The Impact of Block Time Reliability on Scheduled Block Time Setting

Lu Hao, Mark Hansen
Civil and Environmental Engineering
University of California, Berkeley
Berkeley, CA, USA
haolu@berkeley.edu, mhansen@ce.berkeley.edu

***Abstract*—While in ground transportation the concept of reliability has been extensively studied, there is little literature in air transportation. Scheduled block time (SBT) setting is a crucial part in airlines' scheduling. Relevant work in ground transportation have shown that SBT and the historical block time distribution, reflecting block time reliability, have a close relationship. This paper investigates how the change in actual block time distribution will affect SBT and system performance. Firstly this relationship is studied with empirical data and multiple regression models. The distribution of the historical block time for a flight is depicted by the difference between every 10th percentiles. We found that gate delay plays a minor role in setting SBT and that SBTs have decreasing sensitivity to historical flight times toward the right tail of the distribution. With the behavior model results showing that both the median block time and the "inner right tail" of the distribution affect SBT setting, an impact study is conducted to validate these impacts with historical data. The impact of historical block time distribution on SBT is validated with real data in year 2006-2008 and 2009-2011. Furthermore, by studying the flight performance difference based on different changes in SBT, we conclude that ignoring the impact on SBT changes when considering potential benefits of improved block time reliability could lead to inaccurate results.**

Keywords-air transportation system; block time reliability; scheduled block time; behavior analysis; system performance

I. Introduction

The idea of reliability or (inversely) variability is not new in the field of ground transportation, where (un)reliability mainly refers to the unpredictable variations in travel time and is thus directly related to uncertainty of travel time [1]. A rapidly growing body of literature addresses measurement and valuation of travel time variability, and the goal of enhancing reliability seems to be an increasing priority for policy makers [2]. It is now standard that transit operators regularly publish statistics on reliability [2]. In the realm of commercial air transportation, the percentage of flights arriving within 15 minutes of their scheduled times is tracked by DOT and widely published on online flight booking sites. However, there is limited knowledge of how flight time variability affects airline schedule setting. The block time for a flight is defined as the time between when an aircraft moves under its own power for the purpose of flight and when the aircraft comes to rest after landing [3]. Typically the airline scheduler sets scheduled block time, or SBT, for a certain flight more than six months ahead of time [4]. Figure 1 illustrates SBT in the context of flight time decomposition. SBT is the time duration between the scheduled (computer reservation system, or CRS) departure and scheduled arrival time. The actual block time (FT) is the time between actual departure and arrival time and varies from day to day for the same flight. The block time can be further decomposed into taxi-out, airborne and taxi-in time. The time between scheduled and actual departure time is defined as departure delay, or gate delay.

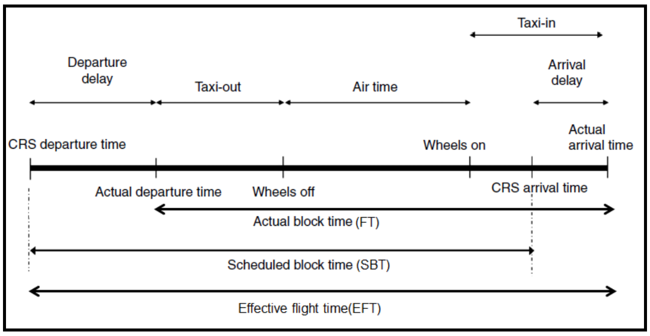

Figure 1. Scheduled Block Time (SBT) in the Context of Flight Time Decomposition. [3]

Airlines face a difficult set of trade-offs in setting SBT. They must balance their cost saving motive against their desire for good schedule adherence. Various researches in ground transportation have shown that travel time variability significantly affects travelers' departure time decisions. Despite the difference in motivation of the two types of decision making, it is natural to assume an analogous relationship between SBT and block-time variability in air transportation.

Understanding this relationship is critical to assessing the consequences of improvements to the NAS, such as those contemplated under NEXTGEN. Such improvements will affect the distributions of realized block times for individual flights. This change in distribution may affect both the SBT and deviations of actual block times from the SBT. It is important to consider both effects, since they have different economic implications. Changes in SBT influence a host of related costs including crew time and aircraft ownership, as well as the earliness and lateness of flights relative to the schedule. In current practice, however, the impact of a NAS improvement on SBT is not explicitly considered. In essence, it is assumed that any reduction in realized block time has the same economic value regardless of its impact on SBT.

In this paper, the potential for changes in the block time distribution to change SBT will be the major focus. To do this, firstly we need to understand the relationship between SBT and actual block times. This leads to our major contribution, which is to analyze the impact of the historical distributions of actual block times on SBT. With the efforts to fully understand the complete cycle of SBT setting, this paper explicitly links the relationship between the changes in actual block time distribution and SBT. This fills the missing piece of the NEXTGEN benefit analysis where the current assumption of benefit from reduced delays and improved block times does not consider the SBT factor.

Beyond this specific focus, our study provides a perspective on how the phenomenon of transport system reliability is manifested in the specific mode of scheduled air transport. As we shall see, these differences cause airlines to focus on a particular part of the block time distribution when setting SBT. The innovative methodology required to reveal this behavior is a further contribution of our work.

The remainder of this paper is organized as follows. Section 2 provides background on relevant concepts in ground travel time reliability as well as the current airline scheduling methodologies. In section 3 a behavioral analysis incorporating block time reliability is conducted to model the SBT setting behavior using historical data. Section 4 conducts an impact analysis where historical data are analyzed to demonstrate the impact of block time distribution changes on SBT setting, and the influence of SBT changes on trends in schedule adherence metrics. Conclusions are presented in section 5.

