CESifo WORKING PAPERS

7875
2019

September 2019

Airline Mitigation of Propagated Delays: Theory and Empirics on the Choice of Schedule Buffers

Jan K. Brueckner, Achim I. Czerny, Alberto A. Gaggero

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7875*
Category 11: Industrial Organisation

Airline Mitigation of Propagated Delays: Theory and Empirics on the Choice of Schedule Buffers

Abstract

This paper presents an extensive theoretical and empirical analysis of the choice of schedule buffers by airlines. With airline delays a continuing problem around the world, such an undertaking is valuable, and its lessons extend to other passenger transportation sectors. One useful lesson from the theoretical analysis of a two-flight model is that the mitigation of delay propagation is done entirely by the ground buffer and the second flight's buffer. The first flight's buffer plays no role because the ground buffer is a perfect, while nondistorting, substitute. In addition, the apportionment of mitigation responsibility between the ground buffer and the flight buffer of flight 2 is shown to depend on the relationship between the costs of ground-and flight-buffer time. The empirical results show the connection between buffer magnitudes and a host of explanatory variables, including the variability of flight times, which simulations of the model identify as an important determining factor.

JEL-Codes: L930.

Keywords: flight delays, delay propagation, buffers.

Jan K. Brueckner
University of California, Irvine
USA - 92697, Irvine, CA
jkbrueck@uci.edu

Achim I. Czerny
Hong Kong Polytechnic University
Hong Kong
achim.czerny@polyu.edu.hk

Alberto A. Gaggero
University of Pavia
Italy – 27100 Pavia
alberto.gaggero@unipv.it

September 2019

Airline Mitigation of Propagated Delays: Theory and Empirics on the Choice of Schedule Buffers

by

Jan K. Brueckner, Achim I. Czerny, Alberto A. Gaggero*

1. Introduction

Flight delays are a worldwide problem, a consequence of the substantial growth in air travel over recent decades. In the US alone, the cost of delays for passengers and airlines was estimated at $32.9 billion in 2010 by Ball et al. (2010). In response to the problem, the US Department of Transportation requires all major US airlines to provide monthly information about delays, which generates widely viewed on-time rankings of the carriers. The European Union has imposed rules for passenger compensation and assistance in the event of long flight delays.

A major source of flight delays is airport congestion, which is the subject of a large literature (see Zhang and Czerny (2012) for a survey). But whether congestion leads to flight delays is largely under the control of the airlines, since they are free to set scheduled flight durations. In other words, the congestion-related lengthening of flight times can be built into airline schedules through a practice known as "schedule padding", whose recent growth is documented by Forbes Lederman and Yuan (2019) and others. While airport congestion may make flights longer, this schedule adjustment prevents them from arriving late.

Despite this overall adjustment in response to broad trends, flight times are still influenced by many random daily factors, including weather, mechanical issues, and unanticipated congestion, which can vary by day and hour around some expected level. Airline scheduling decisions take account of these random influences through the choice of "schedule buffers." One type of buffer is known in the airline industry as a "block-time buffer" (we call it a "flight buffer"), and it equals the amount of time added to the expected (in probabilistic terms) flight time to get the scheduled arrival time. While a longer flight buffer reduces the chance of late arrival,

*We thank Tiziana D'Alfonso for helpful comments and Alex Luttman for comments and for providing us with one of his data sets. Any shortcomings in the paper are our responsibility.

it also makes an early arrival more likely, and while passengers dislike delays, they also do not want flights to routinely arrive early, an outcome that leaves time gaps that must be filled. Longer scheduled flight times also raise airline operating costs, such as the cost of crew time. In setting its flight schedules, an airline will take all three factors (disutilities from lateness and earliness as well as operating costs) into account. Flight buffers are typically positive, reflecting a greater concern about late as opposed to early arrivals on the part of the carrier (following passenger preferences). Note that these same three elements also affect scheduling by other transportation providers, such as passenger railroads and intercity bus lines.

Delays depend on more than just the operating time of a particular flight. If the incoming aircraft arrives late, then the outbound flight is likely to depart late, possibly leading to its late arrival even if operating time is normal. Late-arriving aircraft are in fact the major source of flight delays, as seen in Figure 1, accounting for more delays than mechanical and crew-related delays ("air carrier delays") or weather. This type of delay is known as "propagated delay" since it propagates from one flight to another, and it is also present in the railroad and bus contexts.

