

Building Reliable Air-Travel Infrastructure Using Empirical Data and Stochastic Models of Airline Networks

Mazhar Arıkan

University of Kansas, Lawrence, KS 66045, mazhar@ku.edu

Vinayak Deshpande

University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, Vinayak_Deshpande@unc.edu

Milind Sohoni

Indian School of Business, Hyderabad, India, milind_sohoni@isb.edu

Flight delays have been a growing issue and they have reached an all-time high in recent years, with the airlines' on-time performance at its worst level in 2007 since 1995. The goal of this paper is to build stochastic models of airline networks and utilize publicly available data to answer the following policy questions: Which are the bottleneck airports in the US air-travel infrastructure? How would increasing airport capacity at these airports alleviate delay propagation? What are the appropriate metrics for measuring the robustness of airline schedules? How could these schedules be made more robust? Which flight in an aircraft rotation is a bottleneck flight? The contribution of this paper is two-fold. First, we develop stochastic models, using empirical data, to analyze the propagation of delays through air-transportation networks. Second, our analysis enables us to make policy recommendations regarding managing bottleneck resources in the air-travel infrastructure.

1. Introduction

Flight delays have been a growing issue and they have reached an all-time high in recent years. In 2007, the airlines' on-time performance was at its worst level since 1995 when the U.S. Department of Transportation (DOT) first started to collect detailed on-time performance data¹. According to the DOT, a flight is considered delayed if it arrives at the destination gate 15 minutes or more after its scheduled arrival time. The statistics show that approximately 24% of a total of 7,455,458 commercial flight operations in the US in 2007 were delayed (based on the DOT definition of a delayed flight).

¹ Airlines have reported on-time performance to DOT since 1987. Monthly reports are released in the Air Travel Consumer Report.

Flight delays have a significant impact on the US economy. A recent report by the Joint Economic Committee of the US Congress, chaired by Senator Charles Schumer (Schumer and Maloney 2008), has estimated that the total cost to the US economy due to flight delays was as much as \$41 billion in 2007. This includes an estimated \$19 billion in operating costs to the airlines, as well as \$12 billion in passenger delay costs. The report also estimated that flight delays resulted in consumption of 740 million additional gallons of jet fuel, costing an additional \$1.6 billion in fuel costs, and releasing an additional 7.1 million metric tons of climate-disrupting carbon dioxide in the atmosphere. In another study, Ball et al. (2010) have estimated the total cost of all US air transportation delays in 2007 at \$31.2 billion, including \$16.2 billion in passenger delay costs and a \$2.2 billion cost from loss of welfare incurred by passengers who avoided air travel as a result of delays. The goal of this paper is to propose solutions for improving the reliability of the air-travel infrastructure through Operations Research based models.

Typically, *block-time*, or *travel-time* of a flight, is defined as the elapsed time between pushback from departure gate to arrival at the destination gate. Flight delays are typically attributed to two factors: (i) the randomness in the *intrinsic* travel time for a scheduled flight (which does not include delays caused due to any other flight in the airline's network), and (ii) the propagation of this randomness through the air-travel network and infrastructure. We call the delay caused due to a spill-over from the previous flight of the aircraft as *total propagated delay* (see §3 for precise definitions). We model both of these factors that cause travel delays. The randomness in travel time is affected by independent and collaborative decisions made by multiple agencies within the air-transportation network such as airlines, airports, and regulatory agencies like the Federal Aviation Administration (FAA). For example, airport capacity, which is driven by several factors such as number of runways, weather disruptions, etc., has a significant impact on the total travel time. Airline scheduling and resource utilization decisions impact congestion at airports, which in turn affect the total travel time. Since most airline flight networks are tightly coupled, and resources such as airports are shared across networks, any delay caused within the system (either due to a single flight delay or a ground-stop at an airport) propagates across the entire transportation network and impacts the performance of the air-transportation infrastructure as a whole. While delay propagation models have been studied in the past literature, it is important to note that most of these models are *deterministic* in the sense that a given delay on a specific flight (called the *root flight*) is propagated through the network by assuming that there are no delays on subsequent flights in the network (Kondo 2008, AhmadBeygi et al. 2008, 2010). In other words, these models

replace the block-time random variable by the scheduled block-time for subsequent flights in the network. The key contribution of our modeling approach, however, lies in defining and constructing a stochastic model of delay propagation, which convolves a stochastic root delay with stochastic travel times of subsequent flights in an aircraft rotation to compute the propagated delay in the airline network.

The goal of our research is to answer the following policy questions by building stochastic models and utilizing publicly available data. Which are the bottleneck airports in the US air-travel infrastructure (i.e., airports that cause most delay propagation)? How would increasing airport capacity at these airports alleviate delay propagation? What are the appropriate metrics for measuring the robustness of airline schedules? How could these schedules be made more robust? Which flight in an aircraft rotation is a bottleneck flight (and, hence, deserves managerial attention)?

To answer these questions, we collected publicly available data from several data sources including the FAA, the Bureau of Transportation Statistics (BTS), and aircraft manufacturer web sites. Our primary data source is the on-time performance data set published by the BTS. We collected data on approximately 21 million domestic flights flown by major airlines over a three year period (2005-2007). We first build a stochastic model of intrinsic (i.e., non-propagated) travel time by analyzing factors that affect this travel-time such as route, airport, airline, congestion, etc. This intrinsic travel time distribution is then combined with a stochastic model of delay propagation to obtain the network impact of a delayed flight. We show that the deterministic approach, used in the prior literature, overestimates the root propagated delay and the total propagated delay in an airline network. These stochastic models are then utilized to define several metrics of airline network robustness.

Our analysis reveals that all airlines have negative block-time buffers (i.e., scheduled block-times are not enough to cover actual block-times) built in their schedule, potentially driven by the 15 minutes buffer used by the DOT in reporting on-time performance. We evaluate airlines based on the magnitude of delay propagation in their networks, and recommend that airlines should judiciously allocate schedule block-time and ground-time buffers by analyzing the stochastic propagation of delays. We also construct a metric to identify “bottleneck” flights in airline networks. Airlines should focus their managerial attention on such bottleneck flights to improve the robustness of their schedules. Our airport analysis reveals that the top five airports in the US with the most network impact are Atlanta, Chicago O’Hare, Dallas Fort Worth, Phoenix, and Los Angeles (LAX). Our stochastic model is then applied to study the impact of alleviating congestion at these airports.

Using our stochastic models, we build the marginal network delay impact curve, as a function of the reduction in congestion (increase in capacity at congested airports). Our analysis reveals that a 15% increase in capacity at these airports will result in an average reduction in network delays by approximately 10%. In addition, we show that the highest network delay reduction is obtained within the first 5% capacity increase.

Our paper is organized as follows. In §2, we briefly review the literature on airline flight delays and schedule robustness. §3 explains the data used in the study and the stochastic models developed. Different measures for robustness and service reliabilities based on our stochastic models are introduced in §4. In §5, we develop policy recommendations to improve the reliability of the air-transportation infrastructure. §6 concludes the paper with a brief summary of findings and policy implications.

2. Literature Review

This paper draws on three streams of research on airline flight delays and schedule robustness: *(i)* research identifying factors that affect block-times for better block-time estimation, *(ii)* modeling flight delay propagation, and *(iii)* and developing performance metrics that consider flight schedule robustness.

Factors that have an impact on block-times have been a subject of much discussion recently in both academia and popular press. In a Wall Street Journal article, McCartney (2007, August 13) explains airlines' decision of using smaller airplanes as one of the reasons of increased congestion and delays at airports. The same author reports that scheduled block-times are higher compared to 10 years ago although planes can fly faster and navigate better with more advanced technology (McCartney 2007, May 29), (McCartney 2010, February 4). This block-time inflation (i.e., schedule padding) is explained by increased variability in the air-travel infrastructure, airlines' efforts to improve their on-time rankings, and slower flying speeds of smaller regional jets preferred by airlines. Recently, Ball et al. (2010) show that in 2007, the total schedule buffer of 9,526 million minutes cost the airlines \$3.7 billion and passengers \$6 billion. They define the schedule buffer as the extra time added into the schedule by airlines to reduce the impact of potential delays. In this paper, we provide tools to airline schedulers to intelligently allocate block-times to flights and reduce the overall cost of schedule padding without disrupting the passenger service. Academic research regarding block-time forecasting mostly focuses on individual segments of block-times. Shumsky (1995) uses a linear regression model to predict taxi-out times, while Mueller and Chatterji (2002) compare different distributions for departure, arrival, and en route delays. Hebert and Dietz

(1997), Willemain et al. (2003), and Tu et al. (2008) are among others who study the estimation of different segments of block-times. In contrast, we construct a probabilistic model of total block-time instead of breaking it up into segments such as departure delay, taxi-times, air-time, etc., and propose several explanatory variables to predict the total block-time. Unlike most of the studies mentioned above, which focus on specific airports and shorter time periods, our model uses data on all the domestic flights by all major airlines flown in the US during 2005 through 2007 consisting of more than 21 million observations.

This study contributes to the robust airline scheduling literature (for a review, see Ball et al. (2007) and Barnhart and Cohn (2004)) by developing a stochastic model of delay propagation and using publicly available industry-wide data. Researchers have conducted several studies regarding modeling flight delay propagation. AhmadBeygi et al. (2008) and Lapp et al. (2008) analyze how a flight delay propagates through the network using a tree structure (i.e., propagation tree). Ahmad-Beygi et al. (2010) propose a method that reduces delay propagation by redistributing existing slack in the flight schedule. Lan et al. (2006) and Lan (2003) formulate a delay propagation problem as a mixed integer programming problem. The computational results based on their model reveal a considerable reduction in total propagated delays and in the number of disrupted passengers. The distinction between our study and these studies is in the modeling of block-times of downstream flight legs. We use a stochastic model of block-times of all flight legs, whereas these papers assume deterministic block-times given a random delay for the root flight. In this paper, we show that the deterministic approach used in the prior literature overestimates the impact of expected total propagated delay in the airline network, and, hence, our stochastic model is a better suited for capturing network robustness.

Studies that analyze flight schedule robustness metrics mainly develop new performance measures and compare them to the existing performance measures. Caulkins et al. (1993) argue that the DOT on-time performance metric is deficient in the sense that it does not take the differences across airports into account. In a recent paper, Forbes et al. (2012) find evidence that some airlines misreport some of their flights' arrival times so that the flight is counted as on-time in the DOT on-time ranking. Besides, this misreporting (or gaming) can be aggravated by employee bonus programs which reward employees based on the airline's position on the DOT ranking. In this paper, we develop new on-time performance metrics that will not provide airlines an undue advantage by utilizing the gaming behavior indicated in Forbes et al. (2012). Bratu and Barnhart (2005) propose two alternative flight-based delay metrics that can be more useful in reflecting passenger

disruptions: the percentage of flights that are delayed more than 45 minutes and the percentage of canceled flights. In another study, Wu (2006) generates a schedule reliability index for flights which reflects both inherent and propagated delays. Using a flight schedule information of an anonymous airline, it is shown that the proposed sequential optimization algorithm, which re-allocates ground times between rotation legs through a simulation model, improves the overall schedule reliability. Carey (1999) also develops new measures of reliability from analytical models which can be applicable to various forms of scheduled transportation. Our approach in this area of research is to distinguish the intrinsic delay attributable to the flight itself and the total propagated delay from previous flights, which is then used to define different measures of schedule robustness.

3. Data and Model

The data for our analysis comes from several sources such as the data published by the BTS, the FAA, and aircraft manufacturers. For our analysis we consider Airline On-time Performance data for all domestic flights within the US. The BTS is the primary source for these data. Additional data regarding aircraft types and other details were obtained from Aircraft Registry Database, Deregistered Aircraft file, and Aircraft Reference file for aircraft specific variables. These databases are maintained by the FAA. Besides, we used FAA Operations & Performance Data², the DOT's Domestic Airline Fares Consumer Report Table 6³, T-100 Domestic market data , T-100 Segment (All Carriers) data published by the BTS⁴.

Although we only analyzed domestic flights (and their schedules), we also considered incoming and outgoing international flights while calculating the congestion variables⁵. Inclusion of this data is critically important especially for airports with significant international flight activity. Our data set has approximately 21 million observations over a three year period from 2005-2007 (one observation for each commercial scheduled flight flown by a major carrier in the US).

For every flight in our data set, we compute the *scheduled block-time* (Q_i) as the difference between the scheduled arrival time of the flight and its scheduled departure time. The *actual block-time* (D_i^L) is computed as the difference between the actual arrival time of the flight and its scheduled departure time. The actual block-time is comprised of several components including

² This data contains detailed information regarding departure (arrival) capacities and actual number of departures (arrivals) of ASPM airports (77 US airports that have significant activity).

³ Contains information regarding the average fares for all city-pair markets that average at least 10 passengers each day.

⁴ T-100 data sets contain information regarding seats available and the number of enplaned passengers for each origin destination pair during any month.

⁵ The details of the calculation of congestion variables are available in online appendix E.

taxi-out time, en route time, and taxi-in time. Each of these components is subject to different causes of delay and the total block-time delay is the sum of all individual component delays. It is noteworthy that our definition of actual block-time differs from the traditional definition mentioned earlier in §1 because the latter ignores the impact of flight delays propagated through the system and departure delays associated with the concerned flight.

An airline network is coupled by the resources that are shared across the network. In particular, any given aircraft for an airline typically flies multiple flights over the course of a day, and hence a delay on one flight can potentially spill-over, or propagate, from one flight to the next. Each flight is operated by a particular aircraft, which is identified uniquely by its *tail number*, and an *aircraft rotation* is a sequence of flights operated by a particular tail number. In our model, an aircraft's rotation begins with the first revenue flight after a major maintenance, or a lay-over of more than five hours at an airport, and ends with the last flight operated before the aircraft returns for its next maintenance or remains on the ground for several hours. Throughout our analysis, we assume that airline crews remain with the aircraft. We acknowledge that unavailability of crew schedule information publicly remains a limitation of our analysis. We discuss this assumption in greater detail in §6.

As mentioned earlier, any flight delay is composed of the following two components, (i) that which is attributable (intrinsic) to the flight itself, i.e., *intrinsic delay*, and (ii) that which was caused by a “spill over” from a previous flight in the aircraft's rotation, i.e., *total propagated delay*. A conceptual representation of the coupling of flights on an aircraft rotation and spill-overs is presented in Figure 1. We refer to the actual block-time of a flight, ignoring any spill-overs, as the *intrinsic block-time* of the flight (D_i). The time duration between the next flight's scheduled departure time, on an aircraft rotation, and the earlier flight's scheduled arrival time is referred to as the scheduled *turn-around* or ground-time (G_i). In order to compute G_i , from the on-time performance data set, we first sorted the data by airline, tail number and scheduled departure time so that all aircraft rotations are grouped together. Then, for each flight i , we computed G_i by subtracting the scheduled arrival time of flight $i - 1$ from the scheduled departure time of flight i . A snapshot of one such aircraft rotation flown on September 24, 2007 by Delta Airline's aircraft with tail number N980DL is shown in the online Appendix A. We computed the minimum time to turn an aircraft (T_i) by analyzing ground times at different airports for different types of aircraft for each airline. First, we grouped the actual ground-times for each flight flown in 2005 and 2006 by airline, aircraft model, and departure airport. We then calculated the 5th percentile value

(in minutes) across all actual ground-times for each airline, aircraft model, and departure airport combination. The turn-around time for the corresponding flight i , in 2007, was assumed to be this 5th percentile value. Note that the time to turn an aircraft around shows considerable differences across airports and across aircraft models. In Tables 7 and 8 in the online appendix A, we show the average turn-around times at some airports and some aircraft models respectively. The buffer time available on ground for flight i , B_i , is calculated by subtracting T_i from G_i for all flights except the first flight on the rotation. Note that the B_i value of the first flight of any rotation is assumed to be zero. Thus, the spill-over, L_i , from flight $i - 1$ to flight i is given by

$$L_i = [D_{i-1}^L - (Q_{i-1} + B_i)]^+. \quad (1)$$

Moreover, for any flight i in the rotation, we have

$$D_i^L = \overbrace{D_i}^{\text{no spill-over}} + \overbrace{[D_{i-1}^L - (Q_{i-1} + B_i)]^+}^{\text{spill-over}=L_i}. \quad (2)$$

The first term in (2) captures a flight's own (intrinsic) random block-time and the second term captures any spill-over incurred due to a delay on a previous flight in the aircraft rotation.