II. BACKGROUND

A. Travel Time Reliability in Ground Transportation

While research into surface travel time reliability has followed many branches, we focus here on departure time scheduling when the travel time is uncertain.

In ground transportation, traveler costs due to early or late arrivals are assumed to influence departure time decisions. A traveler's preferred arrival time (PAT) serves as a reference point that determines whether an arrival at a particular time is early or late. Gaver is one of the earliest advocates for this approach. In [5] a framework for explaining variability in trip-scheduling decisions is introduced given a delay distribution and the costs of arriving early or late. Later studies following this stream included more factors influencing travelers' tradeoff into consideration, such as: queue delay and schedule delay [6]; "safety margin" [7], [8]. Vickrey in [9] expressed this idea in terms of flexible functions for the marginal utility of time spent at the origin and at the destination, respectively. The idea was re-introduced and elaborated by Tseng and Verhoef [10], and subsequently applied by Jenelius et al. [11]. Another important contribution is by Small in [12], whose theoretical model, which is typically estimated using a discrete choice model addresses departure time choice in the traveler utility function. The traveler choice model is expanded to include travel time uncertainty in the form of a random variable with given probability density function in [13]. It is found that travelers shift their departure time earlier to avoid late arrivals. More recent work [14] proved mathematically the statement in [15] and [16], that the terms of expected earliness and lateness approximate the impact of standard deviation in the utility function.

B. Block Time Reliability and Scheduled Block Time Setting

One attempt to predict SBT using historical data is by Coy in [17]. A two-stage statistical model of airlines' SBT is applied in the paper. Realized block time is found to be an effective predictor of SBT, having a parameter very close to 1. In addition, arrival times, airport utilizations, poor weather conditions are found to be significant predictors of block time [17]. The variability (inversely reliability) of block time is not directly considered in [17]. Sohoni et al. in [18] defined two service-level metrics for an airline schedule to incorporate reliability and develop a stochastic integer programming formulation to adjust the existing schedule by changing the departure time to maximize expected profit, while ensuring the two service levels. Chiraphadhanakul and Barnhart [19] focused on schedule slack, defined as the additional time allocated beyond the expected time required for each aircraft connection, passenger connection, or flight leg. Considering the complexity of robust scheduling, they studied how to more effectively utilize the existing slacks rather than simply having more slacks to achieve a more robust schedule. Slacks can absorb delay to keep the system more reliable, however at a very high cost per minute. They developed the concept of effective slack (the total aircraft/passenger slack after accounting for the historical arrival delay) with a certain upper bound, as an optimization objective. Mayer, C. and Sinai, T. [20] explored the factors affecting scheduled block time using data on nearly 67 million flights between 1988 and 2000. They showed that average SBT is almost exactly equal to the median block time excluding departure delay, notwithstanding that flights on average leave about 10 minutes after scheduled departure time. As an effort to investigate one step beyond the first-moment metrics focusing on average block time, it is natural to consider measurements of dispersion, such as variance, standard deviation of block time in SBT models. In [21], a traditional mean-variance model is applied to capture both the centrality and dispersion of actual block time. The actual block time is decomposed into taxi-out time, airborne time and taxi-in time. Gate delay is also considered. The mean and standard deviation of each component serve as explanatory variables for SBT setting. Estimation results for the dispersion term are contrary to expectation. The standard deviations are expected to have a positive impact on SBT. However, it is found that the standard deviations of taxi-out time and airborne time, which are the major sources of block time unreliability, both have negative coefficients, suggesting that an increase in unreliability would reduce SBT, all else equal. The authors concluded that while SBT is highly influenced by historical average flight times, when these historical averages are pulled up as a result of high dispersion, the effect of dispersion is discounted. Similar results are found in [20] regarding the impact of standard deviation of historical block time on SBT. Therefore, the distribution of block time must be characterized in a more detailed way than simple second-moment metrics if its impact on SBT setting is to be understood.

Several studies point to the costs of earliness and building extra buffers into scheduled block times. Deshpande and

Arikan [4] calculated a cost ratio of leftover (overage) cost to shortage (underage) cost, representing the relative weight airlines appear to put onto lateness and earliness of a flight. Their results show the implied flight lateness costs are less than early arrival costs for a large fraction of flights [4]. This is different from ground transportation; however, it is consistent with [13], which notes airlines' tendency to shorten SBT in order to save cost. To do this, they are willing to incur more delay and less on-time reliability. The cost impact of SBT is shown in [22], where econometric cost function estimates incorporating a variety of delay-buffer models reveal that both delay and schedule buffer are important cost drivers. The coefficients suggest 0.6% increase in variable cost would occur if there is 1-min increase in average delay against schedule or in schedule buffer. The ability to reduce such buffers by shortening SBTs (without increasing delay against schedule) could thus result in significant cost savings.

III. METHODOLOGY AND AGGREGATE PERCENTILE MODEL

As a measure of travel time variability in ground transportation, most studies have used either the standard deviation or the average delay relative to scheduled arrival time [2]. However, both [20] and [21] find that the standard deviation of actual block time reduces SBT. This is probably because airlines disregard extremely long actual flight times when setting SBTs to maintain competitiveness and efficiency. A interview with a major US airline is conducted in [24] showing the rule for SBT setting seems to be a specific BTR (block time reliability) target, i.e. a certain percentile of the historical block time distribution. Thus, we developed a model with the percentile statistics of the actual flight time. The huge amount of historical data in the field of air transportation allows us to employ this approach to empirically investigate SBT setting behavior.