Propagated delay can be addressed through a long flight buffer, which reduces the chance of a late-arriving aircraft, but another tool is the "ground buffer". This buffer is the difference between the scheduled ground time and the minimum feasible aircraft turnaround time. A long ground buffer can absorb a late arrival of the inbound aircraft, allowing the next flight to depart on time despite this disruption. The flight buffer for the outbound flight can also address delay propagation, allowing the flight to arrive on time even if it departs late. Lengthening this flight buffer is costly, however, and longer ground times are also costly since they require more gate space.[1]

The purpose of the present paper is to analyze an airline's choice of flight and ground buffers, both theoretically and empirically. Our theoretical model, which is more comprehensive than earlier simple models because of its treatment of propagated delay, contains just two flights for tractability. Thus, the airline chooses two flight buffers and one ground buffer,

[1] Our analysis ignores the possibility that ground times may depend on passenger scheduling preferences, which are not present in the model. For instance, airlines may schedule a later departure and accept longer aircraft ground time if the result is a departure closer to the passenger's preferred departure time.

taking passenger distuilities from lateness and earliness into account along with buffer costs. The analysis yields a number of insights. A principal result is that the flight buffer for flight 1 plays no role in mitigating delay propagation, which is instead handled by the ground buffer and flight 2's buffer. The apportionment of the mitigation roles between these two buffers depends on buffer costs, with the ground buffer sometimes (but not always) doing all the work in addressing delay propagation. The numerical examples mainly focus on how buffer magnitudes are affected by the variances of the random factors affecting the operating times of the two flights. Given the parallel to other transportation sectors, the findings of the paper extend beyond the airline industry.

The empirical work relies on US Department of Transportation data showing the daily operations of thousands of commercial aircraft. These data allow computation of flight and ground buffers over an aircraft's operating day, which are then related to exogenous explanatory factors suggested by the model. For example, one set of regressions relates the magnitude of the flight buffer to the standard deviation of operating times for a particular flight (computed across the months of the sample for that flight). The empirical work provides confirmation of some of the hypotheses suggested by the model while providing general insight into the determinants of flight and ground buffers.

The paper is related to several strands of previous work. Earlier papers in operations research, including Deshpande and Arikan (2012) and Arikan, Deshpande and Sohoni (2013), present simple models of the choice of flight buffers, noting the similarity to the classic newsvendor problem of Whitin (1955) (a flight's early (late) arrival is analogous to a vendor ending up with a surplus (shortage) of newspapers). Deshpande and Arikan (2012) consider US airlines and use the newsvendor approach to estimate the airlines' ratios of earliness to lateness costs. They show that flight buffer choices depend on carrier types, route market shares, and route characteristics. Arikan, Deshpande and Sohoni (2013) develop schedule robustness measures for airline networks and use them to show how US carriers use flight and ground buffers to absorb delay propagation. Zhang, Salant and Van Mieghem (2018) present an analysis related to those of Deshpande and Arikan (2012) and Arikan, Deshpande and Sohoni (2013). They show that the historical evolution of flight durations cannot explain increases in scheduled ground and flight times in the US and conclude that these increases instead have strategic motivations.

In the economics literature, Wang (2015) offers a different theoretical approach to the choice of ground buffers, relating this choice to the level of competition while providing empirical evidence of this link. In other empirical work by economists, Forbes Lederman and Yuan (2019) use a much larger dataset to confirm the earlier finding of Shumsky (1993) showing that US carriers have added schedule padding (buffer time) over the years to improve their on-time performance. Forbes, Lederman and Wither (2019) show that this effect is stronger when airlines are large enough for required reporting of on-time performance, a criterion that excludes many regional carriers.[2] See also Hao and Hansen (2014), Kang and Hansen (2017), and Sohoni, Lee and Klabjan (2011) for related studies. Other studies by economists, which are not directly linked to our work, show the connection between market structure (mainly competition) and on-time performance (see, for example, Mazzeo (2003) and Prince and Simon (2015)).

The plan of the paper is as follows. Section 2 analyzes choice of the flight buffer in an introductory model with just one flight, where delay propagation is not an issue, and section 3 analyzes the two-flight model. Section 4 presents numerical examples, while section 5 offers a model extension that incorporates connecting passengers. Section 6 presents the empirical work, and section 7 offers conclusions.

