improve the forecast accuracy. In addition to the inside view approach to forecasting in which forecasters use the relationship between passenger volumes and underlying socioeconomic conditions to build a model, they advocate applying the outside view approach where forecasters use historic empirical errors (the difference between the forecasts and actual values) to inform the current forecast (ACRP Report 76, 2012). For instance, if upon evaluating the past forecasts you discover that they were generally off by +/- 35%, then you assume that the true value of your current forecast may lie somewhere in the similar range of +/- 35%. However, as Ascher (1979) wondered aloud, this outside view approach is rarely adopted in practice.

Third, Bhadra and Schaufele (2007) proposed a probabilistic method of incorporating the demand uncertainty into the aviation demand forecasts. In its simplest form, their method uses a distribution of historic annual growth rates and, using a simulation technique known as Monte Carlo, builds the entire distribution of possible growth rates over the forecast horizon, and converts the simulated growth rates into annual
demand forecasts (ACRP Report 76, 2012). While the probabilistic approach provides realistic ranges of forecasted demand, Bhadra and Schaufele (2007) concede that the method makes the interpretation of the outcomes very difficult. For example, the probabilistic information does not provide insights into what caused the fluctuations in the forecasted demand.

The primary drawback to these methods lies in the fact that the relationship between the demand uncertainty arising from the dynamic changes in the planning environment and aviation demand cannot be established through them and therefore, makes it difficult to interpret and use the outcomes in strategic ways for airport planning. On the other hand, my proposed method, as discussed below, predicts the probability of a severe contraction in passenger volumes (the type of demand uncertainty most detrimental to airport master planning) and treats this uncertainty in a functional form that establishes probabilistic relationship between the underlying socioeconomic patterns and the severe contraction in passenger volumes.

4.4 Methodology

Towards understanding the indicators of a severe future contraction in passenger volumes, I build a binary logistic regression model. In this section, I discuss the methodology and the data used for this research.

A binary logistic regression model takes a binary categorical outcome variable, coded as 1 for the event and 0 for non-event, which in this research is the event of whether
an airport experienced a severe contraction in passenger volumes (1) or not (0). It finds the relationship between the binary outcome and the explanatory variables (or predictors) and produces a probability that the given input (e.g., an airport) belongs to a certain class (e.g., the event of a severe contraction in passenger volumes).

In order to build a robust model that can be generalized beyond the data used in the analysis, I need to 1) construct the binary outcome variable from the data and 2) find explanatory variables that are closely related to airport passenger volumes based on the literature. Because there is no clear definition in the literature of what constitutes as a severe contraction in passenger volumes, I use a data mining technique to leverage the information in the data to construct the binary outcome variable. Then I survey the literature to find meaningful operational and socioeconomic explanatory variables.

**Study Airports**

I scope the sample to the 64 large, medium, and small hub airports (as defined by the FAA based on the share of total traffic moved) that are located within the top 50 metropolitan statistical areas (MSAs) by population (Figure 4-4). These airports served about 90% of total passengers in the US in 2016 (FAA, 2016).
4.4.1 Constructing the Binary Outcome Variable

The literature is relatively scarce on the empirical definition of a severe contraction in passenger volumes. The de-hubbing literature, the literature on the situation where an airline with a predominant presence at an airport scales back or discontinues their service (de-hubs) from the airport, for example, tend to use empirical and qualitative definitions of de-hubbing that are used to identify a narrow timeframe of the de-hubbing event itself (Redondi et al., 2012; Tan & Samuel, 2016). Instead, I am looking at the 10-year window which is the average length of time from the planning and completion of a new runway and the changes in passenger volumes during this period.

Given the lack of a clear definition in the literature, I use a data mining technique to let the data inform the binary outcome. The data in this case are the 10-year percent
changes in the passenger volumes for the 64 airports for all base years from 1995 to 2005 (Table 4-1). The mean and the median are 14.39 and 13.11, respectively (i.e., 14.39\% growth in passenger volumes and 13.11\% growth in passenger volumes). However, there seems to be a wide spread in the data as indicated by the standard deviation of 34.91. There are also some extreme values as big as 395.70 and as small as -80.79.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>10-year percent changes in passenger volumes (N=704)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>14.39</td>
</tr>
<tr>
<td>Median</td>
<td>13.11</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>34.91</td>
</tr>
<tr>
<td>Max</td>
<td>395.70</td>
</tr>
<tr>
<td>Min</td>
<td>-80.79</td>
</tr>
</tbody>
</table>

The left histogram in Figure 4-5 shows the distribution of the 10-year percent changes in passenger volumes. The distribution almost seems normal (i.e., Gaussian) but there are a few spikes in the distribution towards the left tail as well as a very long right tail. Based on this observation, I use a Gaussian Mixture Model, which assumes all the data points are generated from a mixture of Gaussian distributions with unknown parameters. This method uses the Expectation-Maximization (EM) algorithm, an iterative algorithm that estimates the parameters of the distributions, and produces the posterior probabilities, the probabilities of each data point generated from a particular distribution. This method essentially allows me to cluster similar data points together by assigning each data point to a distribution with the highest posterior probability.
Figure 4-5 Histograms of 10-Year Percent Changes in Passenger Volumes

The histogram on the right in Figure 4-5 shows the resulting mixture of four Gaussian distributions ($k=4$). The two blue distributions in the middle include the majority of the data points and the flat red distribution encompasses the large positive outliers. The yellow distribution near the left tail includes the data points of interest, the 10-year percent changes that are negative (i.e., the passenger volumes declined) and extreme (i.e., distinctively larger in magnitude than other negative data points).

Based on these distributions, I cluster or group the data points into 3 distinct clusters (Table 4-2). The first cluster only contains 9 data points that have more than 167% growth on average in passenger volumes in the 10-year period (growth cluster). The second cluster contains the majority of the data points (559) that show a moderate average growth of 18.09% (cyclical cluster). The last cluster contains 136 data points that
showed a distinct pattern of negative growth (contraction cluster); in the 10-year period, the passenger volumes for these airports decline by 29% on average with some airport losing more than 50% of their passengers.

**Table 4-2 Summary Statistics for 10-Year Change Clusters**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Growth Cluster (n=9)</th>
<th>Cyclical Cluster (n=559)</th>
<th>Contraction Cluster (n=136)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>167.20</td>
<td>22.40</td>
<td>-28.61</td>
</tr>
<tr>
<td>Median</td>
<td>138.70</td>
<td>18.09</td>
<td>-22.00</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>87.51</td>
<td>23.08</td>
<td>18.26</td>
</tr>
<tr>
<td>Max</td>
<td>395.70</td>
<td>99.15</td>
<td>-11.23</td>
</tr>
<tr>
<td>Min</td>
<td>110.00</td>
<td>-10.74</td>
<td>-80.79</td>
</tr>
</tbody>
</table>

I decide to code the 559 data points in the cyclical cluster as 0 (non-event) and the 136 data points in the contraction cluster as 1 (event) for the total of 695 data points for the binary outcome variable. The data points in the growth cluster are extreme outliers that may skew the data and because there are only 9 of them, I decide to remove them from the data.

### 4.4.2 Identifying the Explanatory Variables

The literature on air travel demand shows that there is intrinsic relationship between airport passenger volumes and airport’s operational characteristics and the sociodemographic characteristics of its host city/region (J. Brueckner, 2003; Alkaabi & Debbage, 2007; Green, 2007; Bel & Fageda, 2008). I use 9 such metrics as the explanatory variables for the base year figures (i.e., point-in-time numbers in base years) as well as the 5-year average annual percentage change (5AAC) in these 9 variables up to 92
the base year in order to capture the trends within the past 5 years. For example, the population of 2 million in the base year 1995 for an MSA is a point-in-time figure while the 5AAC would be 5% (averaged over 2% change during ‘90-‘91, 3% during ‘91-‘92, 3% during ‘92-‘93, 4% during ‘93-‘94, 3% during ‘94-‘95). In total, I start with 18 explanatory variables (9 point-in-time and 9 5AACs).

