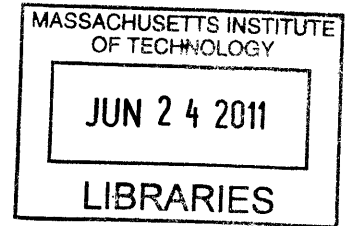# Analysis of Airline Schedule Padding

on U.S. Domestic Routes

by

Gerasimos Skaltsas

Diploma in Civil Enginnering

National Technical University of Athens, Greece (2008)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author...
Department of Civil and Environmental Engineering
May 9, 2011

Certified by...
Peter P. Belobaba
Principal Research Scientist of Aeronautics and Astronautics
Thesis Supervisor

Accepted by.....................
Heidi M. Nepf
Chair, Departmental Committee for Graduate Students

Analysis of Airline Schedule Padding

on U.S. Domestic Routes

by

Gerasimos Skaltsas

Submitted to the Department of Civil and Environmental Engineering on May 9, 2011, in partial fulfillment of the requirements for the degree of
Master of Science in Transportation

Abstract

Every airline passenger faces the risk of arriving late because flight times are subjected to many sources of variability. These can be weather conditions and airspace congestion, imbalances between airport demand and capacity, fleet and crew availability, technical failures and delays in maintenance, and other airline operations such as boarding and fueling. The main objective of this thesis is to explore the most common sources of variability in flight operations and study how U.S. carriers add buffer time (or pad) to scheduled block time to account for them.

Using flight data from FAA Aviation System Performance Metrics, we analyze the scheduled and actual flight times on 2359 directional non-stop domestic routes during 2009. The time of each flight is decomposed to delay at gate, taxi-out time, airborne time and taxi-in time. Then, the buffer time of each flight is computed, using as nominal airborne time the 10th percentile of the actual airborne time distribution. Our study consists of two parts. First, an aggregate statistical analysis is performed, concentrating on trends and correlations among factors such as buffer, flight time components, route distance, seasonality effects, delays caused by Ground Delay Programs, time of day and day of week, a flight's relative position to other flights operated on the same day by the same aircraft, total number of flights operated by the same aircraft during a day, the role of airport and carriers' network structure. Finally, we perform an econometric analysis through linear regression models to estimate how some of the above factors affect carriers' padding and their on-time performance.

The results indicate distance and time of day to be the most important factors that affect schedule padding. While absolute buffer increases with distance, when buffer is measured as a fraction of nominal block time it decreases exponentially. Furthermore, buffer and on-time performance fluctuate strongly over the course of the day, with flights scheduled to arrive during the evening peak having the worst on-time performance, despite the fact that these flights are padded the most. The data reveal that among the studied carriers Southwest pads its schedule more extensively, achieving a very high on-time performance, whereas other low cost carriers pad their flights substantially less, and have a lower on-time performance. Our findings also show that flights destined to the carrier's hub have more buffer than flights destined to spoke airports. Last, competition has a positive effect on schedule buffer and on-time performance.

Thesis Supervisor: Peter P. Belobaba
Title: Principal Research Scientist of Aeronautics and Astronautics

Acknowledgments

First and foremost, I would like to thank my research advisor, Dr. Peter Belobaba, for his inspirational guidance and unfailing support throughout my graduate studies at MIT. I am grateful to him for having transferred his passion about air transportation to me, and for having given me the opportunity to work on this exciting research project. His knowledge and experience about the airline industry made this work a very exciting and enriching experience.

I would also like to thank Professor Amy Cohn for her motivation and encouragement during these last two years. Her stimulating research ideas and her support to present my work at conferences, contributed substantially to the successful completion of this thesis and my overall academic experience.

I am also very grateful to two people who believed in my skills and played a major role in my academic development: my teacher Panos Konstantinou, and my advisor in NTUA Professor Matthew Karlaftis. Dr. Konstantinou had been the first to introduce me to the fascinating world of mathematics and their applicability in engineering, while Professor Karlaftis was a source of inspiration and constant supoport to me throughout my studies in transportation engineering. Without their encouragement, I would not have pursued graduate studies in the U.S. I consider myself very lucky to have been their student.

Looking back at the last two years, I realize how significant has been the direct and indirect contribution of many friends to this work and my broader MIT experience. First, I would like to thank my friend Yannis, for introducing me to the MIT world, and having been a great fellow for many years. I am very happy and thankful for sharing with him the difficult as well as the exciting moments of my two years in Cambridge. A special thank-you note goes to my friend Nikolas, for his research advices, and also for the great fun we had in the lab arguing about our favorite sport teams. My restless labmates, Alex, Alyona, Claire, Diana, Emilio, Himashu, Lishuai, Mehdi, Pierre, Thomas and Varun, enriched my days and nights in ICAT lab with great discussions, fun, and a lot of support. I would also like to thank the rest of my friends in Boston who have made the years I spent here a wonderful and memorable experience. Antonis, Dimitris C., Dimitris N., Stefania, Giorgos, Kimon, Kostas, Michalis, Rimi, Ruben, Spyros, Stella, Vangelis, Vasilis and Yola, you have all been an amazing company. I also want to extend my heartfelt gratitude to my cousin Giorgos, and my friends Alexandros, Katerina, and Mina. Despite the distance, they had always been finding ways to make me smile, keep me sane, and be part of my daily life.

Finally, my biggest thank you goes to my loving family – my siblings, Stella and Dimitris, and my parents, Spyros and Andromachi. Stella and Dimitris, I am so grateful for being the youngest in our family and having you as exemplars in my life. Your great achievements and your constant help, have always been extremely beneficial to me. Mom and dad, I would not have been able to achieve anything without your relentless support and tremendous love. I know that your devotion to my personal and educational development required huge sacrifices from your side, and I hope that this thesis is a small reward for you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Thesis Motivation

Since the 1978 airline industry deregulation, flight delays have been a serious and growing issue in the United States transportation system. Despite major investments in the modernization and expansion of aviation infrastructure in the last decades, the increase in airport and airspace capacity has not followed the substantial growth in traffic, resulting in a rapidly widening imbalance between capacity and demand.

One would expect that airlines, responding to these capacity constraints, would adjust their flight frequencies and timing of operations, in an effort to alleviate aviation infrastructure congestion. However, the highly competitive environment and the congestion externality have had an opposite effect. In practice, carriers tend to operate smaller aircraft offering more frequent flights, and match competitors' schedules at times of high demand. These behaviors lead to peak traffic load, hence flight delays.

Flight delays have a tremendous cost for airlines, passengers, and the environment. First, they increase the direct flight operating cost, resulting in losses for airlines. To recover this cost, airlines might increase passenger fares, making air travel more expensive and less attractive. Moreover, delays reduce the airlines' fleet and crew utilization by inducing carriers to add slack to their schedules, in order to account for the variability in their operations. Furthermore, delayed aircraft increase the need for extra gates and ground personnel. The above impose the additional cost of lost demand from potential airline customers, who may turn to other travel alternatives or telecommunication means. Last, there are revenue losses from re-accommodating passengers who miss their flights (when the airline is the contributing factor) or flight cancellations.

For passengers, the delay impact is mainly expressed in terms of stress and discomfort experienced in the terminal or on board, and the delay in reaching their destination. These costs grow considerably when late aircraft arrivals result in missed connections or cancellations of subsequent flights. In these cases, an additional cost for food and lodging may result. Furthermore, flight delays have a significant environmental impact because delays on tarmac or en route produce more fuel emissions, affecting global climate changes and air quality.

In addition to the direct costs to the airline industry and passengers, flight delays have an indirect effect on the U.S. economy as well, as airlines transfer their costs to their customers, who pay higher fares, thus increasing the travel cost in the business industry. Moreover, flight delays translate into longer trip times for business passengers, resulting in productivity reduction for employers. Just in 2007, the total cost of flight delays to the U.S. economy was estimated at $32.9 billion[1].

In an effort to provide incentives to carriers to enhance the efficiency of their operations and reduce delays, the U.S. Department of Transportation implemented the On-Time Disclosure Rule in 1987. This rule requires U.S. carriers that account for at least one percent of U.S. domestic passenger revenue to submit monthly Airline Service Quality Performance Reports to the Bureau of Transportation Statistics (BTS). BTS releases publicly available monthly reports with the arrival statistics of each carrier. The most widely used performance metric is the on-time arrival rate, where a flight is considered on time if it arrives within 15 minutes of its scheduled arrival time.

---

[1] Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., et al., *Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States*, Final Report, Washington, D.C.: Federal Aviation Administration, October 2010.

In response to the On-Time Disclosure Rule, carriers increased their scheduled gate-to-gate times (block times) by adding buffer, and established longer scheduled flight times compared to those required under optimal conditions. This practice, known as "schedule padding", is used by carriers as a primary means for improving schedule reliability and on-time arrival statistics. A reliable schedule reduces delay propagation throughout an airline's network, increases customers' satisfaction, and improves carrier's operational integrity. Furthermore, a high ranking in BTS's on-time arrival rates may improve the public perception about the carrier's service quality, thus increase the passengers' willingness to pay.

However, carriers face a trade-off between padding their scheduled flights times and improving their on-time performance versus the cost of increased scheduled times. First, typical labor agreements specify that flight crew salaries are based on the maximum of scheduled and actual block time, whichever is higher. Thus, scheduling flights that systematically arrive earlier than planned imposes an extra labor cost to the carriers that could be avoided. Long buffer times also hinder airlines from making optimum use of their limited resources: aircraft, crew, and airport infrastructure. Moreover, long scheduled block times may result in lower demand, revenues and market share, as flights are listed in the global distribution systems according to their scheduled block time, with the shortest flight having the best screen presence.

Recently, schedule padding has attracted public interest in the U.S., as frequent travelers often notice their flights arriving ahead of scheduled time. By lengthening their scheduled block times, thus reducing late flight occurrences, airlines simultaneously increase the likelihood of early arrivals. In other words, the cost of schedule reliability increases with variability in airline operations. These factors are interdependent, meaning that the more variable the conditions under which flights are operated, the more difficult and expensive it is to schedule and operate them in a timely manner.

The aim of this thesis is to analyze the airline schedule padding practices with respect to variability in airline operations. First, we will try to understand what causes this variability, and to what degree it can be predicted by airlines. Then, we will investigate how U.S. airlines adjust their scheduled block times on domestic routes to improve their reliability. Last, we will study how these changes in schedule are reflected to carriers' on time performance.

1.2   Thesis Objectives and Approach

The main scope of this thesis is to study the variability in flight time components on U.S. domestic flights during 2009, and shed light on the padding practices that U.S. carriers use to improve their schedule reliability. Our objective is to examine a number of potential delay causal factors and analyze their magnitude and variation on selected U.S. routes. Furthermore, we want to look for relationships among them, as well as their impacts on buffer and on-time performance.

The data source of this thesis is the Individual Flights Record of the FAA Aviation System Performance Metrics (ASPM). ASPM provides detailed information on individual flight performance for a large number of U.S. carriers and airports, reporting more than 80% of domestic flights. From this data source we extract data for commercial flights on 2359 non-stop routes and 40 carriers, studying in total 59% of the domestic commercial flights in 2009.

The literature review in Chapter 2 suggests that delays at different flight stages vary in terms of predictability, and their impact in flights arrival delays with respect to schedule. To examine their effect on an individual basis, we decompose flight times to delays at gate, taxi-out times, airborne times and taxi-in times. Next, we compute buffer, which is the core metric in our research, using as nominal airborne time the 10th percentile of the actual airborne time distribution during a month period.

Our analysis is separated into two parts. First, we concentrate on the descriptive analysis of our sample, and we look for trends and correlations among our variables. Special attention is given to factors such as flight time components, route distance, seasonality effects, delays imposed by Ground Delay Programs, number of flights operated by the same aircraft during a day, time of day and day of week, the role of airport and carriers' network structure. Finally, we perform an econometric analysis through linear regression models to estimate how some of the above factors affect padding and on-time performance.

The goal of the thesis it to investigate the following issues:

- What is the magnitude of variability in the different flight segments? How do carriers adjust their schedule with respect to them?

- What is the relationship between buffer and stage length? Does stage length influence the likelihood of a flight to arrive on-time?

- Do carriers adjust their scheduled block times over the course of the day?

- What is the impact of a flight's relative position to other flights operated on the same day by the same aircraft on delay components, buffer and on-time performance?

- What are the differences in buffer and scheduled turn-around times between flights in or out from hubs, compared to other airports?

- To what extent do Ground Delay Programs affect a carrier's decision regarding buffer?

- Are there differences in schedule padding practices and on-time performance across carriers?

- How does the number of competitors on a route affect carriers' padding and on-time performance?

1.3   Thesis Structure

This thesis comprises six chapters:

In Chapter 1 we introduce the notion of airline schedule padding and briefly describe the benefits and costs associated with it. We also present the objectives of this thesis and the structure we will follow.

Chapter 2 is a literature review of studies on delays, padding practices and the on-time disclosure rule. First, we review the most important delay metrics and explore their causes and the evolution of delays over time. Then, we present studies addressing the response of the carriers to the on-time disclosure rule and discuss the imposed changes in scheduled block times, arrivals delays, and on-time performance. We also review the impact of competition and hub-and-spoke

networks on airline schedule padding. Last, we look at the cost impacts of delay and buffer on airlines, passengers and the U.S. economy.

Chapter 3 comprises two parts. In the first part, we describe the FAA ASPM database, and introduce in detail the dataset that was extracted for the purposes of this study. In the second part, we present the metrics and the terminology we will use. Furthermore, we explain the methodology we follow for computing the unimpeded airborne time and the buffer time of the flights in the data sample.

Chapter 4 contains an aggregate analysis of our dataset. We present the distributions of delay metrics and buffer on selected non-stop routes, and describe the uncertainty and the variability in airline operations. We also analyze the correlation between the delay components and the buffer time, and we try to understand how uncertainty influences airlines' padding practices. Finally, we further examine the most important factors that affect each flight segment, and formulate assumptions about the way that carriers adjust their schedules with regard to them.

In Chapter 5 we perform an econometric analysis to estimate how certain factors affect buffer and on-time performance. First, we present the explanatory variables that we will study and test for multicollinearity between them. Next, we construct six regression models, we formulate our assumptions, and we discuss how the models' results agree with our expectations. The developed regression models provide us with more information on carriers padding practices by quantifying the impact of the studied variables. Furthermore, they help us understand how changes in padding affect a carrier's on-time performance.

Finally, Chapter 6 summarizes this thesis' findings. It also discusses some of the limitations of this analysis and suggests directions for future research.

Chapter 2

Literature Review

Since deregulation in 1978, when the U.S. airline industry became increasingly competitive and air traffic started reaching the airport and air traffic management system's capacity, a number of studies have investigated the causes of flight delays, the correlation between them, and their impact on the performance and the cost of the air transportation system. This thesis explores the most common sources of variability in the flight time components and studies how U.S. carriers plan their schedules to account for them.

The first part of this literature review illustrates the most important metrics of delays. Based on previous research, we explore the causes of delays and the evolution of delays over time. Looking at the existing literature on flight delays, we hope to get a better understanding of how delays are generated and propagated within the air transportation system. In doing so, we expect to develop the basic framework for studying how airlines select buffer times so as to incorporate the potential for delays in their schedules. Special attention is given to the on-time disclosure rule of 1987. According to this rule, domestic air carriers accounting for at least 1% of U.S. domestic passenger revenues must submit, on a monthly basis, Airline Service Quality Performance Reports to the Bureau of Transportations Statistics (BTS). Furthermore, we review the studies that address the impact of competition in the market and the role of hub-and-spoke networks in airline schedule padding.

2.1 Metrics and Sources of Delay

There are several metrics that describe the delays that a flight can experience. In this thesis, a flight is decomposed into four time segments: the time prior to push-back (gate delay), the taxi-out time, the airborne time and the taxi-in time. Many of the factors that influence these individual flight segments have been extensively studied in the past mostly through simulations and analysis of historical flight data.

There are two government agencies that report air traffic delay statistics in the U.S.; the Federal Aviation Administration (FAA) and the Bureau of Transportation Statistics (BTS). While BTS's main purpose is to provide information to consumers about carriers' operations, FAA's data serve to measure system-wide ATC performance and identify areas for improvement. These two agencies sometimes use alternative ways to track the delays. Therefore, there are differences in the reported data and statistics[2]. All studies described below collect their data from one of these two sources, unless otherwise indicated.

A model to estimate the gate delay distribution was developed by Tu et al. (2008)[3]. The gate delay was defined as the difference between the scheduled departure time and the actual departure time from the gate (push-back time). The gate delay can be influenced by many factors such as weather conditions, airport congestion, airline policies, mechanical problems, baggage handling, demand surges and propagation delays caused by late arrival of the previous flight with the same aircraft. Tu et al. grouped these factors into three main categories: seasonal trend, delay propagation pattern, and random residuals. The model's parameters show a strong seasonal pattern with high delays in summer and winter but low in spring and fall. Moreover, there is a very strong propagation effect with delays building up gradually over the course of the day and decrease only very late at night. It is surprising to notice that, although the propagation delay becomes negative for the few flights scheduled to depart after midnight, the delay of the flights that actually depart after midnight is extremely high compared to the rest. According to the authors, the reason is that the vast majority of these flights have been delayed for a very long time. Furthermore, while the distribution of the scheduled departures follows a very distinct

[2] Mueller, E., Chatterji, G., *Analysis of Aircraft Arrival and Departure Delay Characteristics*, AIAA's Aircraft Technology, Integration, and Operations Forum, Los Angeles, California, 2002.
[3] Tu, Y., Ball, M., Jank, W., *Estimating Flight Departure Delay Distributions - A Statistical Approach with Long-term Trend and Short-term Pattern*, Journal of the American Statistical Association, Vol. 103, No. 481, March 2008.

spiked pattern, the distribution of the actual departures is smoother, mostly due to limitations on the airport departure rate.

In the past, many researchers developed queuing models in order to study the taxi-out delays, as for example Shumsky (1995)[4], Pujet (1999)[5] and Andersson et al. (2000)[6]. These models identified as main causal factors the following: the airline (gate location relative to runway, push-back time, etc.), the departure demand, the airport throughput and the runway configuration. Idris et al. (2001)[7] performed a causal factors analysis for the case of Logan Airport. Their results confirm the findings of previous studies. Moreover, they address the significance of the flow management programs imposed on the departure traffic heading to weather-impacted destination airports, jet routes or exit fixes. Although weather is usually the main reason for having flow management programs, these can be imposed due to other causes, such as high traffic volume or equipment outages (Idris, 2001)[8].

Willemain (2001)[9] developed a method to identify and estimate sources of airborne time variation throughout a regional airspace. In his paper, he studied two types of airborne time variability: deviations from long-run average airborne times, and deviations from the estimated en route times filed in flight plans. He analyzed four sources of variability: the system-wide airspace (day effect), the route's airspace (en-route effects), the airspace around the departure airport (origin effect), and around the arrival airport (destination effect). His methodology shows that airborne time variability is mostly attributable to the en-route effects, followed by the destination effect. The origin effect is limited. A very important finding is that when the variation around the long-run average was studied, the relative directions of two flights exercised a great influence on the correlation of their estimated daily en-route effects. This suggests that en-route effects in fact reflect the impact of jet streams. Furthermore, there is a strong negative correlation between origin and destination effects at the same airports. However, these relationships do not apply in

[4]Shumsky, R., *Dynamic Statistical Models for the Prediction of Aircraft Take-off Times*, Ph.D. Thesis, Dept of Electrical Engineering and Computer Science, MIT, September 1995.

[5] Pujet, N., *Modeling and Control of the Departure Process of Congested Airports*, Ph.D. Thesis, Dept of Aeronautics and Astronautics, MIT, September 1999.

[6] Andersson, K., Carr, F., Feron, E. and Hall, W., *Analysis and Modeling of Ground Operations at Hub Airports*, 3rd USA/Europe Air Traffic Management R&D seminar, Naples, Italy, June 2000.

[7] Idris, H., Clarke, J., Bhuva, R., and Kang L., *Queuing Model for Taxi-out Time Estimation*, Air Traffic Control Quarterly, Vol. 10, No. 1, September 2001.

[8] Idris, H., *Observation and Analysis of Departure Operations at Boston Logan International Airport*, Ph.D. Thesis, Dept of Mechanical Engineering, MIT, February 2001.

[9] Willemain, T., *Estimating Components of Variation in Flight Time,* The National Center of Excellence for Aviation Operations Research (NEXTOR), May 2001. Available at: http://www.nextor.org/pubs/WP-01-2.pdf

the case of deviations from the estimated en route times, where the impact of jet streams is excluded.

Willemain et al. (2003)[10] examined the effect of other factors on the variability in the estimated en-route times filed in flight plans. They demonstrated that the month of the year, the day of the week, the hour of the day, the aircraft type and the carrier were all significant factors. Furthermore, they showed that the effect of the month varies significantly among routes at different latitude, highlighting at the same time the impact of winds aloft. The deviation among airlines is not only caused by the use of different aircraft but also because of the differences in the carrier's flight planning styles. It was noted that the difference in standard deviation of estimated en-route times between two carriers on the same route does not indicate which flight planning process is better. A low standard deviation might mean that the airline puts emphasis on predictability and achieves consistency in the filed flight times. However, this could also be explained in the light of poor planning and rare changes in flight plans.

The effect of en-route weather on flight delays was thoroughly examined by Post et al. (2002)[11]. They developed an en-route weather severity index based on the densities of lightning strikes and flight plan tracks. The relationship between this en-route weather index and the airborne delays was very well described by a linear model with $R^2 = 0.76$.

A spectra analysis by Welch and Ahmed (2003)[12] examined the relation of airport throughput to airborne and arrival delays. It was found that airborne delays at intermediate throughputs occurred mostly in en-route airspace and not on approach to the airport runway. Moreover, the delay relative to scheduled arrival time varied significantly in regard to arrival throughput. At hub airports, the arrival delay decreases with the throughput, because the highest throughput occurs when on-time flights arrive during the connecting banks. On the contrary, at non hub airports, the arrival delay sometimes increases with throughput because the highest throughput occurs when on-time flights merge with delayed flights due to weather.

Mueller and Chatterji (2002)[2] extracted flight data of 2001 from the Post Operations Evaluation Tool (POET) database and studied the departure, en-route and arrival delay characteristics for 10 major U.S. airports. The departure delay measures the discrepancy between

[10] Willemain, T.., Ma, H., Yakovchuk, N., Child, W., *Factors Influencing Estimated Time En Route,* Tech. rep., The National Center of Excellence for Aviation Operations Research (NEXTOR), 2003.
[11] Post, J., Bennett, M., Bonn, J., Knor, D., *Estimation of an En Route Weather Severity Index Using Lightning Strike and Flight Plan Data,* FAA Technical Publication, 2002.
[12] Welch, J., Ahmed, S., *Spectral Analysis of Airport Capacity and Delay,* 5th USA/Europe Air Traffic Management R&D Seminar, Budapest, Hungary, June 2003.

the actual time the aircraft lifts off the runway and the time the aircraft was scheduled to leave the gate, plus an estimated taxi out time. It was shown that the departure delay fits better the Poisson distribution while the en-route and arrival delays are better modeled using Normal distributions. The study revealed a small amount of day to day variation in the average number of departures, percentage of delayed departures and average departure delay (for the delayed flights only). However, there is no evidence of correlation between these three metrics. Xu et al. (2007)[13] argued that the relationship between the delay components (turn-around delay, gate delay, taxi out-delay, airborne delay, taxi-in delay, and arrival delay) and their sources vary greatly among different ranges of the delay components.

El Alj (2003)[14] developed a new metric to estimate delays due to congestion at airports and in airspace, as well as due to other system inefficiencies. This metric was defined as the difference between the actual block time and a baseline time, measured as the 15th percentile of the actual block time distribution of a given month. The true delays based on this metric were about 40% to 60% greater than arrival delays relative to schedule. Her analysis of block and arrival delays suggests that although airlines predict block times with high accuracy, they are not as good at predicting departure times. This might be one of the main reasons that delays relative to schedule are incurred. Furthermore, El Alj showed that there is a strong correlation between the taxi-out delays and the origin airport, and between the taxi-in delays and the destination airport. Last, she estimated that around 60% of the airborne delay on any given route is attributable to airspace congestion, while the remaining 40% is attributable to the destination airport.

In this thesis we will compute the unimpeded airborne time by using an approach similar to El Alj's baseline time. More specifically, we will extract the 10th percentile of the actual airborne time distribution for a given directional route, month and carrier, and then we will compare it to the actual airborne time in order to estimate the delay that a flight experiences en route. The validity of the 10th percentile usage for estimating the unimpeded airborne time has been tested by Vlachou et al.[15].

---

[13] Xu, N., Donohue, G., Laskey, K., Chen, C.H., Williams S. and Sherry L., *Bayesian Network Analysis of Flight Delays*, 86th Transportation Research Board Annual Meeting, Washington D.C, January 2007.

[14] El Alj, Y., *Estimating the True Extent of Air Traffic Delays*, Master Thesis, Dept of Civil and Environmental Engineering, MIT, June 2003.

[15] Vlachou, K., Tripodis, Y., Lovell, D., *A Comparison of Different Approaches to Estimate Unimpeded Flight Times*, 12th Air Transportation Research Society World Conference, Athens, Greece, July 2008.

2.2   The Response of the Carriers to the On-Time Disclosure Rule

In response to a public outcry over the growth in air traffic delays, in 1987, the U.S. Department of Transportation implemented the On-Time Disclosure Rule that requires domestic air carriers that account for at least 1% of U.S. domestic passenger revenues to submit monthly Airline Service Quality Performance Reports to BTS. There are 32 U.S. airports for which these carriers have to report their flight performance data. However, the carriers are voluntarily reporting all domestic operations to the BTS. Based on these data, BTS releases publicly available monthly reports with the arrival statistics of each carrier. This rule aims to prevent carriers from developing two schedules: one that is posted on the reservation systems for marketing purposes, and one that is flown. Furthermore, it gives incentives to carriers to improve their on-time performance either by scheduling times that they can meet in practice or by improving their operations.

BTS's most widely used performance metric is the on-time arrival rate where a flight is defined as delayed if the aircraft fails to set its parking brake or open the passenger door less than 15 minutes after the scheduled arrival time. However, the on-time arrival rate is not a reliable indicator of the impacts on passengers because carriers' operational inefficiencies can be very well hidden through extensive padding. For example, let us assume two carriers A and B that operate on the same route, with the same scheduled departure time, and with scheduled block times of 60 and 70 minutes respectively. If a flight of carrier A arrives with 17 minutes of delay and a flight of carrier B with 10 minutes, BTS will consider the first flight delayed and the second as being on-time. Still, the passengers of carrier A will have arrived at their destination three minutes earlier than those of carrier B.

The desire to achieve a high ranking in the on-time arrival rates can be a strong incentive for carriers to lengthen their scheduled block times so as to improve the public perception about the carrier's service quality and increase the passengers' willingness to pay. Forbes (2008)[16] estimated a fare reduction of $1.42 on average for direct passengers for every minute of additional delay. Her analysis suggests that a decrease in quality has a strong negative effect on the market price in competitive markets, and a much weaker effect in markets with limited competition.

---

[16]   Forbes, S., *The effect of air traffic delays on airline prices*, International Journal of Industrial Organization, Vol. 26, Issue 5, September 2008.

However, in practice there are many constraints in adding an extra buffer to the scheduled block time. First, many labor agreements specify that the flight crews are paid based on the maximum of scheduled and actual block time. Scheduling flights that systematically arrive earlier than planned imposes an extra labor cost to the carriers that could be avoided. In addition, this practice restrains carriers from maximizing the utilization of their resources: aircraft, crew, and airport infrastructure. Furthermore, carriers are competing for passengers, and the flight time can be more critical in passengers' final choice than the on-time performance, about which they are usually not informed. Especially in the past, short block times were offering tremendous competitive advantage to carriers, because the flights are listed in the computer reservation systems according to their scheduled block time, with the shortest flight having the best screen position.

Shumsky (1993)[17] examined how carriers responded to the on-time disclosure rule from 1987 to 1991, and suggested ways for optimal scheduling. He showed that within this five-year period the major carriers consistently increased their scheduled block times (the total increase was about 10%), despite the fact that the actual block times were declining after 1989. It is remarkable that if the carriers had not lengthened their schedules since 1987, the percentage of flights arriving on-time in 1991 would have been 20% lower.

Nevertheless, Shumsky found that the scheduling practices with respect to on-time performance varied among carriers in terms of the time added to their schedules, and also in the way in which this extra time was distributed among flights. For example, in the case of Continental Airlines, the number of minutes added to each flight was well approximated by a simple linear model, using the previous year's on-time performance as the independent variable. On the other hand, American Airlines used a more sophisticated technique, where the buffer time was estimated based on the marginal gain that each flight would have from a potential lengthening. American Airlines lengthened the block time of flights that were most likely to be benefited, and in some cases further shortened the block time of flights with poor on-time performance, if the additional buffer time required for improving it was substantially large. This finding provides significant evidence that carriers take into account the distribution of the actual block time, a hypothesis that will be fully addressed in this thesis.

---

[17] Shumsky, R., *Response of the U.S. Air Carriers to the On-Time Disclosure Rule,* Transportation Research Record, Issue 1379, 1993.

Shumsky explored two strategies for distributing a fixed amount of extra minutes in a carrier's schedule. In the first strategy, the extra time was distributed uniformly among all flights, whereas in the second, a mixed-integer program formulation was used to allocate the extra time in a manner that optimizes the on-time performance of the flights in the previous year. Surprisingly, his results showed that the first strategy performed as well as the second strategy and the carriers' own schedules with respect to on-time performance. Shumsky concluded that the on-time disclosure rule rewards the carriers with less variability in their flight times, and those that have high variability but can afford very large increases.

A longer time period, from 1988 till 2000, was studied by Mayer and Sinai (2003)[18]. While the average annual on-time performance increased between 1989 and 2000 (83%), it then gradually declined reaching the lowest value of 73% in 2000. Similarly, over the 1995-2000 period the average true delay on the 618 routes studied by El Alj[14] increased by 52% from 11.1 minutes to 16.9 minutes. Mayer and Sinai showed that there are systematic differences in on-time performance across months, with the worst on-time performance in the months with bad weather. Because on-time performance has a systemic variation in regard to some factors, they suggested that airlines do not appear to adjust their scheduled block time to compensate for them. In this thesis, we will explore to what extent this suggestion holds true, building on their finding that carriers shorten their scheduled block times when the variability in block time is very large. This seems to support Shumsky's observation about American Airlines marginal gain approach[17]. Last, they showed that there are no substantial differences in the scheduling behavior across carriers, and the reason that Southwest Airlines achieves a better on-time performance is attributed to the fact that it has the shortest turn-around times.

Zhu (2009),[19] in his masters thesis, explored among others the differences and effectiveness of the scheduling practices of two carriers, the legacy carrier A and the low cost carrier B, during two days of operations in 2006, one with low delays and one with high delays. Zhu showed that the legacy carrier A tends to plan more slack in its turn-around times, whereas the low cost carrier B tends to add more slack in its block time. Furthermore, carrier B schedules almost identical amounts of turn-around times at each airport. Zhu argues that the differences in schedule padding practices between the two airlines are directly related to the type of hubs operated by the carrier. In particular, carrier A operates banked hubs in which turn-around times

[18] Mayer, C., Sinai, T., *Why do airlines systematically schedule their flights to arrive late?*, Mimeo, Wharton School of Business, May 2003.
[19] Zhu, Y., *Evaluating Airline Delays: The Role of Airline Networks, Schedules and Passenger Demands*, Master Thesis, Dept of Civil and Environmental Engineering, MIT, February 2009.

must be longer on average to accommodate connecting passengers. On the other hand, carrier B operates de-banked hubs and schedules shorter turn-around times to maximize aircraft and crew productivity. In order to avoid the propagation of delays due to the tighten turn-around times, carrier B schedules more slack in its block times.

2.3 The Impact of Competition on Padding

Market competition can affect a carrier's decision to shorten or lengthen a flight's scheduled block time in several ways: appearance in the computer reservation systems, public perception about quality of service, scheduling at peak travel times, and cost imposed onto passengers.

The benefits that passengers can gain from market competition depend heavily on passengers' accessibility to information about the provided service quality. Foreman and Shea (1999)[20] studied market response to the publication of on-time performance, and showed that the reduction in the search cost of information results to an overall improvement in airlines performance. They also found a positive correlation between on-time performance and competition.

It has been shown by Suzuki (2000)[21] that a passenger is more likely to switch carriers after experiencing a flight delay. His findings also suggest that on-time performance affects carriers' market share more through the passengers' personal experience and perception of the carriers' reliability, rather through the reported on-time arrival statistics. Borenstein and Netz (1999)[22] showed that in competitive markets carriers schedule their flight departures closer together compared to a single carrier that operates the same number of flights. The reason is that by offering schedules at peak travel times, carriers capture more demand. Based on this, one would expect that competition might deteriorate the carriers' on-time performance because it forces them to schedule their departures at peak hours. However, the literature suggests that competition has a positive impact on carriers' on-time performance.

---

[20] Foreman, S., Shea, D., *Publication of Information and Market Response: The Case of Airline on Time Performance Reports*, Review of Industrial Organization, Vol. 14, 1999.
[21] Suzuki, Y., *The Relationship between On-time Performance and Airline Market Share: A New Approach*, Transportation Research Part E: Logistics and Transportation Review, Vol. 36, 2000
[22] Borenstein, S., Netz, J., *Why do all the flights leave at 8 am?: Competition and Departure-time Differentiation in Airline Markets*, International Journal of Industrial Organization, Vol. 17, 1999.

Mazzeo (2003)[23] compared quality outcomes across markets to investigate how competition provides incentives for carriers to improve their service quality. His results render clear that on routes served only by one carrier, or when the carrier's presence prevails, the flight delays are significantly longer and the on-time performance worse, compared to routes with greater competition. According to Mazzeo, this happens because delays will affect airlines profitability less if they are imposed on passengers who have fewer alternatives.

Rupp et al. (2006)[24] reached similar conclusions with regard to the positive relationship between route competition and on-time arrival rates. Moreover, their study showed that hubbing has a greater impact on on-time performance than the route competition. Flights that originate from the airline's hub have on average the worst on-time arrival rates, probably because they are scheduled to depart at peak airport congestion periods. No difference was found in on-time arrival rates between hub and non-hub airlines for flights destined to hub airports. Previous study of Mayer and Sinai (2003)[25] had resulted in similar findings.

---

[23] Mazzeo, M., *Competition and Service Quality in the U.S. Airline Industry, Review of Industrial Organization*, Vol. 22, 2003.