We use several terms and definitions throughout this paper. For ready reference, we provide a detailed list of the terms used and their definitions in Table 5. Keeping these definitions in mind, it is easy to see that there are three cases to consider between any two consecutive flights in an aircraft rotation: (i) previous flight is on-time so there is no spill-over to current flight, (ii) previous flight is not on-time but there is no spill-over to current flight since the lateness of the earlier flight can be absorbed by B_i , and (iii) previous flight is not on-time and B_i is not long enough which creates the spill-over L_i for flight i . The graphical representation of these three cases is shown in Figure 1.

Finally, we define the arrival delay of any flight i (AD_i), including any spill-over from a previous flight in the aircraft rotation, as

$$\begin{aligned} AD_i &= [D_i^L - Q_i]^+ = [D_i + L_i - Q_i]^+, \\ &= \left[D_i - Q_i + [D_{i-1}^L - (Q_{i-1} + B_i)]^+ \right]^+, \\ &= \left[D_i - Q_i + \left[D_{i-1} - (Q_{i-1} + B_i) + \left[\dots + [D_1 - (Q_1 + B_2)]^+ \dots \right]^+ \right]^+ \right]^+. \end{aligned} \quad (3)$$

As can be seen in (3) the arrival delay of flight i is dependent on the intrinsic block-time of all previous flights on the rotation, $D_j, 1 \leq j < i$, through the spillover variables $L_j, 1 < j \leq i$. However, if $L_j = 0$ for some $j < i$, then it is equivalent to treating the rotation as beginning from flight j

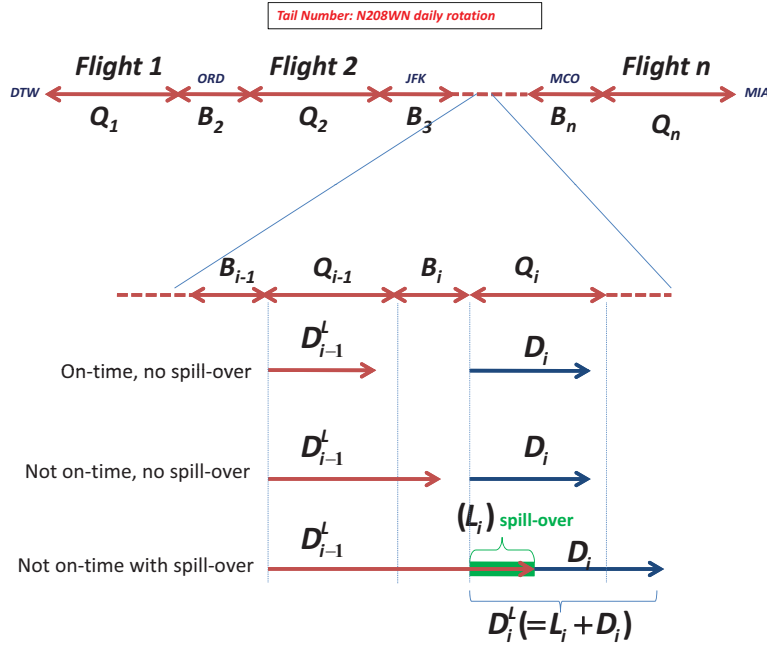


Figure 1 Illustration of spill-overs in an aircraft rotation

instead of the first flight. Using this insight, the arrival delay of flight i can be decomposed into the *arrival delay impact* ($ADI_{j,i}$) of all flights $j < i$ on flight i as follows:

$$ADI_{j,i} = AD_i - AD_i(L_{j+1} = 0). \quad (4)$$

Here, $AD_i(L_{j+1} = 0)$ represents the arrival delay of flight i conditional on a zero spillover L_{j+1} from flight j to $j + 1$. Thus, the *Net Impact* of any flight $j < i$ on the arrival delay of flight i ($NIAD_{j,i}$) can be found as follows:

$$NIAD_{j,i} = ADI_{j,i} - ADI_{j-1,i} \quad (5)$$

Thus, the *total propagated delay* (TPD_i) by any flight i , on an aircraft rotation, is defined as the *additional* delay caused on flight i due to the spill-over from the previous flight $i - 1$, and is computed as

$$TPD_i = [D_i + L_i - Q_i]^+ - [D_i - Q_i]^+ = AD_i - AD_i(L_i = 0) = ADI_{i-1,i}. \quad (6)$$

Therefore, the total arrival delay of a flight, AD_i , is the sum of the intrinsic delay, $[D_i - Q_i]^+$, and the total propagated delay, TPD_i .

As we indicated in §1, in the prior literature, the models are deterministic and focus on computing the *root propagated delay* which is the delay caused by the root flight. Specifically, AhmadBeygi

et al. (2008) develop a heuristic approach to compute the impact of the root propagated delay on all *future/subsequent flights* which share the same aircraft and crew resources. This set of future/subsequent flights, which share the same aircraft and crew resources, form a *propagation tree*. To compute the root propagated delay, their deterministic approach assumes that all subsequent flights in the propagation tree operate according to the planned schedule with no intrinsic delays, i.e., the actual block-times equals the scheduled block-times for these flights. The complexity in their model arises from the fact that they include crew schedule related delays, too. A similar deterministic approach is described in Kondo (2008) where the root propagated delay, due to a root flight, is computed using deterministic block-times for all future flights included in the aircraft rotation. However, Kondo (2008) does not consider crew schedules.

The key contribution of our modeling approach lies in constructing a stochastic model of *total* delay propagation, which convolves a stochastic root delay with stochastic block-times of subsequent flights on an aircraft rotation to compute the total propagated delay in the airline network. To this end, we first show that a deterministic analysis overestimates the root propagated delay and the total propagated delay. We first formally define root propagated delay, (RPD_i), as the net impact of the root flight (which we consider as first flight of the rotation) on the arrival delay of flight i using our notation in (5).

$$RPD_i = NIAD_{1,i}. \quad (7)$$

We state the impact of the assumption of all subsequent flights (j)—following the root flight—operating as per schedule in Theorem 1.

THEOREM 1. *A deterministic analysis, which assumes $D_j = Q_j, 1 < j \leq i$,*

- (i) *overestimates the expected RPD_i , and also*
- (ii) *overestimates the expected total propagated delay TPD_i .*

The proofs of all theorems are provided in the online appendix B.

Similarly, one can use the deterministic approach to estimate the arrival delay of flight i , which consists of the sum of the intrinsic arrival delay and the total propagated delay for flight i . The impact of this approach is stated in Theorem 2 in the online appendix B. Thus the key difference between our approach and past literature lies in (i) a focus on total propagated delay which captures the impact of all previous flights on an aircraft’s rotation, instead of root propagated delay which only captures the impact of a delay on the first flight; and (ii) a stochastic, instead of deterministic, model of delay propagation. Thus, we relax the assumption that the actual block-times on subsequent flights are deterministic, as assumed in prior literature. Theorems 1 and 2

demonstrate that our stochastic approach is well suited for estimating arrival delays and propagated delays within aircraft rotations, and hence developing other network robustness measures. We build on this idea in the rest of this paper.

To elucidate the different notations we introduced in this section and also Theorem 1, we provide an example rotation (with 3 flights) below in Table 1. In this example, the scheduled block-times of all flights are the same whereas the realized intrinsic block-times are different. A root delay of twenty minutes on the first flight is propagated into subsequent flights. However, the net impact of the root delay on flight 3 is only five minutes, since the intrinsic block-time of flight 2 and flight 3 is less than the scheduled block-time which creates an additional buffer. In contrast, the deterministic approach would have estimated the impact of root delay on flight 3 at 20 minutes. This example illustrates how the deterministic approach may overestimate the root propagated delay.

	i	Q_i	B_i	D_i	L_i	AD_i	$AD_3(L_i = 0)$	$ADI_{i,3}$	$NIAD_{i,3}$	RPD_3	RPD_3 (deterministic)
Flight 1	1	100	0	120	0	20	5	5	5		
Flight 2	2	100	0	90	20	10	0	5	0		
Flight 3	3	100	0	95	10	5	0	-	-	5	20

Table 1 Flight rotation example to illustrate Theorem 1

Before using the data for our analysis, we cleaned the data to eliminate several bad observations. Details of the data cleaning methodology, with examples of bad data, are available in online appendix A.

3.1. A stochastic model of actual block-time and delay propagation

In this sub-section we focus on describing two stochastic models that capture the coupling across an airline’s network: (i) a stochastic model of the intrinsic block-time, D_i , for a scheduled flight which captures factors such as airport capacity and schedule congestion, and (ii) a stochastic model of the actual block-time, D_i^L , which includes delay propagation through the airline network.

To do so, first, we build on the model developed in Deshpande and Arıkan (2012), where the authors show that the intrinsic block-time of a flight follows a log-Laplace distribution. Deshpande and Arıkan (2012) provide a methodology for estimating the mean, $\hat{\mu}$, and the standard deviation, $\hat{\sigma}$, of the log-Laplace distribution. Variables that impact the mean of the intrinsic block-time, D_i , distribution include flight route, airline, aircraft type, congestion at origin and destination airport, etc. The variance of the block-time distribution is composed of three components: an origin airport shock, a destination airport shock, and a carrier dependent idiosyncratic shock. We

describe essential details of the regression model developed in Deshpande and Arıkan (2012) in online appendix C.

Second, we develop a new delay propagation model that allows us to capture the ripple effect of flight delays through the network. To analyze the delay propagation in the network, we focus our attention on aircraft rotations. As shown in (2), the total block-time distribution of a flight is a convolution of the *intrinsic* block-time distribution of a flight and the *spill-over* distribution.

However, the distribution of the actual block-time, D_i^L is not easy to characterize, since it is the convolution of a log-Laplace distribution and the spill-over distribution. There does not exist a simple distribution that characterizes D_i^L . Hence, we make the simplifying assumption that D_i^L also follows a log-Laplace distribution with parameters $\tilde{\mu}_i$ and \tilde{b}_i^2 . The challenge, however, is to estimate the parameters of the resulting (assumed) log-Laplace distribution. Next, in Proposition 1⁶, we use a novel two-moment approximation to estimate the parameters $\tilde{\mu}_i$ and \tilde{b}_i^2 of the total block-time distribution.

PROPOSITION 1. *Matching the first two moments of the actual block-time distribution in (2) gives:*

$$\tilde{b}_i^2 = (-1 - 2A) + \sqrt{(4A + 1)(A + 1)} \quad (8)$$

where $A = \frac{\text{Var}[D_i^L]}{\mathbb{E}[D_i^L]^2}$ and

$$\tilde{\mu}_i = \ln \left(\mathbb{E}[D_i^L](1 - \tilde{b}_i^2) \right). \quad (9)$$

Our block-time model presented in online appendix C, captures the correlation in block-times of flight $i - 1$ and i through the common shock term of the common destination/origin airport of flight $i - 1/i$. However, an implicit assumption behind the approximation presented in Proposition 1 is that the block-time distributions of consecutive flights are not correlated. We validate this assumption by conducting a simulation study using our data. The details of this study are available in online appendix D where we show that our independence assumption is computationally reasonable for our large data set.

There are several advantages of using the above two-moment approximation. First, the above approach provides a tractable analytical expression for the actual block-time D_i^L . Second, the above approximation can be applied recursively to a string of flights on an aircraft rotation to compute the distribution of each flight on an aircraft rotation. Thus, it easily captures the cumulative effect of prior accumulated propagated delays on a given flight i by keeping track of a single random

⁶ Explicit representations for $\mathbb{E}[D_i^L]$ and $\text{Var}[D_i^L]$ are available in the online appendix B. $\tilde{\mu}_i$ and \tilde{b}_i^2 for any flight i can be computed using (8) and (9) recursively starting from the first flight on any rotation where $\tilde{\mu}_1 = \hat{\mu}_1$ and $\tilde{b}_1^2 = \hat{b}_1^2$.

variable D_{i-1}^L of the previous flight, instead of having to keep track of the intrinsic block-time distributions of all previous flights on the rotation. Third, the above computational procedure is extremely efficient. Evaluating the actual block-time distributions and the corresponding on-time arrival probability for each of the 7.4 million flights flown in 2007 took less than an hour on our desktop computer. This is very important, since suggesting policy changes requires estimating the impact of changes to the entire network consisting of millions of flights. This would be very difficult to do with other computational approaches such as simulation of the entire system. A detailed discussion on the accuracy of the above approximation is provided in §3.2.

These aforementioned stochastic models allow us to separate any flight delay into the two components mentioned earlier, i.e., the root delay and the propagated delay. To this end, we define the two respective probabilistic metrics OTPF_i^I and OTPF_i^L as follows:

$\text{OTPF}_i^I = \Pr(D_i < Q_i)$: On-time probability without including the late aircraft delay L_i
(i.e., OTP intrinsic to the flight),

$\text{OTPF}_i^L = \Pr(D_i^L < Q_i)$: On-time probability of flight i including the late aircraft delay L_i .

We use the stochastic models described in this section to develop network robustness measures and service levels in §4.

3.2. Accuracy of the Second Moment Approximation

In this sub-section we focus on testing the accuracy of the second moment approximation proposed in the earlier sub-section. Our analysis compares the two-moment model generated on-time arrival probability with the true underlying on-time arrival probability estimated using a simulation experiment. We use randomly selected 14 rotations flown in 2007 from Southwest, Delta, American, Continental, Frontier, JetBlue, and Mesa Airlines (two rotations from each major airline). We picked one short (3 or 4 flights) and one long (between 6 and 9 flights) rotations from each airline, so in total we have 76 flights in this simulation study.

For each flight i in these rotations, 60,000 intrinsic block-times are randomly generated from a log-Laplace distribution with parameters $\hat{\mu}_i$ and \hat{b}_i . Comparing the scheduled arrival time and the simulated actual arrival times, each simulated flight in the rotation is assigned an indicator variable of 0 or 1 depending on whether the flight arrived on time or not (i.e., 1 represents an on-time arrival and 0 represents a delayed flight). For each flight in these rotations, simulated on-time averages are calculated across 60,000 replications using the on-time indicators. Then, these simulated on-time

averages (i.e., simulated on-time probabilities) are compared to the estimated on-time probabilities which are calculated by using $\tilde{\mu}_i$ and \tilde{b}_i obtained from our two-moment approximation.