Table 4-3 Overview of Variables

<table>
<thead>
<tr>
<th>Variables in base year numbers</th>
<th>Unit</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passengers</td>
<td>Persons (millions)</td>
<td>8.42</td>
<td>7.95</td>
<td>FAA</td>
</tr>
<tr>
<td>Airport competition</td>
<td>Unitless</td>
<td>3.74</td>
<td>5.57</td>
<td>FAA</td>
</tr>
<tr>
<td>Connecting passenger share</td>
<td>Proportion</td>
<td>0.47</td>
<td>0.11</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>Avg. number of seats per aircraft</td>
<td>Seats</td>
<td>118.40</td>
<td>26.87</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Avg. ticket price</td>
<td>Dollars</td>
<td>227.70</td>
<td>53.01</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>HHI</td>
<td>Unitless</td>
<td>0.35</td>
<td>0.20</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Population</td>
<td>Persons (millions)</td>
<td>3.56</td>
<td>3.44</td>
<td>Census</td>
</tr>
<tr>
<td>Per capita income</td>
<td>Dollars (thousands)</td>
<td>45.87</td>
<td>7.91</td>
<td>BEA</td>
</tr>
<tr>
<td>Service sector employment</td>
<td>Persons (millions)</td>
<td>0.92</td>
<td>0.91</td>
<td>Census</td>
</tr>
</tbody>
</table>

5-year avg. annual % change up to base year

<table>
<thead>
<tr>
<th>Variables in base year numbers</th>
<th>% change</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passengers (5AAC)</td>
<td>% change</td>
<td>4.00</td>
<td>7.76</td>
<td>FAA</td>
</tr>
<tr>
<td>Airport competition (5AAC)</td>
<td>% change</td>
<td>1.97</td>
<td>3.49</td>
<td>FAA</td>
</tr>
<tr>
<td>Connecting passenger share (5AAC)</td>
<td>% change</td>
<td>0.53</td>
<td>4.51</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>Avg. number of seats per aircraft (5AAC)</td>
<td>% change</td>
<td>-1.68</td>
<td>3.02</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Avg. ticket price (5AAC)</td>
<td>% change</td>
<td>-2.94</td>
<td>3.39</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>HHI (5AAC)</td>
<td>% change</td>
<td>0.44</td>
<td>7.52</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Population (5AAC)</td>
<td>% change</td>
<td>1.12</td>
<td>0.90</td>
<td>Census</td>
</tr>
<tr>
<td>Per capita income (5AAC)</td>
<td>% change</td>
<td>1.77</td>
<td>1.27</td>
<td>BEA</td>
</tr>
<tr>
<td>Service sector employment (5AAC)</td>
<td>% change</td>
<td>4.76</td>
<td>3.45</td>
<td>Census</td>
</tr>
</tbody>
</table>

Descriptions of the Explanatory Variables

Passenger enplanements (Passengers)

In typical aviation demand forecasts, the trend of historic passenger enplanements becomes an essential piece of the model. Because my outcome variable tracks a severe contraction in passenger volumes, essentially representing a disruption in the past trends,
the prior enplanement trends may not serve as a good predictor. However, the literature also shows that the de-hubbed airports served relatively large numbers of passengers prior to the de-hubbing (Redondi et al., 2012).

**Airport competition**

Large metropolitan areas, including Boston, Chicago, Los Angeles, and New York, are typically served by several nearby airports. In these multiple-airport regions, the main airport faces competition from the secondary airports which may offer different types of accessibility, fee charges, and quality of service to compete for passenger traffic. Passengers may also consider tradeoffs between the quality of service offered at competing airports and distance to travel to the airports (Johnson, Hess, & Matthews, 2014). The literature, however, does not provide a consistent empirical definition of airport competition. I decide to create a metric to estimate the magnitude of airport competition; for airport $a_i$, I sum the annual passenger enplanements (in 100,000s) for the neighboring airports within 100 miles divided by their distances from the airport $a_i$. This is to reflect the fact that the neighboring airports serving larger numbers of passengers may pose as greater competitions and that passengers are more likely to consider using the neighboring airports if they are located closer to the airport $a_i$. 
Airport competition \( (AC_{ai}) \) = \( \sum_{n} \left( \frac{E_{ni}}{100000} \right) \frac{1}{d_{na_i}} \)

where \( n \in N \) (set of airports within 100 miles from airport \( a_i \)),

\[ E_{ni} = \text{annual passenger enplanements at airport } a_{ni} \]

\[ d_{na_i} = \text{distance in miles of airport } a_{ni} \text{ from airport } a_{i} \]

**Connecting passenger share**

The key characteristic of a hub airport is that the air traffic is concentrated to foster connections (Redondi et al., 2012). In fact, the airport connectivity is a critical feature in the hub-and-spoke system used in the US and global aviation network, a system that allows airlines to consolidate passengers into the hub airports and distribute them to their final destinations (Paleari, Redondi, & Malighetti, 2010). Airports with high connectivity tend to serve more passengers as a result of this funneling of the passengers model. I estimate the connecting portion of passengers at each airport by using the 10% survey data (DB1B Coupon) from the Bureau of Transportation Statistics (BTS), which details the itinerary information of passengers including origin, destination, and any connections in-between.

**Average number of seats per aircraft**

Fleet mix of aircraft at an airport indicate types of destinations and travel demand at the airport. The literature indicates that small aircraft serving fewer than 100 passengers tend
to be used to provide high frequency service (Givoni & Rietveld, 2009). The high frequency routes are often the busy domestic markets. An inference can be made that the routes using large aircraft are long-distance routes including the international destinations. I estimate the mix of aircraft at an airport by taking the average number of seats per aircraft at each airport as a proxy.

Average ticket price

The literature indicates that the low-cost airlines (such as JetBlue and Southwest) offer lower ticket prices than the larger legacy airlines (e.g., United, American, etc) (Gillen & Lall, 2004). On the other hand, the higher ticket prices may indicate that an airport is largely served by the legacy airlines and/or it serves more passengers travelling long-distances as these tickets tend to cost more.

HHI

The airlines’ decisions about their service level at an airport greatly influence the vitality of the airport. When a few airlines have a large share of airport operations, the impacts of the airlines’ decisions are more powerful than if a large number of airlines use the airport. A hub airline’s decision to de-hub from an airport, for example, has a long-lasting impact on the airport, leaving it with excess capacity and overbuilt infrastructure (Redondi et al., 2012). Airline concentration is measured here by using the Herfindahl–Hirschman index (HHI), a frequently applied economic concept that measures the amount of market
concentration among firms in an industry (Bamberger & Carlton, 2018). HHI is computed as a sum of squared market shares of companies, which in this analysis are airlines at an airport

\[ HH_{a_i} = \sum_l m_{li}^2 \]

where \( m_{li} \) is market share of an airline \( l \) at airport \( a_i \) as estimated by proportion of seats provided by airline \( l \) over total seats by all airlines. A higher HHI indicates a higher concentration, while a lower HHI means greater competition among airlines.

Population, per capita income, and service sector employment

Air traffic has historically been correlated with economic conditions (J. Brueckner, 2003). The study period between 1995 and 2015 captured one of the biggest economic downturns in U.S. history. Therefore, I considered economic conditions in each of the top 50 MSAs as key variables for predicting contraction in demand. The economic variables are population, income, employment in service sectors. Service sector employment as opposed to total employment was used because the literature indicated a stronger systematic relationship between employment in that sector and air passenger volumes (Alkaabi & Debbage, 2007).

Collinearity

For estimating the coefficients in the binary logistic regression model, correlated explanatory variables could result in biased estimates. Before proceeding to fit the model,
I first inspect the correlations among the explanatory variables (Figure 4-6), with darker shades indicating higher correlations. In my data, there is a strong correlation between Population and Service Sector Employment and I decide to remove the Population variable.

**Figure 4-6 Correlation Matrix**
4.4.3 Building a Binary Logistic Regression Model

A binary logistic regression estimates the probability that an event will occur given the values of the explanatory variables. For the binary outcome variable $Y$ and a set of explanatory variables $X = (X_1, X_2, ..., X_k)$, the model takes on the following expression:

\[
\pi_i = \Pr(Y_i = 1|X_i = x_i) = \frac{\exp(\beta x)}{1 - \exp(\beta x)}
\]

\[
\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}
\]

Once the parameters ($\beta$) are estimated, the coefficients are typically interpreted in exponents ($\exp(\beta_1)$), i.e., odds ratio. An odds ratio is a relative measure of effect; if the odds ratio is greater than 1, a unit increase in the explanatory variable increases the odds of the event happening and if the odds ratio is less than 1, a unit increase in the explanatory variable decrease the odds of the event.

Model Selection

I fit the binary logistic regression model to my dataset of 695 observations with 1 binary outcome variable and 17 explanatory variables (the population variable is removed due to the high correlation with the service sector employment variable). Because the explanatory variables are in different units, I first standardized the explanatory variables before fitting the model.
I built the model using backward selection, a variable selection approach under which I start with fitting a model using all of the explanatory variables and remove the least statistically significant variable and continue refitting the model. For each of these models, I use a relative goodness-of-fit measure, AIC, to choose the “best” model. AIC is negative log-likelihood penalized for the number of parameters, in other words, it penalizes the model for including too many explanatory variables and prevents overfitting, a situation where the model explains the nuances in the data very well but performs poorly when applied to other unseen data. As I remove the least statistically significant variable from the model, the AIC value decreases until it starts to increase. This is the model I choose to select as the “best” model.

One of the assumptions of a binary logistic regression is that there is no autocorrelation in the data, that is, each observation is not serially correlated to one another. Certainly, this is not the case for my data because each observation is a delayed copy of another row. The 10-year change for Philadelphia International Airport (PHL) for 1995-2005 is, for example, a delayed copy of the 10-year change for PHL for 1996-2006. Before considering some of the statistical techniques to correct for the auto-correlation, I fit the model using the full data and then compared it to the model using a random sample of the full data, essentially removing auto-correlation from the data. The directions and the magnitudes of the estimates were similar for these models and I concluded that the impact of auto-correlation is minimal in answering my research question. Therefore, the final model uses the full data of 695 observations.
The summary of the binary logistic regression model output is presented in Table 4-4. I only report the odds ratios (exp(\(\beta\))) instead of log-odds (\(\beta\)) for the ease of interpretation.