[24] Rupp, N., Owens, D., Plumly, W., *Does competition Influence Airline On-Time Performance?*, Elsevier, Advances in Airline Economics, Vol. 1, 2006.

[25] Mayer, C., Sinai, T., *Network Effects, Congestion Externalities, and Air Traffic Delays: or Why All delays Are Not Evil*, American Economic Review, Vol. 93, Issue 4, September 2003.

2.4 Assessment of the Cost of Padding

Although there are a few studies that estimate the cost of delays relative to schedule for the carriers and the passengers in the U.S. air travel system, very limited information is available in the present literature for the incremental costs imposed by the buffer times.

The most comprehensive study on the delay impact in the U.S. was conducted by the five NEXTOR universities and the Brattle Group (2010)[1]. The research team derived translog cost functions assuming that each airline minimizes its cost of producing a certain output, given the costs of its input factors for production such as labor and fuel, as well as other factors that influence its production process. One factor among the latter can be delay. The independent variables used in the model estimation were: output (revenue-ton-miles), fuel price, labor price, materials price, capital service, stage length, points served, and delays against the 10th and the 20th percentile of feasible flight time. The monetary values of passenger delays were computed based on the recommended values of travel time in the departmental guidance of the DOT (2003)[26].

In 2007, the total cost of all U.S. air transportation delays was estimated to be $32.9 billion:

- Costs to Airlines: $8.3 billion

- Costs to Passengers: $16.7 billion

- Indirect Impact on Economy: $4.0 billion

- Costs from Lost Demand: $ 3.9 billion

From the $8.3 billion that delays cost to airlines, $3.7 billion (45%) is attributed to scheduled buffer. The fractional cost of scheduled buffer is less for the passenger delay, accounting for the 36% of the total passenger cost ($6 billion out of $16.7 billion). This reduction occurs, because passengers experience a high cost due to flight cancellations ($3.2 billion). In absence of flight cancellations, the fractional cost of scheduled buffer would be the same for both the carriers and the passengers.

---

[26] Department of Transportation (DOT), *Valuation of Travel Time in Economic Analysis*, Revised Departmental Guidance, 2003.

2.5 Summary

The first part of this literature review presented the sources of the main delay components, and highlighted the impact of weather conditions, excess traffic, and delay propagation on on-time performance. Although many studies have shown that variability in these factors has a negative effect on on-time performance, it is still unclear whether airlines shorten their buffer times when variability is very high, or on-time performance is low because airlines failed to predict correctly the extent of this variability. This thesis attempts to provide more insight into this issue by studying the relationship of buffer with every delay component and its causal factors.

Next, the response of U.S. carriers to the on-time disclosure rule was reviewed. Some studies suggest that there have been differences in the scheduling behavior across carriers after the rule was implemented. However, these studies were limited to a small number of U.S. carriers and used small samples of these carriers' yearly operations.

Furthermore, we looked at previous research into the effects of route competition and airports in airline schedule padding. It has been shown that flights on competitive routes have longer buffer, thus achieving better on-time performance. Regarding hub operations, these have a lower on-time performance relative to other flights.

In this thesis, our objective is to explore the schedule padding practices with respect to both route and carrier characteristics. Therefore, we are using a very large data sample including more than half of all domestic operations in 2009.

Chapter 3

Definitions and

Data Presentation

This chapter is separated into two parts. In the first part, we try to familiarize the reader with the data set used. We present the sources of the database and describe explicitly the data used in our analysis. We also describe the requirements we used for constructing the final data sample and analyze its significance and limitations. In the second part, the metrics and the terminology used throughout this thesis are presented and discussed. We also explain the methodology followed for computing the unimpeded airborne time and the buffer time of the flights in the data sample.

3.1 Data Source

The data used in this thesis are derived from the Individual Flights Record of the FAA Aviation System Performance Metrics (ASPM). ASPM uses two primary flight data sources: the Enhanced Traffic Management System (ETMS) and the Out, Off, On, and In (OOOI) data from Aeronautical Radio, Incorporated (ARINC). Additionally, ASPM is updated and enhanced by data from the BTS Airline Service Quality Performance (ASQP) and the Official Airline Guide (OAG).

In ASPM there are data available for 55 U.S. airports after January 2000 and for 20 more U.S. airports after October 2004. For 2009, that is the year we study, ASPM provides detailed information on individual flight performance for 23 ASPM Carriers (Table 3-1) regardless of the airport, and for the vast majority of commercial flights for 77 U.S. airports (Table 3-2) regardless of the carrier. In 2009, these 77 airports accounted for the 88% of all domestic enplanements and the 80% of all scheduled domestic departures (BTS T-100 Domestic Segment).

ASPM reports flight performance data of six aircraft classes that operate at any of the 77 ASPM airports: commercial, air taxi, freight, general aviation, military and other. International flights from U.S. or international carriers operated at these airports are also reported. However, we will limit our analysis to domestic commercial and air taxi flights, operated by U.S. carriers only.

Air Canada	AC	Frontier Airlines	F9
Airtran Airways	FL	Hawaiian Airlines	HA
Alaska Airlines	AS	Jetblue Airways	B6
Aloha Airlines	AQ	Mesa Airlines	YV
American Airlines	AA	Northwest Airlines	NW
American Eagle	MQ	Pinnacle Airlines	9E
Atlantic Southeast Airlines	EV	Skywest Airlines	OO
Comair	OH	United Airlines	UA
Continental Airlines	CO	United Parcel Service	5X
Delta Air Lines	DL	US Airways	US
ExpressJet Airlines	XE	Southwest Airlines	WN
FedEx	FX		

Table 3-1: ASPM Carriers reported in 2009

ATL	Hartsfield-Jackson Atlanta Intl	BNA	Nashville Intl
ORD	Chicago O'Hare Intl	OAK	Oakland Intl
DFW	Dallas/Fort Worth Intl	SDF	Louisville Intl
DEN	Denver Intl	HOU	Houston Hobby
CLT	Charlotte Douglas Intl	SJC	Norman Mineta San Jose Intl
IAH	George Bush Houston Intercontinental	SMF	Sacramento International Airport
LAX	Los Angeles Intl	AUS	Austin-Bergstrom Intl
PHL	Philadelphia Intl	DAL	Dallas Love Field
DTW	Detroit Metropolitan Wayne County	SNA	John Wayne Airport-Orange County
MSP	Minneapolis/St. Paul Intl	SAT	San Antonio Intl
PHX	Phoenix Sky Harbor Intl	SJU	San Juan Luis Munoz Intl
LGA	New York LaGuardia	MSY	Louis Armstrong New Orleans Intl
LAS	Las Vegas McCarran Intl	ABQ	Albuquerque Intl Sunport
EWR	Newark Liberty Intl	BUF	Buffalo Niagara Intl
BOS	Boston Logan Intl	ONT	Ontario Intl
SFO	San Francisco Intl	JAX	Jacksonville Intl
MEM	Memphis Intl	RSW	Southwest Florida Intl
SLC	Salt Lake City Intl	BDL	Bradley Intl
SEA	Seattle/Tacoma Intl	OMA	Omaha Eppley Airfield
MCO	Orlando Intl	OGG	Kahului
JFK	New York John F. Kennedy Intl	BUR	Bob Hope (Burbank/Glendale/Pasadena)
DCA	Ronald Reagan Washington National	PVD	Providence Francis Green State
BWI	Baltimore/Washington Intl	PBI	Palm Beach Intl
IAD	Washington Dulles Intl	DAY	Dayton Intl
CVG	Cincinnati/Northern Kentucky Intl	ANC	Ted Stevens Anchorage Intl
STL	Lambert Saint Louis Intl	BHM	Birmingham Intl
MDW	Chicago Midway	TUS	Tucson Intl
CLE	Cleveland Hopkins Intl	MHT	Manchester
PDX	Portland Intl	HPN	Westchester County
SAN	San Diego Intl	LGB	Long Beach
FLL	Fort Lauderdale/Hollywood Intl	PSP	Palm Springs International
TPA	Tampa Intl	ISP	Long Island Mac Arthur
HNL	Honolulu Intl	SWF	Stewart Intl
MKE	Milwaukee Gnl Mitchell International	RFD	Greater Rockford
MIA	Miami Intl	OXR	Oxnard
IND	Indianapolis Intl	GYY	Gary Chicago Intl
MCI	Kansas City Intl	TEB	Teterboro
RDU	Raleigh/Durham Intl	VNY	Van Nuys
PIT	Pittsburgh Intl		

Table 3-2: List of the 77 ASPM Airports Sorted by Decreasing Domestic Departures in 2009

3.2 Data Sample

Throughout this thesis we will be studying how U.S. carriers adjust their schedules on domestic flights to account for the variability in their operations. In order to do so, we selected a flight sample representative of the U.S. domestic flight network in terms of stage length, carrier competition, time of operations, airport traffic and delays. Furthermore, the number of flights examined on each route are adequately large so that the estimated unimpeded airborne times are not biased and the results of our analysis can be safely expanded to the whole population. The methodology we followed for the construction of the final sample is described below:

- First, we limited our analysis to U.S. carriers and domestic flights that were identified as "Commercial" or "Air Taxi". "Freight", "General Aviation", "Military" and "Other" flights were excluded.

- Only routes between the ASPM airports were studied because the data provided for them are more accurate.

- The flights that were scheduled to depart between midnight and 5am were excluded from our sample. Normally, flights at those hours are not routinely scheduled, and thus could behave as outliers with minimum scheduled and actual flight time components.

- From the final data set we eliminated the routes where no carrier operated on a daily basis. A carrier was considered to operate daily on a given route if it had scheduled flights for at least 300 days of 2009. As a result, Gary Chicago Intl (GYY), Greater Rockford (RFD), Oxnad (OXR), Teterboro (TEB) and Van Nuys (VNY) were excluded.

The final data set consists of 2,359 non-stop routes, on which 40 carriers operate, resulting in 4,250 route carrier pairs. Table 3-3 shows the number of routes and scheduled departures for each studied carrier. The total traffic on these routes served by the selected carriers corresponds to the 75% of the total revenue passengers (BTS T-100 Domestic Segment) and the 59% of the total domestic commercial flights in 2009.

It is worth pointing that for all the legacy carriers (DL, UA, US, NW, CO) and the major Low Cost Carriers (WN, FL, B6, F9, VX) the flights between the studied 72 ASPM airports account for more than 75% of these carriers' total domestic departures. This shows that the U.S. domestic air traffic is very heavily concentrated at the largest commercial airports.

Carrier Name	IATA Code	Routes in Sample	Population Size (Departures)	Sample Size (Departures)	Sample / Population Size	Total Passengers Enplaned
Southwest Airlines	WN	860	1,128,004	951,534	84%	101,374,390
American Airlines	AA	352	547,854	493,295	90%	66,168,794
Delta Air Lines	DL	277	432,033	371,790	86%	55,708,779
United Air Lines	UA	281	375,103	333,549	89%	45,582,670
US Airways	US	309	412,665	379,336	92%	44,554,186
Northwest Airlines	NW	252	299,656	241,709	81%	32,624,283
Continental Air Lines	CO	178	260,172	243,333	94%	31,954,535
AirTran Airways	FL	230	251,441	188,372	75%	23,821,768
Skywest Airlines	OO	494	543,294	195,382	36%	20,025,879
JetBlue Airways	B6	214	193,312	162,665	84%	20,022,359
American Eagle Air.	MQ	334	468,891	167,320	36%	14,966,473
Alaska Airlines	AS	107	129,506	95,647	74%	14,060,609
Atlantic Southeast Air.	EV	203	295,424	38,307	13%	12,815,256
Expressjet Airlines	XE	290	310,792	154,000	50%	11,569,872
Mesa Airlines	YV	256	241,783	88,332	37%	10,816,722
Pinnacle Airlines	9E	338	260,331	71,870	28%	10,279,203
Frontier Airlines	F9	76	86,889	81,076	93%	9,073,307
Republic Airlines	YX	167	142,565	83,957	59%	8,701,157
Hawaiian Airlines	HA	37	73,015	31,314	43%	8,195,675
Mesaba Airlines	XJ	220	180,562	37,488	21%	6,551,978
Comair	OH	161	150,266	87,922	59%	6,019,605
Horizon Air	QX	101	110,243	33,447	30%	5,980,067
Chautauqua Airlines	RP	197	163,433	85,519	52%	5,762,187
Air Wisconsin Airlines	ZW	176	149,792	68,774	46%	5,284,263
Spirit Air Lines	NK	48	41,771	28,098	67%	4,953,800
PSA Airlines	16	120	122,701	27,419	22%	4,796,065
Shuttle America	S5	100	95,215	62,232	65%	4,542,406
Virgin America	VX	26	33,436	32,746	98%	3,747,220
Piedmont Airlines	17	107	128,122	25,066	20%	3,109,515
Compass Airlines	CP	62	50,553	27,049	54%	2,927,289
Colgan Air	9L	103	101,884	19,523	19%	2,531,872
Trans States Airlines	AX	84	68,459	28,204	41%	2,445,344
GoJet Airlines	G7	63	41,758	18,972	45%	2,114,189
Freedom Airlines	F8	66	40,563	18,659	46%	1,492,715
Sun Country Airlines	SY	10	9,581	3,345	35%	1,029,287
Commutair	C5	34	34,280	9,593	28%	818,679
Midwest Airlines	YX	18	10,969	9,755	89%	760,727
Island Air Hawaii	WP	17	18,488	1,334	7%	416,652
Gulfstream Int. Airlines	3M	27	26,414	5,421	21%	250,434
Aloha Airlines	AQ	8	9,521	1,994	21%	

Table 3-3: List of the Airlines in the Data Sample Sorted by Decreasing Passengers Enplaned

The rest of the carriers have lower fractions for two reasons. First, it is common that these carriers do not offer daily scheduled flights on many routes of their network. These flights were excluded from our dataset, based on the daily service requirement we had set. Second, most of these carriers are feeder airlines that operate under contract with the major carriers and serve secondary airports not included in our list.

Figure 3-1 shows that on more than half of the studied routes there was only one carrier offering daily service in 2009. Only 20% of the sample's routes were served directly by three or more carriers. This happens because hub airlines, which are the most dominant in our sample, compete on an origin-destination market and not on a route basis. This demonstrates the substantial role of the hub-and-spoke networks in the U.S. air transportation system and the trend towards traffic consolidation at hubs.
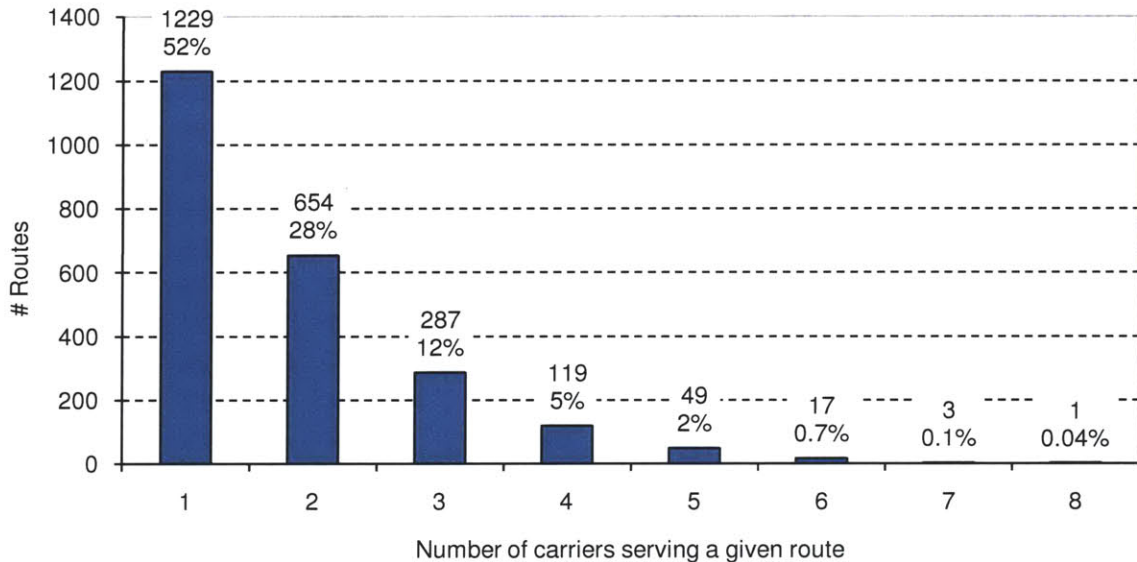
Figure 3-1: Carrier Presence on non-stop Routes

Figure 3-2 describes how the sample routes are distributed in terms of distance. It is easy to notice that the vast majority of the routes range from 250 to 1000 miles. This illustrates a trend towards more frequent short and medium haul flights between spoke cities and hubs, rather than long haul non-stop flights.
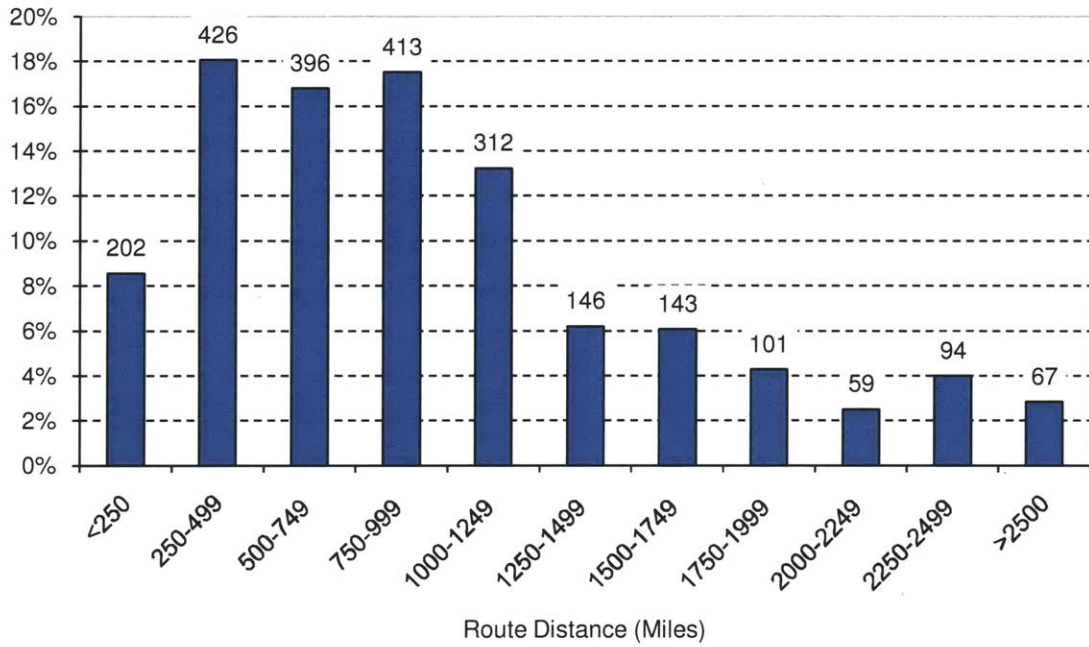
Figure 3-2: Route Distance in Data Sample
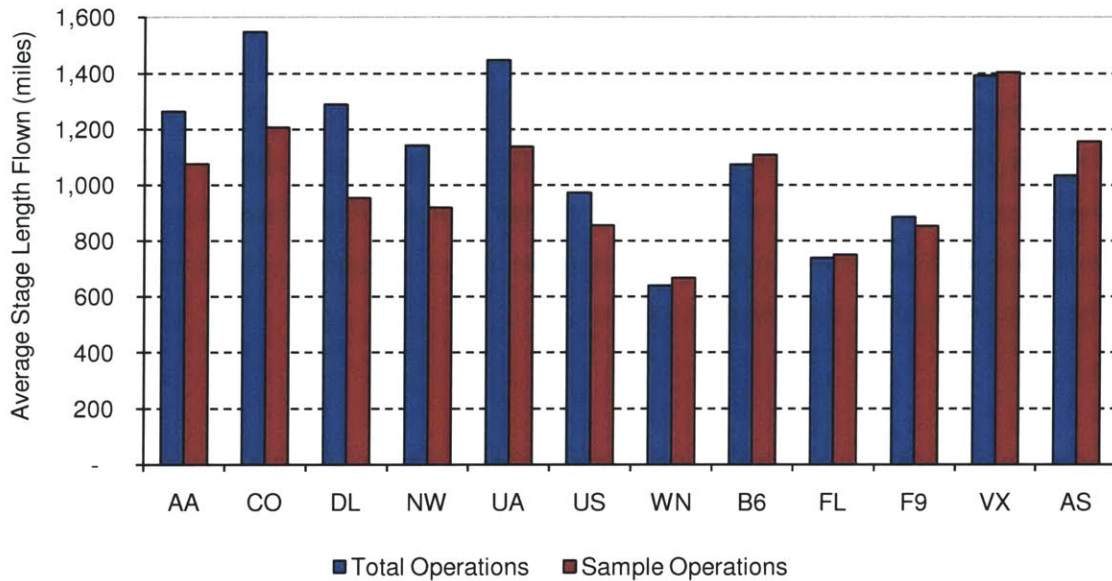
Figure 3-3: Average Stage Length Flown

In order to determine if our final sample is representative of the airlines' networks and operations we computed the sample's average stage length for the each carrier and compared it to the total average[27] (Figure 3-3). First of all, we can notice that for the low cost carriers the two estimates are very close, with Southwest having the shortest sample average stage length. Regarding the legacy carriers, the sample's average is lower than the total, as expected, because our sample does not include the long-haul international flights. However, it is clear that the sample's stage length follows the same pattern with the total average of each carrier across all legacy carriers. Therefore, the sample can be considered representative of the real network, at least for the airlines shown in Figure 3-3.

3.3 Extracted Data

The ASPM database offers a wealth of data that enable the study of carriers' performance for flights departing or arriving at a U.S. airport. For the purposes of our study we extracted 28 pieces of flight information from the ASPM database (Table 3-4). For each selected route and every carrier that was operating daily on it, we extracted all flights with scheduled departure time from 5am until midnight on every day of 2009. The final data set consists of 5,005,348 flights.

Although ASPM provides flight delay data, we do not make use of them. The reason is that it reports only the positive delays, whereas the negative delays are assigned as zero. However, as will be shown in Chapter 4.1, the distribution of early arrivals might have a great impact on carriers' decision to lengthen or shorten their scheduled block times. Therefore, we developed our own methodology for estimating the delay components used throughout this thesis. These are presented in the following section.

---

[27] Source: MIT Airline Data Project,
http://web.mit.edu/airlinedata/www/default.html

Field Description	Data
Scheduled Departure Year and Month (Local Date)	200901
Scheduled Departure Day (Local Day)	1
Scheduled Departure Hour (Local Hour)	7
Scheduled Arrival Year and Month (Local Date)	200901
Scheduled Arrival Day (Local Day)	1
Scheduled Arrival Hour (Local Hour)	11
Flight Carrier Code (ICAO)	AAL
Flight Number	25
Aircraft Tail Number	N5CFAA
IATA Aircraft Equipment Code from ETMS	B752
Departure Airport Code (ICAO)	BOS
Arrival Airport Code (ICAO)	LAX
ARINC OOOI/ASQP Present	Y
ETMS Present	Y
Aircraft Class (Commercial, Air Taxi, Freight, General Aviation, Military, Other)	C
Scheduled Gate Departure Time (Local)	7:45
Actual Gate Departure Time (Local)	9:01
Unimpeded Taxi out Time (minutes)	13
Actual Taxi Out Time (minutes)	19
EDCT Hold Time (minutes)	0
Estimated Airborne Time based on Flight Plan (minutes)	359
Actual Airborne Time (minutes)	352
Unimpeded Taxi In Time (minutes)	7.1
Actual Taxi In Time (minutes)	13
Scheduled Block Time (minutes)	385
Actual Block Time (minutes)	384
Scheduled Gate Arrival Time (Local)	11:10
Actual Gate Arrival Time (Local)	12:25

Table 3-4: Individual Flight Data Sample of ASPM

3.4 Anatomy of Flight Time – Definitions

The main objective of this thesis is to explore how airlines lengthen their planned flight times to account for the variability in flight operations and improve their performance with respect to delays and schedule reliability. Thus, we first develop a methodology for estimating the buffer time based on scheduled and actual flight data, and then we study how this is related to certain variables. To do so, it is essential to understand the structure of a flight and divide it into segments that are easily to define and study. For this purpose we define four time instants that occur during a flight (Figure 3-4):

- The **gate departure time** or **push-back time**: The instance at which the aircraft leaves the gate or the parking position. It is recorded when the pilot releases the aircraft parking brake after aircraft doors have been closed. If the aircraft left the gate and then returned to the gate before departing, the actual gate departure time is the last gate out time[28].

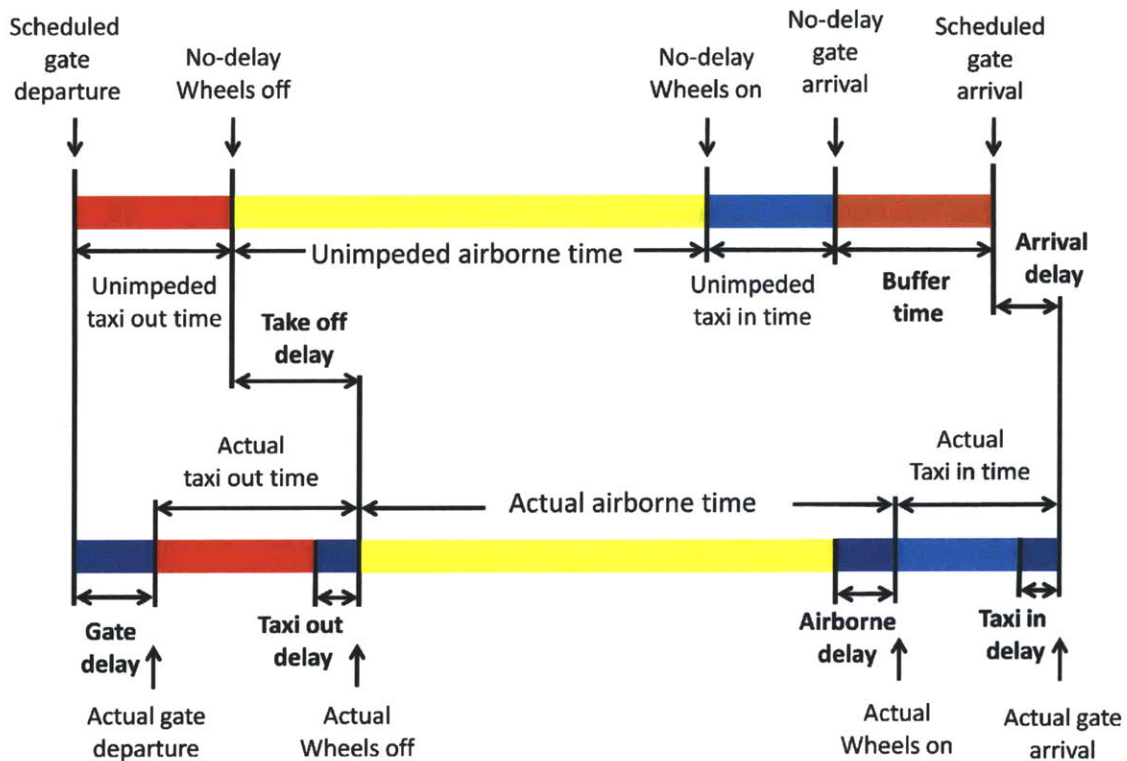

Figure 3-4: Break down of Scheduled and Actual Flight Times

---

[28]Source: BTS Technical Directive: On-Time Reporting, Number 17
http://www.bts.gov/programs/airline_information/accounting_and_reporting_directives/technical_directive_17.html

- The **wheels off time:** The moment at which the aircraft leaves the ground. It is recorded when the air/ground sensor on the landing gear is set to "airborne" state.

- The **wheels on time:** The moment at which the aircraft touches down. It is recorded when the air/ground sensor on landing gear is set to "ground" state.

- The **gate arrival time:** The moment at which the aircraft arrives at the gate or the parking position. It is recorded when the parking brake is applied. If the parking brake is not set, it is recorded when the passenger door opens[28].

The time interval between the gate departure and the gate arrival for a given flight is the **block time.**

Scheduled Block Time = Scheduled Gate Arrival Time – Scheduled Gate Departure Time

Actual Block Time = Actual Gate Arrival Time – Actual Gate Departure Time

Figure 3-4 illustrates how the scheduled and actual block time (upper and bottom bar respectively) can be broken down using the four time instances described above. The scheduled block time can be considered as the unconstrained gate-to-gate transit time under optimal conditions (nominal block time), lengthened by the buffer. Hence, the **buffer time** can be estimated as the difference between the scheduled block time and the nominal block time, where the latter is the sum of the unimpeded taxi-out, airborne and taxi-in times.

Buffer Time = Scheduled Block Time - Nominal Block Time

Nominal Block Time = Unimpeded Taxi-out Time + Nominal Flight Time + Unimpeded Taxi-in Time

The **unimpeded taxi-out time** is defined as the taxi-out time under optimal operating conditions, when neither congestion, weather nor other factors delay the aircraft during its movement from gate to takeoff[29]. ASPM estimates the unimpeded taxi-out time by calendar year and season for each airport and carrier. The reason for this distinction is that unimpeded taxi out time varies by carrier at a given airport depending on the location of that carrier's gates relative to the used runways, by airport depending on the airfield characteristics, and by season depending on the changes in normal operating procedures.

---

[29]Federal Aviation Administration, *Documentation For The Aviation System Performance Metrics (ASPM), Actual Versus Scheduled Metrics*, Washington DC, May 2002.

41

In ASPM the unimpeded taxi-out times are estimated through multiple regressions of available data, where the taxi-out time is the dependent variable and the length of the departure and arrival queues the independent variables. Specifically, the unimpeded taxi-out time is defined as the taxi-out time when the length of the departure queue is equal to 1 and the length of the arrival queue is equal to 0. The departure and arrival queues are estimated by analyzing actual flight data from the Out, Off, On, and In (OOOI) data provided by Aeronautical Radio, Incorporated (ARINC) database. The regression equation used by ASPM is the following:

$$TO_{a,c,s} = b_1 TOQ_{a,c,s} + b_2 TIQ_{a,c,s} + c$$

Where

> TO = taxi-out time
>
> TOQ = number of aircraft in taxi-out queue
>
> TIQ = number of aircraft in taxi-in queue
>
> a = airport
>
> c = carrier
>
> s = season
>
> b_1 = coefficient for TOQ
>
> b_2 = coefficient for TIQ
>
> c = constant

The unimpeded taxi-out time is calculated by setting the departure queue TQQ equal to 1 and the arrival queue TIQ equal to 0. The highest 25% of the values of actual taxi time were excluded from the regression in order to remove the influence of extremely large taxi-out times that cannot occur under optimal operating conditions. Because the queues' length is estimated based on the available data that do not include all actual flights, the actual queues might be longer and the actual unimpeded taxi-out times shorter.

The **unimpeded taxi-in time** is defined and estimated by ASPM similarly to the unimpeded taxi-out time. Again, the actual unimpeded taxi-in times might be shorter, because the actual departure and arrival queues are underestimated.

The estimation of the **unimpeded airborne time** has been one of the most critical steps of this thesis research. Using a very low bound for the unimpeded airborne time would result to an overestimation of delays and buffer. Given the wide range of actual airborne times (Figure 3-5) it was considered appropriate to use a percentile measure in the range of $5^{th} - 20^{th}$. These percentiles would exclude the outliers and provide a good estimate of the time that it takes to fly from A to B under optimal conditions. Sensitivity analysis performed by El Alj[14] had shown that any percentile in this range could be appropriate for use in the nominal airborne time estimation. We decided to use the 10th percentile in order to be consistent with most of the existing literature and the methodology used by ASPM. Additionally, we distinguished by carrier to account for differences in aircraft equipment across carries, and by month because as shown next in Chapter 4.3.1 there is a strong seasonality in airborne times.
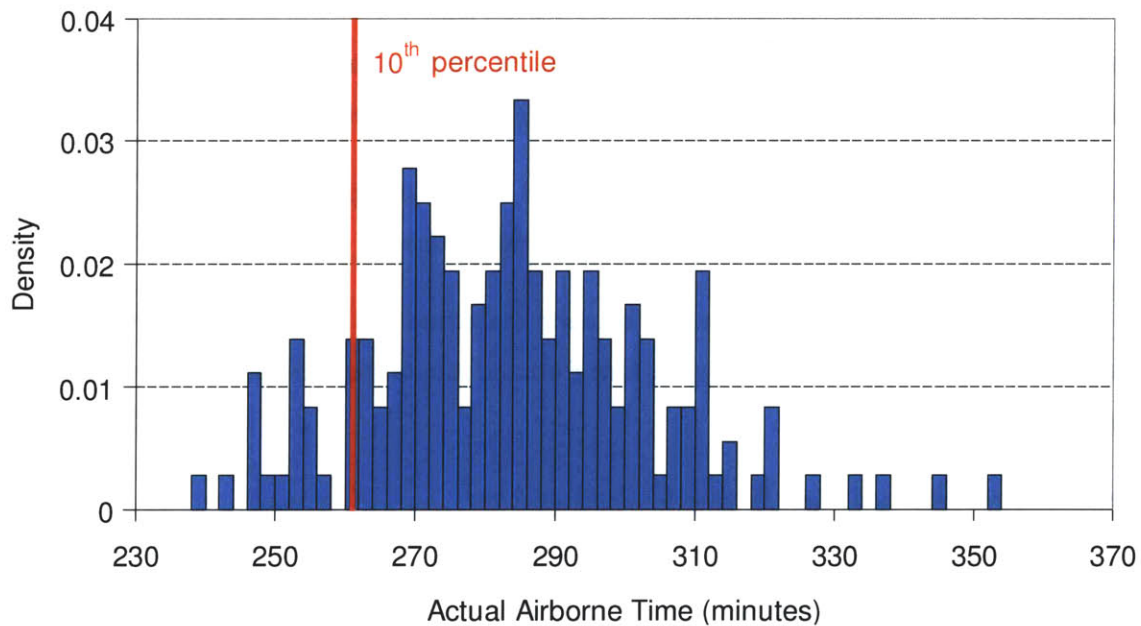
Figure 3-5: Actual Airborne Time distribution on LAX-EWR route in January 2009 by Continental Airlines

As opposed to scheduled block time, the decomposition of the actual block time is straightforward: **actual taxi-out time, actual airborne time, and actual taxi-in time**.