A. Data and Modeling

The model presented here is an extension of the work presented in [24]. For readability, we restate the modeling approach described there.

The variables capture the difference in percentile of historical block time; therefore the model is called the percentile model in this paper. The percentile model is a generalization of the BTR target model, and assumes that, because of the adjustments to the BTR that airlines make based on on-time performance, competition, and other factors, the SBT is influenced by more than a single percentile of the historical block time distribution. For the same reason, other variables than the historical block time distribution that might also affect the SBT decisions are also included in the model.

The data on which the percentile model is estimated are collected from three sources: the Airline On-time Performance dataset, the air carrier statistics data from T-100 Domestic segment with U.S. carrier, Form 41 database, and the aircraft type information from a combined dataset including B43, OAG and FAA registry aircraft. The first two datasets are both acquired from the Bureau of Transportation Statistics (BTS). We employ the Bureau of Transportation Statistics (BTS) Airline On-time Performance data to characterize airline schedule and operations. This database contains detailed performance information for individual flights by major US air carriers between points within the United States. These flight records are aggregated to capture the distribution of historical flight time. An individual flight group is specified by airline, origin-destination pair, 30-minute departure time window, and aircraft type. For instance, ATL BOS 20 B757 DL means the group of flights from ATL to BOS at departure time window between 10:00 to 10:30 am flying Boeing 757 by Delta Air Lines. This is an improvement over [24], in which flights are defined by origin, destination, airline, and flight number. Each flight group is referred to as an individual flight in this paper. The block time distribution for a flight group is based on aggregation over a quarter.

For each quarter, we assume that there is a single SBT for each individual flight, which is the elapsed time between the scheduled departure and the scheduled arrival time. However, since there are occasional variations within the quarter, the median value of SBT in the quarter is used as the dependent variable. The distribution of actual flight time is captured by calculating differences in percentiles of the historical flight time data. In contrast to [24], we explicitly consider two different flight phases, taxi-out (hereafter TO) time and non-taxi out (hereafter NTO) time—whose historical distributions may have different effects on SBT. Also, because gate delay is expected to have a different effect than flight time, we include the difference in percentiles for gate delay separately.

For individual flight f in day t, the deciles of the distribution from 50th to 100th percentiles of the different components of block time and gate delay are calculated. The 50th percentile or median taxi-out time (NTO time, gate delay), denoted as $TO_{0.5}^{f,q,y}$ ($nonTO_{0.5}^{f,q,y}$, $dep_{0.5}^{f,q,y}$) of individual flight $f \in F$ in quarter q of year y are all included in the model. The variability of flight time is further captured by the differences between every successive decile from 50th to 100th percentile. For example, $dTO_{56}^{f,q,y} = TO_{0.6}^{f,q,y} - TO_{0.5}^{f,q,y}$ is the difference between the 50th and 60th percentile of taxi-out time for flight f. This approach depicts the distribution of components of flight time information in a manner that can represent the industry practice of BTR-based block-time setting found in our interview. The different segments of percentiles capture how SBT is influenced by successively rarer but higher realized flight time values, reflecting the reliability of historical flight time.

Competition with other airlines flying the same market may motivate a shorter SBT so that the airline appears to offer faster service, or a longer SBT so that it appears more reliable. Therefore, we include the HHI variables that depict the OD pair competitiveness in the model [24]. For market *od*, the HHI can be calculated as:

$$HHI_{od} = \sum_{i=1}^{N} \left(\frac{s_{odi}}{s_{od}} \right)^2 \tag{1}$$

where S_{odi} is the number of seats provided by carrier *i* flying this OD pair, s_{od} is the total number of seats provided in this OD pair, and *N* is the number of carriers in this OD pair. The US DOT T-100 database provides such information.

Lastly, based on interview with the industry [24], airlines give special consideration to their hubs, where they have the majority of the gates and the traffic. Most airlines adopt some variation of a hub-and-spoke system. Major carriers operate up to five hubs, while smaller ones typically have one hub located at the center of the region they serve [23]. Airlines set shorter SBTs for flights into their major hub airports to avoid early arrivals that can be highly disruptive to their operations. Therefore, in the percentile model we include dummy variables HUB_O and HUB_D that are airline and airport specific. They indicate whether the origin /destination airport is a major operation hub for the specific carrier, for each individual flight. This is an extension building on [24], where hub effects are only considered at individual airline level, rather than over the whole dataset.

In the formulation, we assume that schedulers set SBT for a flight with the knowledge of actual flight information and the HHI competition index of the same quarter in the previous two consecutive years. This implies that schedulers focus on flight experience during the same season for which they are scheduling. In this paper, the years 2009, 2010 and 2011 are chosen to be studied, with the SBT in 2011 modeled based on the actual flight data from the same quarter in 2009 and 2010. The realized flight time information in these two years are aggregated together to calculate the percentiles. The resulting model, with y=2011, and h(y) indicating the two years prior to y, in this case years 2009 and 2010, is:

$$SBT^{q,y}{}_{f} = \alpha_1 \times TO_{0.5}^{f,q,h(y)} + \alpha_2 \times nonTO_{0.5}^{f,q,h(y)} + \alpha_3 \times dep_{0.5}^{f,q,h(y)} + \sum_{i=1}^{5} \beta_i \times dTO_{i+4,i+5}^{f,q,h(y)}$$
$$+\sum_{i=1}^{5} \gamma_i \times dnonTO_{i+4,i+5}^{f,q,y,h(y)} + \sum_{i=1}^{5} \lambda_i \times ddep_{i+4,i+5}^{f,q,h(y)} + \gamma_1 \times HUB_O + \gamma_2 \times HUB_D$$
$$+\chi \times HHI^{h(y)}{}_{od} + \pi \times Dist + \sum_{i=1}^{3} \mu_i \times Quarter_i + const \tag{2}$$