### Table 4-4 Summary Table for the Binary Logistic Regression Model

|                             | Odds ratio | \(P > |z|\) |
|-----------------------------|------------|----------|
| (Intercept)                 | 0.1200     | 0.000*** |
| Airport competition % change (5AAC) | 0.6121     | 0.000*** |
| Connecting passenger share  | 1.5547     | 0.000*** |
| Connecting passenger share % change (5AAC) | 0.9652     | 0.005**  |
| Avg. number of seats per aircraft | 0.7087     | 0.000*** |
| Avg. ticket price           | 0.6123     | 0.000*** |
| HHI                         | 2.2339     | 0.004**  |
| HHI % change (5AAC)         | 1.3456     | 0.003**  |
| Population % change (5AAC)  | 0.2010     | 0.000*** |
| Per capita income           | 1.5385     | 0.001**  |
| Service sector employment   | 0.4056     | 0.001**  |

\(n = 695\)

\(\text{AIC} = 422.66\)

\(\ast p < 0.1 \quad \ast\ast p < 0.01 \quad \ast\ast\ast p < 0.001\)

### Interpreting the Odds Ratios

All the estimated coefficients in the selected model (Table 4-4) are statistically significant at the 5% significance level. As all the explanatory variables were standardized before fitting the model, unit for each selected variable reported in Table 4-4 is one standard-deviation. The following discussion of each of the selected variables assumes that all other variables are held constant (\textit{ceteris paribus}).

101
Airport competition 5-year avg. annual % change (5AAC) (odds ratio = 0.6121)

One unit increase in the 5AAC of airport competition reduces the likelihood of experiencing a severe contraction in passenger volumes by almost a half. In other words, if the airlines in the neighboring airports have been offering more seats or services in the past 5 years, the airport of interest is less likely to experience a severe contraction in passenger volumes. The airlines’ decision to provide more service in the region may indicate that the regional as a whole is a growing market and airports in this region are less likely to experience a sudden disruption in the passenger trends.

Connecting passenger share (odds ratio = 1.5547) and 5-year avg. annual % change (5AAC) (odds ratio = 0.9652)

My model indicates that one unit increase in the connecting passenger shares will increase the likelihood of experiencing a severe contraction in passenger demand by a factor of 1.5. This falls in line with the fact that the airports with high connectivity are typically the hub airports and the hub airports have historically experienced a sudden disruption in passenger volumes due to de-hubbing. On the other hand, one unit increase in the 5-year average annual % change (5AAC) in the connecting passenger shares slightly reduces the likelihood of a severe contraction in passenger volumes. The gain in the connectivity can be interpreted as gaining more passengers in general because the connectivity indicates that the airlines are pooling passengers at the airport.
*Average number of seats per aircraft* (odds ratio = 0.7087)

One unit increase in the average number of seats per aircraft at an airport reduces the likelihood of a severe contraction in passenger volumes by a factor of 0.71. The literature indicates that the small aircraft size is related to the high frequency routes (i.e., short domestic routes) (Wei & Hansen, 2005) and conversely, the large aircraft size can indicate longer routes including international destinations. My model indicates that airports with larger aircraft potentially serving international routes are less likely to experience a severe contraction in passenger volumes.

*Average ticket price* (odds ratio = 0.6123)

One unit increase in the average ticket price decreases the likelihood of a severe contraction in passenger volumes by a factor of 0.61. This is in line with the findings just discussed above. Higher ticket prices often indicate longer routes including international destinations. In addition, the literature indicates that the higher ticket prices may be explained by the mix of leisure and business passengers (Lee & Luengo-Prado, 2005).

*HHI* (odds ratio = 2.2339) and *5-year avg. annual % change (5AAC)* (odds ratio = 1.3456)

One unit increase in HHI, a measure of market concentration, increases the likelihood of a severe contraction in passenger volumes by a factor of 2.2. This result finds support from many case studies in which the dominant airline with a large share of the market at the airport discontinues their service at the airport resulting in a sharp contraction in passenger volumes.
volumes. Similarly, one unit increase in the 5-year average annual % change (5AAC) in HHI also increases the likelihood by a factor of 1.3. This indicates that as fewer and fewer airlines start gaining larger shares of the market at the airport, the airport is more likely to experience a severe contraction in passenger volumes.

*Population 5-year avg. annual % change (5AAC) (odds ratio = 0.2010)*

One unit increase in the 5-year average annual % change (5AAC) in the MSA population reduces the likelihood of a severe contraction in passenger volumes by a factor of 0.20. In other words, as the MSA gains more population over the years, the airports in the region are less likely to experience a sudden disruption in passenger volumes. This result supports the existing literature that links air travel demand and the socio-demographic conditions of the airport’s host cities/regions (J. Brueckner, 2003).

*Per capita income (odds ratio = 1.5385)*

One unit increase in the per capita income increases the odds of a severe demand contraction by a factor of 1.5. This result is somewhat counterintuitive as it indicates that airports in the MSAs with higher per capita income are more likely to experience a sudden disruption in passenger volumes. This may be a case where the MSAs with lower per capita income tend to host hub airports and by definition, hub airports are at a greater risk for “de-hubbing”. The literature, however, is not conclusive on this point.
Service sector employment (odds ratio = 0.4056)

One unit increase in the service sector employment reduces the likelihood of a severe contraction in passenger volumes by a factor of 0.41. This result supports the findings in the literature that show a strong positive relationship between employment in service sector and air passenger volumes (Alkaabi & Debbage, 2007). In other words, airports located in the MSAs with strong service sector employment base have stable passenger volumes bolstered by the service sector employment.

4.5 Discussion of Results

I used a binary logistic regression model to identify the operational and socioeconomic characteristics of an airport on the verge of experiencing a severe contraction in passenger volumes. The findings both confirm the existing literature and provide new insights into assessing the health of an airport.

My research indicates that airports that are more likely to experience a severe contraction in passenger volumes are characterized by a high proportion of connecting passengers, few airlines with large market shares, and host MSAs with high per capita income. This characterization largely stems from the fact that a lot of the airports in the data that experienced a severe contraction in passenger volumes were hub airports that lost their major airlines. The most informative characteristic here is the market share of airlines; my model indicates that as fewer and fewer airlines begin to dominate at the airport, their decision to either reduce or pull their service at the airport can have a significant impact on
the overall passenger volumes. This is a relevant and disconcerting issue especially in today’s landscape of the airline industry where a series of consolidations and mergers have left only a handful of major airlines at most of the major airports in the US. In this environment where the host cities are spending millions of dollars to court airlines (Megan S. Ryerson, 2016a) and trumpeting their airports as the engines of economic growth, airports are increasingly at the mercy of the airlines in terms of ensuring stable passenger volumes.

On the other hand, airports that are characterized by the growing neighboring airports, increasing shares of connecting passengers, larger aircraft, higher ticket prices, growing MSA population, and high service sector employment, are less likely to experience a severe contraction in passenger volumes. This characterization can be summed up into two categories – 1) a growing region with a strong service sector employment and growing population and 2) airports serving diverse types of passengers with a good mix of connecting passengers, domestic and international passengers, and leisure and business passengers. The growing region characterization is well documented in the literature as scholars have shown that there is a strong relationship among population, service sector employment, and air traffic (J. K. Brueckner, 2003; Percoco, 2010). The literature also shows that the diverse passenger mix is an important indicator of how robust airport retail businesses perform (Appold & Kasarda, 2006) and my research indicates that the impact of the passenger mix extends beyond the airport retail business to the overall health of the airport.
These results give airport planners greater insights into how to assess the health of their airports in addition to the traditional forecasting techniques. In a planning environment where airport planners operate under the aspirational assumption of growth, this research provides nuanced understanding of the characteristics that define airports that are likely to experience a sudden disruption in the passenger volumes.
4.6 Chapter Bibliography


FAA. (2016). Passenger Boarding (Enplanement) and All-Cargo Data for U.S. Airports - Previous Years. Retrieved August 22, 2016, from


109


CHAPTER 5. GROUNDING OPTIMISTIC FORECASTS: TESTING THE EFFICACY OF REFERENCE CLASS FORECASTING

5.1 Introduction

Airport planners use the 10-year aviation demand forecasts to prepare the airport master plans and plan for infrastructure investments such as building new runways. Because of the relatively long forecast window of 10 years as well as the growth-oriented mindset of airport sponsors, these forecasts are not only inaccurate but they also show a systematic bias to overestimate. The forecast errors of the 704 10-year aviation demand forecasts for the 64 study airports from 1995 to 2005 (Figure 5-1) reveal this bias in a clear way.
The majority (85%) of these forecast errors are positive errors, meaning the forecasts were higher than the actual passenger volumes. Only a small portion (15%) of the forecasts underestimated the passenger volumes. On average, these forecasts overestimated the passenger volumes by almost 40% but there are also several extremely large and positive forecast errors that are skewing this average. The median, which is less affected by these outliers, is still close to 30%. Because these major airports handle hundreds of millions of passengers per year, the forecast error of 30% means that there are millions fewer passengers that are actually using the airports than estimated.
In this environment of systematic optimism, an obvious question is, as put forth by Ascher (1979), why forecasters do not look back at past forecast errors and adjust their own forecasts. This is what Ascher (1979) and Flyvbjerg (2008) refer to as the “outside view”. Most forecasters and planners adopt the inside view where they are primarily concerned about the factors that are shown to be related to the outcome of interest being forecasted. Econometric model, as in the case of the 2003 Master Plan Update for AUS, is an example of the “inside view” model in which the variables that are believed to be in a relationship with passenger demand are carefully selected and put in multivariate regression models (City of Austin Aviation Department, 2003). Instead, the outsider’s approach aims at “the sociology and psychology of the experts who formulate the forecasts” rather than “the scientific validity of forecasting procedures” (Ascher, 1979). In other words, the outside’s approach incorporates information about optimism bias and “grounds” forecasts by removing forecast errors stemming from this systematic bias.