Actual Block Time = Actual Taxi-out Time + Actual Airborne Time + Actual Taxi-in Time

Actual Taxi-out Time = Actual Wheels Off Time – Actual Gate Departure Time

Actual Airborne Time = Actual Wheels On Time – Actual Wheels Off Time

Actual Taxi-in Time = Actual Gate Arrival Time – Actual Wheels On Time

The actual taxi-out, airborne and taxi-in times can be further broken down to their unimpeded times and their respective delays.

The **taxi-out delay** is the difference between the actual and the unimpeded taxi-out time:
Taxi Out Delay = Actual Taxi-out Time – Unimpeded Taxi-out Time

The **airborne belay** is the difference between the actual and the unimpeded airborne time:
Airborne Delay = Actual Airborne Time – Unimpeded Airborne Time

The **taxi-in delay** is the difference between the actual and the unimpeded taxi-in time:
Taxi-in Delay = Actual Taxi-in Time – Unimpeded Taxi-in Time

The **block delay** is the difference between the actual and scheduled block time. It can also be estimated by subtracting the buffer time from the sum of the taxi-out, airborne and taxi-in delay.

Block Delay = Scheduled Block Time – Actual Block Time

Block Delay = Taxi-out Delay + Airborne Delay + Taxi-in Delay –Buffer Time

However a flight can arrive late at gate not only because of delays experienced during these three flight segments, but also because of delays prior to push-back, named **gate delays**.

Gate Delay = Actual Gate Departure Time – Scheduled Gate Departure Time

It is still unclear if airlines try to incorporate the gate delays into the buffer time. One of the goals of this thesis is to provide insight into this area by studying how gate delays are correlated to buffer time and arrival delays.

The **arrival delays** are measured relative to scheduled arrival time and are defined as the difference between the actual and the scheduled gate arrival time:

Arrival Delay = Actual Gate Arrival Time – Scheduled Gate Arrival Time

Although this delay metric is the most commonly used, it is sensitive to changes in scheduled block time and does not estimate the true extent of a flight's delay. For example, let us assume two flights A and B with the same scheduled times (Figure 3-6). Flight A arrived behind scheduled arrival time and thus has a positive arrival delay. However, there is an additional amount of delay suffered by flight A that is hidden through buffer and is not reflected in arrival delay relative to schedule. The true extent of flight's A arrival delay is the delay relative to schedule plus the buffer time. On the other hand, flight B arrived prior schedule and has a negative arrival delay. Yet, flight B would have had a positive true delay if the scheduled block time had not been lengthened.

For computing the true extent of delays one must first estimate the delay free arrival time and then compute the arrival delay relative to this.

Delay Free Gate Arrival Time = Scheduled Gate Departure Time + Nominal Block Time

True Arrival Delay = Actual Gate Arrival Time –Delay Free Gate Arrival Time
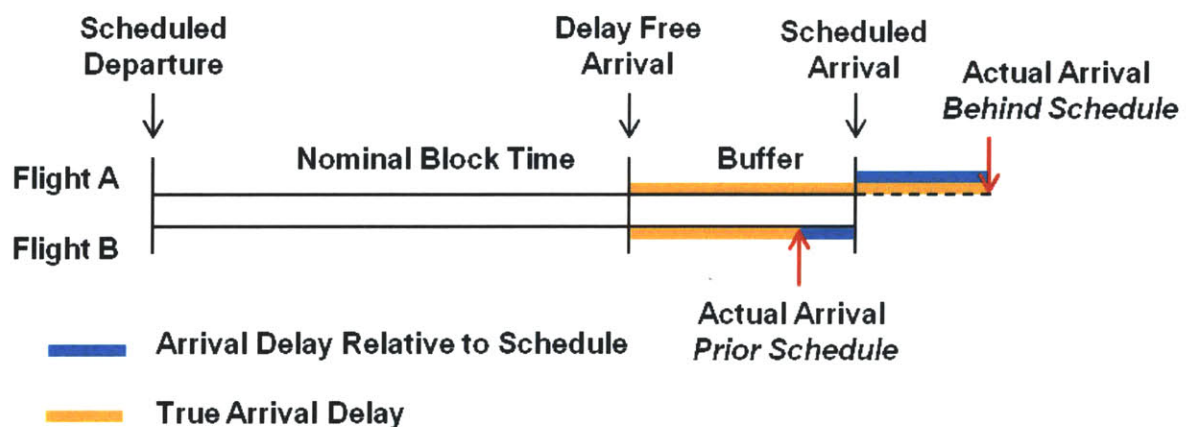
Figure 3-6: True Arrival Delays

3.5 Summary

In the first part of this chapter, we described the FAA ASPM database from which we derived the data used in this thesis. Then, we set the requirements for constructing the final data sample, which consists of 2,359 non-stop routes between 72 airports, served by 40 US carriers. Our final sample accounts for 59% of the total domestic commercial flights in 2009, the year of this study. The majority of these routes are short haul, and competition exists on less than half of them.

In the second part, we detailed the metrics we will use in our analysis and the methodology we will follow to estimate them. Emphasis is given in the description of the unimpeded taxi-out and taxi-in times, as these are defined by ASPM. For the estimation of the unimpeded airborne time, we decided to use the 10th percentile of the actual airborne time distribution in a one month period, a definition that has been used and tested in previous studies. The estimation technique for unimpeded airborne time is of great importance because it is used as the baseline for computing schedule buffer, the core metric in this thesis.

Chapter 4

Data Analysis

This chapter is comprised of three parts. In the first part, the distributions of the delay metrics and the buffer time for selected non-stop routes are presented and discussed, as well as the distributions of the mean and standard deviation on all 4,250 route carrier pairs of the studied database. In the second part, we analyze the correlation between the delay components and the buffer time, and we try to understand how uncertainty can influence airline schedule padding practices. Finally, in the last part, we further examine the most important factors that affect each flight segment, and formulate assumptions about the way that carriers adjust their schedules with regard to them.

4.1 Distributions of Flight Time Components

The time to travel between two airports varies significantly from flight to flight. The sources of variability differ for each stage of flight and can be classified into two categories: 'periodic' and 'stochastic'. Periodic factors are the season of the year, the time of day, and the day of week. Stochastic factors include the weather, flight path, runway configuration, gate assignments, aircraft and crew delays from the previous flight, as well as airline operations such as boarding processing, fueling, baggage handling, catering, and aircraft maintenance issues. As discussed in Chapter 2, each factor influences the flight time components differently, causing distinct levels of variations. Understanding the sources and the magnitude of variation for every component is useful before investigating the practices that US carriers follow to incorporate it in their schedules.
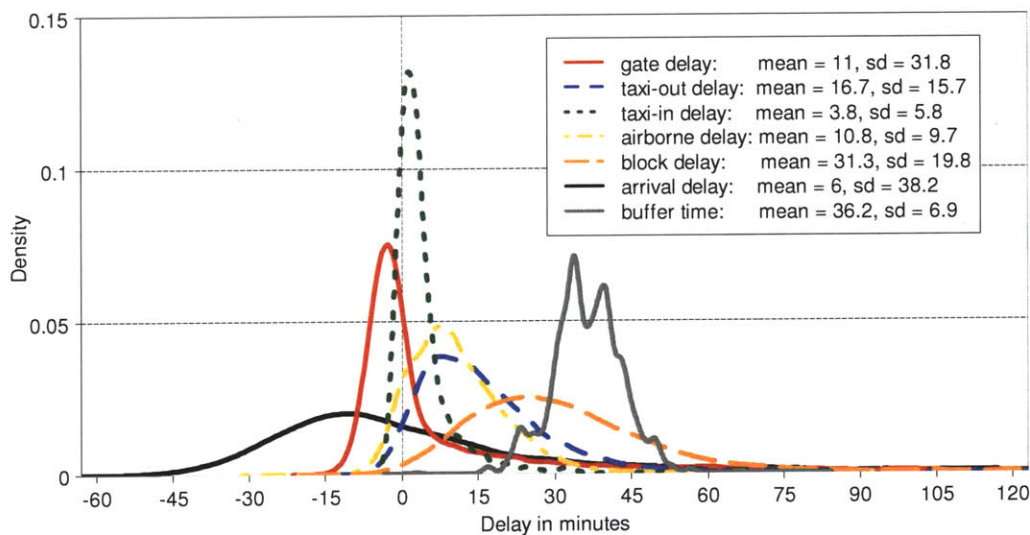
Figure 4-1: Distribution of Delay Components and Buffer Time on LaGuardia – Atlanta Route in 2009 by Delta (5,365 flights)

Figure 4-1 shows the Kernel density functions of buffer time and six delay metrics on the LaGuardia (LGA) - Atlanta (ATL) route for all scheduled flights operated by Delta during 2009. The taxi-out, taxi-in, airborne and block delays are relative to unimpeded times, the gate and arrival delays are relative to scheduled times. It should be noted that the block delay is the sum of the taxi-out, taxi-in and airborne delay, whereas the arrival delay is the sum of the gate and block delay, with the buffer time subtracted.

In analyzing the distributions of gate, taxi-out, and airborne delays, it is notable that they are all right-skewed with very thick right tails that can reach large positive values. Of great

interest is the fact that although 60% of the flights departed the gate before their scheduled time, the mean gate delay is 11 minutes and the standard deviation is 32 minutes. This significant standard deviation value is driven by the large number of flights that left the gate very delayed: 14% later than 30 minutes and 8% later than one hour. The taxi-out and airborne delay density functions have a lower kurtosis and thinner tails, with less than 2% of the flights exceeding the one hour delay for any of these two metrics. The block delay density function is more diffused as a result of the aggregated variation of the taxi-out, airborne and taxi-in delays. Its mean exceeds 30 minutes, and its standard deviation is 20 minutes.

The arrival delay, the metric of the greatest interest to travelers and airlines, has a very wide distribution. Its range starts from minus one hour and exceeds three hours. In contrast to public perception, more than half of the flights (55%) arrived at the gate earlier than scheduled, with 28% arriving more than 15 minutes earlier. On the other hand, 27% of the flights arrived later than 15 minutes after scheduled arrival time, and 9% later than one hour behind schedule.

Given the very high average values of gate and block delays on this route, Delta lengthened its scheduled block times by a significant amount of time to achieve the above mentioned on-time performance. An optimal padding strategy maximizes the number of flights that arrive on time, and at the same time minimizes the total negative delay. The long right tail in the gate and block delay distribution reveals the existence of flights that require a large amount of buffer time to arrive on time. Because these flights are distributed over a wide time range, the gains in on-time performance would be very small compared to the cost of underutilizing the aircraft, the gates and the crew for every minute of early arrival. The buffer added by Delta on this route ranges from 15 to 60 minutes with an average of 36 minutes. Later in this chapter, we study how buffer changes by time of day, season, day of week, and aircraft type.

However, the shape of these distributions can vary significantly for different routes. Figure 4-2 provides a comparison among five non-stop directional routes: Newark (EWR) – Los Angeles (LAX): 2,672 flights; Los Angeles (LAX) – Newark (EWR): 2,657 flights; Denver (DEN) – Seattle (SEA): 5,443 flights; Boston (BOS) – LaGuardia (LGA): 10,563 flights; and Atlanta (ATL) – Miami (MIA): 6,364 flights. It is evident that each route has its own characteristics that determine the thickness and the length of the tails, the skewness, the peak values, and the overall shape of the distributions.
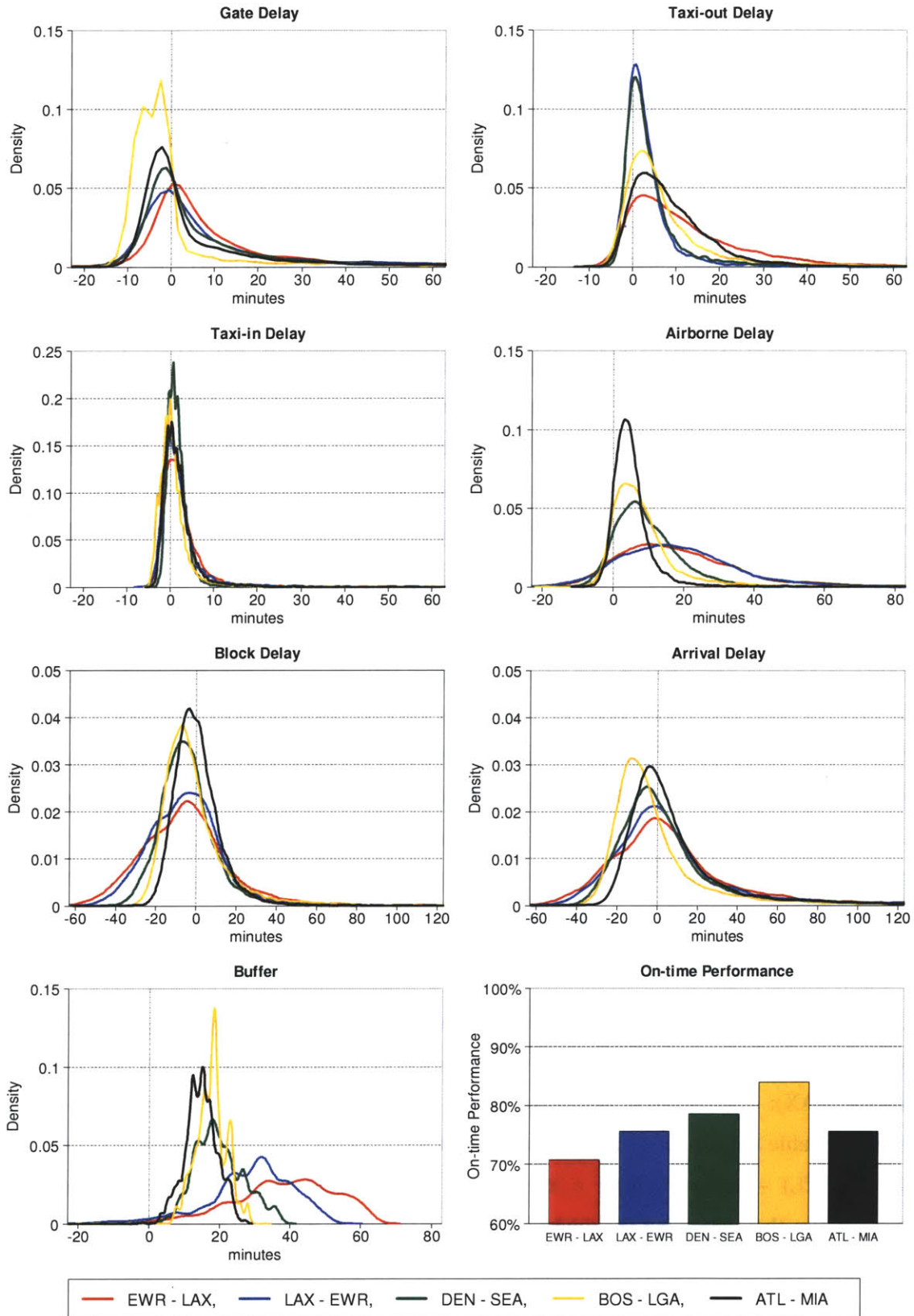
Figure 4-2: On-time Performance and Kernel Density Functions of Delay Components and Buffer Times

The EWR-LAX and LAX-EWR routes have very similar distributions for all metrics except the taxi-out delays and the buffer times. Their commonalities are attributable to some extent, to the similar weather conditions on the two routes. Since on the EWR-LAX route the aircraft fly from east to west and on LAX-EWR from west to east, one might expect that the jet streams would cause opposite effects to the airborne delays on these two routes. However, airborne delays are estimated using as baseline the 10th percentile of the actual airborne times distribution of a given month. Therefore, they measure the dispersion of actual times around the baseline. Large variation in the wind strength causes long airborne delays in both directions, whereas small variation causes short delays, even for very strong winds. For example, on the EWR-LAX route, March had on average longer nominal and actual airborne times compared to January (Figure 4-3). Thus, we can reasonable assume that jet streams in March were stronger. Though, March has shorter airborne delays, meaning that the winds in this month were constantly strong, while in January they were more variable.
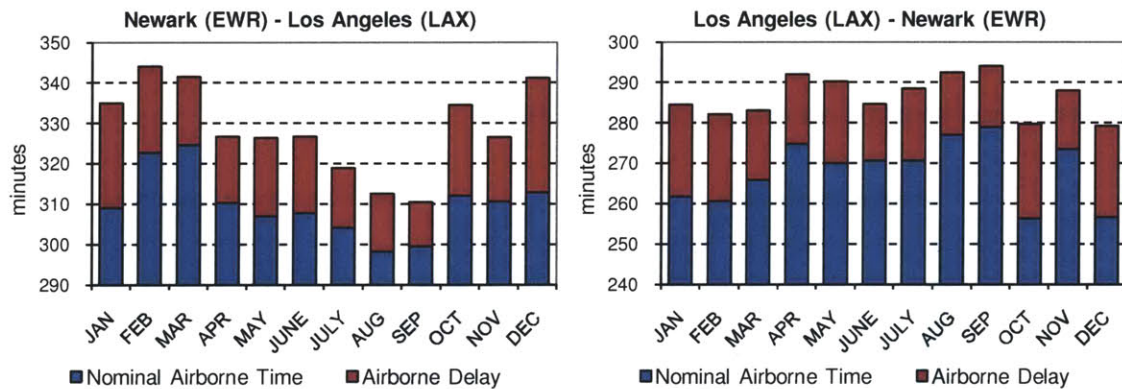
Figure 4-3: Monthly Average Nominal Airborne Times and Airborne Delays on Los Angeles (LAX)-Newark (EWR) and Newark (EWR)-Los Angeles (LAX) Route

The gate delay and the taxi-out delay density functions of the EWR-LAX route have the thickest right tails compared to the other routes. Although the carriers operating on this route have used longer buffer times, this route still has the worst on-time performance (American: 73%, Continental: 70%). Still, this route has the largest percentage of flights that arrived at the gate 20 minutes earlier than scheduled (17%). This illustrates the difficulties in incorporating uncertainty in the schedule, since adding more buffer would cause a disproportional increase in the negative delays relative to the improvement of on-time performance. However, a very important issue is the extent to which this uncertainty is caused by the airline's schedule, operational weaknesses, and poor overall performance, rather by external factors that the airline cannot forecast and handle effectively. This aspect is examined in Chapter 5.

On all five routes, the taxi-in delay density functions are very narrow, with their peak close to zero. Only a very small portion of flights had delays longer than 10 minutes. Taxi-in delays can be caused by congestion on the taxiways and from gate unavailability. Additionally, in the case that gate arrival time is measured by the instance at which the aircraft door opens, taxi-in delays can be the result of lags at attaching the gate airbridge to the aircraft. As shown in Figure 4-4 below, these observations can be expanded to all 4,250 route-carrier pairs studied in this thesis. The average taxi-in delays are, in the vast majority of the cases, negligible and only in four cases they exceed the ten minutes (all at JFK). Similarly, the standard deviation is also very small, indicating that taxi-in time is not a time component that airlines need to take into account when lengthening their schedules.

On the other hand, the average taxi-out delays are more widely distributed, and on 10% of the route-carrier pairs they exceed the 10 minutes. The standard deviation of taxi-out delays is even larger, ranging between 10 and 20 minutes on 21% of the pairs. It is remarkable that JFK is the origin airport on 34 out of the 51 routes on which the standard deviation of taxi-out delay was longer than 20 minutes. This demonstrates the importance of investigating in depth the role of airports in generating taxi-out delays (Chapter 4.3.4)

Figure 4-4: Average and Standard Deviation of Taxi-in and Taxi- out Delays on the 4,250 Route-Carrier Pairs in 2009

The distributions of the mean and the standard deviation of airborne delays look very similar (Figure 4-5), with their values concentrated between 3 and 12 minutes (87% for the average and 92% for the standard deviation). Because different methodologies have been used for the estimation of the taxi-out, taxi-in and the airborne delays, a comparison of their distributions would be misleading. However, it is clear that taxi-out delays and airborne delays are the main contributors to block delays.

Among all flight segments, the time prior to pushback suffers from the highest uncertainty. Although gate delays are on average similar to airborne delays, except that their tails are longer, their standard deviation is substantially higher. On 70% of the route carrier pairs the standard deviation of gate delays is larger than 20 minutes, and on 50 routes it exceeds 40 minutes. It must be noted that EWR is the destination airport on 29 out of these 50 pairs. This large variation in gate delays is driven to some extent by the Ground Delay Programs (GDPs).

A GDP is initiated by FAA every time that a degraded arrival capacity or excess demand is expected for the next few hours at a specific airport. Flights that have scheduled arrival times during the time period of an active GDP and have not departed yet, are delayed on the ground and are assigned new departure and arrival times on a "first-scheduled, first-served" basis. The aim of GDPs is to match arrival demand with arrival capacity, and the rationale behind is that it is more convenient and less expensive for a flight to suffer unavoidable delays on the ground rather than en route.

Despite the existence of many route carrier pairs with lengthy taxi-out and airborne delays, the distribution of the mean block delays relative to schedule suggests that if the aircraft left the gate on-time, the vast majority of the route carrier pairs (86%) would have negative average arrival delays (Figure 4-5). This happens because airlines add large amounts of slack time to their scheduled block times, as indicated by the distribution of average buffer times. The mean buffer of all studied flights in 2009 is 18.5 minutes, and 40% of the route carrier pairs have an average buffer longer than 20 minutes. The standard deviation of block delays varies from 3 to 37 minutes, with an average of 13 minutes. In Section 4.3.2 it is shown that on many routes the gate and the taxi-out delays are a function of the scheduled arrival time and follow an increasing trend through the course of the day. Thus, we would expect that buffer time would also vary by the time of day so as to improve the on-time performance. Nonetheless, the distribution of standard deviation of buffer time does not support this assumption, at least on the majority of the routes, indicating that carriers use almost constant buffer times across all flights on a given route.
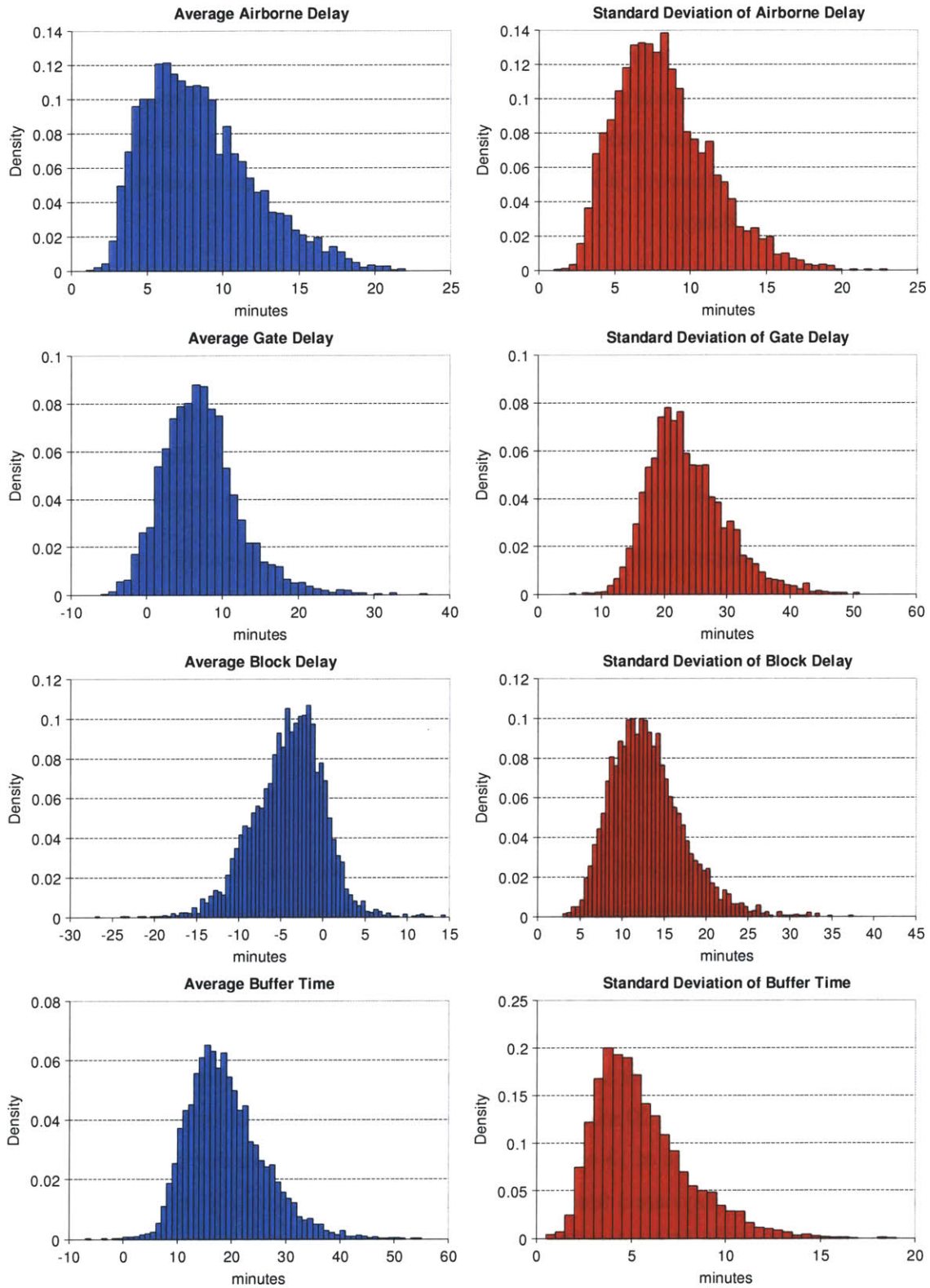
Figure 4-5: Average and Standard Deviation of Airborne, Gate and Block Delays, and Buffer Time on the 4,250 Route-Carrier Pairs in 2009

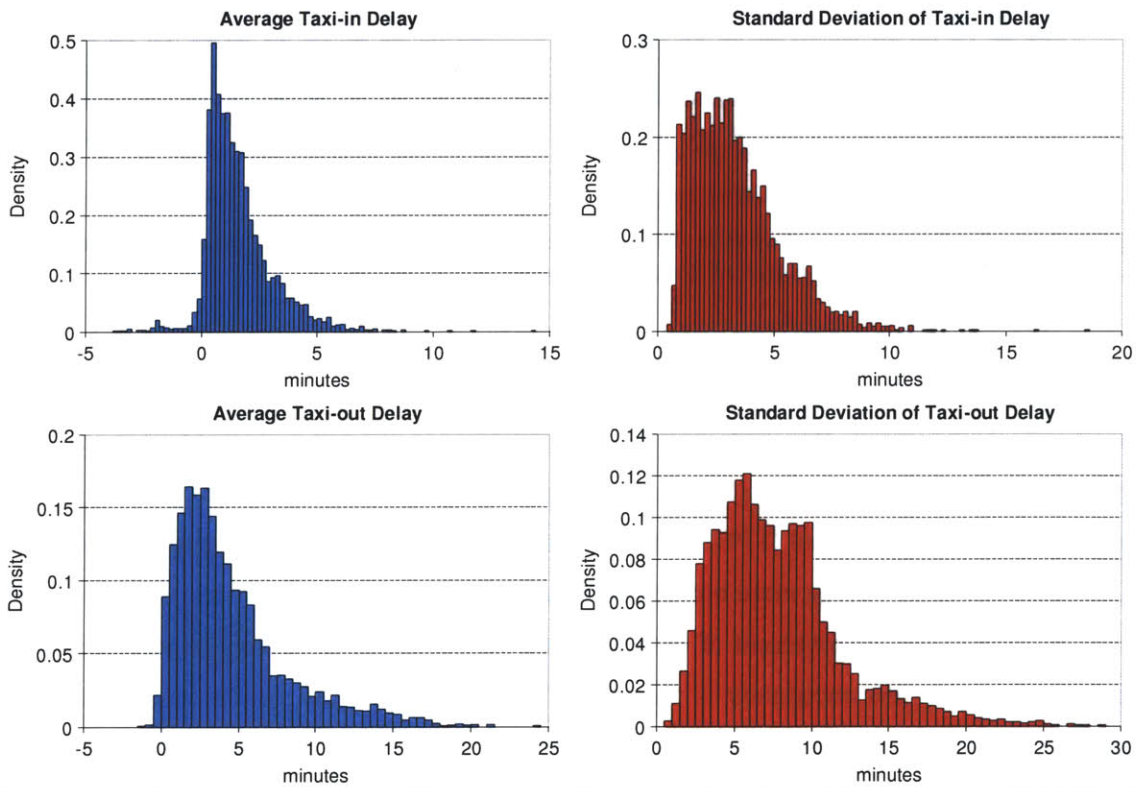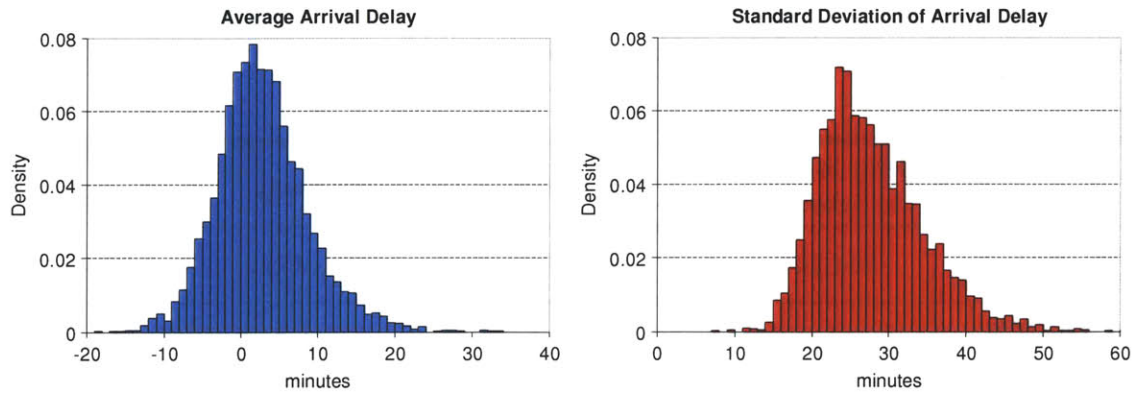*Figure 4-6:* Average and Standard Deviation of Arrival Delay on the 4,250 Route-Carrier Pairs in 2009

The distribution of the average arrival delays (Figure 4-6). approximates the Gaussian, with a mean of three minutes and a range from -18 to 34 minutes However, the average delay is not sufficient to obtain an adequate understanding of airlines' performance, because it does not capture the degree to which these delays are predictable. For example, if all flights on a certain route are delayed by a constant amount of time (e.g. the standard deviation of arrival delays is small), the airline could improve drastically its on-time performance either by rescheduling the departure time if the delays are experienced prior to pushback, or otherwise, by increasing buffer. On the contrary, the average arrival delay might be zero, but the standard deviation very high, with many flights arriving very early and many others arriving very late. The distribution of the standard deviation of arrival delays in Figure 4-6 is the result of the large extent of uncertainty to which flight operations are exposed.

Table 4-1 shows a classification of the delays suffered by the flights in our data sample. Of the 1,846,270 aircraft hours of true delays (as defined in Chapter 3), 1,423,455 hours are attributable to flights that arrived late, and 413,815 hours to flights that arrived before schedule. The first can be further divided to delays relative to schedule (60%) and schedule buffer (40%). The true delays of flights that arrived before schedule are all attributed to excessive schedule buffer. Assuming no change in the actual times of our flights, the extra buffer time that would be required to diminish all delays relative to schedule under optimal time allocation is 17%.

However, these numbers do not capture the delays associated with the disrupted connecting passengers. Their delays are a function of the arrival delays on each leg and their connecting time. In fact, long delays on the first leg can cause a missed connection and

subsequently very large delays for passengers. The shortcomings of the on-time performance metric in reflecting the true extent of passenger delays is addressed by Bratu and Barnhart[30].

	Delay Relative to Schedule	Buffer	True Delay	Buffer / Total True Delay
Flight Arrived behind Schedule	864,249	568,206	1,432,455	30.8%
Flight Arrived prior Schedule	-599,783	used: 413,815 non used: 552,583	413,815	22.4% 29.9%
Total	264,466	1,534,604	1,846,270	83.1%

Table 4-1: Delays in the Data Sample (in aircraft-hours)

4.2 Correlation of Flight Time Components

In this section we perform correlation tests among the delay metrics and the buffer time, using the Pearson product-moment correlation coefficent. These tests provide information on the interaction of the different metrics, and are also necessary for the succesful construction of the linear regression models discussed in Chapter 5. Note that if two variables have a large correlation coefficient, then they should not be used together as explanatory variables in the models.

The Pearson's correlation parameter between two random variables X and Y gives a quantitive measure of how well we can predict the value of one variable, knowing the value of the other. It also indicates the existence of a linear relationship among the two variables, and provides a measure of their dependence. The sample correlation coefficient, denoted by r, is defined as the estimate of the covariance of the two variables divided by the product of their standard deviations:

$$r = \frac{COV(X,Y)}{S_X S_Y}$$

---

[30] Bratu, S., Barnhart, C., *An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data*, Journal of Transportation and Statistics, Vol. 1, 2005

Figure 4-7: Density Functions of Pearson's Correlation Coefficient on the 4,250 Route-Carrier Pairs in 2009

Studying the plots of Figure 4-7, we observe a small correlation between buffer time and arrival delays that is negative on most route-carrier pairs. This result is in accord with our assumption that adding more buffer in the schedule would cause a small reduction in reported delays. Regarding the route-carrier pairs where longer buffer is slightly correlated with increased arrival delays, we expect that these are between highly congested airports that operate at their full capacity, where even the application of additional buffer does not result in better arrival statistics.

As expected, block delays relative to schedule decrease with buffer, but on many pairs the absolute value of the correlation parameter is very small. The reason for this small correlation might be that carriers on these routes do not account for the variability in actual block times and use almost constant buffer times.

Airborne and taxi-out delays are very similarly correlated to buffer, as illustrated by their distributions. In general, the longer the expected airborne or taxi-out delays, the more airlines lengthen their schedules. The area of negative correlation can be attributed to two reasons:

- Airlines used shorter buffer than necessary because of inaccurate forecasting of airborne and taxi-out delays.

- Airlines tried to allocate buffer in a way that would improve their aggregate on-time performance statistics. Therefore, they removed buffer from flights that were operating under large uncertainty (e.g. during peak hours), and allocated it to flights where it was more likely to achieve better performance.

In the previous section we showed that taxi-in delays are negligible and thus we put forward that airlines do not take them into account for adjusting buffer. The distribution of the correlation parameter of taxi-in delays with buffer further supports this argument.