Tables 2 and 3 display the details of the rotations used and the results of the simulation study explained above. The maximum difference between the simulated on-time probabilities and the estimated on-time probabilities is equal to 0.1257 among these 76 flights. The median absolute difference between the simulated on-time probabilities and the estimated on-time probabilities is equal to 0.01765 while the mean absolute difference is equal to 0.0295. The simulated on-time probabilities turn out to be very close (i.e., within 4% range) to the estimated on-time probabilities for most of the flights selected.

The accuracy of the two moment approximation method has been well studied in the multi-echelon inventory literature (Graves 1985). While the airline context is different than multi-echelon inventory management, our model has parallels to multi-echelon inventory models. Given the computational efficiency of our two-moment approximation, as compared to simulation, for performing calculations for millions of flights in our data set, we proceed with this approximation for the rest of the paper.

4. Robustness and Service Reliability Measures for Airline Networks

DOT measures a flight to be on-time as long as the flight arrives at its destination no more than 15 minutes after its scheduled arrival time. Airlines are typically ranked based on the aggregate (average) on-time performance of all the flights in their network. This DOT metric is relatively easy to compute and provides a basis for measuring airline performance. However, with such a measure, as a basis for comparison, several important issues remain unanswered. We list two important issues below:

1. The DOT aggregate measure does not distinguish between the nature of airline operations, network structures, and markets served by individual airlines.
2. The DOT metric focuses on each flight separately and does not distinguish between the delays attributable to the flight itself and the additional delays due to spill-overs propagated (total propagated delays) through the network. Consequently, the aggregate-level DOT metric does not allow us to clearly explain the negative externality imposed by a poorly operated resource on the rest of the network.

In this section our goal is to develop measures that help us describe and quantify the aforementioned effects of (i) intrinsic delays and (ii) total propagated delays. Let $\overline{\text{OTPF}}_a^{\text{DOT}}$ represent the aggregate on-time performance computed, for any airline a 's network, using the DOT definition

Rotation1-Southwest Airlines				Rotation5-American Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.7023	0.7001	0.0022	1	0.5922	0.5909	0.0013
2	0.7086	0.7310	0.0223	2	0.2808	0.3043	0.0235
3	0.3670	0.4080	0.0410	3	0.2712	0.3316	0.0604
4	0.7952	0.7845	0.0107	4	0.5735	0.5819	0.0083
5	0.5400	0.5629	0.0229	5	0.2084	0.2542	0.0458
6	0.7003	0.7000	0.0003	6	0.2425	0.3421	0.0996
Rotation2-Southwest Airlines				Rotation6-American Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.8290	0.8266	0.0024	1	0.5227	0.5228	0.0001
2	0.5451	0.5488	0.0037	2	0.3565	0.3740	0.0175
3	0.2618	0.2986	0.0367	3	0.2157	0.2546	0.0390
4	0.4772	0.5799	0.1027	4	0.2830	0.2862	0.0032
Rotation3-Delta Airlines				Rotation7-Continental Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.4590	0.4597	0.0007	1	0.6397	0.6368	0.0029
2	0.4146	0.4389	0.0243	2	0.4874	0.5298	0.0425
3	0.4146	0.4485	0.0339	3	0.3097	0.3312	0.0215
4	0.2608	0.2734	0.0126	4	0.3636	0.4421	0.0785
5	0.1529	0.1602	0.0072	5	0.3874	0.4454	0.0580
6	0.1241	0.1409	0.0168	6	0.6220	0.6265	0.0045
Rotation4-Delta Airlines				Rotation7-Continental Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.6692	0.6665	0.0028	7	0.2392	0.2516	0.0124
2	0.6574	0.6617	0.0042	8	0.1097	0.1116	0.0018
3	0.3798	0.3888	0.0090				

Table 2 Results of the Simulation Study (First 7 Rotations)

of on-time performance, i.e., using the additional 15 minutes buffer beyond the scheduled arrival time of a flight. Recollect our definitions of $OTPF_i^L$ and $OTPF_i^I$ in §3 which are different from the metrics published by the DOT since we do not use the 15 minutes buffer beyond the scheduled arrival time. Let \overline{OTPF}_a^L represent the aggregate on-time performance computed for any airline a 's network excluding the 15 minutes buffer, and \overline{OTPF}_a^I represent the aggregate intrinsic (excluding the total propagated delays) on-time performance computed for any airline a 's network. Using the distributional parameters $\tilde{\mu}_i$ and \tilde{b}_i developed in (8) and (9), we rewrite the expressions for the corresponding service metrics (represented by these probabilities) as

$$OTPF_i^L = \mathcal{L} \left(\frac{\ln(Q_i) - \tilde{\mu}_i - \hat{\gamma}}{\tilde{b}_i} \right), \quad (10)$$

$$OTPF_i^{DOT} = \mathcal{L} \left(\frac{\ln(Q_i + 15) - \tilde{\mu}_i - \hat{\gamma}}{\tilde{b}_i} \right), \text{ and} \quad (11)$$

$$OTPF_i^I = \mathcal{L} \left(\frac{\ln(Q_i) - \hat{\mu}_i - \hat{\gamma}}{\hat{b}_i} \right) \quad (12)$$

Rotation8-Continental Airlines				Rotation12-JetBlue Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.5717	0.5709	0.0008	1	0.2700	0.2707	0.0007
2	0.4892	0.5121	0.0228	2	0.4011	0.4650	0.0639
3	0.1906	0.2048	0.0142	3	0.6546	0.6751	0.0205
4	0.0614	0.0751	0.0137	Rotation13-Mesa Airlines			
Rotation9-Frontier Airlines				Rotation14-Mesa Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.7309	0.7284	0.0025	1	0.7733	0.7706	0.0028
2	0.7510	0.7522	0.0012	2	0.5624	0.5645	0.0021
3	0.6603	0.6732	0.0130	3	0.4700	0.5103	0.0403
4	0.5019	0.5172	0.0153	4	0.5191	0.5552	0.0360
5	0.3788	0.4045	0.0257	5	0.5020	0.5421	0.0401
6	0.4530	0.5337	0.0808	6	0.5257	0.5833	0.0576
				7	0.3787	0.4372	0.0586
Rotation10-Frontier Airlines				Rotation11-JetBlue Airlines			
Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference	Pos.	On-time Prob. (Estimated)	On-time Prob. (Simulated)	Absolute Difference
1	0.6446	0.6415	0.0032	1	0.8031	0.8014	0.0017
2	0.5921	0.5965	0.0043	2	0.4275	0.4296	0.0022
3	0.5532	0.5787	0.0256	3	0.3371	0.3890	0.0518
4	0.5471	0.6212	0.0741	4	0.3616	0.4602	0.0986
				5	0.4117	0.4998	0.0882
				6	0.2396	0.2746	0.0350
				7	0.3330	0.4344	0.1014
				8	0.2951	0.3649	0.0698

Table 3 Results of the Simulation Study (Last 7 Rotations)

where \mathcal{L} represents the CDF of the Laplace distribution⁷ with parameters $(0,1)^7$ and $\hat{\gamma}$ represents the median of all residuals of the regression model described in the online appendix C. Let \mathcal{A} represent the set of airlines in our data set and \mathcal{F}_a denote the set of flights operated by a particular airline $a \in \mathcal{A}$. Let \mathcal{R}_a denote the set of aircraft rotations that belong to airline $a \in \mathcal{A}$.

Using (10)-(12), we define network measures of on-time performance for each airline a as follows:

$$\overline{\text{OTPF}}_a^L = \frac{\sum_{i \in \mathcal{F}_a} \text{OTPF}_i^L}{|\mathcal{F}_a|}, \quad (13)$$

$$\overline{\text{OTPF}}_a^{DOT} = \frac{\sum_{i \in \mathcal{F}_a} \text{OTPF}_i^{DOT}}{|\mathcal{F}_a|}, \text{ and} \quad (14)$$

⁷ $\mathcal{L}(x) = 0.5[1 + \text{sgn}(x)(1 - \exp(-|x|))]$

$$\overline{\text{OTPF}}_a^I = \frac{\sum_{i \in \mathcal{F}_a} \text{OTPF}_i^I}{|\mathcal{F}_a|}. \quad (15)$$

To begin, we demonstrate the inflation of on-time performance due to the use of the 15 minutes buffer in the DOT definition of on-time performance. In Figures 2 and 3, we illustrate how airline performances would vary if the metric $\overline{\text{OTPF}}_a^L$, i.e., without the use of the additional buffer of 15 minutes, were used instead of the metric using the 15 minutes buffer ($\overline{\text{OTPF}}_a^{DOT}$).

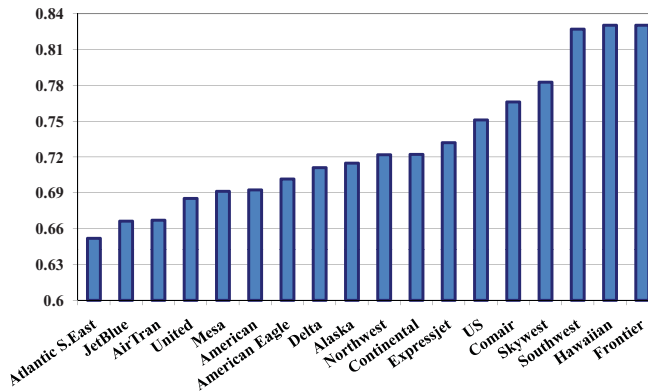


Figure 2 $\overline{\text{OTPF}}_a^{DOT}$ for different US airlines.

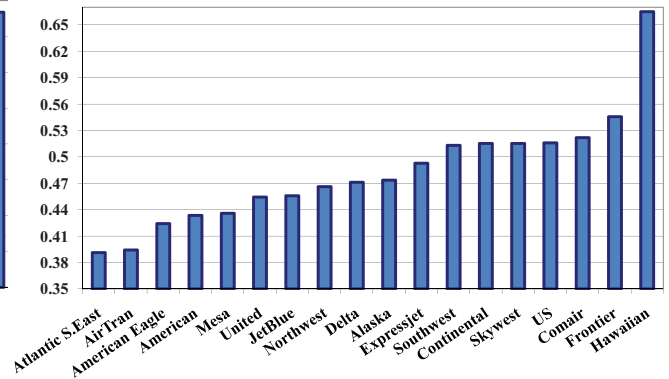


Figure 3 $\overline{\text{OTPF}}_a^L$ for different US airlines.

First, note that there exists a significant difference in the two OTPF metrics as shown in Figures 2 and 3. The average OTPF, without including a 15 minute buffer, across all airlines is 47.75%, while the average OTPF metric including a 15 minute buffer is 73.70%. The DOT OTP metric can be misleading to passengers who are not aware about the definition of OTP used by DOT. Its impact can be particularly severe on passengers who are planning itineraries with short connection times (e.g. 30 minutes). As an example, we found flights in our data set whose OTP performance drops from 80% to 50% by removing the 15 minute buffer, which could have a significant impact on passengers making short connections. A careful analysis needs to be conducted on what metric should be reported by the DOT to the public.

Also, as is observable above, some airlines are able to use the 15 minutes buffer, available to them under the DOT measure, quite effectively. For example, the OTP performance for Southwest Airlines improves under $\overline{\text{OTPF}}_a^{DOT}$. On the contrary, airlines such as JetBlue, improve their position, vis-à-vis other airlines, under $\overline{\text{OTPF}}_a^L$. Our observations suggest that providing an additional buffer does not address the issue of operational challenges faced by airlines, which are inherent to their network structures and markets served. The additional buffer may provide some airlines

a leeway in hiding some of their inefficiencies and improving their position in the DOT on-time performance rankings. Consequently, $\overline{\text{OTPF}}_a^L$ suffices as an on-time performance metric without providing airlines undue advantage.

Next, we define robustness measures for an airline flight network. Each of these measures quantify the delay externality imposed on a flight, or a resource using a set of flights, due to tight coupling of the flights within an airline’s network.

4.1. The average drop in on-time arrivals due to spill-overs

We begin by defining a robustness metric, $\overline{\text{OTPD}}_a$, that measures the average *network* impact (drop) on a flight’s on-time performance due to spill-overs (i.e., late arrival of aircraft from its previous flight) for any airline $a \in \mathcal{A}$.

$$\overline{\text{OTPD}}_a = \overline{\text{OTPF}}_a^I - \overline{\text{OTPF}}_a^L \quad \forall a \in \mathcal{A}. \quad (16)$$

We illustrate the impact of spill-overs (L_i ’s) in Figure 4. Earlier, in Figure 3 we plotted $\overline{\text{OTPF}}_a^L$ for each airline in our data set. In Figure 4, we plot $\overline{\text{OTPD}}_a$ to illustrate the impact of spill-overs. It

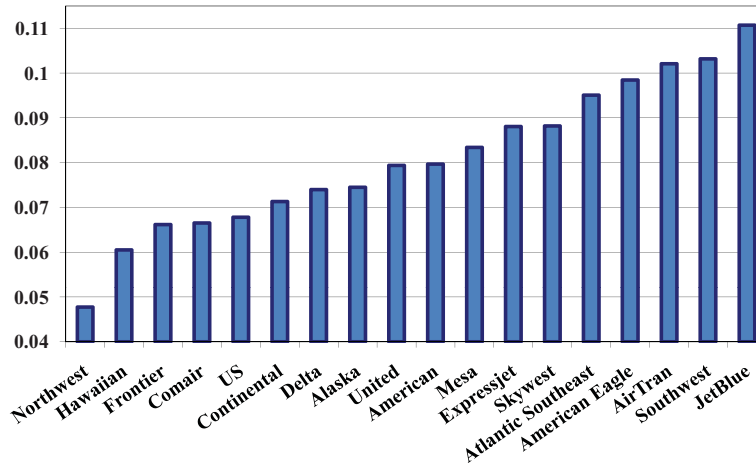


Figure 4 Average drop in on-time performance ($\overline{\text{OTPD}}_a$) due to spill-overs.

can be inferred from Figure 4 that the impact of spill-overs should not be overlooked. The average drop in on-time performance of almost all airlines is more than 5% and it is 8% or more for more than half of all airlines in our data set. Thus, from an airline’s perspective, the need of identifying and controlling total propagated delays, to improve operational performance is an important issue.

This fact is well known amongst practitioners, however, later in §5, we discuss how airlines can address this issue by allocating ground and block-time buffers differentially.