This mode of the outsider’s approach to forecasting has been seriously considered only recently through the application of reference class forecasting, a method of “grounding” forecasts by extracting information on past forecast errors from similar projects/entities, in the area of the demand and cost forecasting for transportation projects (Flyvbjerg, 2008). Yet there have been no known research efforts to apply reference class forecasting to aviation demand forecasts.
In this chapter, I pose the following research questions in order to evaluate the feasibility of applying reference class forecasting to the aviation demand forecasts:

1) Does reference class forecasting produce statistically significant reductions in the forecast errors for the aviation demand forecasts compared to the traditional forecasts?

2) What is the relevant and effective definition of a reference class of the forecast errors?

These two questions are inherently related because the effective implementation of reference class forecasting hinges on the identification of a relevant reference class. Towards addressing these two questions, I develop the following four methodologies through which I construct different sets of reference class forecast errors and test the forecast accuracy of each of these methods compared to that of the original forecasts:

1) Mean Forecast Error (MFE): Use each airport’s own past forecast errors as a reference class

2) Mean Growth-Based Forecast Error (MGBFE): Use the past forecast errors of airports that projected similar level of growth in the passenger volumes as a reference class
3) Mean Peer-Based Forecast Error (MPBFE): Use the past forecast errors of peer airports that experienced similar operational and socioeconomic trends as a reference class

4) Enhanced Mean Peer-Based Forecast Error (EMPBFE): Adjust the past forecast errors in the MPBFE reference class by the predicted probabilities of a dramatic decline

This constitutes, to my understanding, the first attempt to implement reference class forecasting in the field of airport planning. My results indicate that these “grounding” methodologies can potentially improve forecast accuracy but require nuanced approach to implementation. Specifically, my research indicates that the operational and socioeconomic trends can inform airport planners in the selection of a reference class that achieves much more effective and consistent forecast accuracy improvement outcomes.

In the next section, I provide a more detailed overview of reference class forecasting to better contextualize this research. Then, I discuss the four reference class identification methodologies I developed for this research, apply them to FAA’s official 10-year passenger demand forecasts (TAFs), and evaluate their performance. Lastly, I summarize the findings and discuss future research needs.

5.2 Overview of Reference Class Forecasting

The foundational ideas behind reference class forecasting come from the works by Kahneman and Tversky (1977). Their work showed that there is often a type of cognitive
bias in decision-making process and the judgement errors are more systematic and predictable than random. Reference class forecasting builds on this idea and formalizes the process of identifying the predictable cognitive errors in forecasts and removing or “de-biasing” them from the forecasts. It is essentially a formalization of the idea that one needs to learn from the past mistakes, which has been proposed as a simple yet powerful safeguard against optimism bias in forecasting by the likes of Ascher (1979). Specifically, reference class forecasting involves the following three steps (Flyvbjerg et al., 2005):

1) Identification of a relevant reference class of past, similar projects. The class must be broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.

2) Establishing a probability distribution for the selected reference class. This requires access to credible, empirical data for a sufficient number of projects within the reference class to make statistically meaningful conclusions.

3) Comparing the specific project with the reference class distribution, in order to establish the most likely outcome for the specific project.

The efficacy of reference class forecasting was first demonstrated by Lovallo and Kahneman (2003) citing an example of curriculum planning. In this example, a team of teachers were asked to develop a new curriculum for high school and asked to estimate how long the process would take. Their initial estimations for the length of time require to
complete the project ranged from 18 to 30 months. Then one of the team members was asked to recall past project durations involving curriculum planning and concluded a minimum of 7 years. Their project at hand eventually took 8 years to complete, much closer to the forecast using the outsider’s approach (7 years) than the insider’s approach (18-30 months).

It wasn’t until 2005 that reference class forecasting was implemented in practice in the field of planning. In this research that was published in 2008, Flyvbjerg used reference class forecasting for the first time for cost estimates for large transportation infrastructure investments in the UK with generally positive outcomes. His work led the American Planning Association to recommend the use of reference class forecasting for large infrastructure projects (American Planning Association, 2005). Researchers have subsequently applied reference class forecasting to cost estimates of hydroelectric dams (Awojobi & Jenkins, 2016), project management (Batselier & Vanhoucke, 2017), and public school building costs (Bayram & Al-Jibouri, 2017). As of this writing, there is no practical application or research on reference class forecasting in aviation demand forecasts.

I suspect one of the challenges to implementing reference class forecasting in aviation demand forecasting is finding relevant past forecasts. The key task of any serious implementation of reference class forecasting involves identifying a relevant reference class of past forecasts, the forecasts that are truly similar to the forecast of interest in nature and can provide relevant information. A set of irrelevant reference forecasts could
potentially create further forecast errors. This area of inquiry (how to define a reference class) is missing not just in aviation demand forecasting but much of the research on reference class forecasting.

Toward laying the foundational framework for implementing reference class forecasting in aviation demand forecasting, I develop three distinct reference class forecasting methodologies for aviation demand forecasts in the next section.

5.3 Reference Class Forecasting Methodologies for Aviation Demand Forecasting

In this section, I describe my four approaches to reference class forecasting for aviation demand forecasting. These methods differ in their approach to identifying a reference class, a set of relevant past forecasts that could provide useful information for the forecast of interest.

First, I use each airport’s own past empirical forecast errors. For instance, for a 10-year forecast for Philadelphia International Airport (PHL) in 2005, I use the forecast errors (i.e., how much the forecast was off by) of PHL’s past 10-year forecasts (e.g., 10-year forecast in 1995) and calculate the average or Mean Forecast Error (MFE). Then I adjust (i.e., reduce) the current 10-year forecast by the calculated MFE in order to remove the systematic optimism.

Second, in order to reflect the observation that there may exist a correlation between forecasted growth percentage and forecast error, I use the empirical forecast errors of the
past forecasts (of any airport) with forecasted growth percentage that is within a range of
the forecasted growth percentage of interest. In other words, for a given forecasted growth
percentage (e.g., 30% growth for PHL), I find past forecasts of any airport that forecasted
a similar range of growth (e.g., 27.5% - 32.5%) and use the mean of their forecast errors
(Mean Growth-Based Forecast Error or MGBFE) to adjust the forecast of interest.

Third, I use the peer identification methodology developed in CHAPTER 3 to find
airports with similar socioeconomic and airport characteristics and use their past forecast
errors to adjust the current forecast. The main logic behind this approach is that there have
been dynamic socioeconomic changes with disproportionate impacts on airports and
airports that have gone through similar changes in the socioeconomic trends may also share
similar forecast errors. In this approach, I first identify peer groups of airports and calculate
the mean forecast errors for each group (Mean Peer-Based Forecast Error or MPBFE) and
adjust the forecast of interest by its peer group’s MPBFE.

Lastly, I incorporate the predicted probabilities of a severe contraction in passenger
volumes developed in CHAPTER 4 to adjust the MPBFE in the previous method. Because
the predicted probabilities provide the information on how likely it is for an airport to
experience a dramatic drop in the passenger volumes (and thus, a potentially larger forecast
error), I use this additional information to calibrate the MPBFE. I name this approach
Enhanced Peer-Based Forecast Error (EPBFE).

I evaluate the performance of each of these approaches (MFE, MGBFE, MPBFE,
EPBFE) by comparing the forecast errors between the actual forecasts and the adjusted
forecasts (i.e., forecast adjusted by the mean forecast errors from my methods). Specifically, I use the paired Wilcoxon signed rank test to evaluate the null hypothesis that the median of the absolute values of the adjusted forecast errors are less than those of the original forecast errors. By using the absolute values, the rejection of the null hypothesis means that the adjusted forecast errors are closer to zero (i.e., more accurate). I use the Wilcoxon test instead of the t-test for equal means because the Wilcoxon test is a nonparametric test that does not assume normal distributions of the samples and thus suitable for these forecast errors containing several outliers.

I also use the Mean Absolute Percentage Error (MAPE) as an additional measure of forecast accuracy. This is a popular method among forecasters to compare the effectiveness of competing forecast models if the only criteria of evaluation is forecast accuracy. MAPE is calculated as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{X_t - \hat{X}_t}{X_t} \right| \times 100\%$$

where $X_t$ is the observed value at time $t$ and $\hat{X}_t$ is the predicted value at time $t$, with smaller MAPE indicating smaller forecast errors (i.e., higher forecast accuracy). MAPE is a popular measurement of forecast errors because it is scale-independent and easy to implement. One of the major drawbacks of MAPE is that it is undefined if $X_t = 0$, that is, the observed value is zero (Hyndman, 2006); however, this situation is irrelevant for aviation demand forecasting and for this research because all of the passenger demand for the major airports in my dataset is non-zero.
I use the FAA’s official 10-year passenger demand forecasts (Terminal Area Forecast or TAF) from 1995 to 2015 as the forecasts ($\hat{X}_t$) and the enplanement or boardings data (also from FAA) as the actual passenger demand ($X_t$) for the top 64 busiest airports in the top 50 MSAs. Due to the availability of data, I set 2005 as the base year when I assume the forecast of interest is being prepared and evaluate the forecast errors of the 10-year demand forecasts of the target year 2015. This arrangement arises from the fact that I need to have access to past 10-year demand forecasts and because my dataset goes back only as far as 1995, the year 2005 represents the only base year feasible for this research.