To assure robustness in the data, we only include the flights that are frequently flown in a quarter in the two years. We also want to focus more on regularly operated flights in the dataset. Therefore, to be included in the data set, an individual flight must have been flown at least 50 times in a given quarter over 2009 and 2010, not including Saturdays. After this filter is applied, the estimation data set consists of 42,625 observations, each corresponding to an individual flight with a given departure time window, flying a given aircraft type, operated by a given airline, between a given origin and destination.

B. Estimation Results

The estimation results on the whole dataset are shown in Table 1. The R^2 explains almost 100% of the variation in scheduled block-time. Distance is positively related to SBT, suggesting that there is more uncertainty in longer flights that is not reflected in the historical block time distribution and SBTs are set to be longer—at a rate of 4 min per 1000 miles—to take the uncertainty into consideration. We can see that the coefficients for the gate delay distribution are all quite small and some are not significant. SBT increases 0.2 minutes for every 1 minute increase in median gate delay, but, surprisingly, decreases as the difference between the median and the 60th percentile increases. These results confirm that historical gate delay is not a strong consideration in setting SBT, but also suggest that it does have some effect.

TABLE I. ESTIMATION RESULTS: PERCENTILE MODEL

	Percentile Model		HM 1	HM 2
Variable	*Estimate*	*p-Value*	*Coefficient*	*Coefficient*
Intercept	5.838	<.0001	-	-
$dist_{od}$	0.004	<.0001	-	-
$nonTO_{0.5}^{f,q,y}$	0.977	<.0001	0.75	1
$dnonTO_{56}^{f,q,y}$	0.441	<.0001	0.45	1
$dnonTO_{67}^{f,q,y}$	0.353	<.0001	0.3	1
$dnonTO_{78}^{f,q,y}$	0.191	0.0001	0.25	0
$dnonTO_{89}^{f,q,y}$	0.0747	<.0001	0.15	0
$dnonTO_{90}^{f,q,y}$	0.001	0.6813	0.05	0
$TO_{0.5}^{f,q,y}$	0.793	<.0001	0.75	1
$dTO_{56}^{f,q,y}$	0.42	<.0001	0.45	1
$dTO_{67}^{f,q,y}$	0.282	<.0001	0.3	1
$dTO_{78}^{f,q,y}$	0.163	0.0001	0.25	0
$dTO_{89}^{f,q,y}$	0.065	<.0001	0.15	0
$dTO_{90}^{f,q,y}$	0.0018	0.0663	0.05	0
$dep_{0.5}^{f,q,y}$	0.201	<.0001	-	-
$ddep_{56}^{f,q,y}$	-0.126	<.0001	-	-
$ddep_{67}^{f,q,y}$	-0.01	0.2065	-	-
$ddep_{78}^{f,q,y}$	0.0027	0.5666	-	-
$ddep_{89}^{f,q,y}$	0.0054	0.0067	-	-
$ddep_{90}^{f,q,y}$	-0.002	<.0001	-	-
$Quarter_1^y$	-0.913	0.2337	-	-
$Quarter_2^y$	-1.409	0.6249	-	-
$Quarter_3^y$	-0.811	0.2480	-	-
HHI_{od}	-0.882	<.0001	-	-
HUB_O	0.116	0.1396	-	-
HUB_D	-0.812	0.0345	-	-
R-square	0.9958			

The coefficient on median NTO time is 0.977, which is close to 1, indicating that this is a major determinant of SBT. The $d_{i,i+1}^{f,q,y}$ variables are intended to capture the variability of non-taxi-out flight time over the right tail of the distribution where it exceeds the median value. The 1-minute increase in the interval between 50th and 60th percentile generates a 0.44 minute increase in SBT. The coefficient decreases to 0.19 minutes for the interval between 70th and 80th percentile and to only 0.0007 minutes for the far right end tail of the distribution. These results show that SBT is strongly affected by the left tail of the NTO flight time distribution (as reflected by the median), while the "inner right tail" has a moderate effect, whereas the effect of the outer right tail above the 70th percentile has a rather small effect.

The pattern is similar for the taxi-out component of flight time, but the coefficients are somewhat smaller. For example, the median taxi-out time has a coefficient 0.793 (as compared to 0.977 for the NTO time). This is probably because, as indicated in our interview, airlines give more consideration to terminal and gate assignment changes and less to historical data in predicting taxi out times. The right tail of the distribution also has a rapidly decreasing pattern, from 0.42 for *d56* to 0.0018 for *d90*. For the gate delay variables, 1-minute increase in median gate delay increases 0.2 minute in SBT. This impact is relatively small compared to the flight time variables. The right tail of the distribution has even smaller impacts, and also not quite significant.