2. Buffer choice for a single flight

The analysis starts by considering the buffer choice for a situation where the airline operates only a single flight, denoted flight 1. After this analysis is complete, the focus turns to the case where two flights are operated. While delay propagation is not an issue in the single-flight case, it is a crucial factor when two flights are operated.

Flight 1 departs at time 0 and has an uncertain duration. If no outside influences affect the flight's operation, its duration is given by m_1. Outside influences such as weather, mechanical issues, and unanticipated congestion can cause the flight duration to differ from m_1, with the

[2] For analysis of the incentives for integration of regional and mainline carriers and its impacts, see Forbes and Lederman (2009, 2010).

actual duration equal to $m_1+\epsilon_1$, where ϵ_1 is a continuous random variable with support $[\underline{\epsilon}_1, \overline{\epsilon}_1]$. Bad weather would lead to a positive value for ϵ_1, while a strong tail wind would lead to a negative value. Note that while the congestion effect is partly random, appearing in ϵ_1, a persistently high level of airport congestion would lead to a large value for m_1, the expected flight duration.

In scheduling the flight's arrival time, the presence of this random term will lead the airline include a flight buffer, denoted b_1, which is added to m_1. The scheduled arrival time of the flight is thus $t_{a1} = m_1 + b_1$, whereas the actual arrival time, which captures the random effect, is $\widehat{t}_{a1} = m_1 + \epsilon_1$. The flight is late in arriving if $\widehat{t}_{a1} > t_{a1}$, or if $\epsilon_1 > b_1$, and it is early if $\epsilon_1 < b_1$. Passengers dislike being late or early, valuing a minute of late time by the amount x and valuing a minute of early time by the amount y. Therefore, if the flight is late, it generates disutility equal to $x(\epsilon_1 - b_1)$, whereas disutility is $y(b_1 - \epsilon_1)$ if the flight is early.

The probability that flight 1 is late is equal to the probability that ϵ_1 exceeds b_1, which is given by

$$\Pr(\widehat{t}_{a1} > t_{a1}) \;=\; \int_{b_1}^{\overline{\epsilon}_1} f_1(\epsilon_1)d\epsilon_1 \;=\; 1 - F_1(b_1), \tag{1}$$

where f_1 is the density and F_1 the cumulative distribution function of ϵ_1. Conversely, the probability of an early arrival is $F_1(b_1)$. The passenger's expected disutility from earliness or lateness is equal to

$$\Omega_1 \;=\; x\int_{b_1}^{\overline{\epsilon}_1}(\epsilon_1 - b_1)f_1(\epsilon_1)d\epsilon_1 \;+\; y\int_{\underline{\epsilon}_1}^{b_1}(b_1 - \epsilon_1)f_1(\epsilon_1)d\epsilon_1, \tag{2}$$

where the first integral captures lateness and the second captures earliness. Passenger disutility from late and early arrivals affects fare revenue, and assuming that the airline seeks to maximize profit, it takes into account this disutility.[3]

[3] Letting G denote the benefit from travel, the net benefit from a flight, denoted V, is equal to G minus the distutility expression in (2). Letting C denote expenditure on a nontravel good, overall passenger utility equals $U = C + V = Y - P + V$, where Y is income and P is the fare for travel. The airline sets the fare at a level that makes the consumer indifferent between traveling and not traveling, which yields utility Y (in this case $V - P = 0$). The fare for a flight is then equal to V, or G minus (2), and assuming that the passenger count is normalized to 1, fare revenue equals this expression.

But the airline also incurs operating costs from scheduled flight time and ground time. Scheduled flight costs include expenditures on fuel and crew salaries, while ground costs consist mainly of gate rental costs. To facilitate comparison with the two-flight model, suppose that the airline has leased the aircraft for a fixed period T, with scheduled flight time equal to m_1+b_1 and ground time equal to $T-(m_1+b_1)$. The leasing cost is fixed, but letting c_f denote the cost per minute of scheduled flight time and c_g denote the cost of ground time, total operating costs are $c_f(m_1+b_1)+c_g(T-(m_1+b_1))$. The goal of the profit-maximizing airline is to minimize the sum of this expression and Ω_1 by choice of b_1, which (ignoring constant terms) means minimizing Ω_1 from (2) plus $b_1(c_f-c_g)$.