5.3.1 Mean Forecast Errors (MFE)

In this approach, I take Ascher (1979)’s advice in its simplest form and use each airport’s own past forecast errors to adjust the current forecast. First, I identify all available historic 10-year demand forecasts and calculate the Mean Forecast Error (MFE) for each airport. That is,

$$MFE_\alpha = \frac{\sum_{t=1}^{n} \left( \frac{A_{at} - F_{at}}{A_{at}} \right)}{n}$$

where $F_{at}$ is the 10-year forecast of demand for airport $\alpha$ ($\alpha \in \{1, 2, ..., 64\}$) and $A_{at}$ is the actual demand in the target year for airport $\alpha$. As mentioned, due to data availability, I only have access to one set of historic 10-year demand forecasts available, namely, 1995 TAFs for target year 2005. Therefore, $t = 1$ for this analysis. I still keep the notation $t$ for the purpose of use in the future research when more data becomes available.
After identifying and calculating $MFE_\alpha$ for all available historic forecasts, I adjust the current 10-year demand forecast (i.e., 2005 TAF for target year 2015) by $MFE_\alpha$,

$$\hat{F}_{MFE_\alpha} = \frac{F_\alpha}{1 + MFE_\alpha}$$

where $\hat{F}_{MFE_\alpha}$ is the MFE-adjusted 10-year demand forecast. Then I recalculate forecast error for the MFE-adjusted forecast.

$$\hat{e}_\alpha = \frac{A_\alpha - \hat{F}_{MFE_\alpha}}{A_\alpha}$$

**MFE Evaluations**

Now I apply this process to the top 64 airports in the top 50 MSA’s in the US and evaluate whether the adjusted forecasts improve forecast accuracy. First, the summary statistics of the MFE-adjusted forecast errors and the actual errors indicate the MFE method may have produced better results on average (Table 5-1). There were significant reductions both in the mean and the median of the forecast errors using the MFE method. However, the MFE method also resulted in a larger proportion of the forecast errors that now underestimate (45% vs 18%) and the mean and the median might not give a full picture of whether the MFE method produced substantial and statistically significant reduction in forecast errors. For example, a large negative forecast error could distort the mean forecast error for the MFE-adjusted forecasts. In addition, the Mean Absolute Percent Error (MAPE) is higher for the MFE-adjusted forecast errors than that for the actual forecast errors, indicating that the forecast accuracy has declined using the MFE method.
Table 5-1 Summary Statistics for MFE-Adjusted Forecast Errors and Actual Forecast Errors

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Actual Forecast Errors</th>
<th>MFE-Adjusted Forecast Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>41.12</td>
<td>14.66</td>
</tr>
<tr>
<td>Median</td>
<td>25.94</td>
<td>2.19</td>
</tr>
<tr>
<td>MAPE</td>
<td>48.48</td>
<td>53.39</td>
</tr>
<tr>
<td>Proportion above 0</td>
<td>82%</td>
<td>55%</td>
</tr>
<tr>
<td>Proportion below 0</td>
<td>18%</td>
<td>45%</td>
</tr>
</tbody>
</table>

Indeed, a paired Wilcoxon signed ranked test (Table 5-2) indicates that the median forecast error for the absolute MFE-adjusted forecasts, abs(median)=29.16, is statistically significantly higher than the median for the absolute actual forecast errors, abs(median)=28.48. In other words, the MFE-adjusted forecasts in fact produced larger forecast errors than the actual forecasts.

Table 5-2 Paired Wilcoxon Signed Rank Test for MFE-adjusted Forecast Errors

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>The median of the differences between the absolute actual errors and the absolute MFE-adjusted errors is greater than 0</td>
<td>0.5584</td>
<td>Accept the null hypothesis</td>
</tr>
</tbody>
</table>

This result is confirmed by looking at the bar plots of the forecast errors. The actual forecast errors (shown in orange bars) are ranked from the largest positive errors to the largest negative errors along with the corresponding MFE-adjusted forecast errors (shown in blue bars). The plot shows that in some cases, the MFE method over-compensated and resulted in significant underestimations while in others it actually increased the forecast errors. While the net effect on the average is a reduction in the errors (as evidenced by the
reduction in the mean), the MFE method performs poorly because it induces extremely large errors. The bar plots in Figure 5-2 are also informative in terms of understanding the Mean Absolute Percent Error (MAPE) in Table 5-1; MAPE measures the average length of the bars. For example, the MAPE of 48.48 for the actual forecast errors means that the average length of the orange bars is 48.48. The relatively higher MAPE of 53.39 for the MFE-adjusted forecasts can be visually confirmed in Figure 5-2 by noting that there are some extreme lengths for the blue bars.

**Figure 5-2** The MFE-adjusted Forecast Errors and the Actual Forecast Errors (n=64)
5.3.2 Mean Growth-Based Forecast Errors (MGBFE)

In this approach that builds on the previous method, I use growth-based forecast errors of any airport to calculate the Mean Growth-Based Forecast Errors (MGBFE) by which I adjust the forecast of interest. The logic behind this method stems from an observation that there may exist some correlation between the forecasted growth percentage (how much growth is forecasted in the next 10 years) and the forecast errors. In Figure 5-3 I plot the forecasted growth percentage and forecast errors (%) for all 10-year demand forecasts from 1995 to 2005 for the 64 airports (n=704) and fit a LOWESS line (i.e., smooth line through the scatter plot). The LOWESS line indicates that there is a hint of a positive correlation between how much one forecasts to grow and how (in)accurate the forecasts are. In other words, it seems that the more growth is forecasted, the larger the forecast error becomes.

Figure 5-3 Forecasts 10-year growth vs. forecast errors
I leverage this information and select a set of past 10-year demand forecasts that share similar forecasted growth percentages and forecast errors. First, I calculate the forecasted percentage growth \( g_\alpha \) for the forecast of interest for airport \( \alpha \). For instance, if the 10-year demand forecast (base year 2005) for Boston Logan International Airport (BOS) forecasts 35 million passengers in 2015 and their base year passenger demand is 25 million, \( g_\alpha \) for the BOS’ forecast is 40%. Then, I identify all available historic 10-years demand forecasts of any of the 64 airports (not just BOS’ forecasts) within a range of 5 percentage points (\( \pm 2.5\% \)) of the forecasted growth percentage \( g_\alpha \) (e.g., 37.5\% - 42.5\% for \( g_\alpha = 40\% \)) and calculate the Mean Growth-Based Forecast Error (MGBFE) in the following manner:

\[
M_{\alpha} = \frac{\sum_{i=1}^{n} \left[ \frac{A_i - F_{\alpha}}{A_i} \right]}{n}
\]

where \( F_{\alpha} \) is the 10-year demand forecast of any airport \( i \) whose forecasted growth percentage lies in the range of \( (g_\alpha \pm 2.5\%) \) and \( A_i \) is the actual passenger demand for airport \( i \).

After identifying and calculating \( M_{\alpha} \) for all available historic forecasts, I adjust the current 10-year demand forecast (i.e., 2005 TAF for target year 2015) by \( M_{\alpha} \),

\[
\hat{F}_{MGBFE_{\alpha}} = \frac{F_\alpha}{1 + M_{\alpha}}
\]
where $f_{MGBFE_\alpha}$ is the MGBFE-adjusted 10-year demand forecast. Then I recalculate forecast error for the MGBFE-adjusted forecast.

$$\hat{e}_\alpha = \frac{A_\alpha - f_{MGBFE_\alpha}}{A_\alpha}$$

### MGBFE Evaluation

I apply this method to the top 64 airports in the top 50 MSAs. Unlike the previous method, MGBFE narrowed the number of airports down to 52 airports because some of the forecast growth percentages for the 10-year demand forecasts could not be matched to any available historic 10-year demand forecasts. That is, $MGBFE_\alpha$ for some airports could not be calculated because there was no historic demand forecast $F_{tg_\alpha}$ whose forecasted growth percentage was outside the range of $(g_\alpha \pm 2.5)$.