The HHI variable has a negative coefficient. Higher HHI indicates lower competitiveness for the OD pair. Thus, a negative coefficient means that if the OD market is highly competitive, airlines will increase SBT. This suggests that competition drives airlines toward improving on-time performance instead of publishing a shorter flight duration. Regarding the effect of airline hub airports, the dummy variable for origin hub airport is not quite significant, while the effect of destination hub airport is marginally significant, with a coefficient of -0.812. Schedulers thus tend to set a shorter SBT for flights into the airline's hubs. This is consistent with the airlines interviewees' statement that they set shorter SBT for their hub airports in an effort to avoid early arrivals. However, the magnitude of this adjustment is quite small—about 1 minute.

The percentile model represents airlines' composite SBT-setting behavior, in a manner that explicitly shows the weight they place on different regions of the historical distribution of realized block times. To further interpret the results of the percentile model, two hypothetical models for SBT setting are shown in the last two columns in table 1 to compare with our estimation results. The details of the two hypothetical models can be found in [24]. The formulations are:

$$SBT = 0.75\times Q_{0.5} + 0.45\times d56 + 0.35\times d67 + 0.25\times d78 + 0.15\times d89 + 0.05\times d90 \tag{3}$$

$$SBT = 1\times Q_{0.5} + 1\times d56 + 1\times d67 + 0\times d78 + 0\times d89 + 0\times d90 \tag{4}$$

Table 1 also compares the results between percentile model and the hypothetical models. HM1 (equation (3)) only considers the mean value of flight time. In the estimated percentile model, the coefficient for the median NTO flight time ($\mathrm{nonTO}_{0.5}$) is larger in the percentile model. The coefficients for the differences from the 50th to 100th percentile decrease at a faster rate in the estimated model than they do for HM1. This clearly shows that, compared to HM1, SBTs place more weight on the left side of the flight time distribution while down-weighting the far right tail, particularly above the 70^{th} percentile. For the taxi-out component, the median value is closer to HM1. However, every interval above the 50^{th} percentile has a smaller coefficient than under HM1. These findings are broadly consistent with previous literature where the implied flight delay costs are less than the implied costs of early arrivals for a large fraction of flights [4]. Put another way, airlines tend to be "optimistic" when they choose the SBT, tolerating longer delays in order to realize the advantages of shorter SBTs. However, our results go further in showing that airlines specifically discount delays associated with the roughly 20% of flight realizations with the longest durations, while paying some attention to the inner right tail for the block time distribution.

HM2 (equation (4)) assumes SBT is solely based on the 70th percentile of actual block time and thus ignores flights times beyond these values. In the estimated percentile model, the coefficients for the median values are close to 1, as in this hypothetical model. In contrast to that model however, the inner right tail parameters are less than 1 and outer right tail parameters are greater than 0. Thus, compared to HM2, the estimated percentile model shifts weight from the inner right to the outer right tail. One interpretation of this is that the regression model, when estimated for a large diverse set of flights, captures a composite of different BTR standards. Thus, for the NTO component, 97% of flights have a standard at or above 50%, 44% have a standard at or above 60%, and so forth. However, it is also possible that the different regions of the block time distribution are indirectly taken into account through the various adjustments airlines make to the nominal BTR standard. This seems the more likely explanation for the small but significant influence of the far right tail, since we have heard no reports of airlines setting the BTR threshold at 80%, 90%, or 100%.

Returning to the comparison with the morning commute, we observe from these results that airlines are more willing to be late than most workers. While most workers would not want to be late 20% of the time, airlines pay little attention to block times over the 80^{th} percentile. In exchange for this, they reduce earliness and avoid the high costs of setting longer block times.

IV. IMPACT ANALYSIS

The percentile model in section 3 shows the impact of the distributions of historical block time on SBT setting. Different segments of the distribution have varying impacts on the SBT, with left and inner right tails of the distribution the most influential. In the real system block time distributions are constantly changing, and scheduled block times updated in response to the changing distributions, as well as other factors. It is of interest to observe these changes over a period of years, and in particular to observe the contributions of the changing distributions and SBT adjustments to changes in schedule adherence metrics. In this section, we perform such an analysis.

A. Methodology

The purpose of the impact study is to document changes in actual block time distributions and the corresponding change in SBT from historical data. For this comparison, year 2007 is of particular interest because it is an extremely busy year with a large amount of delay. It is a reasonable speculation that the highly dispersed block time, observed in [21], will lead to significantly different behavior in SBT setting compared to years in which the system was less congested. It is also suspected that various flight performance metrics will also exhibit different patterns. Since our percentile model requires two years of data for setting SBT, we include year 2006 and 2007 as one group in the impact study. The data from year 2009 and 2010 are used as the other group for comparison.

For each group of the two-year data, the flights are aggregated in the same manner as before: by OD pair, departure time window, aircraft type, carrier and quarter; each distinct combination of these attributes is treated as an individual flight. Merging the two groups of two-year data together, there are 8,353 observations in total with at least 50 realizations in each period. From section 3, we learnt that the inner right tail of the block time distribution has the main impact on SBT setting, in addition to the median actual block time. Therefore, in a process similar to that described in section 4, we calculated the inner right tail of the historical block time distribution as the difference between the median and the 75th percentile as well as the median value. The average median block time decreased 0.66 minutes between 2006-2007 and 2009-2010, with a standard deviation of 4.92 minutes. For the inner right tail, the average change is -0.21 minutes, with a standard deviation of 2.84 minutes.