Another issue that should be addressed is whether airlines add buffer to compensate for gate delays. One argument against this practice is that if an airline expected gate delays on a certain route, then they should icrease the turn-around times by moving departure times later in the schedule. Longer turn-around times absorb propagated delays, thus improving schedule reliability. On the other hand, they restrain the time that carriers can allocate to flight operations and reduce gate availability. Moreover, a small portion of gate delays occur during the boarding process (e.g. delayed passengers) and can not be absorbed by the turn-around times. Nonetheless, gate delays are very unpredictable and the gains of trying to incorporate them in schedule is questionable. The distributions of the correlations between gate delays and buffer, as well as gate

delays and block delays are symmetric around zero. This indicates that, at least on the majority of the routes, carriers do not adjust buffer based on gate delays.

Furthermore, the relationship between taxi-out and gate delays is very complicated and depends on the sources of delays, the arrival and the destination airport. Overall, their correlation is very small because gate delays can be caused by a large number of factors and suffer from huge uncertainty. On the other hand, taxi-out delays are affected mostly by the weather and the traffic at origin and destination. When bad weather conditions exist at the origin airport, we expect that both delay components are affected in a similar way, resulting in a positive correlation between them. In case of bad weather at the arrival airport, their association depends strongly on the effectiveness of GDPs. Ideally, GDPs should hold the flights at the gate until they absorb all expected delays, so the taxi-out delays are minimized. This implies a negative correlation between taxi-out and gate delays. Indeed, their correlation parameter when only flights under GDPs are studied is -0.14. Although this number is very small, it illustrates the negative relationship we expected. Furthermore, the correlation parameter of the time that flights are held on ground under GDPs with gate delays is 0.77, and with taxi-out delays is 0.30. This suggests that GDPs hold the flights longer times at the gates rather on the taxi-way.

The correlations between taxi-out and airborne delays (Figure 4-8) are very small, close to zero and positive on the majority of the routes. This occurs because the main cause of airborne delays is weather conditions that affect taxi-out delays as well. Nonetheless, we expect that on long haul flights that have departed after a large taxi-out delay pilots might speed up the aircraft, if possible, to recover from the delay. In this case, the correlation should be negative. The latter is further supported by the distribution of the correlation parameters between airborne and gate delays, where more routes have a negative correlation.

As shown in Figure 4-8, block delays relative to schedule are mostly attributed to airborne and taxi-out delays, and less to taxi-in delays. This is expected given the magnitude of the three delay components observed in Figures 4.4 and 4.5. It is important to note that the range of the correlation parameters is very large and therefore it would be wrong to generalize any observations. For example, on routes where long taxi-out delays are associated with long block delays, the effect of airborne delays might be negligible and vice versa. Similarly, airborne and taxi-out delays have in general a moderate correlation with arrival delays relative to schedule, whereas taxi-in delays have a smaller one.
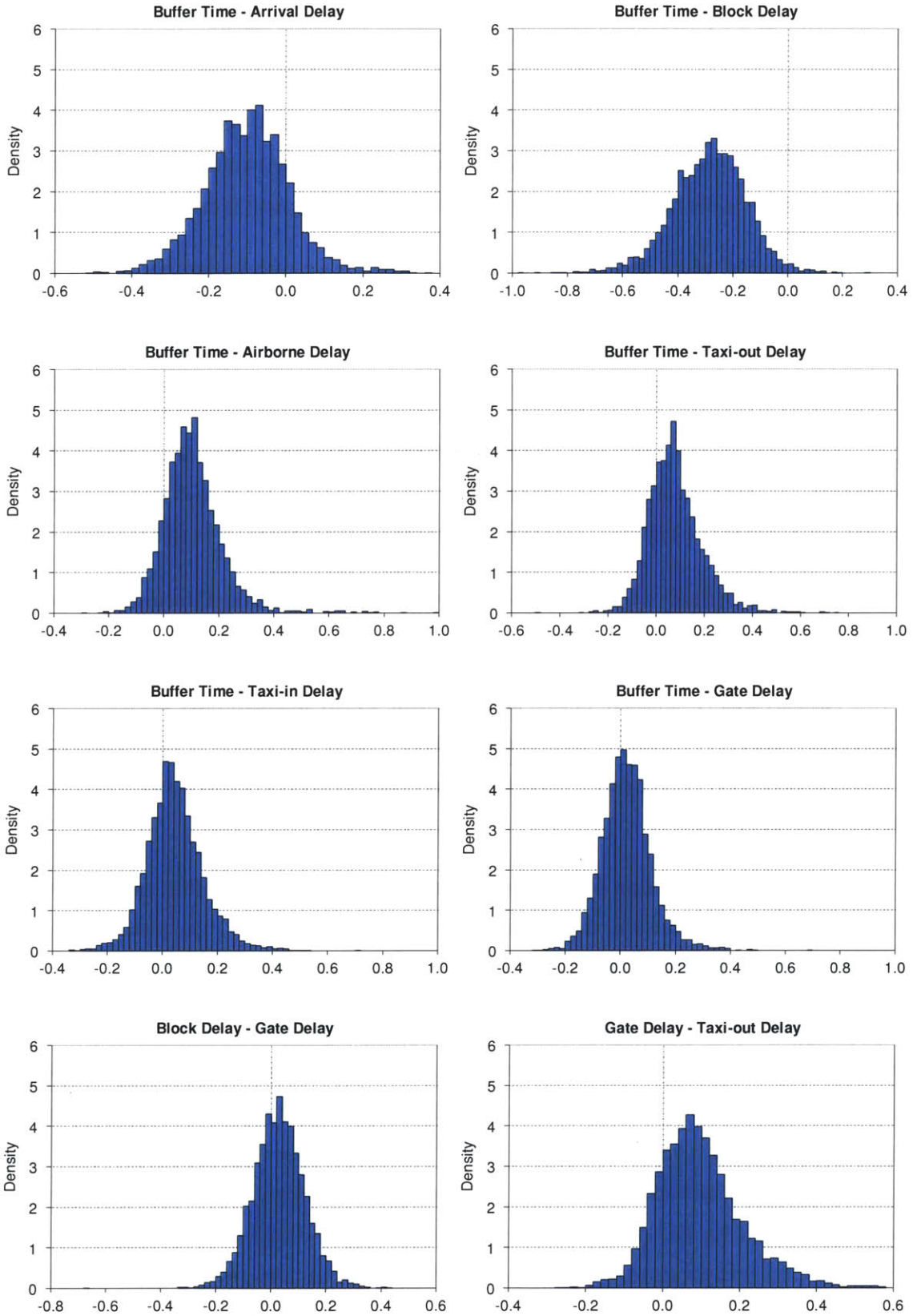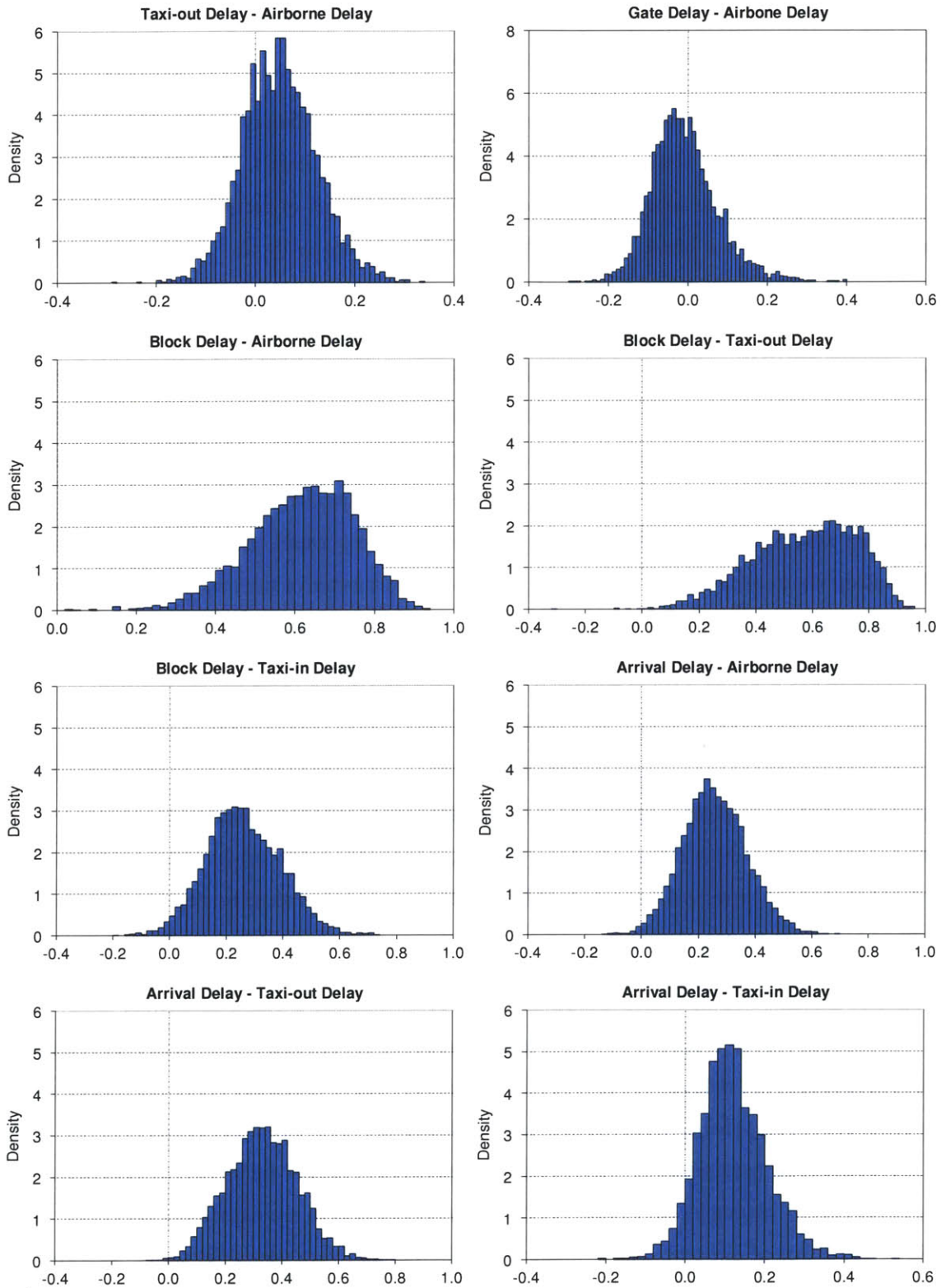
Figure 4-8: Density Functions of Pearson's Correlation Coeeficient on the 4,250 Route-Carrier Pairs in 2009
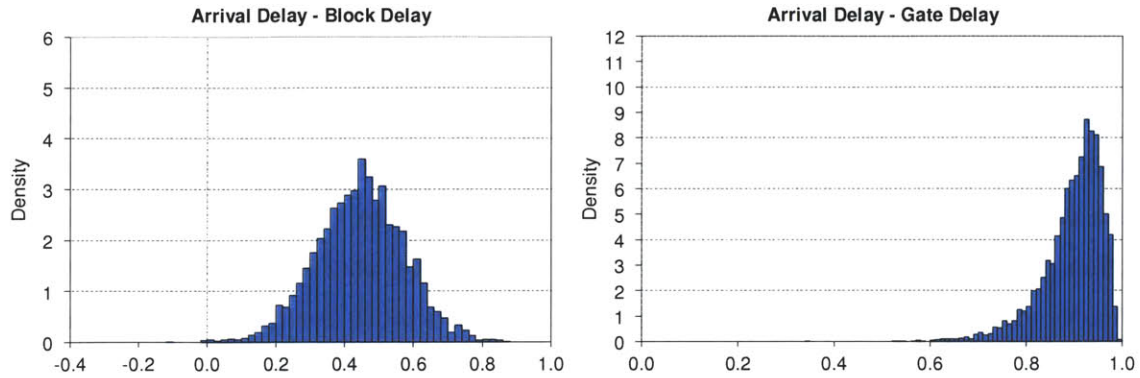
Figure 4-9: Density Functions of Pearson's Correlation Coeeficient on the 4,250 Route-Carrier Pairs in 2009

A comparison between the correlation distributions of block and gate delays with arrival delays (Figure 4-9) illustrates the weakness of airlines in handling effectively the gate delays as opposed to block delays. With a correlation parameter larger than 80% on most route-carrier pairs, longer gate delays result directly in increased arrival delays. The impact of block delays relative to schedule is much smaller, but still substantial.

Summarizing, the correlation among the delay metrics and the buffer time varies significantly by route, and thus the information we can get from these ditsributions is quite limited. Therefore, it is useful to examine closely the impact of each factor on delays and buffer.

4.3 Factors affecting flight time components

In the two previous sections we studied the magnitude and variability of buffer and delay components. We also made assumptions about the causal factors of delays, and discussed how these are validated by the correlation distributions between the several components. In this section we will take a closer look at each factor individually, and will explore how they can influence delays and padding practices, as these are described in our database. This provides guidance as to which variables to include in the regression models developed next.

4.3.1 Seasonality Effect of Airborne Time

Seasonality has a very strong influence in flight operations, and can affect delays through two distinct ways: traffic and weather.

Demand for air travel varies by season, and this does not only affect the load factors but also the number of flights operated. Figure 4-10 shows how domestic US flight departures fluctuated through 2009, with fewer operations in winter and a peak in the summer. Particularly, July had 20% more traffic than February. The traffic volume affects directly the airport and airspace congestion. Based on queuing theory, when a system operates close to its capacity (which is the case in many U.S. airports and airspace sectors), a small variation in demand can have enormous impacts on capacity.
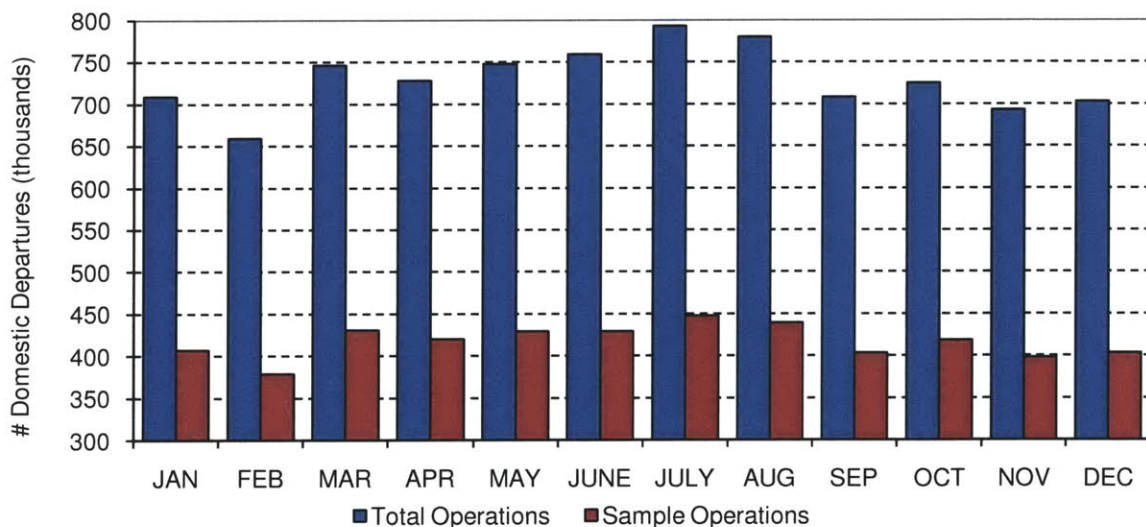
Figure 4-10: Scheduled Domestic Commercial Flights Performed in 2009
Source: Bureau of Transportation Statistics, T-100 Market and Segment

As has been extensively discussed in the relevant literature and in this thesis so far, weather is one of the most significant causes of delays. First, bad weather conditions (such as low visibility, rain, snow, and thunderstorms) at the origin or destination airport area increase the minimum required separations for take-off and landing, and reduce the airport capacity. This has a direct impact on gate, taxi-out, taxi-in, and airborne delays. Second, bad en-route weather conditions reduce the airspace capacity and result in airborne delays through longer flight paths. Lastly, airspace congestion can cause Air Traffic Control (ATC) to hold flights on the ground imposing additional gate and taxi-out delays.

Besides airspace congestion, jet streams have a seasonal effect in airborne times that is a function of wind speed and aircraft direction. Studying the magnitude of this effect allows us to decide if we can use a constant nominal airborne time throughout the year, or if we should limit our analysis to a monthly basis. Furthermore, it provides helpful information on whether carriers should adjust their scheduled times respectively, and if so, on which routes.

To study this effect, we followed a three steps method:

1.   The 2,359 non-stop routes were classified to 36 groups based on the angle of a rhumb line between the origin and the destination airport;

2.   Seasonality indexes were computed for every route-carrier pair and month;

3.   A chi-square homogeneity test was performed between the seasonality indexes on each route-carrier pair and the weighted averaged seasonality indexes of all other routes in the same angle group.

To classify the routes based on the angle of a rhumb line between the origin and the destination airport, the latitude and longitude coordinates of the 73 studied airports were extracted from the National Airspace System Resources (NASR)[31]. Then, for each of the 2,359 directional routes, the angle of the constant bearing path between the origin and the destination airport, i.e. the line that crosses all meridians at the same angle, was computed. The 2,359 directional routes were broken down to 36 groups of $10°$ based on their path-angle. The first group is from $355°$ to $5°$, with $0°$ being the true North. Figure 4-11 shows the distribution of the route classification. Notice that the number of routes increases as we move from $0°$ to $90°$ and then decreases again. The distribution is not perfectly symmetrical around $180°$, because five routes are not studied in both directions.

---

[31] Source: FAA U.S. & Territories Airport Lookup,
https://airports-gis.faa.gov/airportsgis/airportLookup/index.jsp?category=nasr
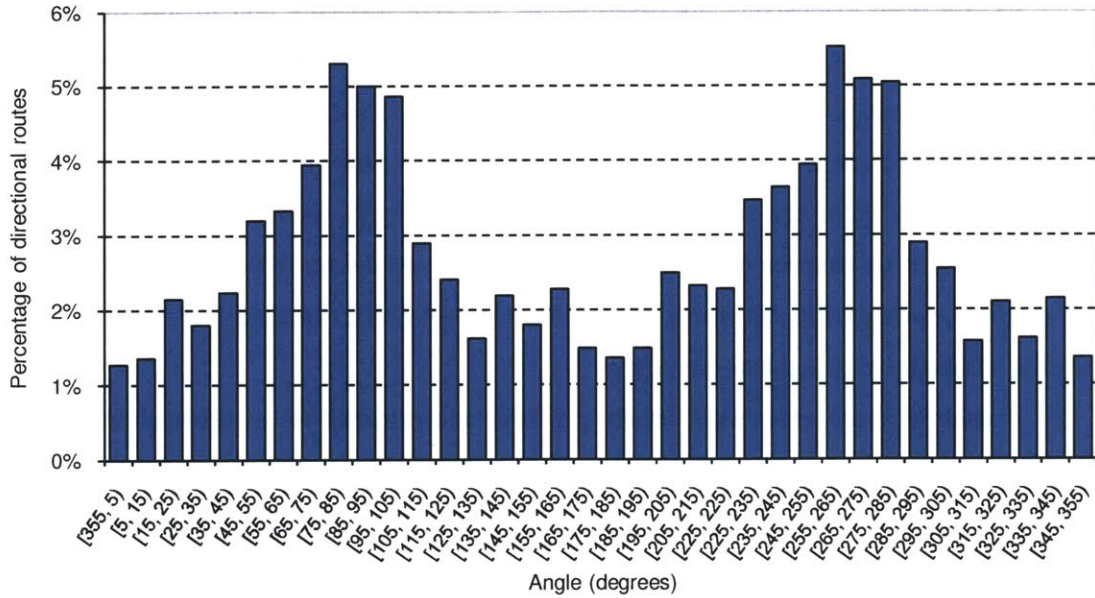
Figure 4-11: Route Classification based on Angle

Next, for every route-carrier pair, we computed the average actual airborne time, first for the whole 2009, and then for every month. We distinguished by carriers to account for potential aircraft differences across them. The seasonality index $T_{i,j}$ for route-carrier pair i and month j is defined as the ratio of the average actual airborne time of route-carrier pair i in month j to the average actual airborne time of the same route-carrier pair i during the whole season:

$$T_{i,j} = \frac{E[Actual\ Airborne\ Times\ on\ Route-Carrier\ i\ in\ Month\ j]}{E[Actual\ Airborne\ Times\ on\ Route-Carrier\ i\ in\ 2009]}$$

Figure 4-12 illustrates the influence of jet streams to the routes of four angle groups: [85°, 95°), [265°, 275°), [355°, 5°) and [175°, 185°). We notice that the seasonal indexes in the west to east direction are almost antisymmetric relative to the indexes of the east to west direction. This result is expected because jet streams cause opposite effects on each direction. On the other hand, flights from south to north and north to south appear to be affected in a slightly similar way by the jet streams. The reason is that on these routes the wind is vertical to the aircraft vector, and therefore it has the same effect independently of the aircraft direction. Differences among the two distributions could be attributed to the different flight paths.
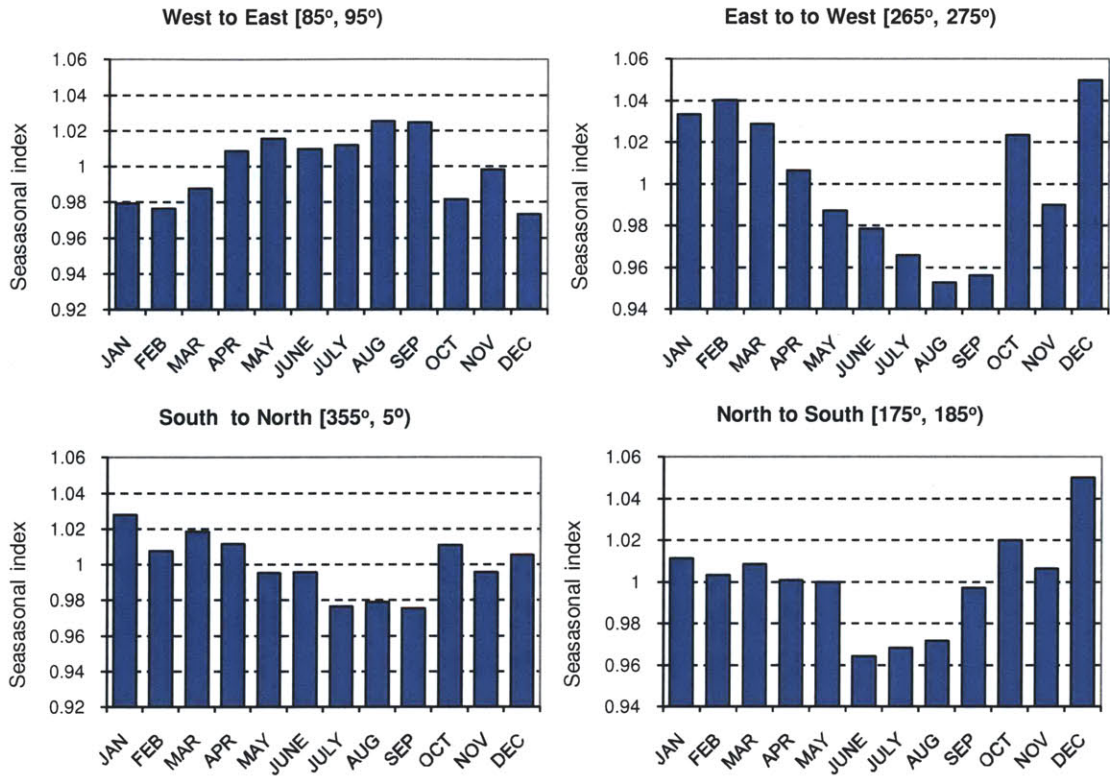
Figure 4-12: Average Seasonality Indexes

Last, chi-square homogeneity tests were performed between the seasonality indexes on every route-carrier pair i and the weighted averaged seasonality indexes of all routes that are in the same angle group with i. In order to ensure that we have enough observations in all 12 months for every route-carrier pair, we excluded those pairs that had less than 20 operations in at least one month. In total, 3,655 directional route-carrier pairs were studied.

Figure 4-13 illustrates the percentage of route-carrier pairs for which the null hypothesis of identically distributed seasonality indexes cannot be rejected at a 20% significance level. It is notable that for the routes with directions close to north-south and south-north, the null hypothesis is more easily rejected than for the west-east and east north directions. This means that on these routes the impact of the jet streams is less significant and their airborne time distributions have fewer seasonal similarities. One other reason that contributes to this result is the fact that the routes from south to north or form north to south have a smaller average distance (Figure 4-14) and thus are exposed for less time to jet streams.

Given the seasonality effect on a large number of routes in our database, we decided to compute separate nominal times for every month.
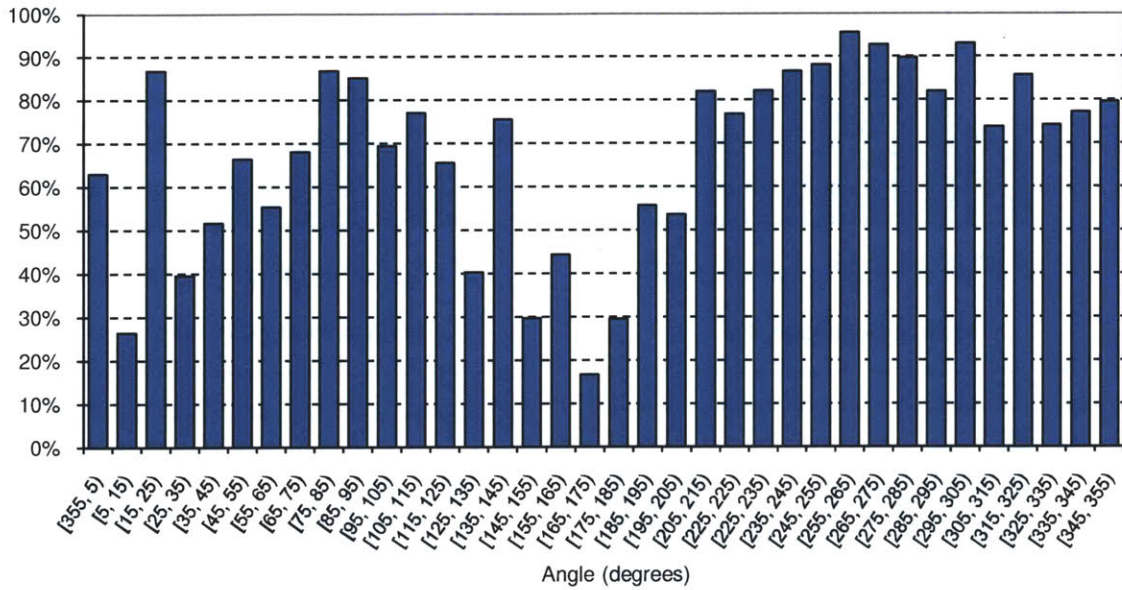
Figure 4-13: Percentage of Route-Carrier Pairs for which the null Hypothesis of Identically Distributed Seasonality Indexes cannot be rejected at a 20% Significance Level
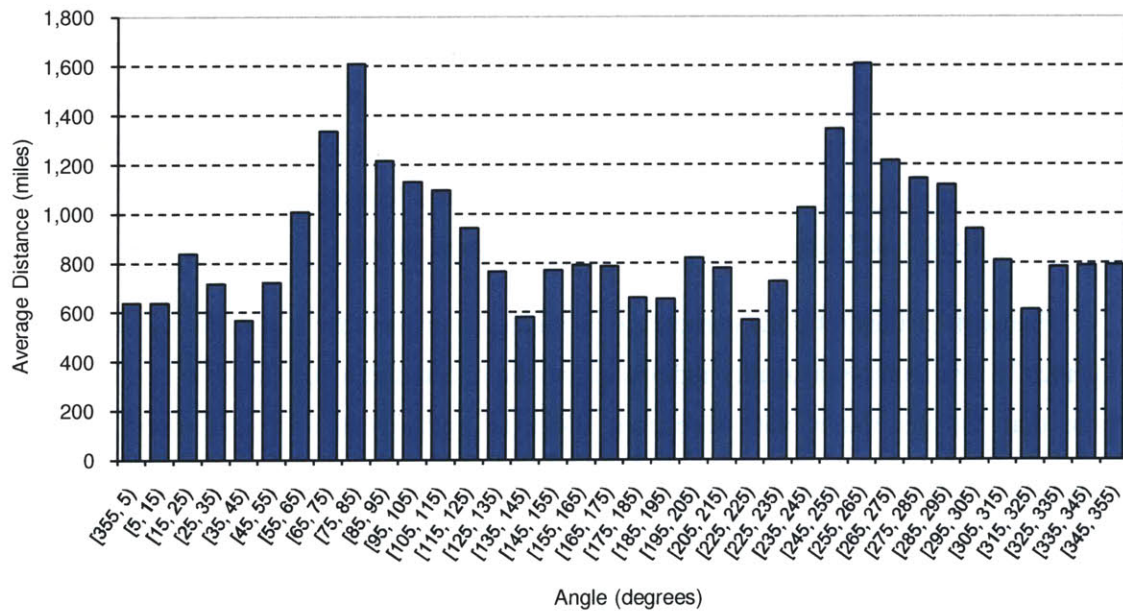
Figure 4-14: Average Distance of Angle Groups

4.3.2 Time of Day

The time of day has a very strong influence on flight operations and their on-time performance through fluctuations in demand, changes in airport capacity, and propagation of delays.

Queuing theory suggests that the relationship between airport utilization ratio (demand-to-capacity ratio) and expected delays is highly nonlinear[32]. In periods that demand exceeds capacity, the expected delay per scheduled movement increases exponentially, often resulting in flight cancellations. Demand may exceed capacity due to:

• Demand surges in peak hours. During periods of very high traffic, there are periods when the scheduled operations exceed the airport capacity even under optimal weather conditions.

• Random deviations from schedule. Although scheduled operations may be below capacity, delays of individual flights can result in intervals with higher demand. These deviations can be attributed to aircraft or crew arrival delays from previous flights, the boarding process, mechanical problems, etc.

• Reductions in airport capacity because of strong winds, low visibility and other weather conditions. This is the most common source of delays.

However, long delays and queues may occur even when the mean demand rate is below capacity but close to it, due to the variability of the intervals between successive requests for runway use, and the variability of the service time for each request[33]. These queues may need a long time to dissipate, especially at airports that operate very close to capacity for extended periods of time.

Furthermore, delays that occur on one flight leg can propagate downstream to the next flights, with respect to aircraft, crew, gate and slot availability. If a short delay suffered by a flight early in the day is not absorbed by adequate buffer in scheduled block time or slack in the planned turn-around time, it will cause schedule disruptions that can result in several hours of delay for subsequent flights. In a network structure, this initial delay can propagate through late flight arrivals and late connections later in the day, affecting other aircraft and airports where the airline operates.

---

[32] Larson, R., Odoni, A., *Urban Operations Research*, Prentice-Hall, NJ, 1982.
[33] De Neufville, R., Odoni, A., *Airports Systems: Planning, Design and Management*, McGraw-Hill, 2003.

The three main airports of the New York area, Newark, LaGuardia and JFK, are excellent examples of airports operating often at their capacity limits. There is a strong interaction between the terminal airspaces of these airports and there exists a common weather impact on its operations. Figure 4-15 shows how scheduled arrivals at the three airports, along with their associated delays and buffer, are spread throughout the day.

Flights that arrive before 10am have on average negative arrival delays and their on-time performance is close to 90%. After that time, taxi-out, gate and arrival delays increase steadily until they reach their peak between 6pm and 7pm. At the peak, only 60% of the flights arrived on-time. These delays are attributed to a great extent to the instituted Ground Delay Programs (GDPs), since the average ground hold time between 6pm and 10pm exceeds the 20 minutes. The ground hold time is defined as the difference between the expected departure clearance time (EDCT) that is assigned to a flight affected by a GDP and its original gate departure time. Ground hold times provide an approximation for the actual gate delay, but are not equal, because a flight can depart before or after the expected departure clearance time. Also, in cases of gate constraints, a large proportion of the ground hold times is spent on the tarmac.

Furthermore, Figure 4-15 shows that the buffer is increasing constantly between 12pm and 6pm. However, it is not clear if this happens because buffer varies through the day on a route basis. Another factor that may drive this increase in average buffer is the stage length of the flights scheduled to arrive during these hours. As will be discussed in Section 4.3.4, buffer has a linear relationship with distance. After 7pm, delays are declining and on-time performance is slightly improved.

Moreover, it is notable that the time when peaks in arrival delay occur lags behind the peak of scheduled demand. The reason for this time lag is that when GDPs are active during periods of high demand, queues of flights requesting an arrival are formed. The flights that are at the end of the period of an active GDP are necessarily placed at the end of the queue, thus experiencing the longest delays.