4.2. The fraction of expected delay attributable to spill-overs

Earlier in §4.1 we analyzed the impact of spill-overs. We next quantify the magnitude of the total propagated delays. For any airline $a \in \mathcal{A}$ and any flight $i \in \mathcal{F}_a$, this robustness metric, FEPD_i , measures the average fraction of total expected arrival delay, observed on flight i , which is attributable to the spill-over effect from previous flights in the aircraft rotation. Essentially, it allows us to separate the average magnitude of the expected delay externality imposed on any flight i .

$$\text{FEPD}_i = \frac{\mathbb{E}(D_i^L - Q_i)^+ - \mathbb{E}(D_i - Q_i)^+}{\mathbb{E}(D_i^L - Q_i)^+}, \text{ where } i \in \mathcal{F}_a. \tag{17}$$

Using (17) we define $\overline{\text{FEPD}}_a$ as the average fraction of expected total arrival delay attributable to propagated delay effects, for each airline $a \in \mathcal{A}$, and is computed as

$$\overline{\text{FEPD}}_a = \frac{\sum_{i \in \mathcal{F}_a} (\mathbb{E}(D_i^L - Q_i)^+ - \mathbb{E}(D_i - Q_i)^+)}{\sum_{i \in \mathcal{F}_a} \mathbb{E}(D_i^L - Q_i)^+} \quad \forall a \in \mathcal{A}. \tag{18}$$

We plot $\overline{\text{FEPD}}_a$ in Figure 5⁸ for all airlines in our data set. Figure 5 shows that the fraction of

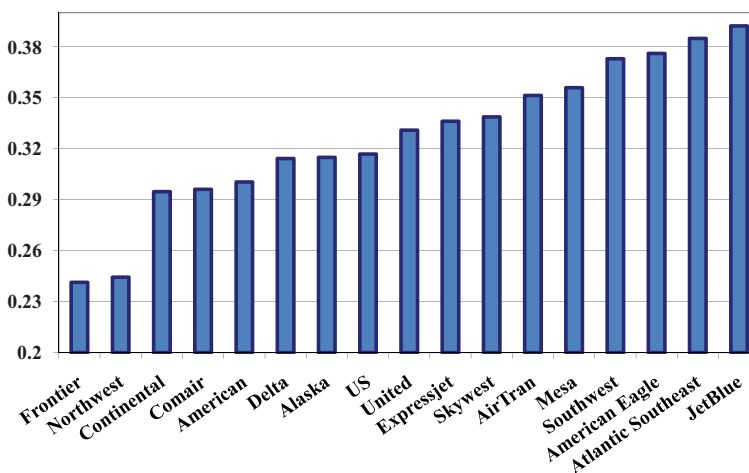


Figure 5 Fraction of expected delay due to propagated delays ($\overline{\text{FEPD}}_a$) for different US airlines.

average delay, attributable to spill-over effects, ranges from 24% to 39% for most airlines.

⁸ We excluded Hawaiian Airlines with $\overline{\text{FEPD}}_a$ value of 0.59 from this figure as it is an outlier on this chart.

4.3. Bottleneck flights

Recollect our definition of an aircraft rotation in §3. Let us denote an aircraft rotation by r . We first define the notion of a *bottleneck* flight within an aircraft rotation. Let $\text{OTPF}_{i,j}^L$ denote the on-time probability of flight j , within a particular aircraft rotation, assuming that the aircraft rotation begins with flight $i < j$ instead of the first flight in that rotation. Consequently, $\text{OTPF}_{1,j}^L = \text{OTPF}_j^L$ and $\text{OTPF}_{1,i}^L = \text{OTPF}_i^L$. Now suppose $N(r)$ denotes the set of flights in the aircraft rotation r (string of flights in rotation r), then the *network impact* of flight i (based on the drop in on-time arrival probability), NIP_i , is defined as the cumulative drop in on-time probability of subsequent flights in an aircraft rotation that is attributable to flight i , and is computed as follows:

$$\text{NIP}_i = \sum_{j=i+1}^{|N(r)|} (\text{OTPF}_{i+1,j}^L - \text{OTPF}_{i,j}^L). \quad (19)$$

We now define the bottleneck flight of any rotation in terms of the network impact based on expected arrival delay. Recollect our definition of the net impact of flight $i < j$ on the arrival delay of flight j that was represented by $\text{NIAD}_{i,j}$ in §3. Let $\mathbb{E}[\text{NIAD}_{i,j}]$ denote the expected value of this net impact on the arrival delay. Then, the network impact of flight i (based on the expected arrival delay), NID_i , is defined as the cumulative increase in the expected arrival delay of subsequent flights in an aircraft rotation that is attributable to flight i , and is computed as follows:

$$\text{NID}_i = \sum_{j=i+1}^{|N(r)|} \mathbb{E}[\text{NIAD}_{i,j}] \quad (20)$$

A *bottleneck* flight, within a rotation aircraft r , is then defined as the flight with the largest network impact. Let BNFL_r denote the position of the bottleneck flight within its rotation, i.e.,

$$\text{BNFL}_r = \arg \max_{1 \leq i < |N(r)|; i \in r} \{\text{NIP}_i\}. \quad (21)$$

Using our data set, we compute the frequency with which a flight at a particular position in any flight rotation, across all rotations for all airlines, is the bottleneck flight, as defined by (21). We denote this frequency value as BNFL_p , where the index p denotes the position of a flight in the aircraft rotation. Anecdotal evidence (Ross and Swain 2007) suggests that a natural question, for airline managers, to ask is: Do these bottleneck flights correspond to earlier flights in the day? We count the bottleneck flights, across all airlines, within a particular departure time slot t and denote the corresponding frequency of occurrence as BNFL_t . Figure 6 plots the frequency of occurrence, BNFL_p , for different flight positions, p , within any aircraft rotation. As is observable in this figure, the first 4 flights in any rotation have a higher probability of being the bottleneck flights. In

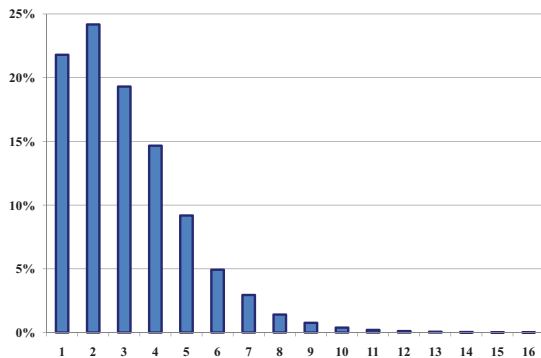


Figure 6 BNFL_p, computed across all airlines, for different flight positions.

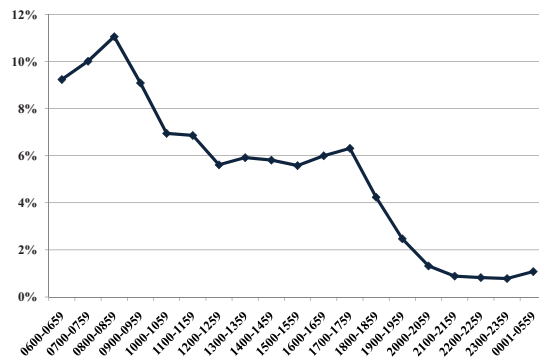


Figure 7 BNFL_t, computed across all airlines, for different departure times.

Figure 7, we plot BNFL_t for different departure times t . Figure 7 suggests that mid-day flights, 10:00 a.m. through 6:00 p.m., also have a reasonable probability (6 to 7%) of being bottleneck flights. Consequently, early morning flights are not significantly more likely to be bottleneck flights. Plausible reasons for mid-day flights becoming bottleneck flights are the increased schedule density and reduced airport capacity. Our observation suggests that airlines must tightly manage almost all their flights throughout a day to reduce their delay impact. This requires well controlled and better managed ground and air operations.

5. Analysis and policy implications

The goal in this section is to compare major airlines and airports in the US based on the robustness measures developed in §4 and to develop a few recommendations for improving the overall operational performance of the domestic air transportation system. We begin by analyzing airline on-time performance and related policy implications.

5.1. Airline on-time performance analysis

On-time performance is a combination of total propagated delays and intrinsic flight delays. Earlier, in §4 we argued that it is important to separate the effect of total propagated delays and showed how we could decompose their impacts. From an airline’s perspective it is important to control both of these effects to improve the overall on-time performance. To capture the impact of such propagated delays, in §4.1 and §4.2, we developed two robustness measures $\overline{\text{OTPD}}_a$ and $\overline{\text{FEPD}}_a$. These metrics allow us to measure airlines’ performances based on the average total propagated delays observed within their networks, i.e., how “tightly coupled” are these airline networks. By “tightly coupled” we mean how much slack is available, in the form of block- and ground-time

buffers within the airline network, to absorb propagated delays.

Thus, the first issue airlines face is the allocation of ground- and block-time buffer within aircraft rotations. Different airlines, depending on their network structures and the nature of service provided, allocate these buffers differentially. In Figures 8 and 9 we show the average ground-time

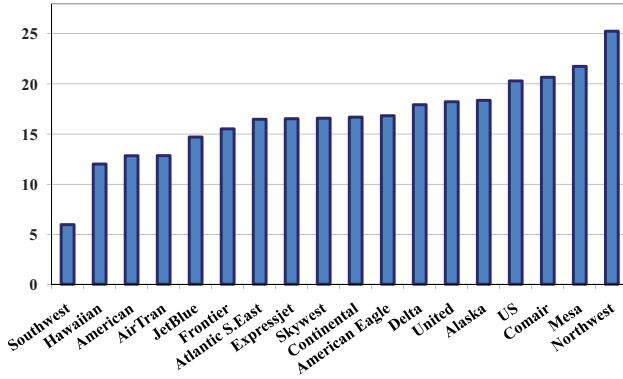


Figure 8 Average ground-time buffer allocation.

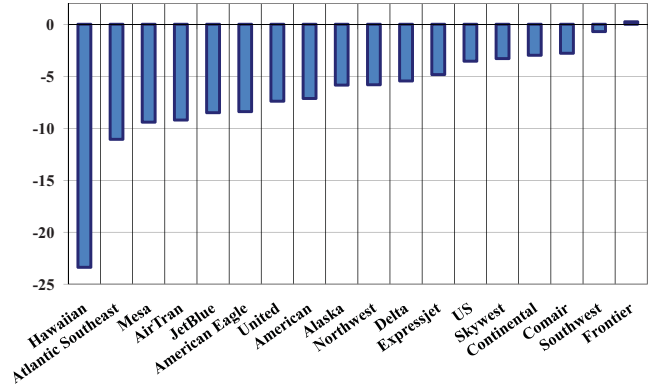


Figure 9 Average block-time buffer allocation.

buffer and block-time buffer for different airlines in our data set. The trade-offs for allocating block-time and ground-time buffers are complex. Increasing block-time buffer can increase planned costs substantially and reduce resource utilization, while decreasing block-time could directly impact the flight's on-time performance, i.e., $OTPF_i^L$. Increasing ground-time buffer is relatively less expensive but decreases resource utilization. Moreover, allocation of these buffers also depends on several other factors pertaining to the airline's flight schedule (network). For example, consider the case of Southwest Airlines. Southwest Airlines allocates low ground-time buffers yet maintains sufficient block-time buffers. Southwest Airlines is predominantly a point-to-point carrier, i.e., does not sell a large fraction of passenger itineraries with connections, and mostly operates to and from lesser congested airports. This allows Southwest Airlines to tightly control its ground-time buffers to maintain high resource (such as aircraft and ground equipment) utilization. Allocating a slightly higher buffer on block-times allows Southwest Airlines to maintain a high $OTPF_a^L$. At the other extreme, consider the example of Northwest Airlines. Northwest Airlines maintains a much larger ground-time buffer yet does not allocate large buffer on its block-time. Plausible reasons are the hub-and-spoke network, congested hub airports, and larger fraction of connecting passenger itineraries at hub airports. For example, hub-and-spoke networks create banks in their schedule to enable passenger connections, which naturally leads to higher ground-time buffers.

The second important issue concerning airlines is the identification of bottleneck flights. Earlier, in §4.3 we developed a metric to identify a bottleneck flight by isolating the impact of a particular flight in propagating delays through the network. As we showed in Figures 6 and 7, bottleneck flights are spread throughout the day, i.e., need not correspond to the earlier flights in the day. Ross and Swain (2007) indicate that airline planners focus on the “first” flights of the day to prevent propagated delays. While that is a good strategy, focusing on bottleneck mid-day flights is essential, too.

From a policy perspective, we make several recommendations based on our analysis. Our analysis shows that airlines carry very small block-time buffers for their flights, with all airlines averaging negative block-time buffers. This implies that, on the average, the scheduled block-time allocated for a flight is not sufficient to cover the actual block-time. We speculate that this scheduling behavior is partly in response to how on-time performance is measured and reported by the DOT, i.e., by adding a 15 minutes buffer to the scheduled arrival time. This analysis shows that the DOT should carefully evaluate which metric gets reported to the public and whether a 15 minute buffer is necessary in measuring OTP.

Moreover, the inclusion of such buffers can be misleading from a passenger’s standpoint. For example, consider a two-flight passenger itinerary wherein both the flights are of short duration and the connection time between flights is relatively small. In this case, it is possible that both individual flights may perform very well when evaluated independently using the DOT on-time performance metric, however, the passenger itinerary on-time performance may be extremely poor due to insufficient connection time. As a percentage of actual flight time, the total passenger delay can be large when both the flights in the itinerary are short haul flights. Consequently, it may be useful to consider a passenger-centric performance metric to compare airline performance in addition to single flight on-time performance measures. For example, comparing airline networks based on the on-time performance (on-time completion probability) of multi-leg passenger itineraries may be useful. Such a metric would also help distinguish carriers serving markets with a larger proportion of multi-leg itineraries versus carriers serving markets with single leg itineraries. In this regard, since our stochastic model generates the entire probabilistic distribution of actual block-time, it can be very useful in computing passenger connection probabilities and multi-leg on-time performance measures.

If providing accurate information to passengers is of importance, particularly to those with short connections, an immediate conclusion is to modify the measurement of on-time performance of

flights so that on-time truly means “on-time”, i.e., a flight should be considered delayed if it arrives past its scheduled arrival time. If this metric were to be adopted by the DOT, airlines could react by simply adding buffers to all flights (which can significantly increase their cost, as shown by Ball et al. (2010)). However, if airlines carefully balance their buffers, by focusing on bottleneck flights as discussed below, a change in the DOT metric can help achieve better OTP performance without increasing buffers and associated cost.

Our model also provides guidelines to airlines on which bottleneck flights they should focus in order to minimize the network impact. We constructed a network impact metric for each flight i using our stochastic model (i.e., NID_i). In order to make their schedules more robust, airlines could allocate larger resources (crew, block-times, etc.) to flights that have large network impact and less resources to non-bottleneck flights. From a policy standpoint, airlines could consider allocating ground- and block-time buffers differentially for bottleneck flights. Since airlines are typically budget constrained, in terms of the overall scheduled block-time minutes across the network, reducing the block- and ground-time buffers around non-bottleneck flights and reallocating the same to bottleneck flights could improve the overall network performance without incurring additional costs due to increased block-time buffers.

Finally, our stochastic model is well suited to isolate the impact of spill-overs on schedule robustness. Separating intrinsic block-time from total block-time helps identify the spill-over accurately for any aircraft rotation. This is important for airline practitioners since it is well known that spill-overs are highly dependent on the sequence of flights flown in an aircraft rotation.