As before, the mean and the median for the MGBFE-adjusted forecast errors showed significant reduction (Table 5-3). However, the MGBFE method also resulted in a larger proportion of the forecast errors that now underestimate (40% vs 15%). But in this case, the Mean Absolute Percentage Error (MAPE) is lower for the MGBFE-adjusted forecast errors than that for the actual forecast errors, indicating that the forecast accuracy may have improved using the MGBFE method.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Actual Forecast Errors</th>
<th>MGBFE-Adjusted Forecast Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>39.18</td>
<td>8.85</td>
</tr>
<tr>
<td>Median</td>
<td>26.58</td>
<td>7.53</td>
</tr>
<tr>
<td>MAPE</td>
<td>43.63</td>
<td>30.16</td>
</tr>
<tr>
<td>Proportion above 0</td>
<td>85%</td>
<td>60%</td>
</tr>
<tr>
<td>Proportion below 0</td>
<td>15%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Table 5-3 Summary Statistics for MGBFE-Adjusted Forecast Errors and Actual Forecast Errors
Indeed, the improvement is tested statistically significant. A paired Wilcoxon signed rank test indicates that at the 5% significance level, I can reject the null hypothesis that the median of the differences between the absolute actual errors and the absolute MGBFE-adjusted forecast errors is greater than 0. In other words, the MGBFE-adjusted forecast errors are statistically significantly closer to 0 than the actual forecast errors.

**Table 5-4** Paired Wilcoxon Signed Rank Test for MGBFE-adjusted Forecast Errors

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>The median of the differences between the absolute actual errors and the absolute MGBFE-adjusted errors is greater than 0</td>
<td>0.0243</td>
<td>Reject the null hypothesis</td>
</tr>
</tbody>
</table>

Again, this can be confirmed visually looking at the bar plots (Figure 5-4). The MGBFE-adjusted forecast errors (again shown in blue bars) do not contain any extreme values like the ones for the MFE-adjusted forecast errors and they are closer to zero than the actual forecast errors (shown in orange bars). The MAPE for the MGBFE-adjusted forecast errors (i.e., the average bar length for the blue bars) is 30.16, which is lower than that for the actual forecast errors (43.63).
Figure 5-4 The MGBFE-adjusted Forecast Errors and the Actual Forecast Errors (n=52)

5.3.3 Mean Peer-Based Forecast Errors (MPBFE)

In this method, I test the proposition that airports with similar characteristics (“peers”) may produce similar forecast errors. I follow the findings in CHAPTER 3 that dynamic changes have brought disproportionate changes to airports and likewise assume that forecast errors for these airports will be disproportionately characterized. In this regard, this method adds a layer of nuance to the previous methods.
First, I use the clustering scheme introduced in CHAPTER 3 to find airport peers using both static and dynamic socioeconomic and airport variables. I begin with the full set of variables shown in Table 5-5 with values in 2005 (2000-2005 values for 5-year annual average change or 5AAC) and select the final set of 14 variables after removing highly correlated variables (these variables are crossed out in Table 5-5). For the detailed discussion of this process as well as description of each variable, please see section 3.3.

**Table 5-5 List of Selected Variables for Peer Identification Clustering**

<table>
<thead>
<tr>
<th>Variables in base year numbers</th>
<th>Unit</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passengers</td>
<td>Persons (millions)</td>
<td>FAA</td>
</tr>
<tr>
<td>Airport competition</td>
<td>Unitless</td>
<td>FAA</td>
</tr>
<tr>
<td>Connecting passenger share</td>
<td>Proportion</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>Avg. number of seats per aircraft</td>
<td>Seats</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Avg. ticket price</td>
<td>Dollars</td>
<td>BTS DB1B</td>
</tr>
<tr>
<td>HHI</td>
<td>Unitless</td>
<td>BTS T-100</td>
</tr>
<tr>
<td>Population</td>
<td>Persons (millions)</td>
<td>Census</td>
</tr>
<tr>
<td>Per capita income</td>
<td>Dollars (thousands)</td>
<td>BEA</td>
</tr>
<tr>
<td>Service sector employment</td>
<td>Persons (millions)</td>
<td>Census</td>
</tr>
</tbody>
</table>

5-year avg. annual % change up to base year
- Passengers (5AAC) % change FAA
- Airport competition (5AAC) % change FAA
- Connecting passenger share (5AAC) % change BTS DB1B
- Avg. number of seats per aircraft (5AAC) % change BTS T-100
- Avg. ticket price (5AAC) % change BTS DB1B
- HHI (5AAC) % change BTS T-100
- Population (5AAC) % change Census
- Per capita income (5AAC) % change BEA
- Service sector employment (5AAC) % change Census

I then cluster airports 64 airports in the top 50 MSAs using k-means clustering algorithm with $k = 7$ (7 clusters or groups), informed by gap statistics. Figure 5-5 shows two-dimensional representation of the 7 cluster partitions of the 64 airports.
For each airport, I identify its peers (airports in the same cluster) and use the available historic 10-year demand forecasts of the peer airports as the reference class. Specifically, for each airport $\alpha$ in cluster $c$ ($c \in \{1,2,\ldots,7\}$), I calculate the Mean Peer-Group Forecast Error (MPGFE) using the forecasts in the same cluster $c$:

$$MPGFE_{\alpha} = \frac{\sum_{j=1}^{n} \left( \frac{A_{c_{\alpha}j} - F_{c_{\alpha}j}}{A_{c_{\alpha}j}} \right)}{n}$$

where $F_{c_{\alpha}j}$ is the 10-year demand forecast of any airport $j$ in cluster $c_{\alpha}$ (cluster to which airport $\alpha$ belongs) and $A_{c_{\alpha}j}$ is the actual passenger demand for airport $j$ in cluster $c_{\alpha}$. Then, I adjust the forecast of interest by MPGFE:

$$\hat{F}_{MPGFE_{\alpha}} = \frac{F_{\alpha}}{1 + MPGFE_{\alpha}}$$
MPGFE Evaluation

I apply the MPGFE method to the top 64 airports in the top 50 MSAs. As before, the mean and the median for the MPBFE-adjusted forecast errors showed significant reduction with the median becoming close to 0 (Table 5-6). The MPBFE method also resulted in an even split of overestimation and underestimation (50%-50%) from the actual forecast errors’ positive error bias (83%-17% split). In this case, the Mean Asolute Percentage Error (MAPE) is also lower for the MPGFE-adjusted forecast errors than that for the actual forecast errors, indicating that the forecast accuracy may have improved using the MPBFE method.

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Actual Forecast Errors</th>
<th>MPBFE-Adjusted Forecast Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>43.71</td>
<td>3.64</td>
</tr>
<tr>
<td>Median</td>
<td>27.38</td>
<td>0.10</td>
</tr>
<tr>
<td>MAPE</td>
<td>48.48</td>
<td>31.05</td>
</tr>
<tr>
<td>Proportion above 0</td>
<td>83%</td>
<td>50%</td>
</tr>
<tr>
<td>Proportion below 0</td>
<td>17%</td>
<td>50%</td>
</tr>
</tbody>
</table>

The improvement is tested statistically significant. A paired Wilcoxon signed rank test indicates that at the 5% significance level, I can reject the null hypothesis that the median of the differences between the absolute actual errors and the absolute MPBFE-adjusted forecast errors is greater than 0. In other words, the MPBFE-adjusted forecast errors are statistically significantly closer to 0 than the actual forecast errors.
Table 5-7 Paired Wilcoxon Signed Rank Test for MPBFE-adjusted Forecast Errors

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>p-value</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>The median of the differences between the absolute actual errors and the absolute MPBFE-adjusted errors is greater than 0</td>
<td>0.0001</td>
<td>Reject the null hypothesis</td>
</tr>
</tbody>
</table>

Again, this can be confirmed visually looking at the bar plots (Figure 5-6). While the MGBFE-adjusted forecast errors (again shown in blue bars) contain a few extreme negative errors, most of them are tighter around zero. The MAPE for the MPBFE-adjusted forecast errors (i.e., the average bar length for the blue bars) is 31.05, which is lower than that for the actual forecast errors (48.48). This represents about 36% reduction in the average absolute forecast errors.
5.3.4 Enhanced Mean Peer-Based Forecast Errors (EMPBFE)

In this last method, I leverage the probabilistic information on the risk of a dramatic decline in passenger volumes in CHAPTER 4 in applying reference class forecasting. Specifically, I use the Mean Peer-Based Forecast Errors (MPBFE), which have shown the best improvement in forecast accuracy so far among the tested methods, and adjust them by the predicted probabilities of the decline. The rationale is that the airports with higher probability of experiencing this decline in passenger volumes can use the MPBFE more...
liberally than those airports with lower probabilities. In other words, this Enhanced MPBFE method (EMPBFE) may prevent over-correcting the forecasts for those airports that are not likely to see their passenger volumes decline in the next 10 years. I apply the predicted probabilities in the following way; using the model built in CHAPTER 4 and the predictors for the study airports in the base year 2005, I predict the probability of a dramatic contraction $p_\alpha$ for each study airport $\alpha$. Then I reduce the MPBFE by this predicted probability and recalculate the EMPGE-adjusted forecasts $\hat{F}_{EMPBFE\alpha}$:

$$\hat{F}_{EMPBFE\alpha} = \frac{F_\alpha}{1 + MPBFE_\alpha \times (1 - p_\alpha)}$$