Although the taxi-out delays of the arrivals to these three airports are driven mostly by the GDPs and have one peak in the evening, the taxi-out delays of flights that depart from these airports have other causes and follow a different distribution (Figure 4-16). The peak of
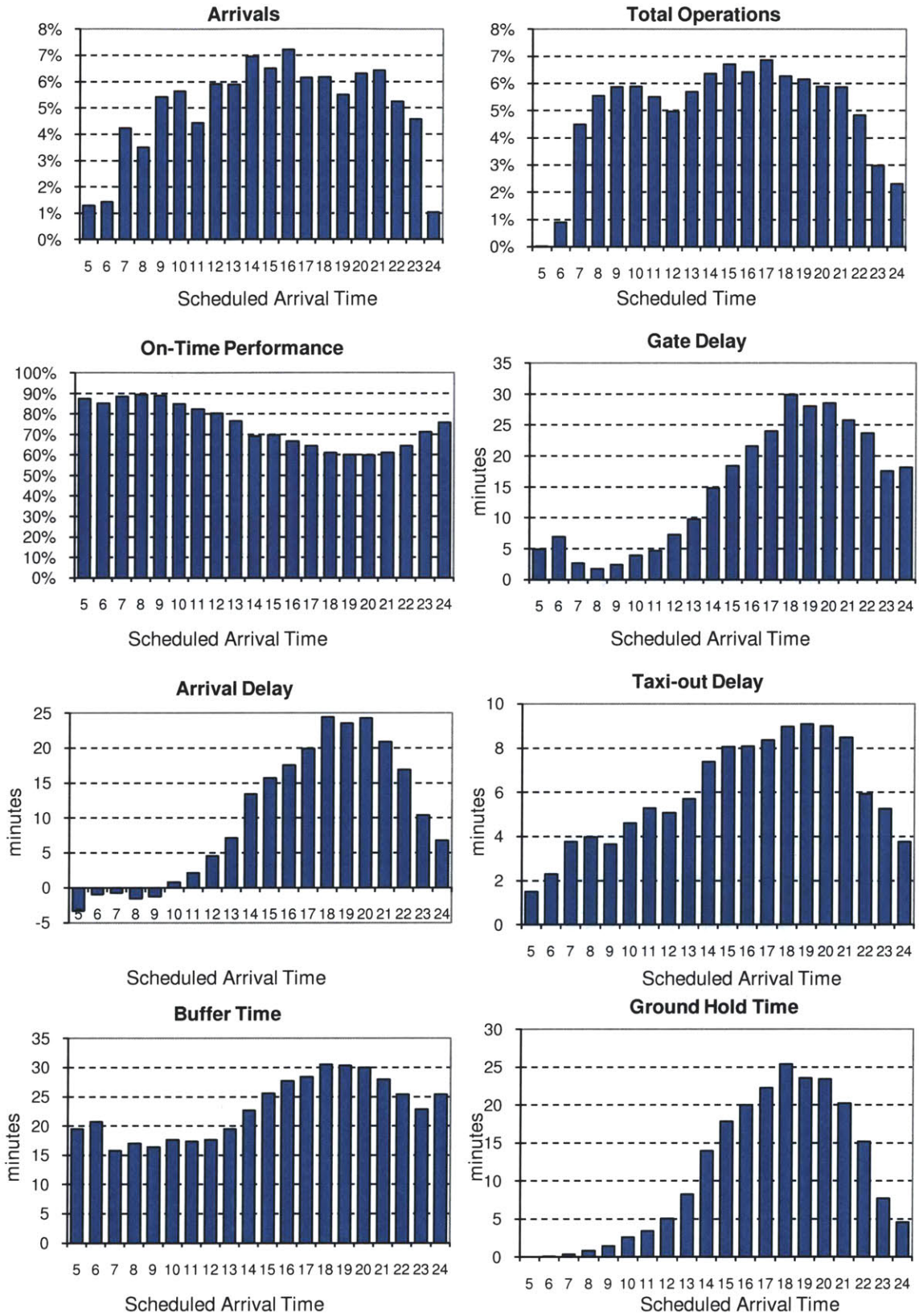
Figure 4-15: Performance of Flights Arriving to EWR, JFK and LGA by Time of Day

scheduled departures occurs at 8am and results in a morning peak in taxi-out delays at 9am. Again, there is a time lag of one hour for the reasons mentioned above. Although the volume of scheduled departures is lower in the evening, the taxi-out delays are higher. The same applies for the gate delays, which cannot be attributed to GDPs because, as shown in Figure 4-16, the ground hold times are negligible. We assume that the evening peak of gate and taxi-out delays is caused by airfield and airspace congestion due to the increased volume of total operation. Gate delays may also be the result of delay propagation and late aircraft arrivals at gate from previous flights.
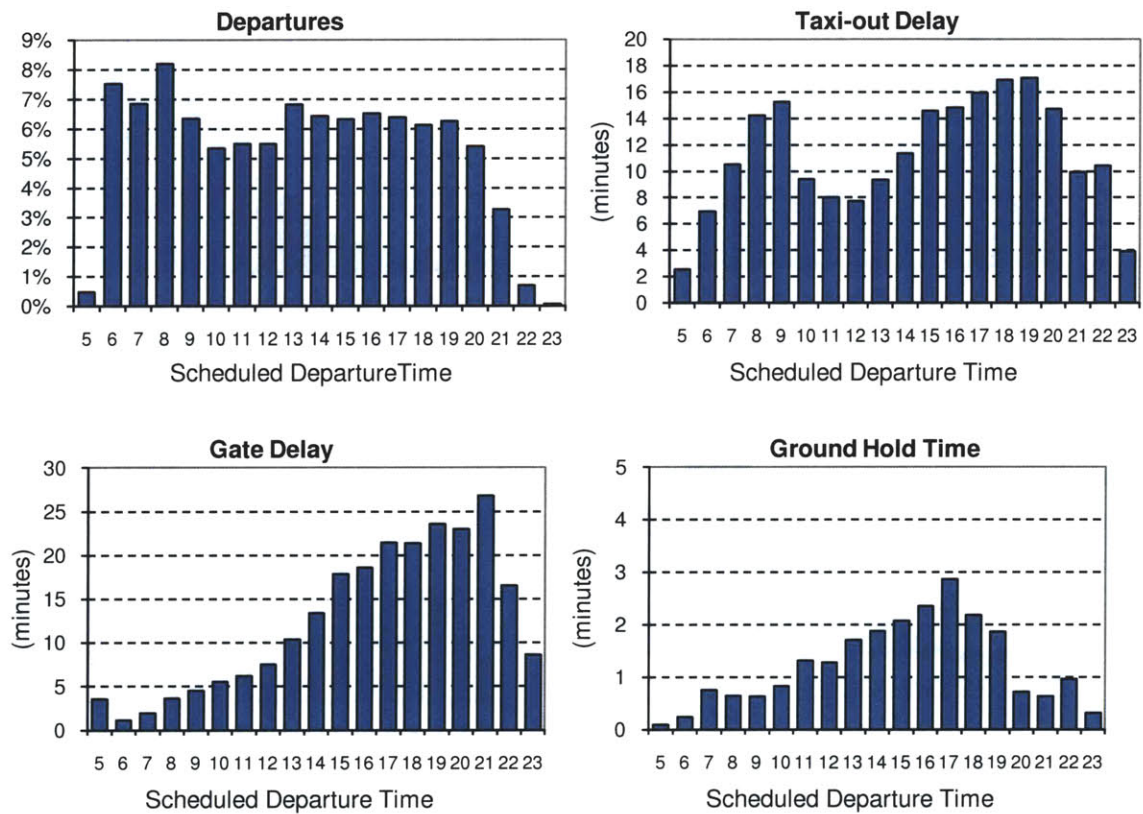
Figure 4-16: Volume and Delays of Flights Departing from EWR, JFK and LGA by Time of Day

In section 4.1 we highlighted the importance of delay variability in schedule padding. We claimed that delays with low variability can be predictable and thus incorporated into schedule through buffer. On the other hand, when delay components are very variable on a given route, the use of sufficient buffer is very costly for airlines, and the gains are limited. In his Master thesis, Morisset[34] studied the arrival delay distribution of four different hours of the day at EWR in 2007. He observed that the average delays increased steadily over the course of the day, and most importantly that the distribution of the delay becomes increasingly dispersed.

Similarly, we studied the arrival and gate delay distribution of six different hours of the day at LGA (Figure 4-17 and Figure 4-18). It is obvious from these graphs that arrival and gate delays are very dispersed at any hour, and their mean and standard deviation increase throughout the day. The situation is improved between 11:00pm and 11:59pm because of traffic reduction, but the reliability of schedule is still very poor.

---

[34] Morisset, T., *Comparing Capacities and Delays at major European and American Airports*, Master Thesis, Engineering Systems Division, MIT, May 2010.
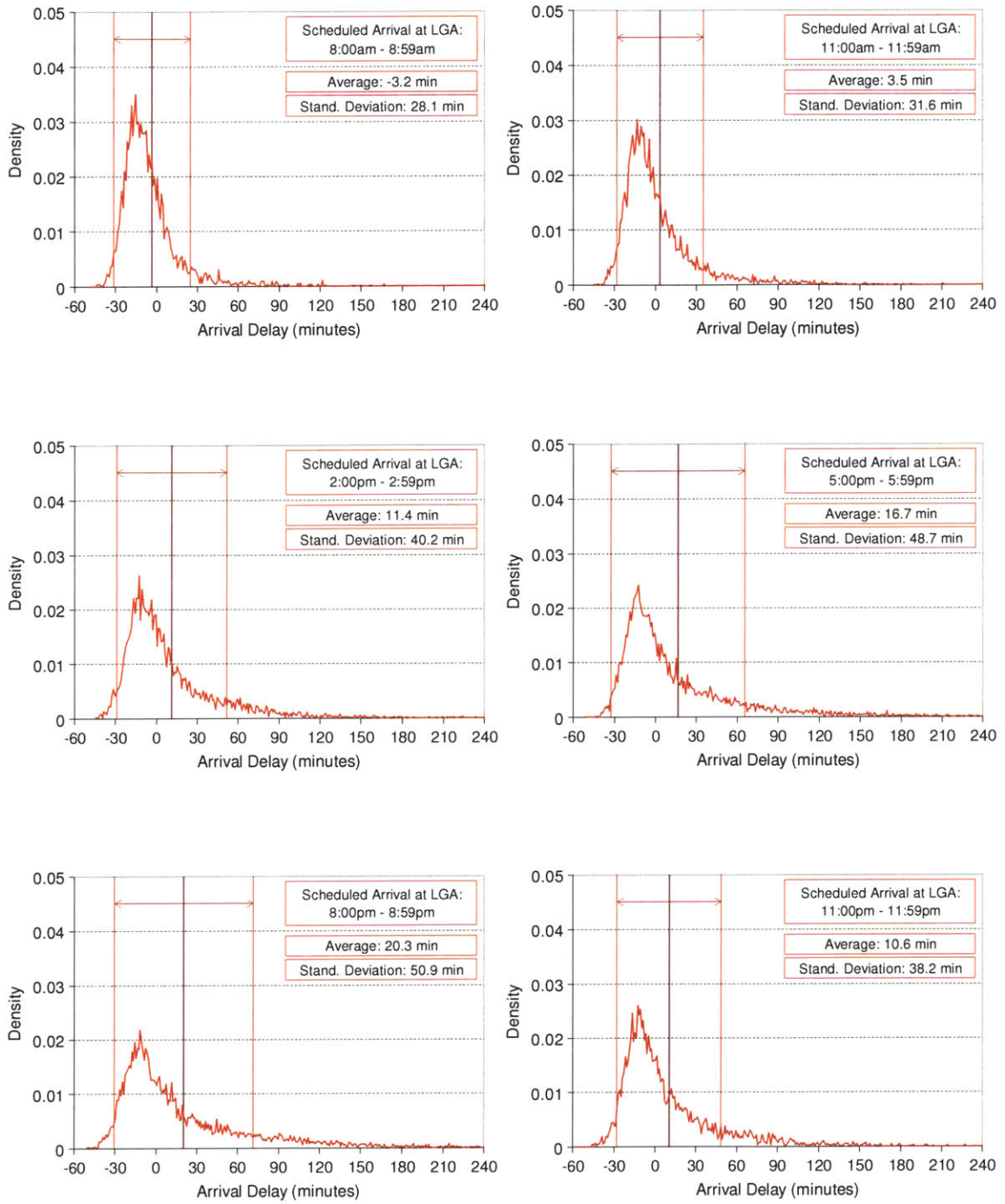
Figure 4-17: Arrival Delay Distributions at LaGuardia
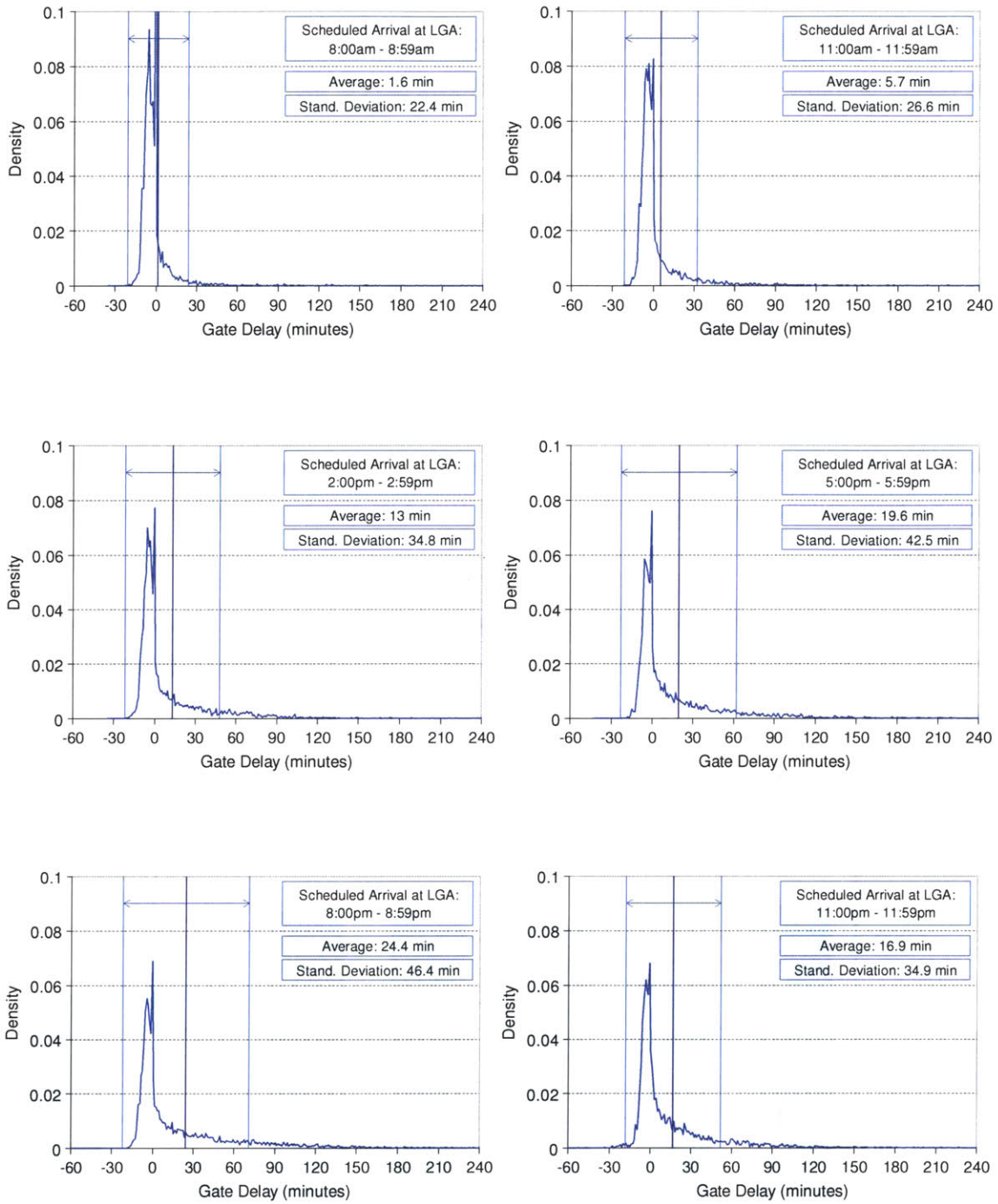
Figure 4-18: Gate Delay Distributions at LaGuardia

4.3.3   Day of Week

The day of the week may influence on-time performance and delay components through the volume of traffic and their distribution over different days of the week. Operations are almost constant at weekdays and slightly reduced on Sundays (Figure 4-19). Saturday is the least busy day. Figure 4-20 shows that delays and on-time performance vary through the week and their distribution indicates a small correlation with the traffic volume.

Thursday and Friday have the longest arrival and gate delays, and the lowest on-time performance. It is interesting to note that although Sunday is the second least busy day, its on-time performance is similar to Monday's, which is the busiest day. A reasonable assumption would be that on Sundays there are more scheduled flights later on the evening hours compared to Mondays, since leisure travelers would prefer to return home late in the evening. However, an hour-by hour comparison of the traffic on these two days (Figure 4-21) does not verify this assumption.

Buffer is very similar for each day of the week, ranging on average between 18.4 minutes and 18.6 minutes. This suggests that carriers do not adjust the scheduled block times based on the day of the week.
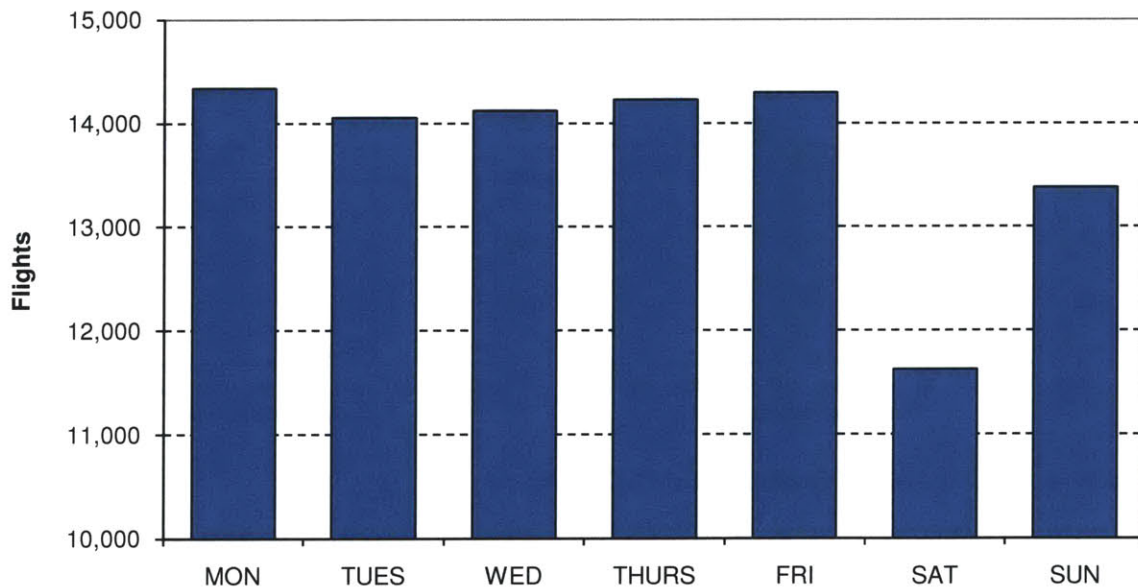
Figure 4-19: Scheduled Flights and Buffer by Day of Week, in 2009
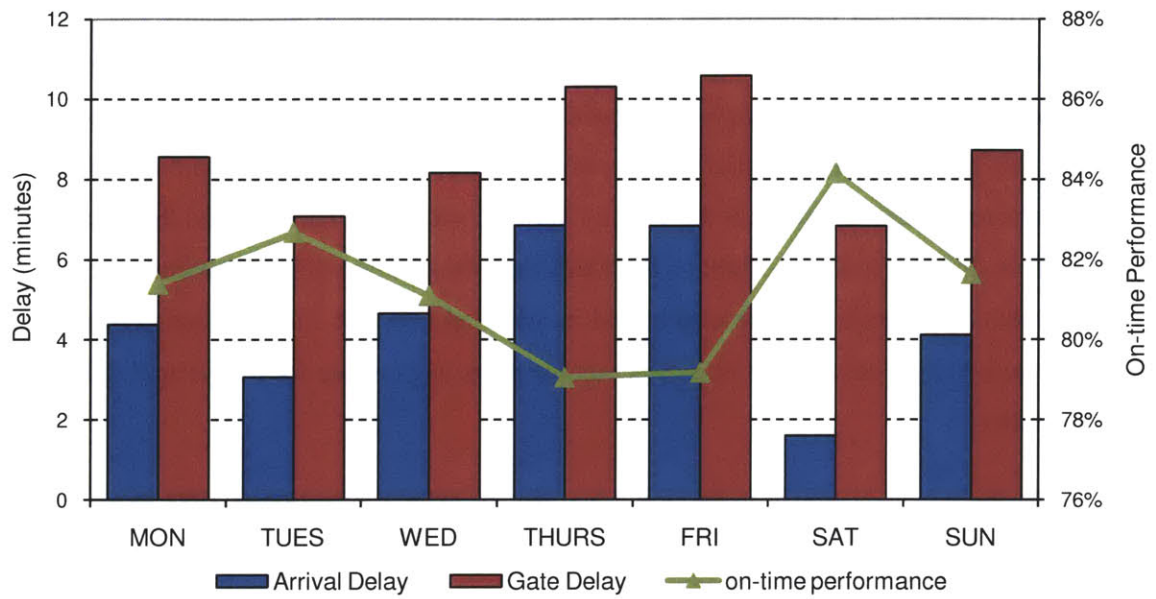
Figure 4-20: Arrival Delays, Gate Delays and On-Time Performance by Day of Week, in 2009
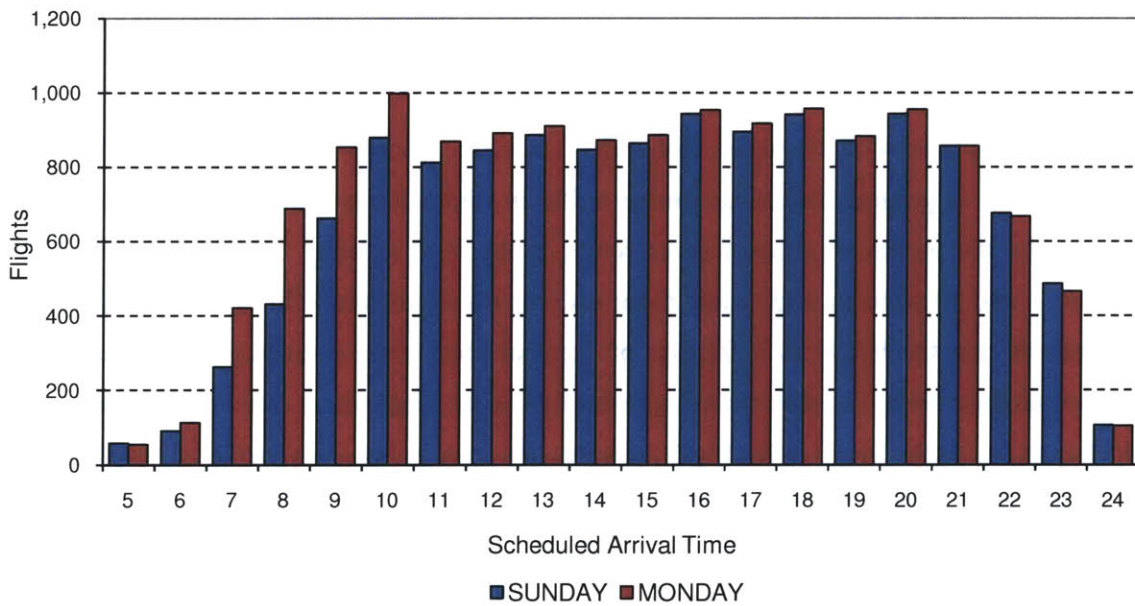
Figure 4-21: Scheduled Arrivals by Time of Day on Mondays and Sundays, in 2009

4.3.4 Route Distance and Airport Effect

Two factors that are expected to be of great importance in determining the schedule buffer are: (i) the flight distance between origin and destination, and (ii), the congestion of airfield and terminal airspace at origin and destination airport. The length of the flight is expected to have a very strong influence on buffer time, as the longer the distance that an aircraft has to fly, the more the time that the aircraft is affected by variable weather conditions, and the more likely that it will be diverted to longer flight paths. Apart from the distance effect, airborne time may be increased due to congestion at the airport and its terminal airspace. Furthermore, constraints at airport capacity either at origin or at destination are often responsible for delays suffered at gate and the taxiway.

We have already shown in Section 4.3.1 that the aircraft direction relative to jet streams has a significant effect on airborne times and their associated delays. Therefore, in analyzing the relationship between buffer and distance, we considered more appropriate to use the nominal airborne time of each directional route-carrier pair, rather than the absolute distance between origin and destination. Figure 4-22 and Figure 4-23 illustrate how buffer, first as absolute value, and second as fraction of nominal block time, is related to nominal airborne time. Every point on these two graphs corresponds to the average values of one of the 4,250 studied route-carrier pairs in September 2009. September was a month with limited airborne delays, and thus the estimated nominal airborne time is thought to be more accurate.

First, Figure 4-22 illustrates that there is a weak linearity between buffer and nominal airborne time. In general, airlines add more buffer in flights that are expected to last longer, to account for the increased uncertainty en-route. In the vast majority of routes, buffer ranges between 5 and 30 minutes, and in few cases exceeds the 60 minutes. On the other hand, Figure 4-23 shows that the buffer as a fraction of nominal block time decreases exponentially with nominal airborne time. On short-haul flights buffer can be a significant fraction of the scheduled block time, but as the length of the flight increases this fraction becomes smaller. This is intuitive, since expected delays do not increase proportionally to distance. However, it is of great importance the existence of few short haul routes where scheduled buffer is more than half of nominal block time. It is expected that these routes are between congested airports that impose lengthy ground delays to users. Indeed, 61 out of these 72 routes depart from or are destined to EWR, LGA, JFK or PHL.
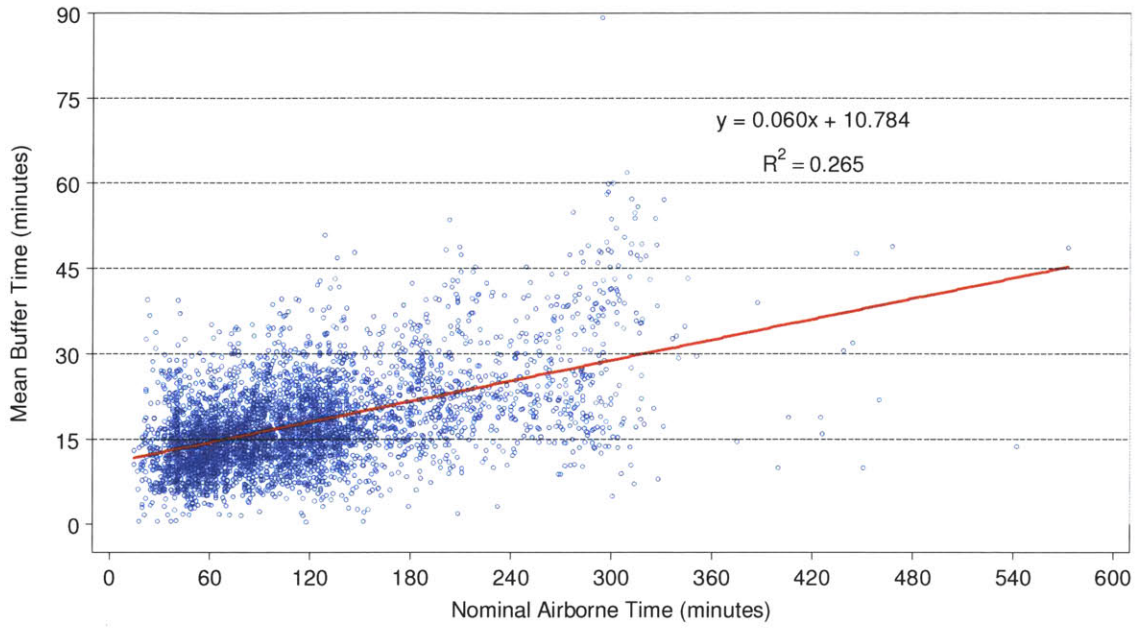
Figure 4-22: Mean Buffer Time vs. Nominal Airborne Time on the 4,250 Route-Carrier Pairs in September 2009
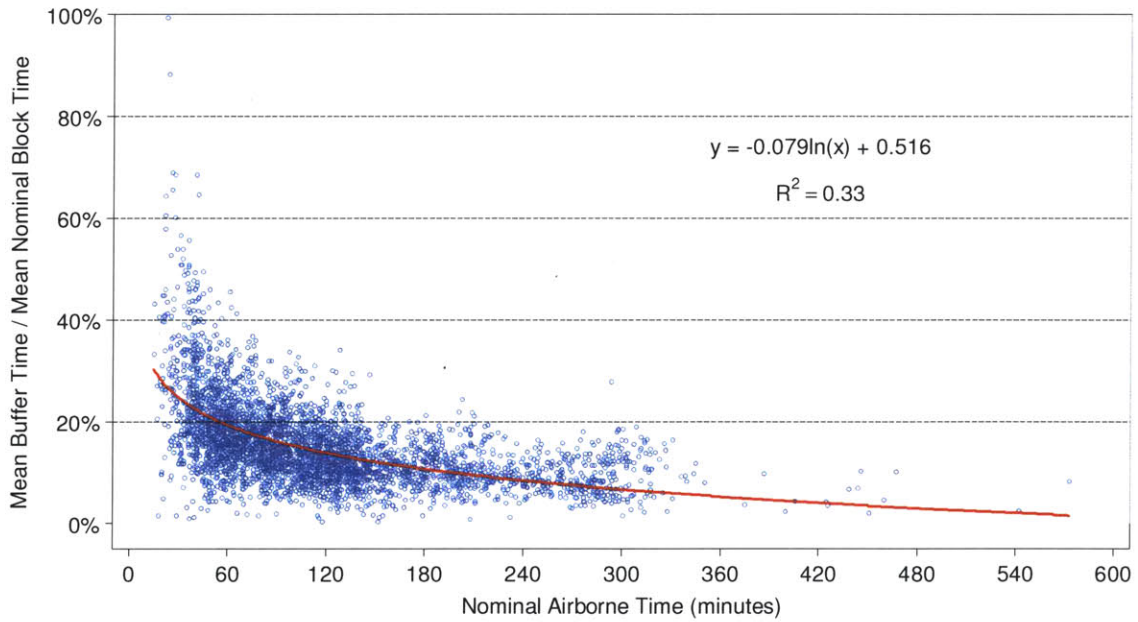
Figure 4-23: Mean Buffer Time / Mean Nominal Block Time vs. Nominal Airborne Time on the 4,250 Route-Carrier Pairs in September 2009

77

The distinction between the effect that origin and destination airport has on the on-time performance of a specific route is complicated and outside the scope of this thesis, because airspace congestion cannot be easily attributed to any specific airport. However, using the residuals from the linear regression model of Figure 4-23 (*Mean Buffer Time/Nominal Block Time* is the dependent variable and *ln (Nominal Airborne Time)* the independent), we can examine if there are airports with routes where flights are more padded. In Table 4-2 and Table 4-3, airports with at least 50 route carrier pairs are ranked based on descending order of the fraction of positive residuals.

It is notable that flights out of airports that have lengthy taxi-out delays tend to be padded more. Particularly, LGA, EWR, JFK, PHL, CLT and ATL have the most positive residuals and the longest average taxi-out delays. The result of increased schedule padding of the departing flights is that their on-time performance is almost equal, and not lower, to flights at the less congested airports. Airborne delays seem to be independent of the origin airport. On the other hand, the ranking of destination airports is independent of their taxi-out delays but is correlated to a small degree with the airborne delays. Airports with a higher fraction of positive residuals have, in general, longer airborne delays. Furthermore, an airport that has a strong effect as origin does not necessarily have a similar effect as destination. Exceptions are EWR, JFK and LGA, as the majority of routes that they serve both at origin as well as at destination are among the most padded.

Origin	Positive Residuals*	Taxi-out Delay		Airborne Delay		On-Time Performance
		Mean	St. Dev.	Mean	St. Dev.	
LGA	98.6%	13.0	13.2	7.0	7.4	86.6%
EWR	94.8%	9.2	10.5	8.1	8.3	89.8%
JFK	92.6%	7.1	14.8	8.8	9.1	89.5%
PHL	88.2%	9.4	11.8	7.2	7.8	82.7%
CLT	76.5%	7.4	9.5	6.5	7.3	87.2%
ATL	67.3%	8.1	9.4	5.9	6.8	82.6%
MDW	65.3%	3.5	6.0	6.6	7.0	86.7%
MSP	63.7%	6.2	7.5	7.7	8.0	87.0%
IAD	56.5%	3.5	8.6	7.1	8.1	89.2%
BOS	53.9%	4.5	7.7	8.2	8.2	89.4%
LAS	52.4%	4.6	6.8	6.0	7.2	90.7%
RDU	50.9%	2.6	8.3	6.0	6.3	89.6%
ORD	48.9%	3.6	6.8	7.0	7.6	89.3%
PIT	47.3%	1.5	5.8	6.3	7.0	89.1%
DCA	43.9%	4.0	8.4	6.4	6.7	89.8%
MCO	42.9%	2.1	5.7	6.7	7.0	89.8%
BWI	42.7%	3.3	6.0	6.0	6.3	87.5%
MCI	41.8%	2.3	5.1	6.3	6.9	89.9%
BNA	41.7%	1.8	6.2	5.9	6.2	87.8%
PHX	36.8%	4.1	6.5	6.4	7.3	89.5%
STL	36.5%	1.1	3.6	6.2	6.9	90.8%
CVG	35.1%	1.8	5.3	6.3	6.9	88.7%
DTW	33.9%	5.8	7.6	6.5	7.0	84.1%
IND	33.9%	1.4	6.2	6.0	6.9	89.6%
LAX	33.7%	2.4	5.1	7.4	8.5	89.8%
CLE	32.9%	1.4	4.8	6.0	6.5	91.5%
MIA	32.0%	2.8	7.5	7.6	8.3	82.7%
DFW	31.9%	3.8	8.3	7.5	7.7	84.9%
MKE	31.0%	2.0	5.3	6.6	7.7	88.0%
DEN	29.7%	2.7	4.9	7.0	7.0	87.0%
IAH	29.3%	6.1	8.3	7.0	7.2	87.5%
SFO	29.3%	4.0	6.4	7.9	9.2	85.4%
SEA	29.2%	3.0	5.1	8.9	9.3	90.0%
MEM	28.8%	3.6	6.9	6.3	7.0	87.9%
TPA	28.1%	2.3	5.3	6.4	6.7	88.1%
SLC	26.9%	5.3	7.0	6.9	7.3	87.9%
FLL	22.6%	2.7	7.3	7.0	7.2	85.3%
PDX	10.2%	3.3	4.6	6.9	7.6	90.4%

Table 4-2: Effect of Origin Airport

* Fraction of positive residuals for the linear regression model of Figure 4-23 with *Mean Buffer Time/Nominal Block Time* being the dependent variable and *ln (Nominal Airborne Time)* the independent

Destination	Positive Residuals[*]	Airborne Delay		Taxi-out Delay		On-Time Performance
		Mean	St. Dev.	Mean	St. Dev.	
SEA	71.2%	9.5	9.4	3.9	7.1	92.5%
LGA	68.1%	8.0	7.9	4.8	9.4	85.6%
SFO	67.5%	8.4	9.6	4.6	8.8	87.4%
JFK	67.4%	10.3	10.8	4.7	9.2	83.3%
IAD	67.0%	7.3	7.6	3.4	7.5	89.1%
EWR	65.6%	8.9	8.8	4.6	9.9	86.0%
ORD	60.8%	6.7	6.8	3.9	7.8	89.7%
CVG	60.0%	5.9	6.3	4.1	7.3	90.3%
LAS	59.2%	6.6	7.2	3.6	6.4	91.6%
PHL	57.0%	8.1	8.7	3.8	8.0	82.9%
MSP	55.4%	9.4	9.9	4.3	8.8	83.7%
LAX	54.6%	6.6	7.9	3.8	7.0	91.4%
PDX	54.0%	6.7	7.0	3.2	6.8	92.8%
MCI	50.0%	6.1	6.1	4.0	6.6	88.0%
PHX	48.3%	6.2	6.7	3.3	6.1	91.3%
SLC	45.0%	6.5	6.6	3.8	6.9	91.0%
ATL	43.3%	7.7	8.1	6.0	10.8	80.2%
DEN	41.2%	7.3	7.2	3.8	6.9	90.0%
DCA	39.5%	5.5	5.4	4.1	7.9	90.1%
STL	39.5%	5.3	5.8	3.3	6.4	89.6%
MDW	38.2%	5.7	6.3	3.1	6.3	89.2%
DTW	38.2%	6.9	6.7	5.0	8.3	86.6%
RDU	37.9%	5.1	5.3	5.6	9.6	88.2%
CLT	37.9%	7.6	7.3	5.0	9.1	88.6%
MKE	36.2%	5.7	6.1	4.3	7.2	89.2%
DFW	36.2%	8.2	9.1	3.8	7.3	87.9%
PIT	35.7%	4.7	4.9	4.6	8.3	88.5%
CLE	34.2%	5.1	5.1	4.1	7.8	91.4%
MCO	34.1%	6.3	6.3	4.2	8.0	89.2%
IND	33.9%	5.0	5.3	4.8	8.2	89.3%
FLL	33.3%	7.0	7.3	5.3	9.0	86.3%
TPA	33.3%	5.7	6.5	4.4	7.8	88.2%
BOS	33.0%	7.7	7.9	5.8	10.7	85.6%
BWI	31.6%	5.4	5.4	3.6	6.9	89.5%
MEM	28.6%	6.4	6.4	4.3	7.4	87.7%
MIA	28.0%	7.6	8.2	4.6	8.0	84.8%
BNA	22.4%	5.1	5.4	4.0	7.2	86.2%
IAH	16.5%	8.1	8.5	3.3	6.9	87.0%

Table 4-3: Effect of Destination Airport

[*] Fraction of positive residuals for the linear regression model of Figure 4-23 with *Mean Buffer Time/Nominal Block Time* being the dependent variable and *ln (Nominal Airborne Time)* the independent

4.3.5 Ground Delay Programs

The impact of Ground Delay Programs on gate, taxi-out and arrival delays has already been addressed in various sections of this chapter. It is intuitive that when a GDP is initiated for a specific destination, gate and taxi-out times at origin increase, and the on-time performance of flights destined for the impacted airport worsens. Figure 4-24 shows the fraction of flights destined for the ten most affected airports that were held on ground by GDPs in 2009. Although this fraction is very high for some airports, e.g. 30% for EWR and 23% for LGA, and around 500 flights are affected in total every day, it is not possible for airlines to predict, at the time of making their schedule, which flights will be delayed and the length of the imposed delays.