5.2. Airport congestion analysis

Airports are a critical resource in air transportation networks and impact airline operations substantially. To understand their impact on the total propagated delays through the entire airline network, we compute the average values of overall on-time performance of flights, and intrinsic on-time performance of flights, within every departure time window t . We represent these by \overline{OTPF}_t^L and \overline{OTPF}_t^I respectively, computed across all airlines in our data set, where t represents a particular departure time window. These plots are shown in Figures 10 and 11 respectively. As is readily observable from Figure 10, the average on-time performance steadily decreases throughout the day until 6:00 p.m. and begins to improve thereafter. However, if we remove the impact of spillovers from prior flights (Figure 11), the performance of flights improves between 9:00 a.m. and 1:00 p.m. These observations suggest two important issues: (i) total propagated delays get worse as the day progresses, and, (ii) even after adjusting for spillovers, on-time performance worsens as the day

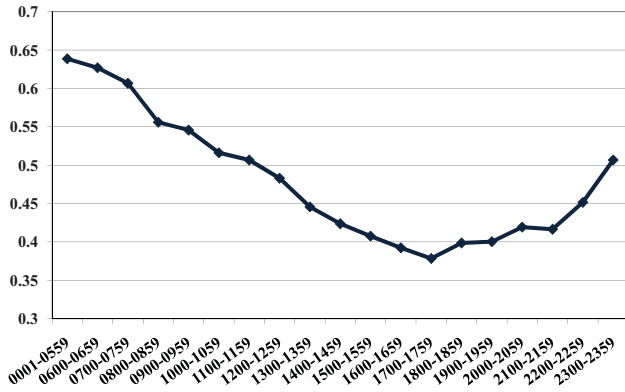


Figure 10 $\overline{\text{OTPF}}_t^L$.

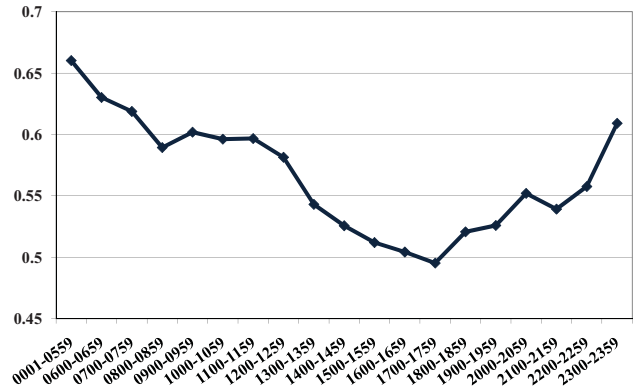


Figure 11 $\overline{\text{OTPF}}_t^I$.

progresses. Airport and airspace congestion together with the scheduling practices of most airlines contribute to this performance. In this section, we focus on the impact of airport congestion. To this end, we compare airports in our data set on two dimensions, (i) largest network impact (volume), and (ii) largest network impact per flight. Let \mathcal{S} denote the set of airports in our data and \mathcal{F}_s represent the set of all flights departed from or arrived to airport $s \in \mathcal{S}$ in 2007. Then, the *total network delay* for each airport s , TND_s , based on the total network impact of all flights $i \in \mathcal{F}_s$, NID_i (see, §4.3 for definition), is computed as follows:

$$\text{TND}_s = \sum_{i \in \mathcal{F}_s} \text{NID}_i. \tag{22}$$

We also normalized the above total network delay metric by considering the total number of flights affected to obtain a *per flight network delay* for each airport s (PFND_s), i.e.,

$$\text{PFND}_s = \frac{\text{TND}_s}{|\mathcal{F}_s|}. \tag{23}$$

We calculated the total network delay (TND_s) and per flight network delay (PFND_s) separately for each airport in our data set using all 2007 domestic U.S. flights and then sorted all airports from highest to lowest in each network delay scale. Figures 12 and 13 display the top 10 airports in each list.

As is observable from Figures 12 and 13, the top five airports in the US based on TND_s are ATL (Atlanta Hartsfield-Jackson), ORD (Chicago O’Hare), DFW (Dallas Fort Worth), PHX (Phoenix Sky Harbor), and LAX (Los Angeles), while the top five airports which have the largest PFND_s are JFK (New York JFK), HNL (Honolulu), EWR (Newark Liberty), SEA (Seattle Tacoma), and ORD (Chicago O’Hare). From a resource allocation perspective, for improving overall operational

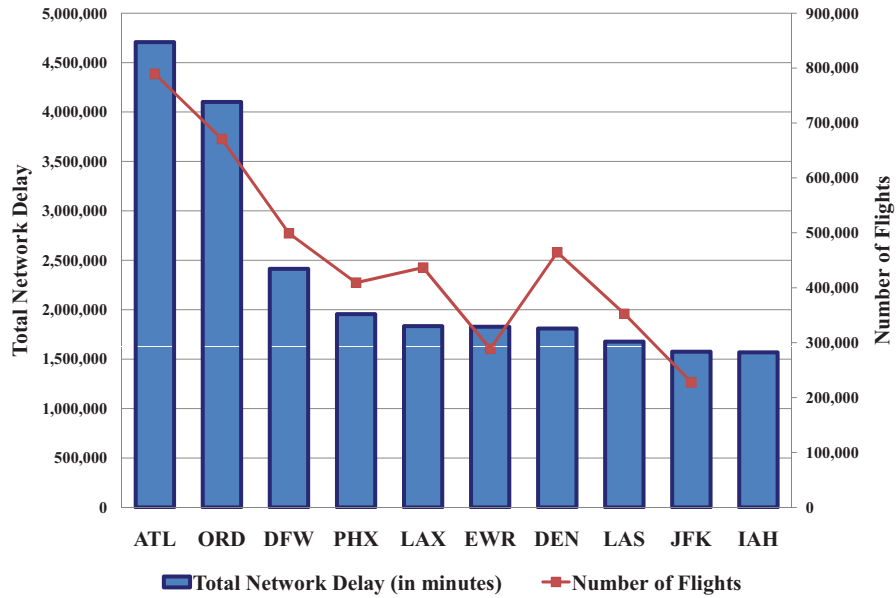


Figure 12 Top 10 U.S. airports with largest TND_s .

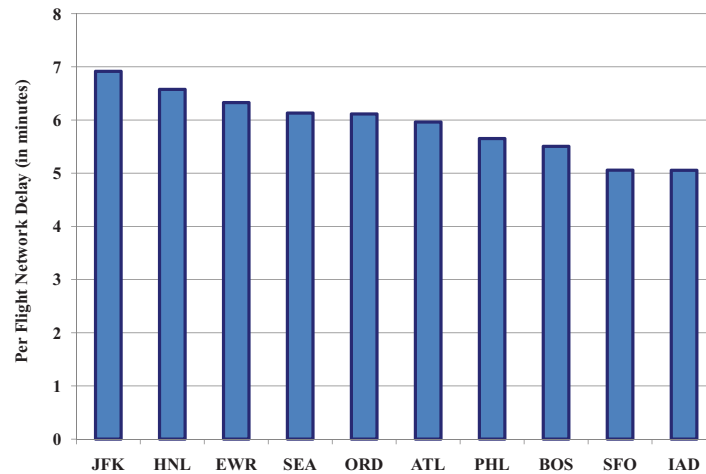


Figure 13 Top 10 U.S. airports with largest $PFND_s$.

performance of the flight networks of all airlines, alleviating the congestion at these airports is essential.

To analyze the impact of airport congestion on the airline networks, we picked top four airports namely ORD, ATL, DFW, and PHX based on TND_s . In addition, we added two more airports, EWR⁹ and JFK¹⁰, considering their position in the $PFND_s$ metric. We increased these airports'

⁹ EWR is the third airport behind JFK and HNL in terms of $PFND_s$ and sixth airport in terms of TND_s .

¹⁰ JFK is top airport based on $PFND_s$ but is not in the top 5 airports list based TND_s .

capacities by 15%, separately one-by-one, and then measured the network impact of increasing their capacities. First, by increasing the capacity at any airport, all flights arriving to and departing from that airport experience reduced delays due to lower congestion at that airport. In addition to this direct impact, our model also captures the indirect impact of increasing the airport capacity, i.e., reduced spill-overs on subsequent flights flown by the aircraft on its subsequent route. Specifically, to compute how increasing capacity affects delays (and network impact) we did the following: We first recomputed the congestion variables for each flight in the network affected by the capacity change. This necessitates re-computation of the actual block-time distributions of all the flights in the network (all airlines), i.e., we need to re-compute the total network impact of each flight i (NID_i) and then re-calculate the total network delay of each airport s (TND_s) based on these new estimates.

In Table 4, we provide the total network delay reduction of increasing the capacity by 15% at these six airports.

Airport	Current Total Network Delay, TND_s (in minutes)	Total Network Delay After 15% Capacity Increase (in minutes)	Total Network Delay Reduction (in minutes)	Network Delay Reduction (%)
ORD	4,101,722	3,703,083	398,639	9.72
ATL	4,707,561	4,207,576	499,985	10.62
DFW	2,414,278	2,226,295	187,984	7.79
PHX	1,956,013	1,800,993	155,019	7.93
EWR	1,827,784	1,653,668	174,116	9.53
JFK	1,574,898	1,411,594	163,305	10.37

Table 4 Change in TND_s of 6 airports with 15% capacity increase.

As can be seen, the total network delay is different at different airports. The total network delay reduction in flight delays by adding capacity at Atlanta Hartsfield-Jackson airport is the largest both in terms of total minutes and percentage reduction. Atlanta Hartsfield-Jackson airport is followed by Chicago, O'Hare airport.

We also conducted a sensitivity analysis for Chicago O'Hare airport (ORD) to examine the effects of different capacity changes on the total network delay. The percentage of capacity increase and corresponding total network delay is displayed in Figure 14. Although a careful observation of this figure reveals diminishing returns for each additional capacity increase, the decreasing trend is not far from linear. Specifically, by using linear interpolation, we observe that for each additional one percent increase in capacity (up to 5% capacity increase), the total drop in expected arrival

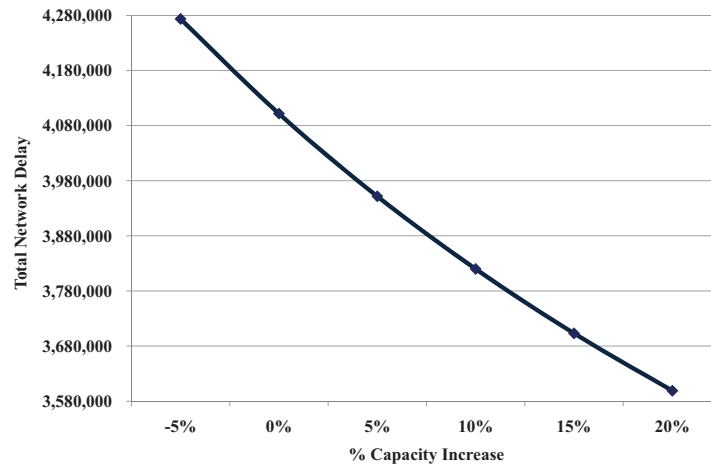


Figure 14 Total network delay of Chicago O'hare airport at different capacity levels.

delays is about 30,000 minutes; while each additional one percent capacity increment between 5% and 10% capacity increase levels leads to about 26,000 minutes total reduction in expected arrival delays. This kind of sensitivity analysis can be replicated for other airports using our model and a cost-benefit model can be established as a decision support tool for policy makers. While we do not include the cost of adding capacity at each airport, our model provides a new tool for evaluating investments in either capacity expansions at each airport, or in evaluating the benefits of new technologies (e.g., modernized landing and tracking technologies, En Route Automation Modernization program (ERAM), and etc.) assessed by the FAA under its Air Traffic Control (ATC) and Next Generation (NextGen) modernization program. Since, our model is at a micro (flight) level, the benefits of such capacity expansion proposals on network delays can be easily evaluated (ATC Modernization and NextGen: Near-Term Achievable Goals 2009).

6. Concluding Remarks

Flight delays have been a growing issue and they have reached an all-time high in recent years, with the airlines' on-time performance at its worst level in 2007 since 1995. A recent report by the Joint Economic Committee of the US Congress (Schumer and Maloney 2008) has estimated that the total cost to the US economy due to flight delays was as much as \$41 billion in 2007.

The goal of this paper was to build stochastic models and use publicly available data to examine the impact of airline network structures and schedules on the reliability of the air-travel infrastructure. Our analysis focused on answering the following questions: Which are the bottleneck airports in the US air-travel infrastructure (i.e., airports that cause most delay propagation)? How would increasing airport capacity at these airports alleviate delay propagation? What are the appropriate

metrics for measuring the robustness of airline schedules? How could these schedules be made more robust? Which flight in an aircraft rotation is a bottleneck flight (and, hence, deserves managerial attention)?

We make several policy recommendations based on our analysis. Our analysis suggests that the DOT on-time metric can significantly inflate true on-time performance and can be misleading, particularly to passengers with short connections. If providing accurate information to passengers is desirable, then the DOT OTP metric should be modified so that on-time really means “on-time”, i.e., the 15 minute buffer provided in the on-time metric should be eliminated. A careful cost benefit analysis of this proposal, as discussed in section 5, needs to be conducted. We also provide an analysis of alleviating the impact of congestion by augmenting airport capacities at different airports. This can be implemented through a variety of ways such as adding runways, using modernized landing techniques, better equipment, and etc. Our analysis shows that alleviating the impact of congestion at Atlanta Hartsfield-Jackson and Chicago O’Hare airports would provide the largest network benefits to the domestic air-travel system.

From the airlines perspective, our stochastic model is well suited to isolate the impact of spillovers on schedule robustness. Separating intrinsic block-time from total block-time helps identify the spillover accurately for any aircraft rotation. This is important for airline practitioners since it is well known that spillovers are highly dependent on the sequence of flights flown in an aircraft rotation. Our model also provides guidelines to airlines on which bottleneck flights they should focus to minimize the total network impact. In order to make their schedules more robust, airlines should allocate larger resources (e.g., crew, block-times, etc.) to flights that have a large network impact.

As indicated in Ball et al. (2010), airlines have been padding their schedules and this consequently increases the planned costs of airline operations. Airlines should intelligently allocate their resources (e.g., block-times) to flights. As a future avenue of research, our stochastic model can be used to allocate ground-time and block-times *optimally* across different flights without exceeding a budget constraint (i.e., total block-time that can be allocated across the entire network). This optimization model can provide potential benefits in terms of reducing total network arrival delays. Another potential future research idea is to compute the passenger on-time arrival probability which is driven by airline network structures, as well as connecting itineraries taken by passengers. Since a sample of this data is also publicly available, a new metric that computes passenger on-time arrival probability can be developed and a passenger-centric schedule robustness measure can be

computed.

Our modeling approach has a few limitations. We did not explicitly model delays due to crew and ground resources/gates/baggage. The lack of available data precludes us from incorporating such constraints into our models. This limitation may not be as severe as one thinks for three reasons. First, our block-time probabilistic distribution models are agnostic to the reason behind a flight disruption and hence the *aggregate* impact of crew related disruptions are captured in our model. Second, most domestic carriers have reserve crews readily available to cover for crew disruptions (see Sohoni et al. (2006)). Finally, many airlines, recognizing the potential for crew related disruptions, now often build crew rotations so that crew members stay with the same aircraft for most of their duty periods (see Mercier et al. (2005) and Ehrgott and Ryan (2002)). Our analysis, however, can be sharpened at the flight level if crew information was available.