We want to relate the changes in bock time distribution to changes in SBT. Therefore, the dataset is divided into nine separate "scenarios" where the median block time and the inner right tail could increase, decrease, or remain the same across these two time periods. If the change in median or inner right tail is less than one standard deviation from the mean change, then the flight is assigned to the "Same" group for that variable. If there is a change greater than one standard deviation above the mean change, then the observation is categorized as an "Increase", and conversely for "Decrease" (We use the terms "Same," "Increase" and "Decrease" somewhat loosely, since they are defined in terms of deviation from the average change; however, since the average change is close to zero they are appropriate).

For each scenario, we are interested in the changes in SBT and various schedule adherence metrics years after 2006-07 and 2009-10: i.e. the years 2008 and 2011. Firstly we studied the changes in SBT. For each scenario in both years, the median SBT for each individual flight is calculated, as defined in section 3 and the average SBT for each scenario is recorded and compared for the two years. Moreover, the on-time performance metrics—A0 and A14—and average negative deviation (ND) and positive deviation (PD) of the flights for each scenario is also calculated and compared.

We consider two on-time performance metrics: A0 and A14. A14 is based on the DOT definition: a flight is on time if it arrives at its destination gate less than 15 minutes after its scheduled arrival time. Let $o^{q,y}_{f,i} = 1$ if the ith realization of the flight $f \in F$ that flew a total of N times in quarter q of year y is on time, and 0 other wise. Then the variable depicting the quarterly on-time performance is: $A14^{q,y}_f = \sum_i o^{q,y}_{f,i} \Big/ N^{q,y}_f$. The A0 on-time performance is stricter than A14, counting a flight as on time only if it arrives no later than its scheduled arrival time. The calculation for A0 is similar to that of A14, and is denoted as $A0^{q,y}_f$. In addition to on-time performance, positive and negative deviations between realized and scheduled block times are also investigated. The negative deviation (ND) for the ith realization of the flight $f \in F$ that flew in quarter q of year y is $\max(SBT^{q,y}_f - ABT^{q,y}_{f,i}, 0)$ where $ABT^{q,y}_{f,i}$ indicates the actual block time of this single realized flight. Similarly, the positive deviation (PD) is $\max(ABT^{a,y}_{f,i} - SBT^{a,y}_f, 0)$ for the ith realization of flight f. Note that these deviations only reflect the block time, not gate delay. The latter is, however, captured in the on-time performance variables A0 and A14.

Lastly, we investigate how the change in SBT affected these metrics, by calculating their values under the counter-factual scenario in which the SBT in 2011 is the same as that in 2008.

B. Results

Table 2 below shows the impact analysis results. The nine scenarios are numbered as 1 to 9. The second major column in table 2 presents the results from comparing SBTs in year 2008 and year 2011. The upper half shows the average SBT values, where the bottom part shows the average change in SBT and its standard deviation. The largest increase (decrease) of SBT happens where both median and inner right tail of previous years' actual block time increased (decreased). Median block time change is clearly the major determinant of whether SBT increases or decreases. However, the effect of the inner right tail is also significant. Comparing scenario 1 and 7, where median BT increased and the inner right tail increased and decreased respectively, the change in SBT has a 3.5 minutes greater in the former case. These differences are 3.3 and 3.2 minutes when we make similar comparisons for the scenarios with the same median BT (scenarios 2 and 8) and a reduced median BT (scenarios 3 and 9). Differences among the various median scenarios for a given inner right tail scenario are also fairly consistent—around 9 minutes. Finally, in scenario 5, where both the median and the inner right tail are in the "same" category, average SBT decreases 0.35 minutes. This is in line with the small reductions in the average median and average inner right tail between the two periods.

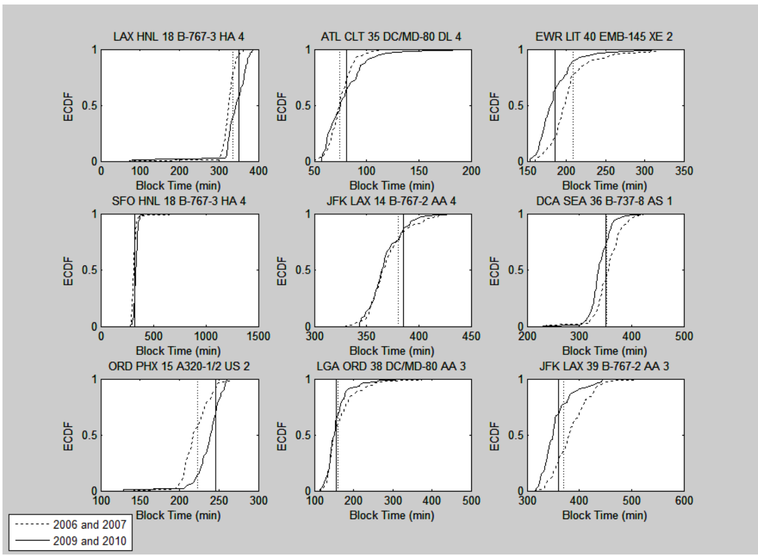

Figure 2. Empirical CDF of Actual Block Time for Representative Flight under Each Scenario