For example, assume that flights A and B both having the same origin, destination and departure time. Flight A has no buffer, whereas flight B has a buffer of 20 minutes, and thus it is scheduled to arrive 20 minutes later than flight A. An early forecast for bad weather at the terminal airspace of the arrival airport is issued, and a GDP is applied to both flights. In this case, flight A will be assigned an earlier Expected Departure Clearance Time (EDCT) because of the "first scheduled, first served" policy followed for the new arrival sequence.

This example illustrates that additional buffer may lead to further flight delays. It must be noted that the above example would be applicable only in case of GDPs that are initiated several hours before scheduled departure. For ground hold delays that are imposed on flights when the pilot requests clearance for pushback or take off, the placement of the flight in the arrival sequence is computed based on flight plan. The ASPM database does not provide any information about when ground hold times were assigned to the affected flights.

Figure 4-25 shows that in 2009 more than 50% of the flights that were scheduled to arrive at EWR between 5pm and 8pm were held by a GDP, independent of the delay interval. For ground delays longer than 30 minutes this fraction drops to 35%, and for longer than 60 minutes to 23%. However, a day by day analysis reveals that GDPs at EWR were initiated on 259 days in 2009, with an average number of 141 held flights, and most importantly, a standard deviation of 78 flights. This implies that flights were not affected by GDPs on a systematic basis and that delays imposed by GPDs are very unpredictable. Therefore, it is unlikely that carriers use ground delay data in their buffer selection process.
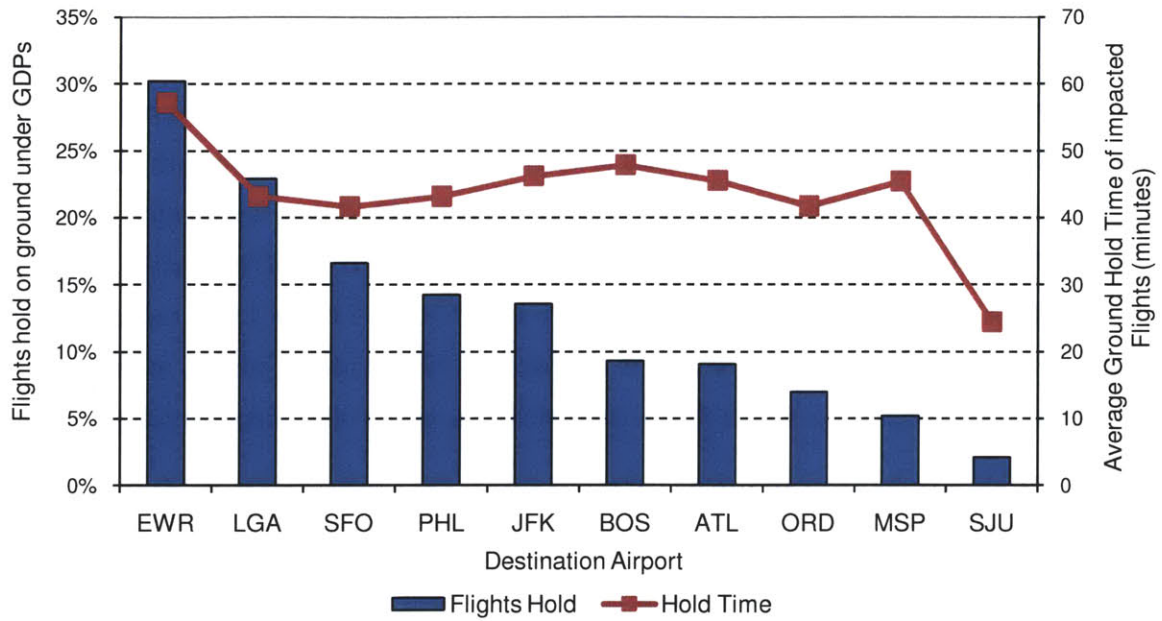
Figure 4-24: Airports affected by Ground Delay Programs
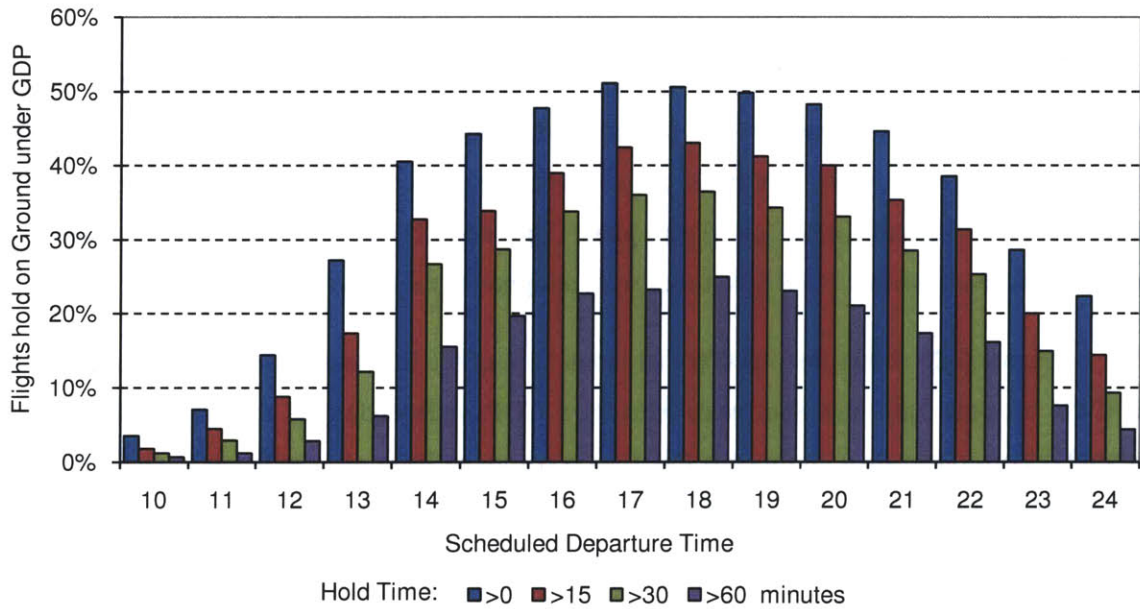
Figure 4-25: Ground Delay Programs at EWR by the Time of Day

4.3.6 Aircraft Type

Aircraft type has an impact on airborne time, because of the differences in cruising speed. To account for any potential differences in aircraft types across carriers, we computed the nominal airborne time separately for each carrier on a given route. Furthermore, the size of the aircraft may influence gate delays, through the time required for the boarding process, and taxi-out times because of the taxiing speed. At congested airports aircraft size can further affect airborne and taxi-out times, if air traffic controllers are optimizing the sequences of landing and taking-off aircraft based on wake vortex separation requirements. The relation of aircraft size with flight delays is discussed further in Section 4.3.8.

4.3.7 Gate Assignment and Runway Configuration

Gate assignment and runway configuration have a direct impact on taxi-out and taxi-in times by decreasing or increasing the length of taxiing for a given flight. Furthermore, the use of different runway configurations and gates affects the estimation of unimpeded taxi-out and taxi-in times, and consequently their delays. ASPM does not encompass these differences in the estimation of unimpeded times, because no data is available on the runway configurations and gates used for each flight. Therefore, small deviations in actual taxi-out and taxi-in times or delays may be attributed to different configurations even under optimal conditions. The runway configuration may also affect the airborne time by changing the flight path and the length of the flight.

4.3.8 Flight Sequence

In this final section of this chapter, we analyze on-time performance, buffer, flight delays and turn-around times on the basis of the number of flight segments that the aircraft has already flown at the same day. As has been discussed, delays that occur on one flight leg can propagate downstream to the next flights. To study the effect of delay propagation both in scheduling and in actual flight performance, we identified the sequence of flights performed by each aircraft in September 2009.

The ASPM database uses two data sources, ETMS and OOOI (see Section 3.1). Only the latter reports systematically the aircraft tail number, and thus has been the only source used for this section's analysis. Even so, there were data entries that reported the same aircraft at two simultaneous flights. Therefore, we cleaned the database and kept only flight sequences satisfying the following criteria:

1.   All flights in the sequence are performed by an aircraft of the same IATA equipment code and tail number.

2.   The origin of every flight segment in the sequence is the destination of the preceding flight.

3.   All actual turn-around times in the sequence are positive.

In total, we analyzed 31,343 sequences having from two up to eight flights per day (Table 4-4). The results of this analysis are plotted in Figures Figure 4-26 to Figure 4-34. Every line of these graphs corresponds to the maximum number of flight segments flown by an aircraft in a day and is averaged across all sequences in September 2009.

Fights in Sequence	2	3	4	5	6	7	8
Sequences in Sample	9,382	7,131	7,145	3,508	2,953	734	490

Table 4-4: Number of Flight Sequences

Looking at the plot of on-time performance (Figure 4-26), we observe two trends:

1. For the same sequence, on-time performance decreases linearly over the course of the day. This happens because of propagation of delays, and traffic increase in the evening hours. The plots of arrival and gate delays (Figure 4-27 and Figure 4-28) follow a similarly increasing trend along each sequence, despite the fact that block delays relative to schedule decrease.

2. For flights that have the same number of preceding flight segments, but belong to sequences of different length, on-time performance decreases as sequence length decreases. For example, the second flight of a 5-flight sequence has a worse on-time performance than the second flight of a 6-flight sequence. The reason is that aircraft in short sequences operate on longer routes (Figure 4-29), and thus suffer longer airborne delays (Figure 4-30).

Another interesting remark is the evidence of correlation between aircraft size, stage length, gate delays and taxi-out times. Given that short sequences consist of long-haul flights, we can assume that the aircraft in these sequences have a larger size. Operating a larger aircraft in a long flight reduces the unit cost and increases profits when there is sufficient demand. Moreover, frequency competition in short-haul markets results in more flights with smaller aircraft. Our assumption about aircraft size is supported by the fact that flights in short sequences have in general, longer gate delays and taxi-out times than flights in long sequences. The boarding process lasts typically longer for larger aircraft, since the boarding time is proportional to the number of boarding passengers. Furthermore, push-back and taxiing speed is slower for larger aircraft, resulting in longer taxi-out times.

The most important observation in the plot of taxi-out times is that the first flight of the day has the longest taxi-out time, with the exception of the 6 and 8-flight sequence. This can be explained by the morning peak in departing operations. Similarly, block delays relative to schedule of the first flights are substantially longer than of the later flights (Figure 4-33). Since airborne delays and buffer time are almost constant for each sequence, this trend must be driven by delays in taxiing. Regarding the 6 and 8-flight sequence, the large systematic fluctuation in taxi-out times may be attributed to a hub-effect and the excess demand during bank periods in hub airports. Mayer and Sinai[18] had found that flights departing from the airline's own hub are 2.9% to 5.4% less likely to be on time relative to flights on the same route by non-hub airlines.

On-Time Performance

Number of Flight in Sequence

◆ 2 flights ■ 3 flights ▲ 4 flights ✕ 5 flights ✳ 6 flights ● 7 flights ╀ 8 flights
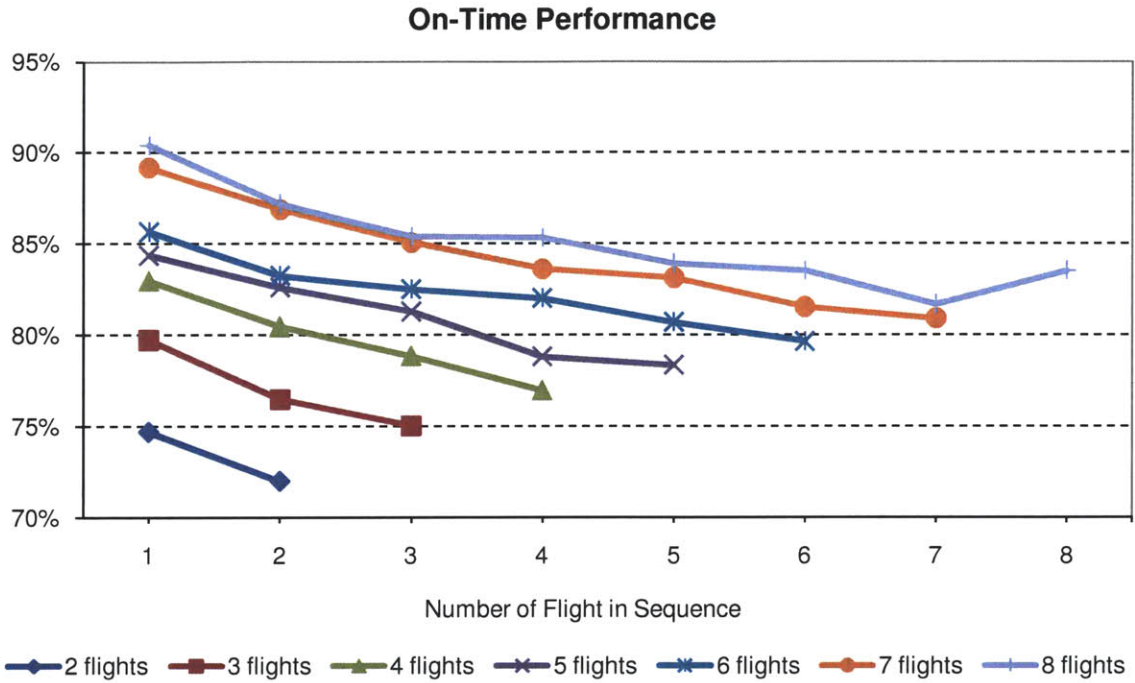
Figure 4-26: On-Time Percentage by Flights per Aircraft

Arrival Delay

minutes

Number of Flight in Sequence

◆ 2 flights ■ 3 flights ▲ 4 flights ✕ 5 flights ✳ 6 flights ● 7 flights ╀ 8 flights
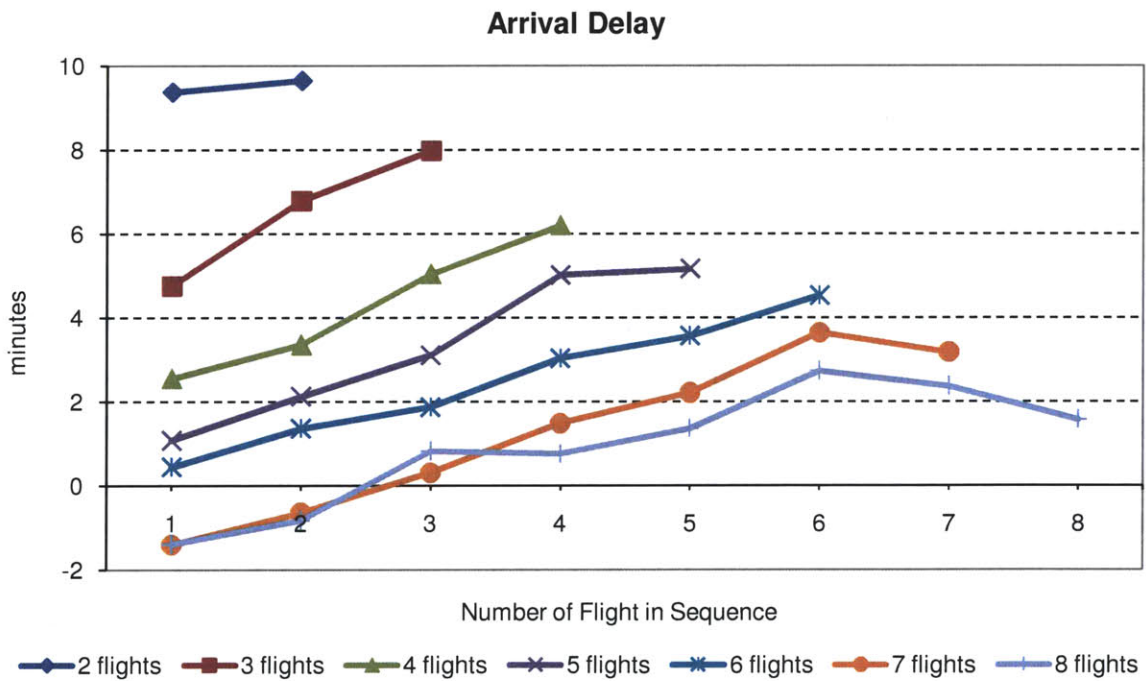
Figure 4-27: Arrival Delay by Flights per Aircraft
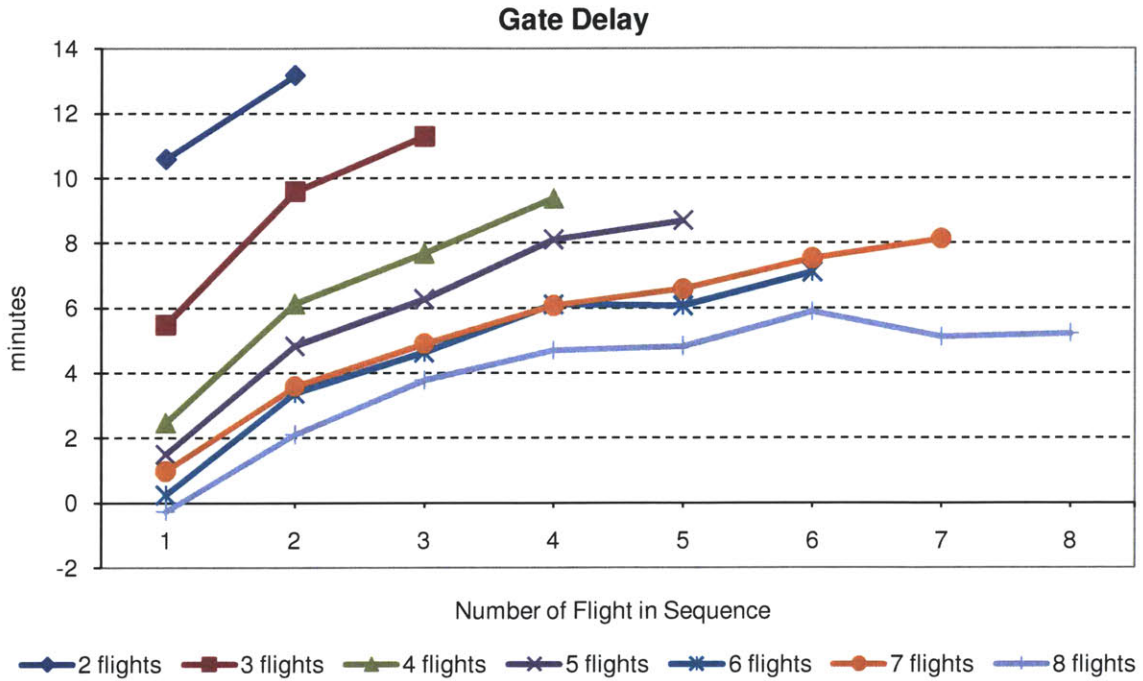
Gate Delay

Figure 4-28: Gate Delay by Flights per Aircraft

Scheduled Block Time
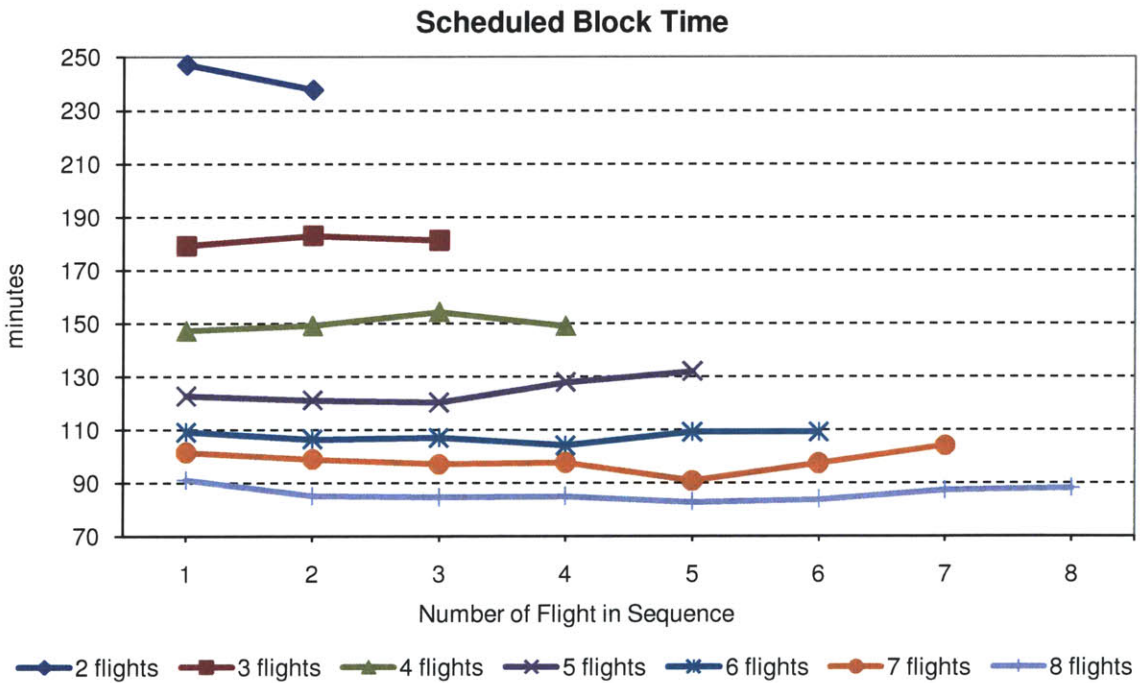
Figure 4-29: Scheduled Block Time by Flights per Aircraft
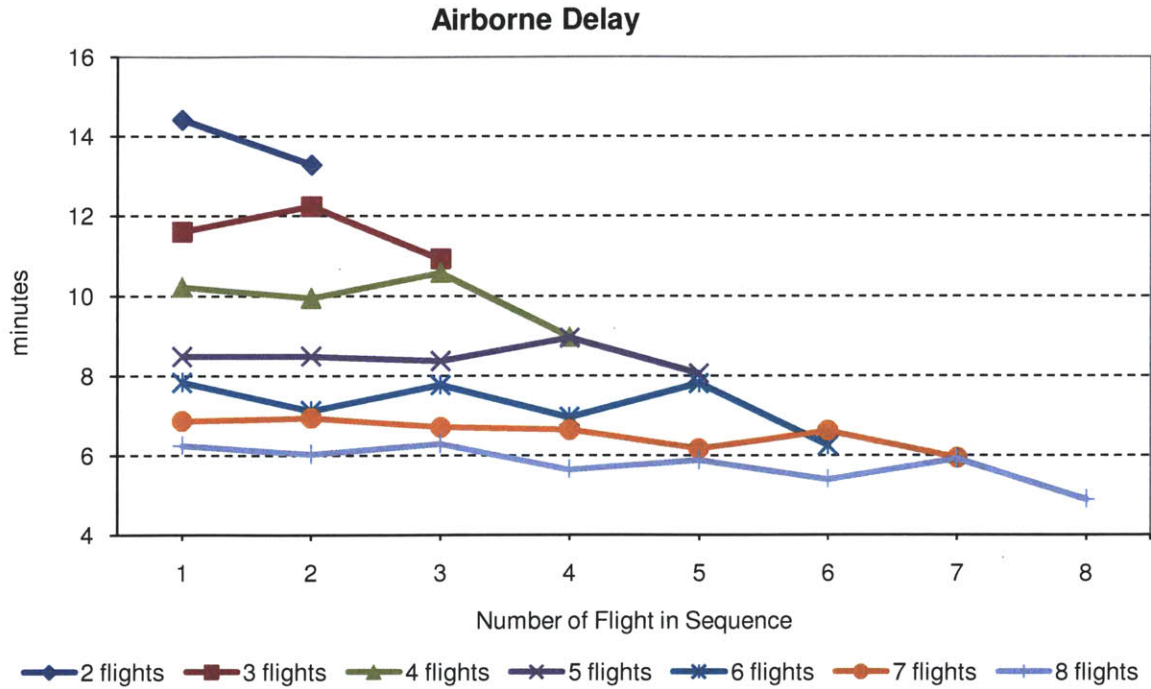
Figure 4-30: Airborne Delay by Flights per Aircraft
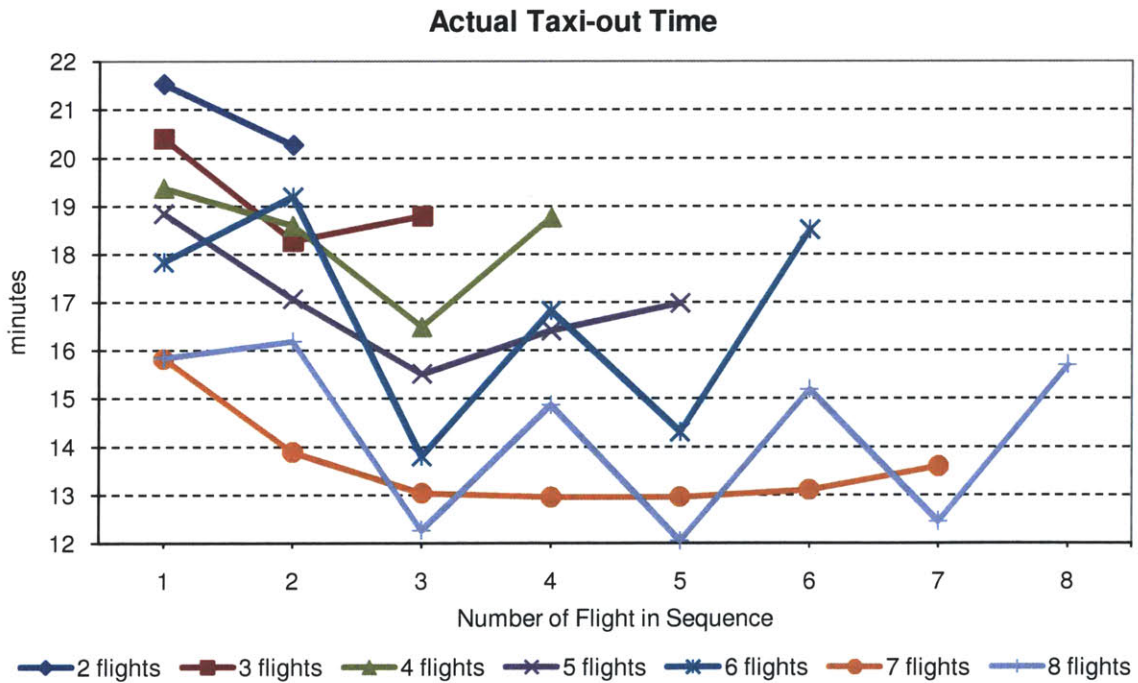
Figure 4-31: Actual taxi-out Time by Flights per Aircraft
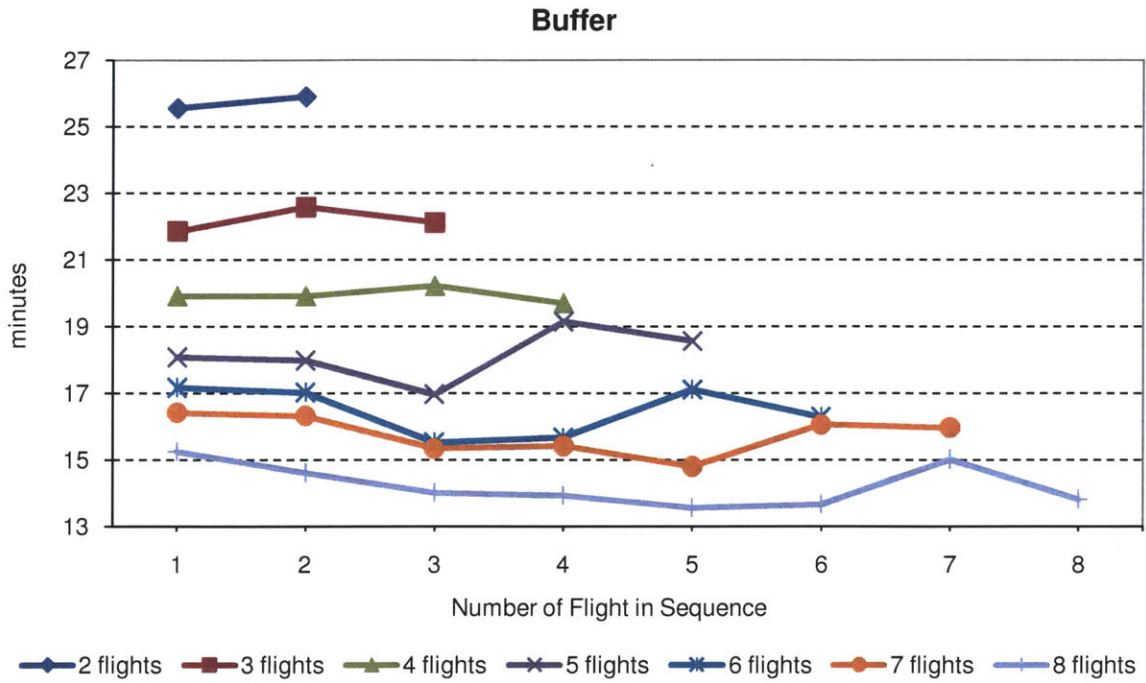
Figure 4-32: Buffer by Flights per Aircraft
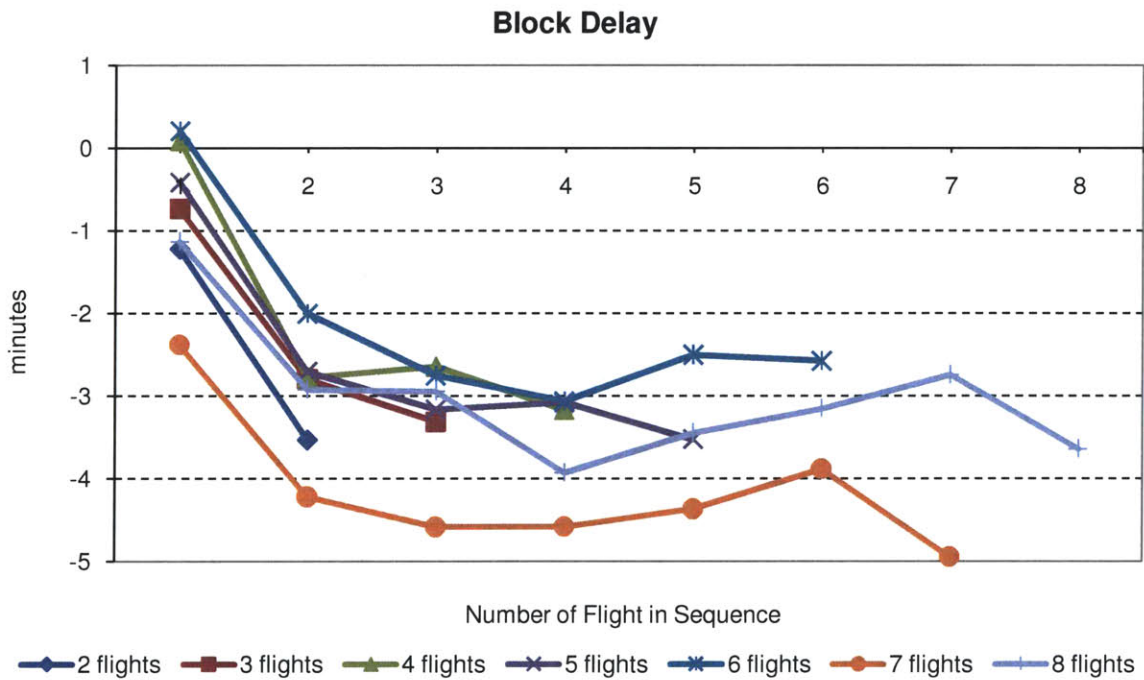
Figure 4-33: Block Delay by Flights per Aircraft
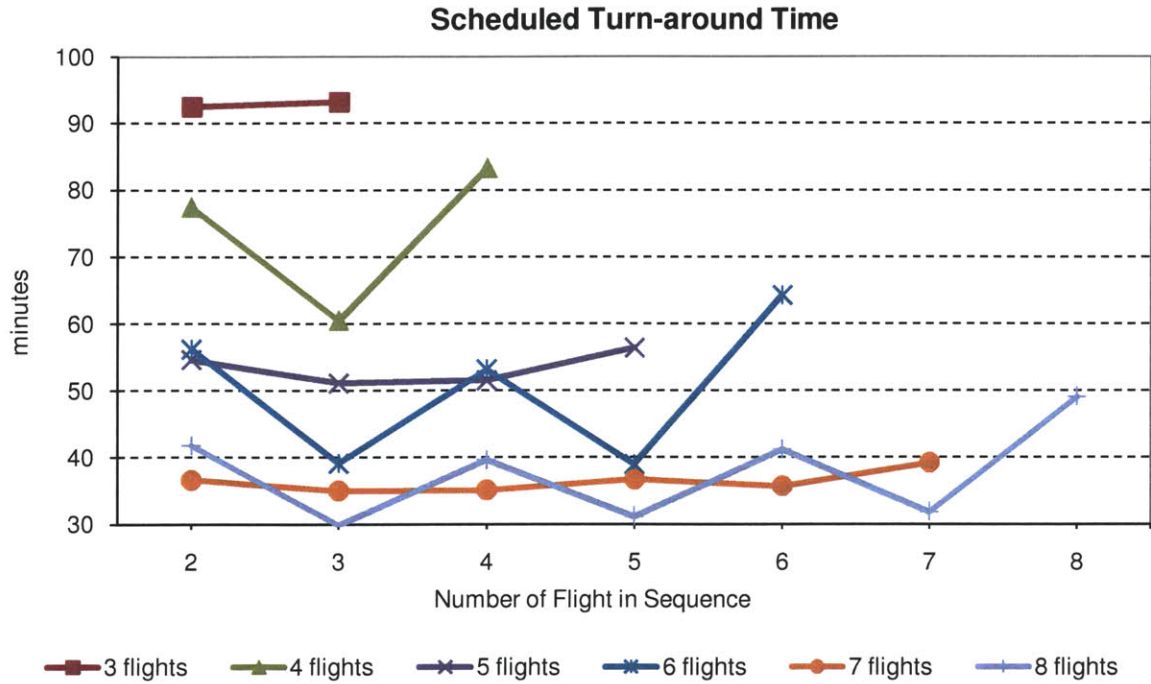
Figure 4-34: Scheduled Turn-around Time by Flights per Aircraft
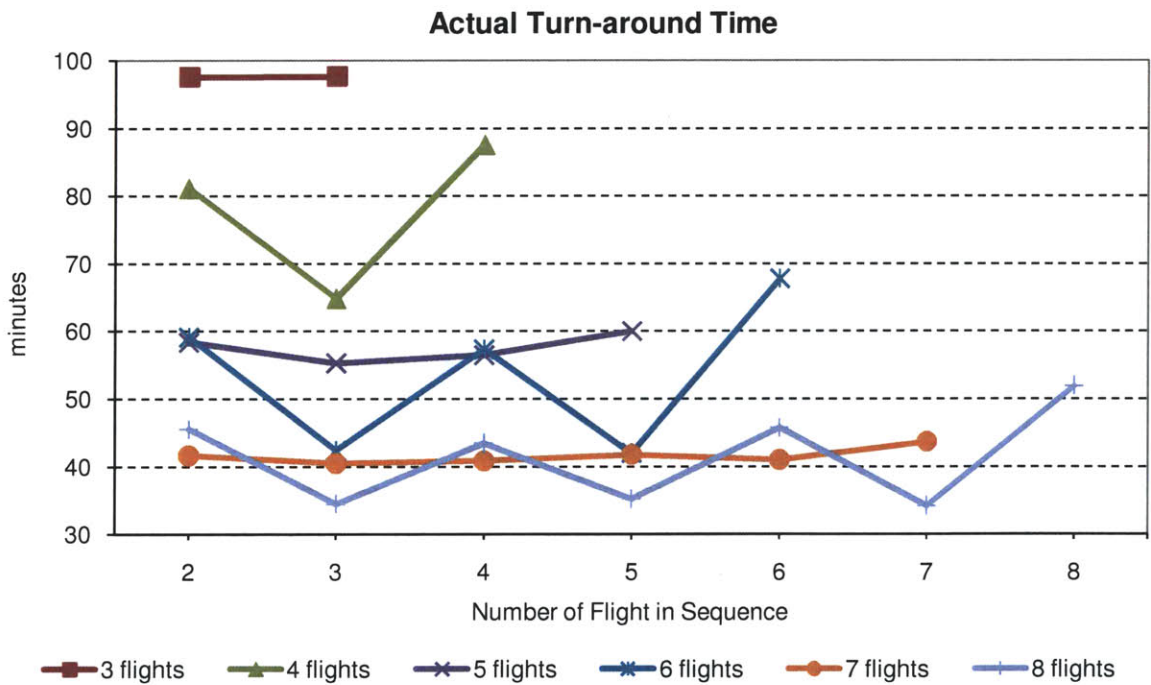
Figure 4-35: Actual Turn-around Time by Flights per Aircraft