Overall, the contribution of this paper is two-fold. First, we develop stochastic models, using empirical data, to analyze the propagation of delays through air-transportation networks. A key feature of our analysis is that we utilize data available on the entire US domestic network consisting of all airports and all major airlines, instead of focusing on a few airports or a few airlines. Another key feature of our model is that we conduct analysis at a micro (individual flight) level, instead of an aggregate level. Such a model is useful in evaluating the impact of changes made to one part of the system on the entire network. Also, our stochastic models allow us to develop three important robustness measures for airline networks. Second, our analysis enables us to make policy recommendations regarding managing bottleneck resources in the air-travel infrastructure, which if addressed, could lead to improvement in air-travel reliability.

References

- AhmadBeygi, S., A. Cohn, Y. Guan, P. Belobaba. 2008. Analysis of the potential for delay propagation in passenger airline networks. *Journal of Air Transport Management* **14** 221–236.
- AhmadBeygi, S., A. Cohn, M. Lapp. 2010. Decreasing airline delay propagation by re-allocating scheduled slack. *IIE Transactions* **42** 478–489.
- Arikan, M. 2011. The impact of airline flight schedules on flight delays: An analysis of block-time, delay propagation, and schedule optimization using stochastic models. Ph.D. thesis, Purdue University.
- ATC Modernization and NextGen: Near-Term Achievable Goals. 2009. 111th Congress Subcommittee on Aviation Hearing.

-
- Ball, M., C. Barnhart, M. Dresner, M. Hansen, K. Neels, A. Odoni, E. Peterson, L. Sherry, A. Trani, B. Zou. 2010. Total delay impact study. Tech. rep., NEXTOR.
- Ball, M., C. Barnhart, G. Nemhauser, A. Odoni. 2007. *Air transportation: Irregular operations and control*. C. Barnhart, G. Laporte, eds., *Handbooks in Operations Research and Management Science: Transportation*, vol. 14, chap. 1. North-Holland, 1–61.
- Barnhart, C., A. Cohn. 2004. Airline schedule planning: Accomplishments and opportunities. *Manufacturing & Service Operations Management* **6**(1) 3–22.
- Bratu, S., C. Barnhart. 2005. An analysis of passenger delays using flight operations and passenger booking data. *Air Traffic Control Quarterly* **13**(1) 1–27.
- Carey, M. 1999. Ex ante heuristic measures of schedule reliability. *Transportation Research Part B: Methodological* **33** 473–494.
- Caulkins, J.P., A. Barnett, P.D. Larkey, Y. Yuan, J. Goranson. 1993. The on-time machines: Some analyses of airline punctuality. *Operations Research* **41**(4) 710–720.
- Deshpande, V., M. Arikan. 2012. The impact of airline flight schedules on flight delays. *Manufacturing & Service Operations Management* forthcoming.
- Ehrgott, M., D.M. Ryan. 2002. Constructing robust crew schedules with bicriteria optimization. *Journal of Multi-Criteria Decision Analysis* **11**(3) 139–150.
- Forbes, S. J., M. Lederman, T. Tombe. 2012. Quality disclosure programs with thresholds: Misreporting, gaming, and employee incentives. *Working paper*. University of California, San Diego, CA.
- Graves, S. C. 1985. A multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science* **31**(10) 1247–1256.
- Hebert, J.E., D.C. Dietz. 1997. Modeling and analysis of an airport departure process. *Journal of Aircraft* **34**(1) 43–47.
- Kondo, Akira. 2008. Delay propagation: Tail-number tracking methodology. *Office of Aviation Policy and Plans, Federal Aviation Administration INFORMS Annual Meeting*, Washington D.C.
- Lan, S. 2003. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to achieve minimum passenger disruptions. Ph.D. thesis, Massachusetts Institute of Technology.

- Lan, S., J. Clarke, C. Barnhart. 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Science* **40**(1) 15–28.
- Lapp, M., S. AhmadBeygi, A. Cohn, O. Tsimhoni. 2008. A recursion-based approach to simulating airline schedule robustness. S. J. Mason, R. Hill, L. Moench, O. Rose, eds., *Proceedings of the 2008 Winter Simulation Conference*.
- McCartney, S. 2007, August 13. Frequent flying: Small jets, more trips worsen airport delays. *The Wall Street Journal*, pp. A1.
- McCartney, S. 2007, May 29. Why flights are getting longer. *The Wall Street Journal*, pp. D1, D5.
- McCartney, S. 2010, February 4. Why a six-hour flight now takes seven. *The Wall Street Journal*, pp. D1.
- Mercier, A., J.F. Cordeau, F. Soumis. 2005. A computational study of benders decomposition for the integrated aircraft routing and crew scheduling problem. *Computers & Operations Research* **32**(6) 1451–1476.
- Mueller, E.R., G.B. Chatterji. 2002. Analysis of aircraft arrival and departure delay characteristics. Tech. rep., AIAA Aircraft Technology, Integration, and Operations (ATIO).
- Ross, A., A. Swain. 2007. Fighting flight delays. *OR/MS Today* **34**(2).
- Schumer, C. E., C. B. Maloney. 2008. Your flight has been delayed again: Flight delays cost passengers, airlines, and the us economy billions. *The US Senate Joint Economic Committee* .
- Shumsky, R. A. 1995. Dynamic statistical models for the prediction of aircraft take-off times. Ph.D. thesis, Massachusetts Institute of Technology.
- Sohoni, M.G., E.L. Johnson, T.G. Bailey. 2006. Operational airline reserve crew planning. *Journal of Scheduling* **9**(3) 203–221.
- Tu, Y., M. Ball, W.S. Jank. 2008. Estimating flight departure delay distributions—a statistical approach with long-term trend and short-term pattern. *Journal of the American Statistical Association* **103**(481) 112–125.
- Willemain, T. R., H. Ma, N. V. Yakovchuk, W. A. Child. 2003. Factors influencing estimated time en route. Tech. rep., The National Center of Excellence for Aviation Operations Research (NEXTOR).
- Wu, C.L. 2006. Improving airline network robustness and operational reliability by sequential optimisation algorithms. *Networks and Spatial Economics* **6**(3) 235–251.

Appendix

Table 5 Description of Variables

Label	Definition	Description
Q_i	Scheduled block-time of flight i	Scheduled arrival time of flight i –
G_i	Scheduled ground/turn-around time of flight i	Scheduled departure time of flight i
T_i	Minimum time to turn the aircraft around from its previous flight for flight i	Scheduled departure time of flight i –
B_i	Buffer time available on ground for flight i	Scheduled arrival time of the previous flight $i - 1$
L_i	Spill-over caused by the previous flight on the aircraft's rotation	$G_i - T_i$
D_i^L	Actual block-time of flight i including the spill-over L_i	$[D_{i-1}^L - (Q_{i-1} + B_i)]^+$
D_i	Time to complete the flight excluding the spill-over L_i	Actual arrival time of flight i –
$OTPF_i^L$	On-time arrival probability of flight i including the spill-over L_i	Scheduled departure time of flight i
$OTPF_i^J$	On-time arrival probability of flight i excluding the spill-over	$D_i^L - L_i$
AD_i	Arrival delay of flight i	$\Pr(D_i^L < Q_i)$
$AD_{j,i}$	Arrival delay impact of all flights $j < i$ on flight i	$\Pr(D_i < Q_i)$
$NIAD_{j,i}$	Net impact of flight $j < i$ on the arrival delay of flight i	$[D_i^L - Q_i]^+$
TPD_i	Total propagated delay of flight i (the additional delay caused on flight i due to spill-over L_i)	$AD_i - AD_i(L_{j+1} = 0)$
RPD_i	Net impact of the first flight of the rotation on the arrival delay of flight i (root propagated delay)	$AD_{j,i} - AD_{j-1,i}$
$OTPF_a^{DOT}$	Aggregate on-time performance of airline a 's network using DOT definition of on-time arrival performance	$NIAD_{1,i}$
\overline{OTPF}_a^L	Aggregate on-time performance of airline a 's network excluding the 15 minutes buffer	See (14)
$OTPF_a^J$	Aggregate intrinsic on-time performance of airline a 's network	See (13)
$OTPD_a$	Average network impact (drop) on a flight's on-time performance due to spill-overs for airline a	See (15)
$FEPD_a$	Average fraction of expected total arrival delay attributable to propagated delays for airline a	$OTPF_a^J - \overline{OTPF}_a^L$
$OTPF_{i,j}^L$	On-time arrival probability of flight j assuming that the aircraft rotation begins with flight i instead of the first flight in that rotation	See (17)
$N(r)$	Set of flights in the aircraft rotation r	$\sum_{j=i+1}^{ N(r) } (OTPF_{i+1,j}^L - OTPF_{i,j}^L)$
NIP_i	Cumulative drop in the on-time arrival probability of subsequent flights in an aircraft rotation that is attributable to flight i	$\sum_{j=i+1}^{ N(r) } \mathbb{E}[NIAD_{i,j}]$
NID_i	Cumulative increase in the expected arrival delay of subsequent flights in an aircraft rotation that is attributable to flight i	$\arg \max_{1 \leq i < N(r) } \{NIP_i\}$
$BNFL_r$	Position of the bottleneck flight within its rotation r	$\sum_{i \in \mathcal{F}_s} NID_i$
$BNFL_p$	Cumulative frequency of the position of the bottleneck flight (p) within its rotation across all flights	TND_s / \mathcal{F}_s
$BNFL_i$	Cumulative frequency of occurrence of a bottleneck flight within a particular one-hour time slot t	
\mathcal{F}_s	Set of all flights departed from or arrived to airport s	
TND_s	Total network delay attributable to airport s	
$PFND_s$	Per flight network delay attributable to airport s	

Online Appendices for Manuscript

Building Reliable Air-Travel Infrastructure Using Empirical Data and Stochastic Models of Airline Networks

Appendix A: Tables and Cleaning Up the Data

ORIGIN	DEST	CRS DEP-TM	ACTUAL DEP-TM	CRS ARR-TM	ACTUAL ARR-TM
BHM	ATL	6:00 AM	5:57 AM	7:50 AM	7:54 AM
ATL	IND	8:45 AM	8:40 AM	10:18 AM	10:14 AM
IND	ATL	10:58 AM	10:50 AM	12:33 PM	12:15 PM
ATL	DCA	1:20 PM	1:17 PM	2:59 PM	2:54 PM
DCA	ATL	4:00 PM	3:57 PM	5:53 PM	5:35 PM
ATL	IND	6:55 PM	6:52 PM	8:32 PM	8:22 PM
BLOCK-TM (Q_i)	ACTUAL BLOCK-TM (D_i^L)	TURN- AROUND-TM (T_i)	BUFFER-TM (B_i)	GROUND-TM (G_i)	POSITION
50 min	54 min	30 min	0 min	-	1
93 min	89 min	38 min	17 min	55 min	2
95 min	77 min	31 min	9 min	40 min	3
99 min	94 min	38 min	9 min	47 min	4
113 min	95 min	30 min	31 min	61 min	5
97 min	87 min	38 min	24 min	62 min	6

Table 6 A snapshot of aircraft rotation data: Delta Airline’s aircraft with tail number N980DL flown on September 24, 2007

Airport	Avg. Turnaround Time
New York JFK	42.36 min
Miami	38.17 min
Chicago O’Hare	36.44 min
Baltimore	31.73 min
Indianapolis	31.14 min
Chicago Midway	26.9 min
Houston William P. Hobby	24.02 min
Dallas Love Field	19.3 min

Table 7 Average turn-around times of some airports across all aircraft types

Model	Num. of Seats	Avg. Turnaround Time
Boeing 737-2	104-118	26.68 min
Boeing 757-2	200-234	41.05 min
Boeing 767-3	218-350	52.39 min
Airbus A-319	124-156	33.87 min
Airbus A-320	148-180	35.47 min
McDonnell Douglas DC-9	90-125	26.47 min

Table 8 Average turn-around times of some aircraft models across all airports

We cleaned up bad data and eliminated some flights with the following conditions: (i) scheduled departure time of an aircraft from an airport is earlier than its scheduled arrival of the previous flight to the same airport, (ii) same aircraft flying on the same route successively on the same day (duplicate records, e.g., Flight 1: IND-PHL, Flight 2: IND-PHL, and Flight 3: PHL-JFK), and (iii)

an aircraft arrives to an airport and the next flight of the same aircraft is from a different airport in less than 5 hours. By excluding these bad observations (probably due to data entry errors), the total number of valid flights between 2005 and 2007 dropped down to 20,684,721 from 21,735,733 (See Tables 9, 10, and 11 below for examples of bad data). In addition, we assumed that if an aircraft stays idle at an airport for more than 5 hours, the next flight of the same aircraft is the first flight of its next rotation.

ORIGIN	DEST	DATE	ACTUAL DEP-TM	ACTUAL ARR-TM	ACTUAL-DIFF	POSITION
ATL	LAS	1/30/2005	8:36	13:25	35:48	1
LAS	ATL	1/30/2005	14:20	21:00	0:55	2
ATL	FLL	1/30/2005	16:15	17:52	0:00	.
ATL	FLL	1/30/2005	22:33	0:15	1:33	3
FLL	ATL	1/31/2005	6:50	8:54	6:35	1

Table 9 A duplicate record

ORIGIN	DEST	DATE	ACTUAL DEP-TM	ACTUAL ARR-TM	ACTUAL-DIFF	POSITION
VPS	ATL	1/15/2005	6:10 AM	8:40 AM	.	1
ATL	CHA	1/15/2005	1:35 PM	2:10 PM	4:55	2
CHA	ATL	1/15/2005	3:00 PM	3:55 PM	0:50	3
PFN	ATL	1/15/2005	6:20 PM	8:21 PM	.	.
ATL	VPS	1/15/2005	10:00 PM	10:20 PM	6:05	1

Table 10 Wrong origin or destination

ORIGIN	DEST	DATE	ACTUAL DEP-TM	ACTUAL ARR-TM	ACTUAL-DIFF	POSITION
ORD	IAH	2/09/2007	13:00	15:21	0:55	1
IAH	ORD	2/09/2007	16:27	.	.	2

Table 11 Missing data

Appendix B: Proofs

Proof of Theorem 1.

PROOF. Recollect that deterministic models assume that all the downstream flights operate as planned, i.e., do not incur any intrinsic delays. However, our stochastic model allows for incorporating such delays. Consider a rotation with n flights. Our goal is compute the additional delay, caused on flight i due to the root propagated delay, i.e., $NIAD_{1,i}$.

We first consider a deterministic analysis wherein we assume that all future flights do not incur any intrinsic delays. Thus, the the deterministic estimate of root propagated delay can be written as,

$$\text{RPD}_i = \left(\text{AD}_1 - \sum_{k=2}^i B_k \right)^+, \quad (24)$$

where B_k is the buffer, i.e. scheduled ground-time, before flight k . However, with a stochastic model, the propagated delay is computed as

$$\text{RPD}_i = \left(\text{AD}_1 - \sum_{k=2}^i B_k + \sum_{k=2}^i (D_k - Q_k) \right)^+.$$

Now notice that if $D_k > Q_k$ for any consequent flight, then this is flight k 's intrinsic delay. The additional impact of flight k 's intrinsic delay on flight i 's arrival delay is attributed to flight k itself (i.e., $\text{NIAD}_{k,i}$) and is not attributed to the root propagated delay (i.e., $\text{NIAD}_{1,i}$). Consequently, we ignore such cases and rewrite the stochastic propagated delay as

$$\begin{aligned} \text{RPD}_i &= \left(\text{AD}_1 - \sum_{k=2}^i B_k + \sum_{k=2}^i (D_k - Q_k)^- \right)^+ \\ &= \left(\text{AD}_1 - \sum_{k=2}^i B_k - \sum_{k=2}^i (Q_k - D_k)^+ \right)^+. \end{aligned} \quad (25)$$

It is readily verifiable that in (25), $\sum_{k=2}^i (Q_k - D_k)^+ \geq 0$. Consequently, RPD_i computed using (25) is smaller than RPD_i computed using (24), for every realization of the random variable $D_k, 1 < k \leq i$. Hence, the deterministic model overestimates the impact of (root) propagated delay of the root flight 1. This proves part (i) of Theorem 1.