**EMPGE Evaluation**

I apply the EMPBFE method to the top 64 airports in the top 50 MSAs. While the mean and the median for the EMPBFE-adjusted forecast errors declined compared to those for the actual forecast errors (Table 5-8), the magnitudes of reduction in the mean and the median are relatively small compared to the other methods. The MPBFE method, for example, reduced both the mean and the median close to zero (Table 5-7). Likewise, the proportion of negative forecast errors (i.e., underestimations) for the EMPBFE has increased very little (17% to 28%) compared to all other methods (45%, 40%, and 50% for MFE, MGBFE, and MPBFE, respectively). The Mean Absolute Percent Error (MAPE) for the EMPBFE-adjusted forecast errors showed reduction, indicating improvement in forecast accuracy. Together, this indicates that as expected, the EMPBFE method induced relatively conservative corrections.
The improvement in forecast accuracy is tested statistically significant. A paired Wilcoxon signed rank test indicates that at the 5% significance level, I can reject the null hypothesis that the median of the differences between the absolute actual errors and the absolute EMPBFE-adjusted forecast errors is greater than 0. In other words, the EMPBFE-adjusted forecast errors are statistically significantly closer to 0 than the actual forecast errors.

| Table 5-8 Summary Statistics for EMPBFE-Adjusted Forecast Errors and Actual Forecast Errors |
|---|---|---|
| Statistics | Actual Forecast Errors | EMPBFE-Adjusted Forecast Errors |
| Mean | 43.71 | 26.95 |
| Median | 27.38 | 17.28 |
| MAPE | 48.48 | 36.41 |
| Proportion above 0 | 83% | 72% |
| Proportion below 0 | 17% | 28% |

The bar plots of the EMPBFE-adjusted forecast errors tell the above statistics in a very compelling manner. Unlike the other methods, the EMPBFE method produced little to no extreme negative errors. On the other hand, some of the corrections for the large positive forecast errors (the bars in the top portion of the plot) are more conservative than the other methods as well. But this represents the overall improvement in forecast accuracy.
as indicated by the MAPE of 36.41 for the EMPBFE-adjusted forecast errors (i.e., the average bar length for the blue bars), compared to the MAPE of 48.48 for the actual forecast errors. This represents about 25% reduction in the average absolute forecast errors.

Figure 5-7 The EMPBFE-adjusted Forecast Errors and the Actual Forecast Errors (n=64)
5.4 Discussion of Results

In this chapter, I developed the four methods to identify a relevant reference class to implement reference class forecasting for the aviation demand forecasts and answer the following research questions:

1) Does reference class forecasting produce statistically significant reduction in the forecast errors for the aviation demand forecasts compared to the traditional forecasts?

2) What is the relevant and effective definition of a reference class of the forecast errors?

The summary statistics for the four reference class forecasting methodologies in Table 5-10 show that for three of the methods, there is a statistically significant reduction in the forecast errors. The only method that did not show a statistically significant result is also the simplest form of reference class forecasting tested in this chapter. Specifically, the Mean Forecast Error (MFE) method simply took each airport’s own past forecast errors and applied them to adjust the current forecast. On one hand, the MFE method represents the purest form of learning from the airport’s past but it also runs into the same question of whether an airport’s own past trajectories of passenger volumes can be a good indicator of its future passenger volumes, a tenuous assumption in most current aviation demand forecasting techniques. As more nuanced approaches for identifying a relevant reference
class were employed in the MGBFE, MPBFE, and EMPBFE methods, these methods all produced statistically significant and substantial amount of reduction in the forecast errors (Table 5-10).

Table 5-10 Summary Statistics for MFE, MGBFE, MPBFE, and EMPBFE

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Actual (n=64)</th>
<th>Actual (n=52)</th>
<th>MFE (n=64)</th>
<th>MGBFE (n=52)</th>
<th>MPBFE (n=64)</th>
<th>EMPBFE (n=64)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>43.71</td>
<td>39.18</td>
<td>14.66</td>
<td>8.86</td>
<td>3.64</td>
<td>26.95</td>
</tr>
<tr>
<td>Median</td>
<td>27.38</td>
<td>26.58</td>
<td>2.19</td>
<td>7.53</td>
<td>0.10</td>
<td>17.28</td>
</tr>
<tr>
<td>MAPE</td>
<td>48.48</td>
<td>43.63</td>
<td>53.59</td>
<td>30.16</td>
<td>31.05</td>
<td>36.41</td>
</tr>
<tr>
<td>Proportion above 0</td>
<td>83%</td>
<td>85%</td>
<td>55%</td>
<td>60%</td>
<td>50%</td>
<td>72%</td>
</tr>
<tr>
<td>Proportion below 0</td>
<td>17%</td>
<td>15%</td>
<td>45%</td>
<td>40%</td>
<td>50%</td>
<td>28%</td>
</tr>
<tr>
<td>% Change in MAPE</td>
<td></td>
<td></td>
<td>+10.5%</td>
<td>-30.9%</td>
<td>-36.0%</td>
<td>-25%</td>
</tr>
<tr>
<td>Paired Wilcoxon</td>
<td></td>
<td></td>
<td>0.5584</td>
<td>0.0243</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
<tr>
<td>Signed rank test (p-value)</td>
<td>-</td>
<td>-</td>
<td>0.0001</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

The definition of a relevant class for the aviation demand forecast errors depends on the desired goal of error reduction. For example, if the goal of error reduction is simply to minimize the overall forecast errors, the use of the Mean Absolute Forecast Error (MAPE) can distinguish the best method. Using this criterion, it seems that the Mean Peer-Based Forecast Error (MPBFE) produced the most significant reduction in the forecast errors (-36% change in MAPE). If, on the other hand, the goal of error reduction is to minimize the forecast error without over-correcting the forecasts, the probabilistic information incorporated in the Enhanced Peer-Based Forecast Error (EPBFE) method can reduce the forecast errors without over-correcting. This type of a situation may arise if airport planners
would like to ground their forecasts but they also do not want to underestimate the passenger volumes for the fear of providing inadequate amount of infrastructure based on the forecasts.

My research indicates that reference class forecasting could be not only a viable option but an effective one for grounding optimistic aviation demand forecasts. The four methods of reference class identification in this research are by no means the only ways to apply reference class forecasting to aviation demand forecasting. For instance, the use of median instead of the mean of the forecast errors of a reference class, can result in much more conservative corrections on the forecast errors. Because there are several outliers in the data (i.e., extremely large forecast errors), the mean of the forecast errors of a reference class may be skewed while the median would represent a more balanced central tendency. In addition, this research was limited in the scope because of the data availability. As more data becomes available, future researchers can employ even more nuanced approaches to reference class identification by employing a wider variety of metrics.
5.5 Chapter Bibliography


CHAPTER 6. CONCLUSIONS AND IMPLICATIONS

6.1 Summary of Conclusions

In this dissertation, I posed and answered the following research questions in order to produce valuable insights into the landscape of airport planning in the post-2000 era and build methodologies that leverage these insights into measurable improvements in the practice of peer-group learning and aviation demand forecasting:

1) What are the dynamics of socioeconomic and operational changes affecting the airports in the post-2000 era of airline mergers and consolidations?

2) What are the operational and socioeconomic characteristics of an airport on the verge of experiencing a severe contraction in passenger volumes?

3) Does reference class forecasting (a method of calibrating a forecast based on past forecast errors) produce statistically significant reductions in forecast errors for aviation demand forecasts?

   a. What is the relevant and effective definition of a reference class of the forecast errors?

In Chapter 2, I argued that without proper considerations of demand uncertainty and optimism bias in airport planning, these airport planning techniques (peer-group learning and aviation demand forecasting) will perpetuate the risk of unwise and wasted investments in unwarranted airport infrastructures. After surveying the literature on the ways to tackle
the problem of demand uncertainty and optimism bias, I found that the proposed alternative planning frameworks lack empirically proven efficacy and remain theoretical. Instead, I argued that the aviation industry is ripe for methodological improvements to peer-group learning and aviation demand forecasting and such improvements can have much more immediate and meaningful impacts on airport planning. Specifically, I found three areas of inquiry that are left unanswered in the literature; 1) developing a peer identification technique (finding airports that share similar experiences) in the post-2000 era of demand uncertainty that has disrupted the existing airport hierarchy, 2) addressing the demand uncertainty of a severe contraction in passenger volumes that has plagued some of the major hub airport in the US, and 3) leveraging the systematic past forecast errors in order to reduce optimism bias in the aviation demand forecasts. I developed new methodologies to fill these areas of missing inquiry and test their efficacy in each of the subsequent chapters.

In Chapter 3, I argued that the current airport peer identification method based on a single metric of passenger enplanements cannot capture the dynamic nuances introduced to the system since the deregulation of the airline industry in the 1970s and intensified in the post-2000 era of airline mergers and consolidations. I further argued that the current peer classification system could even be detrimental to the practice of peer-group learning because airport planners could either push aspirational goals by using a set of peers that are in fact unparalleled in terms of business model or learn very few usable lessons from peers that have undergone rather different operational and socioeconomic changes. As an
alternative, I conducted a cluster analysis using dynamic metrics that track changes over time in various operational and socioeconomic indicators in a significant departure from the current peer-group learning practice by expanding on the metrics as well as timeframe. I showed that traditional hub designations distort the level of similarities among airport peers and mask the trends that allow for much more meaningful peer-to-peer exchange of information. Instead, the cluster analysis showed that the trajectories of airports on multiple dimensions over time can reveal meaningful peer groups with distinct shared experiences that enrich the practice of peer-group learning. In the era of demand uncertainty, I showed that airport planners need to be flexible and adaptable to dynamic changes and capture these trends by using variations of the methodology proposed here.