The hub effect on flight sequence becomes more evident in the plots of scheduled and actual turn-around times (Figure 4-35 and Figure 4-34). It must be noted, that the first point of every line on these two figures corresponds to the time that an aircraft stays at the gate after the first arrival and prior to the second departure. In sequences with four, six and eight flights, both the actual and the scheduled turn-around times, alternate between long and short values. We assume that the flights in these sequences are between hub and spoke airports. We expect that high turn-around times occur at hubs during "connecting banks" and short times at spoke airports. The rationale behind this is that aircraft arriving at connecting banks must stay longer on ground to allow all connecting passengers adequate time to disembark their arriving aircraft and board the aircraft for their next flight. However, one could argue against this assumption that aircraft departing from spoke cities often have to stay longer on the ground so as to arrive closer to the next connecting bank at the hub.

To examine the validity of our assumption, we separated flights to three groups: (i) flights departing from a specific hub, (ii) flights destined to this hub, and (iii) flights departed from or destined to other airports. A high fraction of departures from a hub, for a certain flight number in the sequence, means a departing bank from this hub. We limited our analysis to six carriers, Delta, American Airlines, United, Continental, Northwest, and US Airways, and to one hub per carrier. The results for 4-flight and 6-flight sequences are summarized in Table 4-5. The shaded entries in the HUB – Spoke column indicate a departing bank, and those in the Spoke-HUB column an arriving bank.

The most important observation in the table below is that the turn-around times during connecting banks at hubs, are always longer than the turn-around times at spokes. Furthermore, turn-around times are longer in 4-flight sequences than in 6-flight sequences. Except for the longer time required for passengers to disembark and board, larger aircraft also need more time for fueling. But most importantly, a delay of an aircraft that carries more connecting passengers on a long-haul route with less frequent service is associated with a much higher reaccommodation cost for disruptive passengers. Therefore, we expect that at connecting banks airlines are usually inclined to schedule large aircraft to arrive first and depart last.

		Delta						American Airlines					
		ATL – Spoke		Spoke – ATL		Other		DFW – Spoke		Spoke – DFW		Other	
		mean	flights	mean	flights	mean	flights	mean	flights	mean	flights	mean	flights
4-flight	2nd	72.2	55%	52.2	9%	53.8	36%	58.5	48%	51.5	10%	57.9	42%
	3rd	67.8	9%	52.4	57%	53.7	34%	57.2	10%	48.5	46%	58.5	44%
	4th	75.9	57%	58.2	10%	58.8	33%	56.8	47%	51.5	9%	60.1	44%
6-flight	2nd	64.4	89%	41.9	4%	50.8	7%	54.5	72%	47.8	3%	54.5	26%
	3rd	58.7	4%	44.3	91%	53.4	6%	56.5	3%	45.5	68%	50.4	30%
	4th	60.7	91%	43.2	4%	48.9	6%	50.9	67%	50.2	3%	55	30%
	5th	67.5	3%	44.4	90%	52.8	7%	47.5	3%	45.6	70%	53.2	28%
	6th	65.5	89%	48	3%	55.1	8%	52.6	70%	43.9	4%	53	27%

		United Airlines						Continental					
		ORD – Spoke		Spoke – ORD		Other		IAH – Spoke		Spoke – IAH		Other	
		mean	flights	mean	flights	mean	flights	mean	flights	mean	flights	mean	flights
4-flight	2nd	74	31%	55.3	4%	65.6	65%	71.9	61%	56.7	11%	66.1	28%
	3rd	54.1	3%	53.1	37%	53.1	59%	66.9	11%	55.5	61%	63.4	28%
	4th	63.2	38%	63.1	5%	71.8	58%	74.9	61%	57.3	7%	70.7	32%
6-flight	2nd	55.7	56%	54.7	3%	61.2	41%	65.2	92%	55	1%	73.8	7%
	3rd	55.9	3%	46.6	50%	45.4	47%	57	1%	51.9	95%	48	4%
	4th	57.5	50%	34	2%	53.7	48%	62.7	95%	50	1%	69.5	4%
	5th	29.3	2%	46.9	42%	49.2	56%	76	1%	48.4	95%	45	4%
	6th	55.4	42%	68	4%	63.8	53%	65.3		–	–	77.6	6%

		Northwest						US Airways					
		DTW – Spoke		Spoke – DTW		Other		CLT – Spoke		Spoke – CLT		Other	
		mean	flights	mean	flights	mean	flights	mean	flights	mean	flights	mean	flights
4-flight	2nd	79	33%	64.6	7%	67.9	61%	71.4	30%	71.7	5%	63.1	65%
	3rd	64.3	6%	61	32%	65.8	62%	66.8	5%	61.6	30%	61.3	65%
	4th	75.7	32%	64.7	6%	81.2	62%	74.3	30%	83.4	4%	71.2	66%
6-flight	2nd	68.9	56%	57	10%	61.7	34%	65.2	55%	65.3	2%	54.1	44%
	3rd	80.1	10%	48.5	58%	49.2	32%	60.9	1%	54.6	58%	54.4	41%
	4th	63.2	57%	54.8	7%	58	37%	64	57%	72.5	1%	59.6	42%
	5th	67.7	7%	49.1	59%	49.7	35%	56.7	1%	52.6	58%	51.8	40%
	6th	66.8	59%	61.3	8%	76.8	33%	67.4	57%	60	1%	59.5	42%

Table 4-5: Actual Turn-around Times (minutes), in 2009

4.4 Summary

In this chapter we explored the variability in flight time components and buffer time, and examined their correlation. Moreover, we analyzed some of the major causal factors of delays, and discussed their impact on schedule padding. The main findings of the analysis are summarized below.

Airline operations experience large uncertainty, and their flight time components are very variable. The magnitude and the effects of this variability differ for each flight component and route because there are many different delay sources. Therefore, it is very difficult to attribute arrival delays to specific causes.

Among all delay components, gate delays have the strongest correlation with arrival delays. Given also that the correlation of gate delays with buffer is minimal, we can conclude that carriers do not take gate delays into account when determining the amount of schedule buffer.

On-time performance is highly affected by the traffic volumes at airports and the delay propagation from previous flights. Taxi-out times and block delays are longer for the first flight of an aircraft at a given day, because of the morning peak in departures. However, on-time performance deteriorates over the course of the day, until it reaches its minimum late in the evening, and then improves again. To minimize the time that aircraft are delayed on the tarmac, Ground Delay Programs are initiated by the FAA and revised departure times are assigned to flights. Despite the fact that a large proportion of flights at several major U.S. airports are affected by GDPs, our analysis suggests that airlines do not adjust their buffer to account for them because it is impossible to predict their impact on individual flights on specific days.

Although on-time performance can be improved by the airline adding more buffer, our findings show that there are cases of flight operations under large uncertainty where airlines use shorter buffer. This practice worsens the on-time arrival statistics for these routes, but enables carriers to allocate buffer to flights where the gains in on-time performance can be larger.

Furthermore, the seasonality test that we performed revealed a very strong seasonal effect in airborne times. This effect is a function of the route direction and increases by flight distance. Similarly, absolute buffer increases almost linearly by distance, whereas buffer as a fraction of nominal block time follows an exponentially decreasing trend by distance.

Aside from distance, congestion at destination and origin airports also has a very strong correlation with the amount of added buffer. Flights out of airports that have lengthy taxi-out delays tend to be padded more, resulting to similar on-time performance with flights departing from less congested airports. However, an airport that has a strong effect on schedule padding as origin does not necessarily have a similar effect as destination.

Additionally, an analysis at six major hubs showed that turn-around times at connecting banks at hubs are larger than turn-around times at spoke airports. On the other hand, we did not find any evidence for a hub-effect on buffer. Furthermore, turn-around times increase by the size of the aircraft.

To further study these results, and validate our assumptions about them, we perform an econometric analysis by constructing several regression models in the next chapter.

Chapter 5

Regression Analysis

In the previous chapter, we examined the most important factors that affect the operations on each flight segment, and formulated assumptions about the way that carriers adjust their schedules based on these factors. In this chapter, we perform an econometric analysis to validate these assumptions, by constructing several regression models to study how (1) buffer and (2) on-time performance depend on the combination of several factors, such as the flight time components, the route competition, the hub effect, the carrier type, and the scheduled arrival time. This chapter is separated in two sections. First, we present the different explanatory variables and test for multicollinearity between them. Then, we run six regression models, we elaborate on the results, and try to explain the effect of the studied parameters in our models.

5.1 Model and Variables

5.1.1 Linear Models

The regression models used for the purposes of our analysis are all linear and take the form of the following equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon$$

where:

y is the dependent (endogenous) variable,

$x_1 \ldots x_n$ are the independent, explanatory (exogenous) variables,

$\beta_0 \ldots \beta_n$ are the regression coefficients (parameters), estimated by the method of Ordinary Least Squares (OLS),

and ε is the error (disturbance) term.

To determine whether the explanatory variables of our models have a statistically significant effect on the dependent variable, we test if each of the coefficients of the explanatory variables is significantly different from 0. Particularly, we perform the following significance test:

Null hypothesis H_0: $\beta_n = 0$

Alternative hypothesis H_1: $\beta_n \neq 0$

Test statistic: (t-value) $t_{\beta_n} = \dfrac{\beta_n}{s_{\beta_n}}$, where s_{β_n} is the standard error of the coefficient β_n.

The decision rule for rejecting the null hypothesis in favor of the alternative is $\left| t_{\beta_n} \right| > t_{crit}$, where t_{crit} depends on the level of significance α and the associated degrees of freedom. For a large number of observations, such as our sample, and a significance level $\alpha = 0.05$, the t_{crit} is equal to 1.96. Thus, a t-value greater than 1.96 (or less than -1.96) indicates that there is sufficient evidence, within a 95% confidence interval, to support the studied parameter as a statistical significant parameter for explaining variation in the dependent variable.

5.1.2   Variables

In Chapter 4, we illustrated that there is a strong seasonality effect in flight time components, with airborne time being the most affected by weather. Figure 5-1 shows the distribution of on-time performance by month in 2009. It is surprising that the difference in on-time performance between November (highest) and December (lowest) is 16.6%. Because of this large variability, we decided to limit the analysis in this chapter to specific months. Particularly, the months that are used in the regression models are January (low on-time performance), August (on-time performance almost equal to average), and November (high on-time performance).
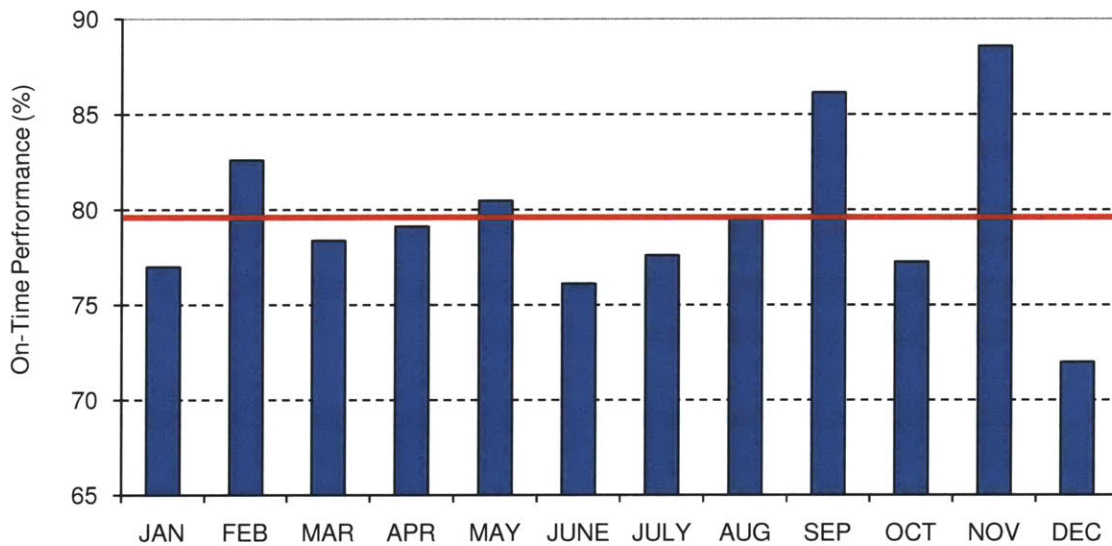
Figure 5-1: On-time Performance by Month, in 2009

In the regressions models discussed in this chapter, the following variables are used:

BUFFER: The buffer time of a single flight operation.

ON_TIME_PERFORMANCE: The on-time performance of all flights on a given non-stop route, carrier and month.

ACT_AIR_AV: The average actual airborne time of all flights on a given non-stop route, carrier and month. The actual airborne time of a single flight has been defined in Section 3.4.

ACT_TO_AV: The average actual taxi-out time of all flights on a given non-stop route, carrier and month. The actual taxi-out time of a single flight has been defined in Section 3.4.

ACT_TI_AV: The average actual taxi-in time of all flights on a given non-stop route, carrier and month. The actual taxi-in time of a single flight has been defined in Section 3.4.

GATE_DELAY_AV: The average gate delay of all flights on a given non-stop route, carrier and month. The gate delay of a single flight has been defined in Section 3.4.

COMPET_1, 2, 3, 4, 5: Dummy variables indicating the route competition; 1 = one carrier operates in the route, 2 = two carriers operate on the route, 3 = three carriers operate on the route, 4 = four carriers operate on the route, 5 = at least five carrier operate on the route.

HUB_DEP: A dummy variable that takes on value 1 if the origin airport is a hub for the carrier operating the flight, and 0 otherwise. Hub airports are shown in Table 5-1.

HUB_ARR: A dummy variable that takes on value 1 if the destination airport is a hub for the carrier operating the flight, and 0 otherwise.

ARR_7,8, ..., 23: Dummy variables indicating the scheduled arrival time of a flight; 7 = scheduled arrival time before 8am, 8 = scheduled arrival time between 8:00am and 8:59 am, ..., 23 = scheduled arrival time between 11pm and 11:59 pm.

SWA: A dummy variable that takes on value 1 if the flight is operated by Southwest, and 0 otherwise.

NLC: A dummy variable that takes on value 1 if the flight is operated by a legacy carrier, and 0 otherwise. The legacy and low cost carriers are shown in Table 5-2.

LCC: A dummy variable that takes on value 1 if the flight is operated by a low cost carrier, and 0 otherwise.

OTHER: A dummy variable that takes on value 1 if the flight is operated by a carrier other than Southwest and the airlines listed in Table 5-2, and 0 otherwise.

Carrier	Designated Hubs
American Airlines (AA)	DFW, ORD, MIA
Alaska Airlines (AS)	SEA
Continental Airlines (CO)	IAH, EWR, CLE
Delta Airlines (DL)	ATL, CVG, SLC
Frontier Airlines (F9)	DEN
JetBlue Airways (B6)	JFK
Northwest (NW)	MSP, DTW, MEM
Southwest (WN)	LAS, MDW, PHX, BWI, DEN
AirTran Airways (FL)	ATL
United Airlines (UA)	ORD, DEN ,SFO, IAD
US Airways (US)	PIT, CLT, PHX, PHL

Table 5-1: Hub Airports

Legacy Carriers (NLC)	Low Cost Carriers (LCC)
American Airlines (AA)	AirTran Airways (FL)
Continental Airlines (CO)	JetBlue Airways (B6)
Delta Air Lines (DL)	Frontier Airlines (F9)
Northwest Airlines (NW)	Spirit Air Lines (NK)
Unites Airlines (UA)	Virgin America (VX)
US Airways (US)	

Table 5-2: Legacy and Low Cost Carriers

5.1.3 Multicollinearity

Before we construct any regression models, it is important that we test for multicollinearity, which occurs when two or more independent variables are highly correlated. When multicollinearity between independent variables exists, the estimated regression parameters have large sampling variability, and this may result in misinterpretation of the relationship between variables. Moreover, multicollinearity increases the standard errors of the coefficients for individual independent variables. Thus, variables that in reality are significant may appear to be insignificant in the regression model.

To detect multicollinearity, we estimated the pairwise correlation between the independent variables (Table 5-3). We notice that *BUFFER* is highly correlated with *ACT_AIR_AV*. This is intuitive, as longer flights have larger airborne delays, and thus carriers add more buffer in their scheduled block times. In the models where *BUFFER* is used as the dependent variable, this high correlation between the two variables is desirable. However, in the model with *ON_TIME_PERFORMANCE* as the dependent variable, we should use only one of the two variables.

Moreover, there is a moderate negative correlation between SWA and *ACT_TO_AV*, indicating that Southwest has shorter taxi-out times than other airlines. The same likewise applies for the taxi-in times, but the correlation coefficient is smaller.

Furthermore, *NLC* and *OTHER* are correlated with *HUB_DEP* and *HUB_ARR*. *NLC* has a positive correlation with these two variables, whereas *OTHER* has a negative. This occurs because hub airports have been defined for all carriers designated as *NLC*, while from *OTHER* carriers only Alaska is studied for its hub operations. Additionally, there is a small positive correlation between *NLC* and *ACT_AIR_AV*, *ACT_TO_AV*, and *ACT_TI_AV*. The first relation shows that on average, legacy carriers fly longer distances (Figure 3-3). The second and the third relations suggest that legacy carriers either operate on more congested airports than the rest of the carriers, or operate during periods of very high demand. Last, we observe a small correlation between *ACT_TI_AV* and *HUB_ARR*. This is likely the result of the large number of very closely scheduled arrivals at hubs during connecting banks. Despite the indication of correlation between the above mentioned variables, its magnitude is relative small, and their simultaneous use in the regression models is not expected to cause multicollinearity issues.

	BUFFER	ACT_AIR_AV	ACT_TO_AV	ACT_TI_AV	GATE_AV	SWA	LCC	NLC	OTHER	COMPET_2	COMPET_3	COMPET_4	COMPET_5	HUB_DEP	HUB_ARR
BUFFER	1.000	0.626	0.337	0.184	0.246	-0.045	-0.023	0.185	-0.147	0.012	0.009	-0.009	0.014	0.114	0.089
ACT_AIR_AV		1.000	0.183	0.134	0.104	-0.126	0.072	0.295	-0.258	0.095	0.027	-0.061	-0.076	-0.056	0.142
ACT_TO_AV			1.000	0.213	0.211	-0.435	0.026	0.316	0.109	0.051	0.097	0.069	0.074	0.181	-0.052
ACT_TI_AV				1.000	0.044	-0.354	-0.007	0.277	0.014	0.002	0.089	0.058	0.092	-0.178	0.365
GATE_AV					1.000	-0.046	-0.034	-0.118	0.192	0.028	0.043	0.021	0.008	0.055	-0.172
SWA						1.000	-0.164	-0.422	-0.313	0.237	-0.076	-0.122	-0.069	-0.037	-0.009
LCC							1.000	-0.281	-0.208	0.065	0.056	0.051	-0.003	0.003	0.004
NLC								1.000	-0.536	-0.003	0.022	-0.006	0.032	0.346	0.343
OTHER									1.000	0.027	0.046	0.034	-0.001	-0.372	-0.372
COMPET_2										1.000	-0.351	-0.220	-0.149	0.002	-0.005
COMPET_3											1.000	-0.160	-0.108	-0.014	0.003
COMPET_4												1.000	-0.068	0.003	-0.015
COMPET_5													1.000	-0.020	-0.020
HUB_DEP														1.000	-0.239
HUB_ARR															1.000

Table 5-3: Correlation Matrix for the Independent Variables

5.2 Buffer Time Regression Models

In this section, we construct five regression models with *BUFFER* as the dependent variable to test the effect of various variables upon schedule padding. All models were run separately for three months of 2009: January, August and November. This was done to study whether the estimated coefficients are consistent across months.

5.2.1 Model 1: Average Flight Times

Objective:

Study to what degree buffer is associated with the average flight times on a certain route, for a given carrier and month.

Hypotheses:

We expect that actual airborne time will have a very strong positive relationship with buffer because as has been shown in Section 4.3.4, buffer increases proportionally to distance. Similarly, actual taxi-out time should have a significant positive effect, since taxi out delays are responsible for a large extent of block delays and can be absorbed through buffer.

On the other hand, actual taxi-in time and gate delay are expected to have a smaller effect. In Chapter 4, we showed that taxi-in times are very short and have a limited variability. Thus, we assumed that airlines do not take them into account for adjusting buffer. This argument was further supported by the distribution of the correlation parameter of taxi-in delays with buffer (Figure 4-7). Also, we expect that carriers do not adjust buffer based on gate delays because these are very unpredictable. Moreover, if a route has consistent lengthy gate delays, it would be more effective for carriers to adjust their turn-around times rather their scheduled block times.

Results:

Table 5-4 shows the results for Model 1. The adjusted R^2 varies from 40.6% to 48.1% meaning that the actual performance over a period of time explains a substantial portion of the variation in padding. Contrary to our hypothesis, all variables are statistically significant at a 1% significance level, with *ACT_AIR_AV* being the most significant and *ACT_TI_AV* the least.

The constant parameter varies from -1.28 to 1.95. This is the amount of buffer that a flight would have if all flight time components were zero. However, this can never be the case. The regression coefficients represent the increase in buffer for every additional minute in the mean of each flight time component. For example, a 30 minutes average airborne time in January will increase buffer by 2.4 minutes. Similarly, 10 minutes of average gate delay, taxi-out time and taxi-in time will increase buffer by 1.9, 3.6 and 2.2 minutes respectively, resulting in a total buffer of 12.1 minutes. If the airborne time were 240 minutes, the buffer would be 18.9 minutes.

Although the coefficients vary across months, their aggregate effect is almost constant. Particularly, a flight with 30 minutes average airborne time will have 11.6 minutes of buffer in August and 13.3 minutes in November, whereas for an average airborne time of 240 minutes, buffer will increase to 22.1 and 28 minutes respectively. We notice that the effect of airborne time is limited in August. The reason is that airborne delays are shorter in August, most probably due to better weather conditions, meaning that the variability in airborne time is lesser, and thus carriers add less buffer to account for them.

	JANUARY		AUGUST		NOVEMBER	
	Coeffic.	t-Stat	Coeffic.	t-Stat	Coeffic.	t-Stat
Intercept	1.95	37.867	-0.71	-15.889	-1.28	-22.582
ACT_AIR_AV	0.08	511.144	0.05	307.135	0.07	425.966
GATE_AV	0.19	99.846	0.22	128.254	0.17	61.526
ACT_TO_AV	0.36	161.589	0.58	296.127	0.72	256.630
ACT_TI_AV	0.22	45.799	0.28	63.143	0.36	65.401
Adjusted R^2	0.4805		0.4058		0.4514	
Sample Size (flights)	400,793		434,532		388,297	

Table 5-4: Model 1

5.2.2 Model 2: Average Flight Times from 2008

Objective:

(1) Test if the use of the average values of the flight times from 2008 instead of 2009 is more effective in explaining the variability in buffer, (2) and study if the effect of flight time components is similar in different years.

Hypotheses:

One of the parameters that airlines take into account when determining the schedule buffer on a route, is their past performance in previous months and years. In Section 4.3.1 it was shown that there is a strong seasonality in flight times. The actual flight times of the same month in the previous year can be used for forecasting the actual flight times in the current year. Thus, we expect that there is a strong correlation between the buffer times in 2009 and the actual flight times in 2008. If this assumption holds true, using 2008 data, instead of 2009, for the independent variables of Model 1 should improve the goodness of fit of the model. Furthermore, because of the seasonality in flight times and their associated delays across years, we do not expect large changes in the coefficient parameters.

Results:

Comparing the R^2 for the regressions in Model 1 (2009 data) with those in Model 2 (2008 data), we notice a small improvement in the fit of the latter for all three months. This supports our hypothesis that flight data from the previous year provide more information about the amount of buffer used, compared to flight data of the same year.

Furthermore, the estimated parameters of the four independent variables, with the exception of *GATE_AV*, are very similar to those of Model 1. This can be attributed to the fact that the mean values and the magnitude of variation of the dependent variables are very similar across consecutive years.

Nonetheless, it is very important that carriers forecast any changes in air traffic and take into account special events to adjust their schedule buffer accordingly. Otherwise, the use of historical data can result in an over or underestimation of optimal buffer, causing either an unnecessary increase in operational cost or a reduction in on-time performance.

	JANUARY		AUGUST		NOVEMBER	
	Coeffic.	t-Stat	Coeffic.	t-Stat	Coeffic.	t-Stat
Intercept	2.00	38.752	0.44	10.120	0.08	1.550
ACT_AIR_AV	0.08	481.835	0.04	282.427	0.07	394.333
GATE_AV	0.11	82.390	0.11	74.922	0.13	60.256
ACT_TO_AV	0.41	168.426	0.57	303.171	0.67	255.829
ACT_TI_AV	0.16	31.506	0.25	56.687	0.18	33.564
Adjusted R^2	0.4860		0.4263		0.4557	
Sample Size	367,399		395,866		362,710	

Table 5-5: Model 2

5.2.3   Model 3: Competition and Hub Effects

Objective:

Study how competition and hub operations affect padding.

Hypotheses:

We expect that carriers use more buffer on flights destined to their hub, compared to other flights, so as to allow adequate time for passenger connections. The reason is that a late arrival at a connecting bank is associated with a higher cost because of the re-accommodation of connecting passengers. Flights departing from a carrier's hub are expected to be padded similarly to spoke-to-spoke flights.

Competition may affect a carrier's decision to shorten or lengthen a flight's scheduled block time in several ways: screen presence in the global distribution systems, public perception about quality of service, and scheduling at peak travel times.

From the quality of service perspective, competition should have a positive impact on buffer, since we expect that all else being equal, passengers would prefer the airline that reports the best on-time performance on a certain route. This agrees with the literature, which suggests that competition has a positive impact on carriers' on-time performance (see Section 2.3). Additionally, competition affects buffer time by forcing competitive airlines to match their schedules and offer departures during peak hours.

On the other hand, in competitive routes, shorter buffer can offer a competitive advantage to carriers, because flights are listed in the global distribution systems according to their scheduled block time, with the shortest flight having the best screen presence. Copeland et al.[35] illustrated that, on average, 70% of all bookings were made from the first six options presented to a travel agent. Thus, from a marketing perspective, it might be more valuable for an airline to have a better screen presence in the distribution systems rather than a better on-time performance. However, the emergence of direct booking from airlines websites and other internet travel intermediaries limits the benefits associated with shorter block times, because users can customize the ranking criteria, with price being the most common criterion. Since the effects of competition in block time length are contradictory, we cannot make any reasonable hypothesis towards a specific direction.

Results:

The positive coefficient estimates of *HUB_ARR* (Table 5-6) support our hypothesis that flights destined to the carriers' hub are padded more, when compared to non-hub flights. Particularly, the increase in buffer for hub departures ranges from 0.47 to 1.61 minutes. However, this trend cannot be generalized for all hub operations, because *HUB_DEP* has opposite signs in different months. Although the coefficient of *HUB_DEP* is negative only in August (and even then it is very close to zero), it could be interpreted as a result of less flexibility in lengthening the block time of flights departing from connecting banks at hubs, since these flights have significantly longer turn-around times (Section 4.3.8). Another possible explanation for the opposite signs, is that the origin airport has very limited effect on delays and consequently on buffer.

Although all competition indicator variables are statistically significant at a 1% significance level, the effect of competition on a certain route is not captured in the models. Coefficient estimates of competition indicator variables are quite small and vary in sign, depending on the month. As stated in our hypothesis, competition can have contradictory effects in padding. These may depend heavily on the flight demand, the offered frequency, the traffic (local vs. connecting), the competing carriers, etc.

---

[35] Copeland, D., Mason, R., McKenney, J., *Sabre: The development of information-based competence and execution of information-based competition*, IEEE Annals of the History of Computing, Vol. 17, No. 3, 1995.

The above results suggest that airlines allocate buffer in a way that not only improves individual flights' on-time performance, but also maximizes market share and minimizes delay cost at a network level. This means that legacy carriers, with hub operations and a large fraction of connecting passengers, might use more buffer on certain flights that are considered more important for the airline profitability, operational integrity and public perception, even if the marginal gains in on-time performance from this practice are limited.

	JANUARY		AUGUST		NOVEMBER	
	Coeffic.	t-Stat	Coeffic.	t-Stat	Coeffic.	t-Stat
Intercept	1.92	36.182	-0.64	-13.829	-1.66	-28.848
ACT_AIR_AV	0.08	492.925	0.05	291.616	0.07	402.659
GATE_AV	0.1942	101.498	0.2239	130.759	0.1811	66.688
ACT_TO_AV	0.36	153.243	0.59	290.666	0.69	236.986
ACT_TI_AV	0.18	34.323	0.23	47.702	0.28	47.205
COMPET_2	-0.19	-6.14221	-0.16	-5.54775	0.33	10.32842
COMPET_3	0.24	6.675	-0.36	-10.730	0.72	18.996
COMPET_4	1.07	22.264	0.16	3.710	1.65	31.267
COMPET_5	1.35	20.104	-0.50	-9.592	1.50	22.396
HUB_ARR	0.47	16.009	0.92	34.936	1.61	52.084
HUB_DEP	0.18	6.454	-0.15	-5.600	1.22	39.808
Adjusted R^2	0.4822		0.4082		0.4577	
Sample Size	400,793		434,532		388,297	

Table 5-6: Model 3

5.2.4 Model 4: Differences in Padding across Carriers

Objective:

Study if the extent of padding varies across carrier groups (i.e. legacy carriers, low cost carriers, Southwest, and others). Southwest is studied separately from the other low cost carriers because it has a different network structure with very short stage length.