To prove part (ii), recollect that for a given $L_i = l_i$, the total propagated delay is given by $\text{TPD}_i = \text{AD}_i - [D_i - Q_i]^+$ where $\text{AD}_i = [D_i + l_i - Q_i]^+$. Then,

$$\mathbb{E}(\text{TPD}_i | L_i = l_i) = \mathbb{E}[D_i + l_i - Q_i]^+ - \mathbb{E}[D_i - Q_i]^+ \quad (26)$$

Note that $[D_i + l_i - Q_i]^+ - [D_i - Q_i]^+ \leq l_i$ for every realization of the random variable D_i . Also, a deterministic approximation to the expected total propagated delay substitutes $D_i = Q_i$ in equation

26. Thus the deterministic approximation estimates the total propagated delay as l_i . Since, the expected stochastic total propagated delay cannot exceed l_i , the deterministic model overestimates the expected total propagated delay.

□

It is also noteworthy that while computing the root propagated delay, the deterministic models typically assume that the scheduled block-time $Q_i = \mathbb{E}(D_i)$ in addition to the assumption that all subsequent flights operate as per schedule. One issue with doing so is that the model may underestimate the expected arrival delay of subsequent flights in an airline network. Note that the arrival delay of a flight is the sum of the intrinsic delay of a flight and the total propagated delay. While the deterministic approach overestimates the total propagated delay, as shown in Theorem 1, the deterministic approach underestimates intrinsic delays since $E(D_i - Q_i)^+ \geq 0$. We next show that the net result of the deterministic approach is to underestimate arrival delays. Moreover, the magnitude of the underestimation of the expected arrival delay is maximum for “small” spill-overs (which is the case for most spill-overs in our data set). We formalize these results in Theorem 2 below.

THEOREM 2. *Consider any flight i in an aircraft rotation. For any given $L_i = l_i$, let the magnitude of difference of the expected arrival delay of flight i under a stochastic model and an equivalent deterministic model be denoted by the difference $\Delta(l_i) = \mathbb{E}_{D_i}(AD_i | L_i = l_i) - \max\{0, \mathbb{E}(D_i) + l_i - Q_i\}$ where $\mathbb{E}_{D_i}(\cdot)$ denotes the expectation over the intrinsic block-time D_i . Then,*

- (i) *a deterministic model of delay propagation underestimates $\mathbb{E}(AD_i)$, i.e., $\Delta(l_i) \geq 0$, and*
- (ii) *$\Delta(l_i)$ is decreasing in l_i if $Q_i = \mathbb{E}(D_i)$.*

Proof of Theorem 2.

PROOF. Let $f(\cdot)$ and $F(\cdot)$ denote the pdf and cdf of D_i . To prove part (i) consider an aircraft rotation with two consecutive flights $i - 1$ and i . We prove the result for any value of Q_i . Using our notations defined earlier let $\mathbb{E}(AD_i)$ denote the expected arrival delay of flight i . From (3) we have $AD_i = [D_i + l_i - Q_i]^+ = \max\{0, (D_i + l_i - Q_i)\}$. Notice that AD_i is a piece-wise linear

convex function of the random variable D_i . In a deterministic model of delay propagation, D_i would be replaced by $\mathbb{E}(D_i)$. From Jensen's inequality we know that if $f(X)$ is a convex function of X , then $\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$. Thus, we must have $\mathbb{E}(AD_i|L_i = l_i) \geq \max\{0, (\mathbb{E}(D_i) + l_i - Q_i)\}$ implying that any deterministic model underestimates arrival delays of subsequent flights. This result holds for any flight i on an N flight rotation for a given realization $L_i = l_i$.

Now to prove part (ii) notice that

$$\mathbb{E}_{D_i}(AD_i|L_i = l_i) = \int_{Q_i - l_i}^{\infty} (x - Q_i + l_i) f(x) dx, \quad (27)$$

$$= \mathbb{E}(D_i) - Q_i + l_i - \int_0^{Q_i - l_i} (x - Q_i + l_i) f(x) dx. \quad (28)$$

Therefore, when $Q_i = \mathbb{E}(D_i)$

$$\begin{aligned} \Delta(l_i) &= \mathbb{E}(D_i) - Q_i + l_i - \int_0^{Q_i - l_i} (x - Q_i + l_i) f(x) dx - (\mathbb{E}(D_i) + l_i - Q_i), \\ &= - \int_0^{Q_i - l_i} (x - Q_i + l_i) f(x) dx. \end{aligned}$$

Evaluating the partial derivative with respect to l_i we get

$$\frac{\partial}{\partial l_i} \Delta(l_i) = -F(Q_i - l_i), \quad (29)$$

$$\leq 0. \quad (30)$$

Thus, $\Delta(l_i)$ is decreasing in l_i when $Q_i = \mathbb{E}(D_i)$. \square

Proof of Proposition 1.

PROOF. Deshpande and Arıkan (2012) showed that $\log(D_i) \sim \text{Laplace}(\hat{\mu}_i, 2\hat{b}_i^2)$, so

$D_i \sim \text{log-Laplace}\left(\frac{e^{\hat{\mu}_i}}{1 - \hat{b}_i^2}, e^{2\hat{\mu}_i} \left(\frac{1}{1 - 4\hat{b}_i^2} - \frac{1}{(1 - \hat{b}_i^2)^2}\right)\right)$. Given (2) and assuming two consecutive flights are independent, we have

$$\mathbb{E}[D_i^L] = \mathbb{E}[D_i] + \mathbb{E}[D_{i-1}^L - (Q_{i-1} + B_i)]^+, \quad (31)$$

$$\text{Var}[D_i^L] = \text{Var}[D_i] + \text{Var}[D_{i-1}^L - (Q_{i-1} + B_i)]^+. \quad (32)$$

We assume that D_{i-1}^L follows a log-Laplace distribution with parameters $\tilde{\mu}_{i-1}$ and \tilde{b}_{i-1} so that $D_i^L \sim \text{log-Laplace}\left(\frac{e^{\tilde{\mu}_i}}{1-\tilde{b}_i^2}, e^{2\tilde{\mu}_i}\left(\frac{1}{1-4\tilde{b}_i^2} - \frac{1}{(1-\tilde{b}_i^2)^2}\right)\right)$. Let $\tilde{f}_{i-1}(x)$ denote the PDF of D_{i-1}^L . Then,

$$\mathbb{E}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ = \int_{\tilde{Q}_{i-1}}^{\infty} (x - \tilde{Q}_{i-1}) \tilde{f}_{i-1}(x) dx$$

where

$$\tilde{f}_{i-1}(x) = \exp(-\tilde{\mu}_{i-1}) \frac{1}{2\tilde{b}_{i-1}} \begin{cases} x^{\frac{1-\tilde{b}_{i-1}}{\tilde{b}_{i-1}}} \exp\left(-\tilde{\mu}_{i-1} \frac{1-\tilde{b}_{i-1}}{\tilde{b}_{i-1}}\right), & 0 < x < \exp(\tilde{\mu}_{i-1}) \\ x^{-\frac{1+\tilde{b}_{i-1}}{\tilde{b}_{i-1}}} \exp\left(\tilde{\mu}_{i-1} \frac{1+\tilde{b}_{i-1}}{\tilde{b}_{i-1}}\right), & x \geq \exp(\tilde{\mu}_{i-1}). \end{cases}$$

Then,

$$\begin{aligned} \mathbb{E}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ &= \int_{\tilde{Q}_{i-1}}^{\infty} (x - \tilde{Q}_{i-1}) \tilde{f}_{i-1}(x) dx \\ &= \begin{cases} \frac{\exp(\tilde{\mu}_{i-1})}{1-\tilde{b}_{i-1}^2} - \tilde{Q}_{i-1} + \frac{\tilde{b}_{i-1} \exp\left(\frac{-\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \tilde{Q}_{i-1}^{\frac{1+\tilde{b}_{i-1}}{\tilde{b}_{i-1}}}}{2(1+\tilde{b}_{i-1})}, & \text{if } \exp(\tilde{\mu}_{i-1}) > \tilde{Q}_{i-1} \\ \frac{\tilde{b}_{i-1} \exp\left(\frac{\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \tilde{Q}_{i-1}^{\frac{\tilde{b}_{i-1}-1}{\tilde{b}_{i-1}}}}{2(1-\tilde{b}_{i-1})}, & \text{otherwise.} \end{cases} \end{aligned}$$

So, we can rewrite (31) as follows:

$$\mathbb{E}[D_i^L] = \frac{\exp(\hat{\mu}_i)}{1-\hat{b}_i^2} + \begin{cases} \frac{\exp(\tilde{\mu}_{i-1})}{1-\tilde{b}_{i-1}^2} - \tilde{Q}_{i-1} + \frac{\tilde{b}_{i-1} \exp\left(\frac{-\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \tilde{Q}_{i-1}^{\frac{1+\tilde{b}_{i-1}}{\tilde{b}_{i-1}}}}{2(1+\tilde{b}_{i-1})}, & \text{if } \exp(\tilde{\mu}_{i-1}) > \tilde{Q}_{i-1} \\ \frac{\tilde{b}_{i-1} \exp\left(\frac{\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \tilde{Q}_{i-1}^{\frac{\tilde{b}_{i-1}-1}{\tilde{b}_{i-1}}}}{2(1-\tilde{b}_{i-1})}, & \text{otherwise.} \end{cases} \quad (33)$$

Let $K = \mathbb{E}[D_{i-1}^L - \tilde{Q}_{i-1}]^+$, then

$$\begin{aligned} \text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ &= \mathbb{E}\left[\left(D_{i-1}^L - \tilde{Q}_{i-1} - K\right)^2\right]^+ \\ &= \mathbb{E}\left[\left(D_{i-1}^L - \tilde{Q}_{i-1}\right)^2 - 2K(D_{i-1}^L - \tilde{Q}_{i-1}) + K^2\right]^+ \\ &= \int_Q^{\infty} ((x - Q)^2 - 2K(x - Q) + K^2) \tilde{f}_{i-1}(x) dx \\ &= \int_Q^{\infty} (x - Q)^2 \tilde{f}_{i-1}(x) dx - 2K^2 + K^2 \Pr(D_{i-1}^L \geq Q). \end{aligned}$$

Let's find the expression for $\Pr(D_{i-1}^L \geq Q)$:

$$\Pr(D_{i-1}^L \geq \tilde{Q}_{i-1}) = \begin{cases} 1 - \frac{1}{2} \tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{-\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right), & \text{if } \exp(\tilde{\mu}_{i-1}) > \tilde{Q}_{i-1} \\ \frac{1}{2} \tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right), & \text{otherwise.} \end{cases}$$

Now, we need to calculate $\int_{\tilde{Q}_{i-1}}^{\infty} (x - \tilde{Q}_{i-1})^2 \tilde{f}_{i-1}(x) dx$. This expression can be written as follows:

$$\int_{\tilde{Q}_{i-1}}^{\infty} (x - \tilde{Q}_{i-1})^2 \tilde{f}_{i-1}(x) dx = \begin{cases} \frac{\exp(2\tilde{\mu}_{i-1})}{1-4\tilde{b}_{i-1}^2} - \frac{2\tilde{Q}_{i-1} \exp(\tilde{\mu}_{i-1})}{1-\tilde{b}_{i-1}^2} \\ + \tilde{Q}_{i-1}^2 - \frac{\tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{-\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \tilde{b}_{i-1}^2}{(1+2\tilde{b}_{i-1})(1+\tilde{b}_{i-1})}, & \text{if } \exp(\tilde{\mu}_{i-1}) > \tilde{Q}_{i-1} \\ \tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \frac{\tilde{b}_{i-1}^2}{(1-2\tilde{b}_{i-1})(1-\tilde{b}_{i-1})}, & \text{otherwise.} \end{cases}$$

Recall that $\text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ = \int_Q^{\infty} (x - Q)^2 \tilde{f}_{i-1}(x) dx - 2K^2 + K^2 \Pr(D_{i-1}^L \geq Q)$, as a result when $\exp(\tilde{\mu}_{i-1}) > \tilde{Q}_{i-1}$,

$$\begin{aligned} \text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ &= \frac{\exp(2\tilde{\mu}_{i-1})}{1-4\tilde{b}_{i-1}^2} - \frac{2\tilde{Q}_{i-1} \exp(\tilde{\mu}_{i-1})}{1-\tilde{b}_{i-1}^2} + (\tilde{Q}_{i-1}^2 - K^2) \\ &\quad - \tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{-\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \left(\frac{\tilde{Q}_{i-1}^2 \tilde{b}_{i-1}^2}{(1+2\tilde{b}_{i-1})(1+\tilde{b}_{i-1})} + \frac{K^2}{2}\right). \end{aligned}$$

And otherwise when $\exp(\tilde{\mu}_{i-1}) \leq \tilde{Q}_{i-1}$,

$$\text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ = \tilde{Q}_{i-1}^{-\frac{1}{\tilde{b}_{i-1}}} \exp\left(\frac{\tilde{\mu}_{i-1}}{\tilde{b}_{i-1}}\right) \left(\frac{\tilde{Q}_{i-1}^2 \tilde{b}_{i-1}^2}{(1-2\tilde{b}_{i-1})(1-\tilde{b}_{i-1})} + \frac{K^2}{2}\right) - 2K^2,$$

where $K = \mathbb{E}[D_{i-1}^L - \tilde{Q}_{i-1}]^+$. We know that $\text{Var}[D_i^L] = \text{Var}[D_i] + \text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+$. Then,

$$\text{Var}[D_i^L] = \exp(2\hat{\mu}_i) \left(\frac{1}{1-4\hat{b}_i^2} - \frac{1}{(1-\hat{b}_i^2)^2} \right) + \text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ \quad (34)$$

To develop analytic expressions for $\tilde{\mu}_i$ and \tilde{b}_i , we perform a second-order matching. To do so, we solve the following expressions for the mean and variance together:

$$\mathbb{E}[D_i^L] = \frac{\exp(\tilde{\mu}_i)}{1-\tilde{b}_i^2}, \quad (35)$$

$$\text{Var}[D_i^L] = \exp(2\tilde{\mu}_i) \left(\frac{1}{1-4\tilde{b}_i^2} - \frac{1}{(1-\tilde{b}_i^2)^2} \right). \quad (36)$$

Let $A = \frac{\text{Var}[D_i^L]}{\mathbb{E}[D_i^L]^2} = \left(\frac{(1-\tilde{b}_i^2)^2}{1-4\tilde{b}_i^2} - 1 \right)$ and $S = \tilde{b}_i^2$. Then,

$$\begin{aligned} A + 1 &= \frac{(1-S)^2}{1-4S} \\ (A+1)(1-4S) &= 1-2S+S^2 \\ 0 &= S^2 + S(2+4A) - A. \end{aligned} \tag{37}$$

By solving the quadratic equation (37) we obtain:

$$\begin{aligned} S_1 &= \frac{-(2+4A) + \sqrt{(2+4A)^2 + 4A}}{2} \\ S_2 &= \frac{-(2+4A) - \sqrt{(2+4A)^2 + 4A}}{2} \end{aligned}$$

Since \tilde{b}_i cannot be negative, $S = \tilde{b}_i^2 = -(1+2A) + \sqrt{(4A+1)(A+1)}$.