In Chapter 4, I showed that a sudden disruption in passenger volumes (e.g., an airline de-hubs from the airport) has a particularly detrimental effect on airport planning especially in planning for capacity expansion. These airports are left with excess capacity and end up spending even more money to recapture the lost demand. This type of a severe contraction in passenger volumes is traditionally not modeled into the forecasts. In this chapter, I built a binary logistic regression model to predict the probability of an airport experience a severe contraction in passenger volumes in the next 10 years in order to identify and understand the operational and socioeconomic trends that characterize the airports vulnerable to this type of disruption. My research showed that the regional health of the host cities and metropolitan areas is a good indicator of stability or instability in the passenger demand. Airports in the metropolitan areas that are gaining population with
strong service sector employment are far less likely to experience a severe contraction in passenger volumes. In addition, airports with diversified demand (types of trips) and an even distribution of airline market shares are more likely to have a stable passenger demand. These results not only give airport planners greater insights into assessing the risk of a sudden disruption in passenger volumes, but also bring into question the current landscape of the airline industry where only a handful of dominant airlines are investing and using the capacity at major airports.

Lastly in Chapter 5, I tested the feasibility and efficacy of applying reference class forecasting, a method of incorporating the systematic patterns of forecast errors into the current forecast to “ground” or reduce the forecast errors. This constitutes, to my understanding, the first such attempt in the field of airport planning. The key element of reference class forecasting is identifying a relevant reference class of similar entities (in the case of my research, past forecast errors). I developed four ways to identify reference classes and evaluated whether there is a statistically significant reduction in the forecast errors. My results indicate that these reference class methodologies can dramatically reduce forecast errors but also require nuanced approach to implementation. As more complexities were introduced into the reference class identification (going from a class of each airport’s own past forecast errors to a class of forecast errors of airports that share similar operational and socioeconomic trends), the reduction in forecast errors significantly improved, indicating that the operational and socioeconomic trends that define airport peers with
shared experiences are correlated with the level of forecast errors in their aviation demand forecasts.

6.2 Implications for Practice

This dissertation fills an important gap in the airport planning literature by contributing valuable insights into the landscape of airports and the airline industry in the post-2000 era and developing methodological approaches to enhancing the practice of peer-group learning and aviation demand forecasting. My research shows that the traditional techniques used to identify airport peers and to forecast future use of airports produce large discrepancies between the assumed or expected outcomes (e.g., true peers and forecasted passenger volumes) and actual outcomes. By significantly expanding airport planners’ understanding of the operational and socioeconomic trends that create these discrepancies in a systematic way, my research enables airport planners to consider these dynamic trends in a manner that helps them reduce the risk of making unwarranted or unwise infrastructure investment decisions.

The key of this dissertation is that airport planners should prioritize identifying and understanding the various trends that disrupt and recreate airport peers because the benefits of peer-group learning extends beyond the exchange of qualitative information and can greatly enhance aviation demand forecasting as well. Typically relegated to a category of tools that airport planners use to benchmark their airports’ performances and to compare and set planning goals, the peer-group learning framework based on robust peer identification methods can enable airport planners to collect and analyze relevant
data on the shared experiences and chart the future courses both quantitatively and qualitatively. For example, airports with excess capacity can use the past forecast errors of their peers to reduce their own forecast errors and can also calibrate the level of fee charges (landing fees and terminal use fees that airports can adjust in order to incentivize airline service) based on the outcomes of similar experiments of their peer airports. This framework also prevents airport planners from over-generalizing the success stories of airports that are characterized by unparalleled operational and socioeconomic trends. Every major airport aspires to be like Atlanta International Airport (ATL), the busiest airport in the world and also a home to one of the largest airlines in the world. The success stories of ATL may be unique to the scale and magnitudes of ATL’s operations and may not generalize well to other airports.

6.3 Limitations and Further Research

There are largely two areas where further research can help further improve the airport planning practice and infrastructure investment decisions; re-evaluation of these methods using a larger set of available data and investigating the political dimension of airport expansions.

All the data in this dissertation are publicly available data. Even so, some of the data, for instance, the FAA’s Terminal Area Forecasts (TAF), are not easily accessible and required contacting FAA personnel to acquire them. In addition, because I used 10-year demand forecasts in my research, this limited the number of data points within the available data range (1995-2015). As future researchers will have access to more historic
data for longer ranges of time, I leave the task of re-evaluating these methods using more robust data sets to the researchers in the future. In particular, more research is needed to validate the many of the results in this dissertation using different time frames. In the long run, I suspect many of the results will vary because the dynamics that define the post-2000 era will change in the future. This goes back to the research implication that airport planners and future researchers should prioritize identifying and understanding trends that define each planning era.

This dissertation was mostly concerned with forecast accuracy as a measure of improvement in the airport master planning process. But there are a political dimension to airport planning that is largely ignored in this dissertation. There is consensus in the transportation planning literature that an infrastructure investment decision is as much a political exercise as a technical one (Goetz & Szylowiucz, 1997b; Kane & Del Mistro, 2003; May & Hill, 2006). In other words, forecast accuracy alone will not guarantee that no unwarranted infrastructure investments will be made. While the discussion of policy evaluations and implications is beyond the scope of this dissertation, future research can investigate the nature of political pressures influencing airport infrastructure investment decisions. May and Hill (2006) touch upon this issue in the context of airports in Australia and future research building on their approach to evaluating airport expansions in light of both political and technical forecasts can contribute significantly to the airport planning literature.


INDEX

Adaptive Airport Strategic Planning, 6, 35, 38, 45
Adaptive-Policy Making, 35
airline mergers and consolidations, 4, 7, 27, 52, 56, 73, 143, 144
airlines, 1, 3, 4, 10, 15, 17, 18, 19, 26, 27, 50, 55, 58, 59, 61, 69, 71, 84, 95, 96, 97, 102, 104, 105, 110, 146, 148
airport capacity, 3, 15, 60, 62
Airport Cooperative Research Program, 5, 54, 81
airport expansion, 16, 19, 20, 29, 34, 36, 46, 54, 82, 109, 148, 149
Airport Improvement Program, 19, 26, 30, 44, 75
airport infrastructure investment, 1, 15, 78, 149
Airport master planning, vi, 2, 3, 6, 7, 9, 11, 15, 16, 21, 35, 36, 38, 42, 80, 82, 87, 149
airport planning, iii, vi, 1, 2, 3, 4, 5, 6, 8, 12, 15, 16, 17, 18, 21, 26, 30, 33, 35, 36, 37, 38, 39, 42, 43, 49, 53, 57, 73, 74, 77, 80, 87, 115, 143, 145, 146, 147, 148, 149
Airport sponsors, vi, 1, 9, 21
Atlanta, 50, 51, 54, 55, 69, 70, 148
aviation demand, vi, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 16, 17, 20, 21, 22, 23, 25, 26, 28, 30, 31, 33, 34, 35, 39, 40, 41, 42, 71, 77, 79, 80, 81, 82, 85, 86, 87, 93, 111, 113, 114, 117, 118, 120, 138, 139, 140, 143, 147, 149
aviation demand forecasting, 1, 10, 15, 16, 20, 21, 23, 28, 41, 42, 81, 118, 140, 144
binary logistic regression, 80, 87, 97, 99, 100, 101, 105, 145
City of Austin, 23, 25, 28, 32, 44, 79, 108, 113, 141
center analysis, 56, 57, 62, 70, 71, 72, 145
centers, 56, 63, 64, 65, 67, 69, 70, 76, 91, 130
congestion, 2, 3, 18, 20
Demand management, 18
demand uncertainty, vi, 3, 4, 6, 7, 8, 16, 28, 31, 32, 33, 34, 35, 36, 38, 39, 40, 42, 57, 79, 80, 81, 85, 86, 87, 143, 145
Denver International Airport, 28, 45, 51, 54, 70, 74
deregulation of airline industry, 17, 27
Dynamic Strategic Planning, 6, 35, 36, 41
econometric modeling, 22, 23, 24, 28
economic development, vi, 5, 19, 26, 44, 45, 54, 74, 75, 109
enplanement, 9, 49, 54, 57, 63, 93, 94, 95, 121, 144
FAA, vi, 1, 9, 12, 20, 26, 30, 31, 32, 40, 44, 47, 49, 50, 53, 54, 55, 57, 58, 62, 65, 69, 71, 72, 75, 77, 81, 82, 88, 108, 115, 121, 148
Federal Aviation Administration, vi, 1, 9, 30, 49, 77, 82, 109
Federal funding, vi, 5, 26, 30
Flexible Strategic Planning, 6, 35, 36
Forecast accuracy, 40
forecasting methods, 21, 48
Gaussian Mixture Model, 90
hub, 3, 8, 13, 18, 26, 27, 40, 47, 49, 50, 53, 55, 60, 61, 65, 67, 69, 71, 72, 75, 84, 88, 89, 94, 95, 96, 102, 104, 105, 110, 144, 145
hub-and-spoke, 18, 27, 55, 95
insider’s approach, 24, 41, 117
metropolitan statistical areas, 57, 88
MSA, xi, 30, 33, 34, 77, 82, 83, 84, 109, 115, 121, 122, 126, 148
Terminal Area Forecast, 30, 77, 78, 82, 109, 121, 148