Hypotheses:

For flights with the same stage length, one would expect that Southwest uses less buffer than other airlines, because it flies mostly between secondary non-congested airports. However, Southwest's on-time performance is among the highest in the industry (87%), despite the fact that it has very short scheduled turn-around times. According to Gittell[36], this is attributed, to some extent, to the excellent relational coordination of the company. However, it might also be that Southwest allocates more buffer in scheduled block times, whereas other carriers add more slack in scheduled turn-around times, rather than in block times. Model 4 sheds some light on this strategy.

We also expect that legacy carriers pad their schedules more than low cost and *OTHER* carriers because they carry more connecting and business passengers, who are more sensitive to schedule reliability.

Results:

The coefficient estimates of dummy variables for carrier groups reveal some differences in padding practices among carriers, especially between Southwest and the other groups. The base of comparison is the flights operated by airlines other than Southwest and those included in Table 5-2.

First, Southwest's flights have 4.2 to 9.6 minutes additional buffer than the rest of the carriers. This explains how Southwest achieves a very high on-time performance, although it schedules very short turn-around times. On the other hand, low cost carriers pad their flights the least, and this results in a poor on-time performance (78.5%). Legacy carriers pad their flights

---

[36] Gittell, J., *Supervisory Span, Relational Coordination and Flight Departure Performance: A Reassessment of Postbureaucracy Theory,* Organization Science, Vol. 12, Issue 4, 2001.

very similarly to *OTHER* carriers, and they have almost the same on-time performance, 80.9% and 81.2% respectively.

First of all, these findings show that carriers with more buffer have on average a better on-time performance. Furthermore, Southwest's example shows that when a carrier does not operate at congested airport, it might be more effective for absorbing delays to add more slack in block times, rather than in turn-around times. However, this practice has a higher cost because usually crews are paid on the maximum of scheduled and actual block time, whichever is higher.

	JANUARY		AUGUST		NOVEMBER	
	Coeffic.	*t-Stat*	*Coeffic.*	*t-Stat*	*Coeffic.*	*t-Stat*
Intercept	-4.31	-69.506	-3.66	-70.419	-7.05	-105.902
ACT_AIR_AV	0.08	502.421	0.05	304.675	0.07	421.214
GATE_AV	0.16	87.853	0.17	99.617	0.12	44.786
ACT_TO_AV	0.58	234.842	0.69	317.904	0.93	301.254
ACT_TI_AV	0.46	96.473	0.44	96.398	0.59	106.977
NLC	-0.10	-3.321	-0.38	-13.285	0.71	21.500
LCC	-0.83	-18.926	-1.94	-48.246	-2.46	-53.675
SWA	6.99	163.712	4.16	104.579	7.16	159.302
Adjusted R^2	0.5226		0.4293		0.4989	
Sample Size	400,793		434,532		388,297	

Table 5-7: Model 4

5.2.5   Model 5: Time of the Day

Objective:

Study the impact of the time of the day (scheduled arrival time) on padding practices.

Hypotheses:

Following our discussion in Section 4.3.2, we expect that the time of day has a strong influence on the amount of buffer that carriers add to their block times. Particularly, we expect that flights scheduled to arrive in the evening peak are padded more, so as to absorb more effectively the propagated delays from previous flights, and any further delays caused by airport and airspace congestion.

On the other hand, our analysis in Chapter 4 did not provide enough evidence for the existence of a morning peak in padding.

Results:

The coefficient estimates of the dummy variables for scheduled arrival time (shown in Figure 5-2 and in Table 5-8) reveal a large fluctuation in padding practices over the course of the day. All coefficients are positive, indicating that flights scheduled to arrive before 8am are padded the least.

Flights scheduled to arrive between 8am and 12am, as well as between 5pm and 9pm, have more buffer compared to the rest of the flights. This is in agreement with our initial hypothesis about the effect of congestion and delay propagation on schedule padding. Furthermore, the regression coefficients are almost the same in January and November. On the other hand, August flights have substantially more buffer in the evening peak compared to the morning peak. This could be attributed to the fact that thunderstorms which disrupt a large proportion of flights at U.S. airports during the summer months, occur most likely in the evening.

Another important observation is that all competition indicator variables have positive coefficients. Furthermore, most of the coefficient estimates increase with the number of competitors on a route, suggesting that buffer in general increases with the number of carriers competing on the same route.
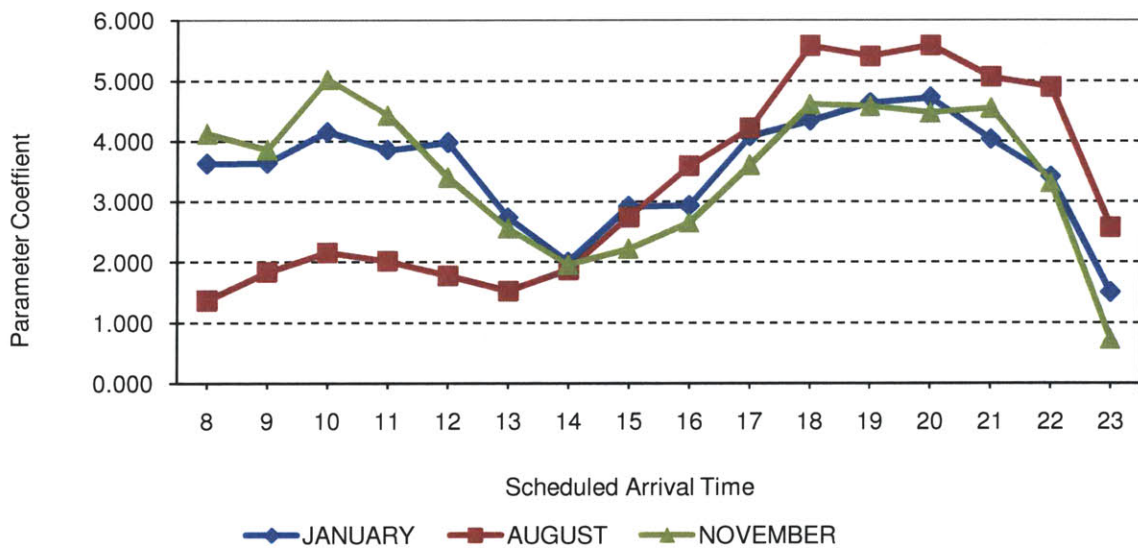

Figure 5-2: Parameter Coefficients of Dummy Variables for Scheduled Arrival Time

	JANUARY		AUGUST		NOVEMBER	
	Coeffic.	t-Stat	Coeffic.	t-Stat	Coeffic.	t-Stat
Intercept	-8.97	-108.944	-7.29	-106.174	-11.85	-136.566
ACT_AIR_AV	0.08	508.137	0.05	305.825	0.08	431.728
GATE_AV	0.16	86.803	0.16	94.866	0.12	47.901
ACT_TO_AV	0.55	221.574	0.69	319.094	0.89	286.928
ACT_TI_AV	0.43	84.614	0.38	80.873	0.51	87.757
COMPET_2	0.50	17.343	0.46	16.588	1.19	39.098
COMPET_3	1.08	32.132	0.43	13.339	1.79	49.341
COMPET_4	1.66	36.170	0.91	22.301	2.35	47.139
COMPET_5	1.71	26.813	0.04	0.753	1.88	29.737
HUB_ARR	0.02	0.773	1.06	35.210	0.96	27.637
HUB_DEP	0.02	0.626	-0.07	-2.165	0.61	17.793
NLC	-0.14	-3.545	-0.83	-23.014	0.08	1.848
LCC	-0.90	-19.773	-2.32	-55.953	-3.09	-65.225
SWA	6.93	153.225	3.78	90.404	6.66	139.108
ARR_8	4.65	57.380	1.96	28.077	5.07	60.815
ARR_9	4.86	64.682	2.52	39.580	4.82	60.651
ARR_10	5.27	76.220	2.89	46.973	5.93	81.101
ARR_11	4.85	66.627	2.57	40.488	5.37	70.390
ARR_12	4.95	69.421	2.47	39.156	4.24	56.590
ARR_13	4.04	56.502	2.35	37.487	3.72	50.034
ARR_14	3.12	43.343	2.50	39.624	2.89	38.375
ARR_15	3.96	55.156	3.33	52.415	3.02	39.938
ARR_16	4.10	58.173	4.31	70.223	3.79	51.390
ARR_17	5.28	74.181	4.87	78.902	4.56	60.841
ARR_18	5.46	75.872	6.22	100.680	5.54	74.007
ARR_19	5.63	79.064	6.03	96.075	5.49	72.870
ARR_20	5.50	79.204	6.18	100.772	5.30	72.249
ARR_21	4.84	67.411	5.56	88.753	5.20	68.199
ARR_22	4.64	60.405	5.60	83.347	4.44	54.635
ARR_23	2.53	31.232	3.22	44.117	1.63	18.659
Adjusted R^2	0.5380		0.4645		0.5204	
Sample Size	400,793		434,532		388,297	

Table 5-8: Model 5

5.3   On-Time Performance Regression Models

In this section, we construct a regression model with *ON_TIME_PERFORMANCE* as the dependent variable. In contrast to the previous models, where each entry is one flight, every entry of this model corresponds to one flight number. Only flight numbers that had at least ten flights in the studied month were included in the sample. *ON_TIME_PERFORMANCE, BUFFER, AIRBORNE_AV, GATE_AV, ACT_TO_AV,* and *ACT_TI_AV* were estimated as the average values over all flights with the same flight number in a month. In the cases where flights with the same flight number had different scheduled arrival hour, the indicator variables for the scheduled arrival time were determined based on the median arrival time.

Objective:

Study to what degree the variables studied in Models 1 through 5 affect the on-time performance.

Hypotheses:

First, we expect that longer buffer will be associated with better on-time performance. Although adding more buffer increases the likelihood that a flight will arrive on-time, it does not follow that the flights with the longest buffer have also the best on-time performance. The flights that need substantial buffer are those that suffer from the largest uncertainty, and thus are more likely to arrive late with respect to their nominal arrival time. However, the gains from adding more buffer to these flights are small, as shown in Chapter 4. On the other hand, flights with less variability in their flight time components can achieve a good on-time performance without having excessive buffer.

Considering that long-haul routes are associated with a larger variability in airborne times, one would expect that distance has a negative effect on on-time performance. On the other hand, carriers often absorb some of the previously suffered delays by increasing the cruising speed of the aircraft, with long flights being benefited at most from this practice. Also, given that buffer increases with distance, we cannot make any reasonable hypothesis about the correlation of distance with on-time performance.

Regarding gate delays, taxi-out times, and taxi-in times, we expect that longer times are associated with worse on-time performance. Gate delays should be the most important because, as has been shown in Chapter 4, they are very highly correlated to arrival delays.

Furthermore, according to existing literature (see Section 2.3), competition should have a positive impact on on-time performance. We also expect that flights that are scheduled to arrive at peak hours, both morning and evening, will have a worse on-time performance, although they are padded more. The rationale is that the likelihood of a late arrival increases exponentially with traffic.

In section 5.2.3, it was shown that, on average, airlines pad more the flights destined to their hubs, compared to non-hub flights. However it is not clear yet if this practice also results to higher on-time performance. The reason is that connecting banks at hubs consist of a large number of very closely scheduled arrivals followed by many very closely scheduled departures. This often results to short periods of large congestion, and thus to excess taxi-in and taxi-out delays. Model 6 is expected to show whether carriers use adequate buffer to absorb these delays.

Results:

The variable coefficient estimates in Table 5-9 represent the change in on-time performance (in percentage points) that each variable is associated with. The fields in bold are the statistically significant variables at a 5% significance level. The most significant variables are *BUFFER, GATE_AV, ACT_TO_AV*, and *ACT_TI_AV*. From the remaining variables only *COMPET_4, LCC, SWA, ARR_19*, and *ARR_21* are significant at a 5% level for all three months. *COMPET_2*, and *COMPET_3* are not significant for any month, whereas the significance of the other variables varies across months.

First, we notice that buffer is associated with an increase in on-time performance. Particularly, 10 additional minutes of buffer increase the percentage of flights that arrive on time by approximately 2 percentage points, depending on the month. The largest effect of every extra minute of buffer is observed in August and the smallest in January.

At this point we should mention that *BUFFER* is highly correlated with *AIRBORNE_AV* (see Section 5.1.3). In order to avoid multicollinearity in Model 6, *AIRBORNE_AV* was not included in the regression. To estimate its effect we run separate regressions with *AIRBORNE_AV* instead of *BUFFER*. Based on the results, *AIRBORNE_AV* is statistically significant at a 1% level for the three months, and its per minute impact on on-time performance varies from -0.003 percentage points in January, to 0.011 in August, and 0.016 in November. For a 60 minutes flight, this is translated to an absolute change in on-time performance of -0.16, 0.63, and 0.98 percentage points, respectively. This result contradicts our hypothesis about the negative effect of airborne

time, and suggests that long haul flights often have better on-time performance than short haul flights. Although this is explained to some extent by the increase of buffer with distance, it could also be attributable to the fact that on long flights carriers can absorb some of the suffered delays by increasing the cruising speed of the aircraft.

As expected, *GATE_AV, ACT_TO_AV,* and *ACT_TI_AV* have all negative effects on on-time performance. It is surprising that every minute of average gate delays reduces on-time performance by almost 1 percentage point. The effect of the taxi-out time is approximately the half. This difference can be attributed to the fact that taxi-out delays are more related to buffer than gate delays are, and thus can be absorbed easier, resulting in less arrival delays.

Regarding competition, there is no evidence of differences between the on-time performance of non competitive routes and routes with two or three competitors. However, when more than three airlines compete on the same route their on-time performance increases. This could be explained by the results of Models 3 and 5, which show that on routes with four or five competitors the used buffer is longer compared to less competitive routes. The only exception to this result, were routes with five competitors in August, but in this case competition does not have a statistically significant effect on on-time performance (Table 5-9).

The sign and the significance of *HUB_DEP* and *HUB_ARR* vary across months. In January, which is the month with the worst on-time performance, hubs are associated with a reduction in on-time performance by approximately 1 percentage point for both arrivals and departures. In August, the same effect applies also to hub departures, whereas hub arrivals are associated with a slightly better on-time performance. Both variables are insignificant in November.

As far as the carrier groups are concerned, low cost carriers are consistently associated with a worse on-time performance. On the other hand, the coefficient signs of *SWA* and *NLC* vary across months, suggesting that any differences in on-time performance between Southwest, legacy carriers and *OTHER* carriers might be attributed to other factors, such as network structure, airports, and allocated buffer.

Last, the indicator variables for the scheduled arrival time illustrate a high reduction in on-time performance of the flights that are scheduled to arrive during the evening peak. This is expected based on our discussion in Section 4.3.3.

	JANUARY		AUGUST		NOVEMBER	
av(ON_TIME_PERFORMANCE)	77.02%		79.68%		88.59%	
	Coeffic.	t-Stat	Coeffic.	t-Stat	Coeffic.	t-Stat
Intercept	0.9836	311.552	0.9905	452.394	0.9775	451.756
BUFFER	0.0017	30.454	0.0026	53.208	0.0022	57.708
GATE_AV	-0.0096	-126.784	-0.0096	-167.915	-0.0097	-145.013
ACT_TO_AV	-0.0074	-68.657	-0.0048	-58.350	-0.0047	-52.940
ACT_TI_AV	-0.0034	-16.509	-0.0065	-39.956	-0.0033	-21.079
COMPET_2	0.0006	0.502	0.0014	1.593	-0.0004	-0.510
COMPET_3	0.0021	1.578	0.0006	0.635	0.0012	1.320
COMPET_4	0.0103	5.983	0.0031	2.473	0.0033	2.778
COMPET_5	0.0153	6.626	0.0005	0.322	0.0101	6.383
HUB_DEP	-0.0124	-9.667	-0.0146	-15.747	0.0002	0.219
HUB_ARR	-0.0112	-8.496	0.0026	2.710	-0.0008	-0.860
LCC	-0.0120	-6.605	-0.0213	-15.026	-0.0281	-23.173
NLC	0.0144	9.575	-0.0062	-5.473	0.0005	0.507
SWA	0.0090	5.256	-0.0036	-3.108	0.0142	12.092
ARR_8	-0.0009	-0.298	-0.0077	-3.685	-0.0061	-3.003
ARR_9	0.0009	0.319	-0.0049	-2.453	-0.0034	-1.739
ARR_10	-0.0020	-0.720	-0.0103	-5.254	-0.0055	-2.927
ARR_11	-0.0015	-0.516	-0.0096	-4.766	-0.0021	-1.090
ARR_12	-0.0018	-0.655	-0.0084	-4.202	-0.0060	-3.157
ARR_13	0.0019	0.686	-0.0080	-3.978	-0.0023	-1.195
ARR_14	-0.0005	-0.179	-0.0086	-4.283	-0.0037	-1.947
ARR_15	0.0009	0.333	-0.0094	-4.663	-0.0038	-2.002
ARR_16	0.0022	0.780	-0.0120	-6.095	-0.0058	-3.096
ARR_17	-0.0026	-0.938	-0.0158	-8.027	-0.0054	-2.813
ARR_18	0.0002	0.076	-0.0147	-7.502	-0.0064	-3.413
ARR_19	-0.0071	-2.528	-0.0151	-7.493	-0.0075	-3.896
ARR_20	-0.0007	-0.263	-0.0168	-8.511	-0.0049	-2.603
ARR_21	-0.0062	-2.187	-0.0187	-9.300	-0.0085	-4.387
ARR_22	0.0022	0.721	-0.0129	-5.969	-0.0018	-0.856
ARR_23	-0.0012	-0.368	-0.0096	-3.938	-0.0016	-0.695
Adjusted R^2	0.6921		0.7237		0.6846	
Sample Size	13,964		19,038		14,802	

Table 5-9: Model 6

5.4 Summary

In this chapter we performed an econometric analysis through linear regression models using buffer and on-time performance as the dependent variables. Our objective was to study how these two variables are related with each other and with a set of factors, such as the flight time components, the route competition, the hub effect, the carrier type, and the scheduled arrival time. The main findings of the analysis are summarized below.

Concerning the impact of the four flight time components on buffer, airborne and taxi-out times are estimated to be the most significant, with airlines adding on average 0.5 – 0.8 minutes of buffer for every ten minutes of airborne time, and 3.6 – 7.2 minutes for every ten minutes of taxi-out time. Gate delays and taxi-in times have a smaller impact in determining the schedule buffer, with airlines adding approximately 2 to 3 minutes for every ten minutes of taxi-in times or delays at gate. Nonetheless, the effect of each time component on buffer varies strongly across months and routes.

Additionally, flights that are destined to the carrier's hubs have more buffer than flights destined to other airports. This suggests that flights with passengers connecting to succeeding flights might be associated with a higher delay cost for airlines than flights with local passengers. However, the effect of the carrier's hubs on the buffer of departing flights is variable across months.

The data reveal that Southwest pads its flights more extensively than other carriers. This practice results to one of the highest on-time arrival ratings in the industry. On the other hand, the other low cost carriers pad their flights the least, and have a lower on-time performance. The group of legacy carriers pads its flights similarly to OTHER carries.

Furthermore, the regression models provide evidence that buffer increases with the number of carriers competing on the same route. On routes where more than three airlines compete, there is also a positive effect on on-time performance. These results suggest that competition affects carriers' padding practices more through on-time arrival statistics and passengers perception of carriers' reliability, rather than through screen presence in the reservation systems.

The estimates of the arrival time indicator variable coefficients show that flights scheduled to arrive during the morning and the evening peaks have more buffer compared to the

rest of the flights. While morning and evening flights are similarly padded in January and November, flights in August have more buffer during the evening peak compared to the morning. We attribute this difference to two reasons: (1) to the increase in the number of evening operations during August, and (2) to the frequent evening thunderstorms during the summer season. However, despite the fact that evening flights are padded the most, they are also less likely to arrive on time.

Last, out findings suggest that buffer is associated with an increase in on-time performance. Particularly, 10 additional minutes of buffer increase the percentage of flights that arrive on time by approximately 2%. Although longer flights suffer from a larger uncertainty, usually on-time performance does not decrease with distance. This is explained, first, by the fact that carriers add more buffer in longer flights, and second, because pilots often speed up the aircraft when a flight departs late. The benefits of the latter can be significant in long-haul flights, but are negligible in short flights.

Summarizing, this chapter's analysis confirmed most of the assumptions we formulated in chapter 4. Moreover, it helped us quantify the impact of the studied factors on padding and understand how these are associated with carriers' on-time performance.

Chapter 6

Conclusions

6.1 Summary of Findings

The main objective of this thesis was to study the variability in flight time components and shed light on the padding practices that U.S. carriers use to improve their schedule reliability. For this purpose, we extracted and analyzed individual flights on 2,359 U.S. domestic non-stop directional routes, served by 40 U.S. carriers. Our dataset accounted in total for 59% of the domestic commercial flights in 2009. In Chapter 1, we presented a set of key questions on which we wanted to concentrate our research. After performing and presenting the analysis in Chapters 4 and 5, we can now respond to these questions and summarize our findings:

- *What is the magnitude of variability in the different flight segments? How do carriers adjust their schedule with respect to them?*

Flight times experience a very large uncertainty, which is driven by periodic and stochastic factors. Periodic factors are the season of the year, the time of day and the day of week. Stochastic factors include the weather, flight path, runway configuration, gate assignments, aircraft and crew delays from previous flights, as well as airline operations such as boarding processing, fueling, baggage handling, catering, and aircraft maintenance issues. Each factor influences the flight time components differently, causing distinct levels of variation.

Gate delays are the most variable, followed by airborne and taxi-out times. Delays during the taxi-in process are very limited, with the exception of few airports such as JFK and LGA. The results of the regression models in Chapter 5 suggest that airlines schedule their block times based on their past performance and their expectations of the four flight time components. Airborne and taxi-out times are estimated to be the most significant, with airlines adding on average 0.5 – 0.8 minutes of buffer for every ten minutes of airborne time, and 3.6 – 7.2 minutes for every minute of taxi-out time. Gate delays and taxi-in times have a smaller impact in determining the schedule buffer, with airlines adding approximately 2 to 3 minutes for every ten minutes of taxi-in times or delays at gate. However, the effect of each time component on buffer varies strongly across months and different routes.

- *What is the relationship between buffer and stage length? Does stage length influence the likelihood of a flight to arrive on time?*

Stage length is one the most significant parameters affecting buffer. In general, carriers increase buffer with route distance. In other words, all else being equal, a long-haul flight will have more buffer than a short-haul flight. The reason is that the longer the flight distance, the larger the exposure of the aircraft to variable weather conditions, thus the more likely that the flight will suffer airborne delays. On the other hand, buffer as a fraction of nominal block time decreases exponentially with distance. This happens for two reasons: First, airborne delays do not increase proportionally to distance. Second, on longer flights, some delays can be absorbed by increasing the aircraft speed.

Moreover, the effect of distance to airborne delays, and consequently to buffer, is highly related to route direction and season. Jet streams have a larger impact on west-east and east-west

flights compared to south-north and north-south flights. In addition, this results in bigger seasonal differences in airborne times on west-east and east-west routes, meaning that carriers have to make larger adjustments to their scheduled block times from month-to-month to account for changes in jet streams.

Although the regression models in Chapter 5 do not provide enough evidence of correlation between distance and on-time performance, the analysis we performed in Section 4.3.8 based on the number of flights that an aircraft flies on the same day, suggests that on-time performance decreases with distance.

- *Do carriers adjust their scheduled block times over the course of the day?*

During a day, airline operations can be affected by delay propagation throughout airline networks and changes in traffic volume. To minimize the effect of these factors on their on-time performance, carriers often schedule distinct block times for flights on the same route that operate at different hours. This results in large buffer fluctuation over the course of the day. Particularly, flights scheduled to arrive during the morning and the evening peaks have more buffer compared to the rest of the flights. Despite the adjustments in block times, on-time performance deteriorates over the course of the day, with evening flights being the least likely to arrive on time. The reason is that the dispersion of delay distributions increases with time of day, hence it is more difficult to achieve high reliability.

- *What is the impact of a flight's relative position to other flights operated on the same day by the same aircraft on delay components, buffer and on-time performance?*

Gate and arrival delays increase almost linearly as the number of flights previously flown on the same day by the same aircraft grows. This happens because, first, delays propagate downstream to next flights, and second, because there is more congestion in the evening hours. Airborne delays, scheduled block times and buffer remain almost constant, whereas taxi-out times fluctuate strongly. Moreover, the first aircraft flight has on average a longer taxi-out time and block delay than subsequent flights, due to the morning peak in departing operations. Regarding on-time performance, it is highly correlated to gate delays and deteriorates as the flight's relative position in the aircraft's daily schedule becomes later in the day.

- *What are the differences in buffer and scheduled turn-around times between flights in or out from hubs, compared to other airports?*

Poor on-time performance in hub networks may be associated with a higher cost for airlines and passengers, as it often results in more missed connections. Our analysis shows that flights destined to the carrier's hubs have more buffer than flights destined to other airports. Moreover, flights during connecting banks at hubs have longer turn-around times, as opposed to flights departing from spoke cities. Regarding departures from the carrier's hubs, there is no evidence of consistent schedule padding differentiation from other flights. Last, on-time performance does not appear to be related to whether an airport is hub or not.

- *To what extent do Ground Delay Programs affect a carrier's decision regarding buffer?*

Ground Delay Programs are responsible for a large proportion of the variation in gate delays and taxi-out times on many domestic routes, with the rationale behind them being that it is more convenient and less expensive for flights to experience unavoidable delays on the ground rather than en route. Our analysis suggests that airlines do not adjust buffer to account specifically for them because it is impossible to predict the impact of GDPs on individual flights on specific days. However, when GDPs are initiated frequently on a route, they cause a significant increase in gate delays and taxi-out times. Therefore, they indirectly affect the amount of schedule buffer when carriers take into account the length and variability in taxi-out times and gate delays in determining scheduled block times.

- *Are there differences in schedule padding practices and on-time performance across carriers?*

Among the four carrier groups we studied, i.e. Southwest, Legacy carriers, LCCs and OTHERs, Southwest pads its schedule the most, achieving one of the highest on-time arrival ratings in the industry (at least in 2009, the focus year). On the other hand, other low cost carriers use on average the shortest buffer, a practice that results in lower on-time performance. Legacy carriers follow similar padding practices with OTHERs. Their on-time performance is close to the industry's average and fluctuates strongly during the year, driven mostly by changes in buffer, weather and traffic.

- *How does the number of competitors on a route affect carriers' padding and on-time performance?*

According to the developed regression models, buffer increases with the number of carriers competing on the same route. When more than three airlines compete on the same route, there is also a positive effect on on-time performance. These results support the findings of previous studies that on-time performance is positively correlated to competition. Furthermore, they suggest that competition affects carriers' padding practices more through on-time arrival statistics and passengers' perception of their reliability, rather than through screen presence in the reservation systems.

6.2   Future Research

Flight delays and on-time performance are very sensitive to weather conditions and vary significantly across years and seasons. Airlines adjust constantly their scheduled block times based on their past on-time performance and changes in their fleet, schedule and network. Hence the results of the regression models discussed in Chapter 5 should not be extended over time. Updating this research periodically is very essential for tracking changes in airlines schedule padding practices and understanding how airlines respond to past on-time arrival statistics.

One of the limitations of this research is the fact that we did not account for the use of different aircraft types on the same route. If a carrier consistently assigns distinct aircraft types on flights scheduled at different hours on one route, it would be more appropriate to compute separate nominal airborne times for each type. However, given the very large size of our dataset, the impact of aircraft speed on the estimation of buffer was considered insignificant and thus was not taken into account.

Furthermore, a more extensive econometric analysis may be performed by introducing weather information, such as VFR versus IFR conditions, and airport concentration levels, like the Herfindahl-Hirschman Index. We expect that both these factors affect significantly carriers' padding practices, and thus incorporating them into the regression models could explain much of the variability in buffer. Beyond that, it would also be of great interest to examine the implications of on-time performance to market share and revenues.

Last, every minute of buffer is associated with different costs and benefits for each carrier, even if there are commonalities among them, e.g. legacy versus low cost carriers. Therefore, for evaluating carriers' choices regarding padding, one should also examine their network structure, operating cost, passengers' type, and restrictions in fleet, crew and gates. Unfortunately, not all of these data are publicly available, and such an analysis could only be conducted at an individual carrier level.

References

[1] Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A., et al., *Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States*, Final Report, Washington, D.C.: Federal Aviation Administration, October 2010.

[2] Mueller, E., Chatterji, G., *Analysis of Aircraft Arrival and Departure Delay Characteristics*, AIAA's Aircraft Technology, Integration, and Operations Forum, Los Angeles, California, 2002.

[3] Tu, Y., Ball, M., Jank, W., *Estimating Flight Departure Delay Distributions - A Statistical Approach with Long-term Trend and Short-term Pattern*, Journal of the American Statistical Association, Vol. 103, No. 481, March 2008.

[4] Shumsky, R., *Dynamic Statistical Models for the Prediction of Aircraft Take-off Times*, Ph.D. Thesis, Dept of Electrical Engineering and Computer Science, MIT, September 1995.

[5] Pujet, N., *Modeling and Control of the Departure Process of Congested Airports*, Ph.D. Thesis, Dept of Aeronautics and Astronautics, MIT, September 1999.

[6] Andersson, K., Carr, F., Feron, E. and Hall, W., *Analysis and Modeling of Ground Operations at Hub Airports*, 3rd USA/Europe Air Traffic Management R&D seminar, Naples, Italy, June 2000.

[7] Idris, H., Clarke, J., Bhuva, R., and Kang L., *Queuing Model for Taxi-out Time Estimation*, Air Traffic Control Quarterly, Vol. 10, No. 1, September 2001.

[8] Idris, H., *Observation and Analysis of Departure Operations at Boston Logan International Airport*, Ph.D. Thesis, Dept of Mechanical Engineering, MIT, February 2001.

[9] Willemain, T., *Estimating Components of Variation in Flight Time*, The National Center of Excellence for Aviation Operations Research (NEXTOR), May 2001. Available at: http://www.nextor.org/pubs/WP-01-2.pdf

[10] Willemain, T.., Ma, H., Yakovchuk, N., Child, W., *Factors Influencing Estimated Time En Route,* Tech. rep., The National Center of Excellence for Aviation Operations Research (NEXTOR), 2003.

[11] Post, J., Bennett, M., Bonn, J., Knor, D., *Estimation of an En Route Weather Severity Index Using Lightning Strike and Flight Plan Data*, FAA Technical Publication, 2002.

[12] Welch, J., Ahmed, S., *Spectral Analysis of Airport Capacity and Delay*, 5th USA/Europe Air Traffic Management R&D Seminar, Budapest, Hungary, June 2003.

[13] Xu, N., Donohue, G., Laskey, K., Chen, C.H., Williams S. and Sherry L., *Bayesian Network Analysis of Flight Delays*, 86th Transportation Research Board Annual Meeting, Washington D.C, January 2007.

[14] El Alj, Y., *Estimating the True Extent of Air Traffic Delays*, Master Thesis, Dept of Civil and Environmental Engineering, MIT, June 2003.

[15] Vlachou, K., Tripodis, Y., Lovell, D., *A Comparison of Different Approaches to Estimate Unimpeded Flight Times*, 12th Air Transportation Research Society World Conference, Athens, Greece, July 2008.

[16] Forbes, S., *The effect of air traffic delays on airline prices*, International Journal of Industrial Organization, Vol. 26, Issue 5, September 2008.

[17] Shumsky, R., *Response of the U.S. Air Carriers to the On-Time Disclosure Rule*, Transportation Research Record, Issue 1379, 1993.

[18] Mayer, C., Sinai, T., *Why do airlines systematically schedule their flights to arrive late?*, Mimeo, Wharton School of Business, May 2003.

[19] Zhu, Y., *Evaluating Airline Delays: The Role of Airline Networks, Schedules and Passenger Demands*, Master Thesis, Dept of Civil and Environmental Engineering, MIT, February 2009.

[20] Foreman, S., Shea, D., *Publication of Information and Market Response: The Case of Airline on Time Performance Reports*, Review of Industrial Organization, Vol. 14, 1999.

[21] Suzuki, Y., *The Relationship between On-time Performance and Airline Market Share: A New Approach*, Transportation Research Part E: Logistics and Transportation Review, Vol. 36, 2000

[22] Borenstein, S., Netz, J., *Why do all the flights leave at 8 am?: Competition and Departure-time Differentiation in Airline Markets*, International Journal of Industrial Organization, Vol. 17, 1999.

[23] Mazzeo, M., *Competition and Service Quality in the U.S. Airline Industry, Review of Industrial Organization*, Vol. 22, 2003.

[24] Rupp, N., Owens, D., Plumly, W., *Does competition Influence Airline On-Time Performance?*, Elsevier, Advances in Airline Economics, Vol. 1, 2006.

[25] Mayer, C., Sinai, T., *Network Effects, Congestion Externalities, and Air Traffic Delays: or Why All delays Are Not Evil*, American Economic Review, Vol. 93, Issue 4, September 2003.

[26] Department of Transportation (DOT), *Valuation of Travel Time in Economic Analysis*, Revised Departmental Guidance, 2003.

[27] Source: MIT Airline Data Project, http://web.mit.edu/airlinedata/www/default.html

[28] Source: BTS Technical Directive: On-Time Reporting, Number 17,
http://www.bts.gov/programs/airline_information/accounting_and_reporting_directives/technical_directive_17.html

[29]Federal Aviation Administration, *Documentation For The Aviation System Performance Metrics (ASPM), Actual Versus Scheduled Metrics*, Washington DC, May 2002.

[30] Bratu, S., Barnhart, C., *An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data*, Journal of Transportation and Statistics, Vol. 1, 2005

[31] Source: FAA U.S. & Territories Airport Lookup,
https://airports-gis.faa.gov/airportsgis/airportLookup/index.jsp?category=nasr

[32] Larson, R., Odoni, A., *Urban Operations Research*, Prentice-Hall, NJ, 1982.

[33] De Neufville, R., Odoni, A., *Airports Systems: Planning, Design and Management*, McGraw-Hill, 2003.

[34] Morisset, T., *Comparing Capacities and Delays at major European and American Airports*, Master Thesis, Engineering Systems Division, MIT, May 2010.

[35] Copeland, D., Mason, R., McKenney, J., *Sabre: The development of information-based competence and execution of information-based competition*, IEEE Annals of the History of Computing, Vol. 17, No. 3, 1995.

[36] Gittell, J., *Supervisory Span, Relational Coordination and Flight Departure Performance: A Reassessment of Postbureaucracy Theory*, Organization Science, Vol. 12, Issue 4, 2001.