By definition, $A > 0$, so

$$\begin{aligned} (1+2A)^2 &= (1+4A+4A^2) + A > (1+4A+4A^2) \\ (4A+1)(A+1) &> (1+2A)^2 \\ \sqrt{(4A+1)(A+1)} &> 1+2A. \end{aligned}$$

Hence,

$$\tilde{b}_i^2 = -(1+2A) + \sqrt{(4A+1)(A+1)} > 0. \tag{38}$$

We know that $\mathbb{E}[D_i^L] = \frac{\exp(\tilde{\mu}_i)}{1-\tilde{b}_i^2}$, so $\tilde{\mu}_i$ can be found using

$$\tilde{\mu}_i = \ln \left(\mathbb{E}[D_i^L](1-\tilde{b}_i^2) \right). \tag{39}$$

□

The reader is referred to Arıkan (2011) for the details of the expressions in-between steps some of which were skipped above. The two-moment approximation works as follows:

1. For the first flight of the rotation, let $\tilde{b}_1 = \hat{b}_1$ and $\tilde{\mu}_1 = \hat{\mu}_1$.
2. Starting from the second flight of the rotation, calculate $\mathbb{E}[D_i^L]$ and $\text{Var}[D_i^L]$ using the recursive formulations (33) and (34). Then, calculate $\frac{\text{Var}[D_i^L]}{\mathbb{E}[D_i^L]^2}$ and find $\tilde{b}_i^2 = -(1+2A) + \sqrt{(4A+1)(A+1)}$ and $\tilde{\mu}_i = \ln \left(\mathbb{E}[D_i^L](1-\tilde{b}_i^2) \right)$.
3. Go back to step 1 once you start a new rotation.

Appendix C: A Stochastic Model of Intrinsic Block-time

Deshpande and Arıkan (2012) identify several variables that affect the intrinsic block-time of a flight. These include route flown by the flight, air-carrier flying the flight, origin and destination airports, congestion at the origin/destination airports at the time of flight departure/arrival, age of aircraft, the number of passengers on the flight, and month, day-of week, and arrival/departure time block of the flight. Note that all of these variables—except congestion, average passengers, and aircraft age variables—are readily available in airline on-time performance data. A unique measure of congestion is used to understand its impact on block-times. For each flight in the on-time performance data set, the arrival/departure congestion is calculated by counting the total number of flights scheduled to arrive/depart in an adjacent one-hour time block around the scheduled arrival/departure time at the destination/origin airport. This number is then normalized by dividing by that airport’s arrival/departure capacity which was obtained from the ASPM database. This congestion measure is unique in the sense that it measures congestion for an *individual* flight and is not an average congestion measure which has been typically used in prior literature. By merging airline on-time performance data with aircraft registry database, deregistered aircraft file, and aircraft reference file, Deshpande and Arıkan (2012) obtain aircraft specific variables (e.g., aircraft age, number of seats) associated with each flight. In order to calculate the number of average passengers on each flight, Deshpande and Arıkan (2012) multiply the total seats of an aircraft by the load factor they obtain from T-100 domestic market data (load factor is the proportion of available seats that were actually filled during a specified period).

Deshpande and Arıkan (2012) show that the log-Laplace distribution provides a good-fit to model the intrinsic block-time of a flight. They also provide a method to estimate the parameters μ_i and b_i^2 of the log-Laplace distribution for the intrinsic block-time of a flight i . $\hat{\mu}_i$ represents the best estimate of μ_i computed by using the above-mentioned data sets.

The first estimator, $\hat{\mu}_i$, can be found by regressing $\ln(D_i)$ on X_i , where X_i is the vector of variables that impact the intrinsic block-time distribution. The estimation model is given by

$$\ln(D_i) = \beta_0 + \beta_1 * \text{Route}_i + \beta_2 * \text{Origin}_i + \beta_3 * \text{Destination}_i + \beta_4 * \text{Carrier}_i + \beta_5 * \text{Month}_i +$$

$$\begin{aligned}
& \beta_6 * \text{Day-of-Week}_i + \beta_7 * \text{Departure-Time-Block}_i + \beta_8 * \text{Arrival-Time-Block}_i + \\
& + \beta_9 * \text{Arrival-Congestion}_i + \beta_{10} * \text{Departure-Congestion}_i + \\
& \beta_{11} * \text{Aircraft-Age}_i + \beta_{12} * \text{Average-Passengers}_i + \varepsilon_i
\end{aligned} \tag{40}$$

The above regression model is run on 2005-2006 data consisting of 13,227,718 observations to obtain an estimate of β , $\hat{\beta}$. Note that the estimates obtained for all variables are significant at the %99.99 level. These estimates are then applied to 2007 data consisting of 6,736,521 flight observations to obtain $\hat{\mu}_i$ as follows.

$$\hat{\mu}_i = X_i^T \hat{\beta} \tag{41}$$

To estimate the parameter b_i^2 , it is assumed that the variance of the block-time across all flights is heteroscedastic and the error terms can be split into three components¹¹:

Total Error = Origin Airport Shock + Destination Airport Shock + Carrier Shock.

For any flight i , in any month m in the given data, denoting the origin airport shock as $\hat{\sigma}_{\text{org},m}^2$, the destination airport shock as $\hat{\sigma}_{\text{dest},m}^2$, and the carrier specific shock as $\hat{\sigma}_{i,c}^2$, the estimate of the variance of the log intrinsic-block-time, $2\hat{b}_i^2$, is

$$2\hat{b}_i^2 = \hat{\sigma}_{\text{org},m}^2 + \hat{\sigma}_{\text{dest},m}^2 + \hat{\sigma}_{i,c}^2. \tag{42}$$

See Deshpande and Arıkan (2012) for the methodology to obtain the estimates $\hat{\sigma}_{\text{org},m}^2$, $\hat{\sigma}_{\text{dest},m}^2$, and $\hat{\sigma}_{i,c}^2$.

Appendix D: Covariance Between Intrinsic Block-time and Spill-over

In the approximation presented in Proposition 1, we assume that the covariance between D_i and $(D_{i-1}^L - (Q_{i-1} + B_i))^+$ is equal to zero. However, since the estimate of the variance of the log intrinsic block-time of each flight i is equal to $\hat{\sigma}_{i,\text{org},m}^2 + \hat{\sigma}_{i,\text{dest},m}^2 + \hat{\sigma}_{i,c}^2$, for any two arbitrary consecutive

¹¹ Deshpande and Arıkan (2012) also tried different variance models including a *route* and a *month* shock term but *origin*, *destination* and *carrier* shock model showed the best fit.

flight in the same rotation, we have $\widehat{\sigma}_{i-1,\text{dest},m}^2 = \widehat{\sigma}_{i,\text{org},m}^2$; and therefore, D_i^L and D_{i-1}^L are correlated.

This means that, based on (2), variance of actual block-time can be written as follows:

$$\text{Var}[D_i^L] = \text{Var}[D_i] + \text{Var}[D_{i-1}^L - \tilde{Q}_{i-1}]^+ + 2\text{Cov}[D_i, (D_{i-1}^L - \tilde{Q}_{i-1})^+], \quad (43)$$

where $\tilde{Q}_{i-1} = Q_{i-1} + B_i$. Note that if $\ln(D_i^L) = \mu_i + \xi_{i,\text{org},m} + \xi_{i,\text{dest},m} + \xi_{i,c}$ where the random error components $\xi_{i,\text{org},m}$, $\xi_{i,\text{dest},m}$, and $\xi_{i,c}$ are normally distributed with means zero and variances $\widehat{\sigma}_{i,\text{org},m}^2$, $\widehat{\sigma}_{i,\text{dest},m}^2$, and $\widehat{\sigma}_{i,c}^2$ respectively, then $\ln(D_i^L)$ has a normal distribution¹² (i.e., D_i^L has a log-normal distribution). We compute $\mathbb{E}[D_i]$, $\text{Var}[D_i]$, $\mathbb{E}[(D_{i-1}^L - \tilde{Q}_{i-1})^+]$, $\text{Var}[(D_{i-1}^L - \tilde{Q}_{i-1})^+]$, and $\mathbb{E}[D_i(D_{i-1}^L - \tilde{Q}_{i-1})^+]$ with the assumption that the variance components $(\xi_{i,\text{org},m}, \xi_{i,\text{dest},m}, \xi_{i,c})$ are normally distributed (these components are analytically not tractable for the Laplace distribution).

It is plausible that we could construct empirical distributions from the historical data and compute the covariance measures based on these empirical distributions. However, (i) to do so is computationally prohibitive (extremely time consuming in the least), and (ii) such an approach need not result in better results because the empirical distributions may not be representative of the flights performance in the future (just as our continuous approximation, this too is an approximate distribution). Let $f_i(x)$ and $g_{i-1}(x)$ denote the probability distribution functions of D_i and D_{i-1}^L respectively, hence

$$\begin{aligned} \mathbb{E}[D_i] &= \int_0^\infty x f_i(x) dx \\ \text{Var}[D_i] &= \int_0^\infty (x - \mathbb{E}[D_i])^2 f_i(x) dx \\ \mathbb{E}[(D_{i-1}^L - \tilde{Q}_{i-1})^+] &= \int_{\tilde{Q}_{i-1}}^\infty (x - \tilde{Q}_{i-1}) g_{i-1}(x) dx \\ \text{Var}[(D_{i-1}^L - \tilde{Q}_{i-1})^+] &= \int_{\tilde{Q}_{i-1}}^\infty \left(x - \tilde{Q}_{i-1} - \mathbb{E}[(D_{i-1}^L - \tilde{Q}_{i-1})^+] \right)^2 g_{i-1}(x) dx. \end{aligned}$$

Let $\widehat{\sigma}_{i,\text{dest},m}^2 + \widehat{\sigma}_{i,c}^2 = \sigma_y^2$, $\widehat{\sigma}_{i-1,\text{org},m}^2 + \widehat{\sigma}_{i-1,c}^2 = \sigma_z^2$, and $\widehat{\sigma}_{i-1,\text{org},m}^2 = \sigma_x^2$, then the probability distribution functions of D_i and D_{i-1}^L can be written as follows:

$$f_i(x) = \exp\left(\frac{-(\ln(x) - \mu_i)^2}{2(\sigma_x^2 + \sigma_y^2)}\right) \frac{1}{x\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \quad (44)$$

¹² We cannot specify that $\ln(D_i^L)$ has a Laplace distribution if $\xi_{i,\text{org},m}$, $\xi_{i,\text{org},m}$, and $\xi_{i,c}$ are Laplace distributed since the sum of Laplace random variables is not a Laplace random variable.

$$g_{i-1}(x) = \exp\left(\frac{-(\ln(x) - \mu_{i-1})^2}{2(\sigma_x^2 + \sigma_z^2)}\right) \frac{1}{x\sqrt{2\pi(\sigma_x^2 + \sigma_z^2)}}. \quad (45)$$

The expression $\mathbb{E}[D_i(D_{i-1}^L - \tilde{Q}_{i-1})^+]$ can now be stated as follows:

$$\begin{aligned} \mathbb{E}[D_i(D_{i-1}^L - \tilde{Q}_{i-1})^+] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{(\ln \tilde{Q}_{i-1} - \mu_{i-1}) - x}^{\infty} \exp(\mu_i + y + x) \\ &\quad \times \left(\exp(\mu_{i-1} + z + x) - \tilde{Q}_{i-1} \right) \\ &\quad \times \frac{\exp(-z^2/2\sigma_z^2)}{\sigma_z\sqrt{2\pi}} \frac{\exp(-x^2/2\sigma_x^2)}{\sigma_x\sqrt{2\pi}} \frac{\exp(-y^2/2\sigma_y^2)}{\sigma_y\sqrt{2\pi}} dz dx dy \end{aligned}$$

We randomly select 1205 flights from our 2007 data set. After excluding the first flight of each rotation, we end up with 989 valid flights. For each valid flight, we calculate $\text{Var}[D_i]$, $\text{Var}[(D_{i-1}^L - \tilde{Q}_{i-1})^+]$, $\text{Cov}[D_i, (D_{i-1}^L - \tilde{Q}_{i-1})^+] = \mathbb{E}[D_i(D_{i-1}^L - \tilde{Q}_{i-1})^+] - \mathbb{E}[D_i]\mathbb{E}[(D_{i-1}^L - \tilde{Q}_{i-1})^+]$, and $\text{Var}[D_i^L]$ using (43). Then, we compute the effect of covariance in total variance by calculating the ratio $\frac{2\text{Cov}[D_i, (D_{i-1}^L - \tilde{Q}_{i-1})^+]}{\text{Var}[D_i^L]}$. The mean value of this ratio across all of these 989 flights is 3.69% (0.0369) with a maximum value of 19.72% and a minimum value of 0.00%. Thus, ignoring the covariance term leads to a small underestimation of the total variance in our model. However, computing the exact covariance term for each of the millions of flights in our data set is computationally not feasible. Therefore, we continue our analysis with the assumption that the terms D_i and $(D_{i-1}^L - (Q_{i-1} + B_i))^+$ are independent.

Appendix E: Including International Flights in the Congestion Calculation

To accurately measure the congestion variable, in addition to domestic flights, we also consider international flights. For this purpose, we use T-100 Segment (All Carriers) data which combines monthly domestic and international T-100 segment data. It includes departures scheduled and departures performed for each origin-destination pair served by different carriers. Flights with both origin and destination in a foreign country are not included. To calculate the new congestion variables, we follow the steps outlined below:

1. For each origin-month (*destination-month*) combination, we calculate the number of domestic and international flight departures (*arrivals*) separately. Let's call these variables

T100-domestic-flights-out (*T100-domestic-flights-in*) and T100-international-flights-out (*T100-international-flights-in*) respectively.

2. Next, we merge these origin-month-departures (*destination-month-arrivals*) triplets with our main data set.

3. From Airline On-Time Performance data set, we also calculate the number of domestic flight departures (*arrivals*) per month from each airport. Let's call this variable Ontime-domestic-flights-out (*Ontime-domestic-flights-in*).

4. Finally, we calculate the percentage of flight departure (*arrival*) increase by using the ratio of
$$\frac{\text{T100-domestic-flights-out} + \text{T100-international-flights-out}}{\text{Ontime-domestic-flights-out}} \left(\frac{\text{T100-domestic-flights-in} + \text{T100-international-flights-in}}{\text{Ontime-domestic-flights-in}} \right)$$

5. As the last step, we multiply our old normalized departure (*arrival*) congestion variable by the percentage of flight departure (*arrival*) increase calculated in Step 4.

In summary, we normalized the congestion variable as follows: we counted the total number of scheduled domestic and international departures (*arrivals*) around a one hour window of each flight and then divided this number by the hourly departure (*arrival*) capacity of each airport.