TABLE II. CHANGES IN SCHEDULED BLOCK TIMES AND VARIOUS PERFORMANCE METRICS FOR DIFFERENT SCENARIOS

		SBT (min)		A0			A14		
Scenario	**Scenario Description**	***2008***	***2011***	***2008***	***2011***	***2011'***	***2008***	***2011***	***2011'***
1	Med +, IRTail +	150.591	156.425	0.526	0.679	0.557	0.762	0.842	0.803
2	Med same, IRTail +	153.834	155.107	0.546	0.631	0.598	0.776	0.822	0.820
3	Med –, IRTail +	218.986	214.877	0.517	0.599	0.675	0.716	0.794	0.828
4	Med +, IRTail same	150.112	153.750	0.550	0.669	0.571	0.797	0.857	0.832
5	Med same, IRTail same	118.388	118.040	0.593	0.632	0.638	0.818	0.845	0.853
6	Med –, IRTail same	181.443	176.533	0.554	0.540	0.647	0.757	0.777	0.828
7	Med +, IRTail –	199.307	201.574	0.509	0.657	0.584	0.757	0.851	0.817
8	Med same, IRTail –	158.923	156.873	0.539	0.581	0.621	0.742	0.788	0.810
9	Med –, IRTail –	184.304	176.956	0.491	0.511	0.620	0.668	0.705	0.763
		SBT Change (min)		**ND (min)**			**PD (min)**		
Scenario	***Scenario Description***	***mean***	***s.t.d***	***2008***	***2011***	***2011'***	***2008***	***2011***	***2011'***
1	Med +, IRTail +	5.834	6.323	4.939	8.966	5.414	6.397	3.579	5.861
2	Med same, IRTail +	1.273	6.429	5.456	7.322	6.624	5.069	3.642	4.218
3	Med –, IRTail +	-4.109	7.968	8.313	8.871	11.833	6.139	4.524	3.377
4	Med +, IRTail same	3.638	6.513	4.654	7.186	5.148	5.336	3.219	4.818
5	Med same, IRTail same	-0.348	5.113	4.722	5.293	5.595	3.554	2.870	2.825
6	Med –, IRTail same	-4.909	6.637	7.596	6.352	9.879	4.343	4.312	2.930
7	Med +, IRTail –	2.267	7.995	6.817	9.132	7.933	6.250	3.744	4.812
8	Med same, IRTail –	-2.050	6.733	6.739	7.196	8.704	5.078	4.050	3.508
9	Med –, IRTail –	-7.348	8.024	9.396	7.629	13.219	5.827	5.030	3.272

To further illustrate the changes in actual block time distribution and the change in SBT, one representative individual flight is picked from each scenario and the empirical CDF of historical block time in year 2006-2007 and 2009-2010 is plotted in figure 2. The vertical lines in each graph denote the SBT for the earlier year (2008—dashed line), and the newer year (2011—solid line). The graph shows patterns similar to those found table 2. The middle column is especially interesting as it illustrates scenarios where median BT does not change while the inner right tail does. In the top graph where inner right tail increased, the SBT also increased, and conversely for the graph in the middle bottom.

Regarding the performance metric results in table 2, we note first that there was overall improvement in on-time performance and reductions in average PD between 2008 and 2011. Average ND also increased for most scenarios. These overall results derive from changes in block time distributions (and in the case of the on-time metrics, gate delay) combined with changes in SBT. The impact of the latter is isolated by comparing the 2011 results with the 2011' results, which show what the 2011 performance would have been if SBTs had not changed from their 2008 values. We see that changes in SBT have sizable impacts. For example in Scenario 1—median increase and right tail increase—a large increase in on-time performance between 2008 and 2011, as well as a large reduction in average PD (and increase in average ND) are mainly due to a 6 minute increase in SBT. On the other hand, in Scenario 9—median and right tail decrease—reductions in SBT virtually eliminate what would otherwise be substantial increases in on-time performance and decreases in average PD. More generally, the magnitudes of differences in on-time performance between 2011 and 2011', which reflect only the impact of SBT change, are comparable to the magnitudes of the overall differences 2008 and 2011. In the case of average PD, changes resulting from SBT adjustment are of somewhat larger magnitude than the overall changes observed. In sum, changes in SBT were as large or larger a driver of change in schedule adherence between 2008 and 2011 as changes in underlying operational performance.

V. CONCLUSION

In this paper, we study the impact of historical block time distribution (reflecting block time reliability) on SBT, as well as how changes in both affect schedule adherence metrics. We developed the percentile model in order to capture airlines' BTR-based practice. The variability in block time is captured by increments between every 10^{th} percentile above the 50^{th}. This enables us to observe how different regions of the historical block-time distribution are considered in SBT setting. Other variables, such as gate delay, distance, airport hub status, and competiveness are also included in the model. To investigate changes in block time distributions and associated scheduled block times that actually result from a change in operating conditions in the NAS, we compare the high traffic period around 2007 with the period of curtailed traffic around 2010.

The behavioral model suggests that the entire right tail of the block time distribution is considered when setting SBT, but that the inner right tail up to 70^{th} percentile receives by far the most consideration. In general, airlines are willing to experience occasional severe delays in exchange for a shorter

SBT. Among other factors, notable results include that historical gate delay is virtually ignored, that historical taxi-out time is given somewhat less weight in the setting SBT compared to the non-taxi-out component of flight, that airlines with hubs tend to set shorter SBTs for their hub-bound flights, and that competition encourages longer SBTs. Our impact analysis demonstrates that when changes in block time distributions—in particular the median and inner right tail—occur significant adjustments in SBTs often result, and that the impacts of these adjustments on schedule adherence is great or greater than the changes in the underlying operational performance.

This research has sought to understand the connection between SBT and the historical distribution of realized block times and gate delay. We have also considered how changes in SBT influence changes in schedule adherence metrics such as on-time performance. In the future, NEXTGEN is expected to be a major source of change in operational performance, and therefore SBTs, and finally of deviations between scheduled and realized times. It is clear from our results that knowledge of the change in average block times is not sufficient to understand these impacts, since a given change in the average can arise from many different changes in the distribution. This suggests that business cases for NAS improvements should pay more attention to impacts on the distribution of block times, instead of the average. Broadly speaking, improvements that push in the far right tail of the distribution will affect delays and on-time performance but not the SBT, while improvements that shift the inner right tail will effect scheduled block times but have limited impact on on-time performance. There is benefit from either change, but the nature of the benefit is fundamentally different, and it is important that NEXTGEN business cases recognize this.

As a final remark, we note that airlines virtually ignore historical gate delay in setting SBTs. This is curious since gate delay is such a major contributor to variation in effect flight time. It may be possible to improve on-time performance significantly by adjusting scheduled block time to reflect expected gate delay, to the extent it can be predicted from past experience.

ACKNOWLEDGEMENT

An earlier version of part of the paper has been presented at the Tenth USA/Europe Air Traffic Management Research and Development Seminar, June 2013, Chicago, IL. Substantial improvements have been made to the model as well as content in this paper. The earlier contents are included only to insure readability of this paper. Detailed descriptions of the earlier contents can be found in the corresponding reference.

REFERENCES

[1] Carrion, C., Levinson, D., 2012. Value of Travel Time Reliability: A Review of Current Evidence. *Transportation Research Part A* 46[4], 720-741.

[2] Börjesson, M., Eliasson, J., Franklin, J.P., 2012. Valuations of Travel Time Variability in Scheduling versus Mean-variance Models. *Transportation Research Part B* 46 [7], 855-873

[3] Electronic Code of Federal Regulations, Title 14: Aeronautics and Space, section 1.1. http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=91ac073c1ab333fde1e47f9d87fa244f&rgn=div8&view=text&node=14:1.0.1.1.1.0.1.1&idno=14. Last accessed Jul 5, 2012.3

[4] Deshpande, V., Arikan, M., 2012. The Impact of Airline Flight Schedules on Flight Delays. *Manufacturing & Service Operations Management* 14[3], 423-440.

[5] Gaver, D., 1968. Headstart Strategies for Combating Congestion. *Transportation Science*, 2[2], 172-181.

[6] Vickrey, W., 1969. Congestion Theory and Transportation Investment. *The American Economic Review* 59[2], 251-260.

[7] Knight, T., 1974. An Approach to the Evaluation of changes in Travel Unreliability: A Safety Margin Hypothesis. *Transportation* 3[4], 393-408.

[8] Pells, S., 1987. The Evaluation of Reductions in Travel Time Variability. PhD thesis, University of Leeds [UK].

[9] Vickrey, W.,1973. Pricing, Metering and Efficiently using urban transportation facilities. Highway Research record 476, 36-48.

[10] Tseng, Y.-Y., Verhoef, E.T., 2008. Value of time by time of day: a stated-preference study. *Transportation Research Part B* 42[7-8], 607-618.

[11] Jenelius, E., Mattsson, L.-G., Levinson, D., 2011. Traveler delay costs and value of time with trip chains, flexible activity scheduling and information. *Transportation Research Part B* 45 [5], 789–807.

[12] Small, K., 1982. The Scheduling of Consumer Activities: Work Trips. *American economic Review* 72[3], 467-479.

[13] Noland, R., Small, K., 1995. Travel-time uncertainty, Departure Time Choice, and the Cost of Morning Commutes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1493, Transportation Research Board of the National Academies, Washington, D.C., 150-158.

[14] Fosgerau, M., Karlstrom, A., 2010. The Value of Reliability. *Transportation Research Part B* 44[1], 38-49.

[15] Bates, J., et al., 2001. The Valuation of Reliability for Personal Travel. *Transportation Research Part E* 37[2], 191-229.

[16] Fosgerau, M., Engelson, L., 2011. value The of travel time variance. *Transportation Research Part B* 45 [1], 1–8.

[17] Coy, S., 2006. A Global Model for Estimating the Block Time of Commercial Passenger Aircraft. *Journal of Air Transport Management* 12[6], 300-305

[18] Sohoni, M., et al., 2011. Robust Airline Scheduling Under Block-Time Uncertainty. *Transportation Science* 45[4], 451-464

[19] Chiraphadhanakul, V., Barnhart C., 2011. Robust Flight Schedules through Slack Re-Allocation. Paper submitted to Anna Valicek Medal. www.agifors.org/award/submissions2011/VirotChiraphadhanakul.pdf. Last accessed Mar.10, 2013.

[20] Mayer, C., Sinai, T., 2003. Why do airlines systematically schedule their ights to arrive late?" Working paper, The Wharton School, University of Pennsylvania

[21] Hao, L., Hansen M., 2013. Flight Time Predictability: Concepts, Metrics and Impact on Scheduled Block Time. 92nd Annual Meeting of the Transportation Research Board Carrier Cost Structure: Evidence from US Airlines. *Transportation*, January 2013, Washington, DC.

[22] Zou, B., Hansen M., 2012. Impact of Operational Performance on Air *Research Part E* 48[5], 1032-1048.

[23] Bazargan, Massoud, 2010. Airline Operations and Scheduling.

[24] Hao, L., Hansen M., 2013. How Airlines Set Scheduled Block Times. Tenth USA/Europe Air Traffic Management Research and Development Seminar, June 2013, Chicago, IL.

[25] Liston-Hayes, C., Pilkington A., 2004. Inventive Concentration: An Analysis of Fuel Cell Patents. *Science and Public Policy* 31[1], 15-25.