The first-order condition for this minimization problem is

$$\begin{aligned}\frac{\partial\Omega}{\partial b_1} &= -[1-F_1(b_1)]x \;+\; F_1(b_1)y \;+\; c_f-c_g \\ &= (x+y)\left[F_1(b_1)-\frac{x}{x+y}\right] \;+\; c_f \;-\; c_g \;=\; 0, \end{aligned} \tag{3}$$

and the second-order condition, which requires $(x+y)f_1(b_1)>0$, is satisfied. Rearranging (3), the optimal buffer, denoted b_1^*, satisfies

$$F_1(b_1^*) \;=\; \frac{x+c_g-c_f}{x+y}, \tag{4}$$

a formula similar to one derived by Deshpande and Arikan (2012). Differentiation of (4) yields

$$\frac{\partial b_1^*}{\partial x} \;>\; 0, \quad \frac{\partial b_1^*}{\partial y} \;<\; 0, \quad \frac{\partial b_1^*}{\partial c_g} \;>\; 0, \quad \frac{\partial b_1^*}{\partial c_f} \;<\; 0. \tag{5}$$

with a greater disutility from lateness raising the buffer (thus reducing the chance of lateness) and a greater disutility from earliness reducing it. Similarly, a higher c_f (c_g) reduces (increases) b_1.

If the distribution of ϵ_1 is symmetric around zero, then $F_1(0)=1/2$. As a result, if $(x+c_g-c_f)/(x+y)=1/2$, then $b_1^*=0$. Rearranging, $b_1^*>(<)\;0$ will then hold as $(x-y)/2>(<)\;c_f-c_g$ in this symmetric case. If x is much larger than y, as seems realistic, and c_f is close to c_g, then b_1^* is positive. This conclusion makes sense given that a much greater disutility from a late arrival will lead the airline to lessen the chance of this outcome via a positive buffer.

3. The two-flight model

3.1. The setup

The airline is now assumed to operate two flights using the same aircraft, so that a delay for flight 1 can cause lateness of flight 2. While the existence of two flights introduces the possibility that some passengers connect from one flight to another, we assume initially that connections are absent, showing later in the paper how they affect the analysis.

Following the single-flight assumptions, m_2 denotes the undisrupted flight time for flight 2, with the actual flight time given by $m_2 + \epsilon_2$. The random term ϵ_2 has density f_2, cumulative distribution function F_2, and support $[\underline{\epsilon}_2, \overline{\epsilon}_2]$. The flight buffers are b_1 and b_2, so that the scheduled arrival times of flights 1 and 2 are $t_{a1} = m_1 + b_1$ and $t_{a2} = t_{d2} + m_2 + b_2$, where t_{d2} is flight 2's scheduled departure time. The scheduled aircraft ground time is denoted t_g, and the ground buffer, given by $b_g = t_g - \overline{t}_g$, is the excess of ground time over the minimum feasible aircraft turnaround time, denoted $\overline{t}_g$. The size of the ground buffer is thus set by choice of t_g. The scheduled departure time of flight 2 is then $t_{d2} = t_{a1} + t_g = m_1 + b_1 + t_g$, and the flight's scheduled arrival time is

$$t_{a2} \;=\; t_{d2} + m_2 + b_2 \;=\; t_{a1} + t_g + m_2 + b_2 \;=\; m_1 + b_1 + t_g + m_2 + b_2. \quad (6)$$

As before, the actual arrival time of flight 1 is $\widehat{t}_{a1} = m_1 + \epsilon_1$. The actual arrival time of flight 2 equals $\widehat{t}_{a2} = \widehat{t}_{d2} + m_2 + \epsilon_2$, where $\widehat{t}_{d2}$ is the actual departure time of flight 2. To find $\widehat{t}_{d2}$, note that if flight 1 is late in arriving, then the ground time will be reduced below the scheduled time t_g in an attempt to prevent delay propagation via late departure (and possible late arrival) of flight 2. However, ground time cannot be reduced below the minimum feasible time, equal to $\overline{t}_g$. Therefore, the actual departure time of flight 2 is given by

$$\widehat{t}_{d2} \;=\; \max\{m_1 + \epsilon_1 + \overline{t}_g,\ t_{d2}\} \;=\; \max\{m_1 + \epsilon_1 + \overline{t}_g,\ m_1 + b_1 + t_g\}. \quad (7)$$

Flight 2 departs on time if $\epsilon_1 + \overline{t}_g < b_1 + t_g$ or

$$\epsilon_1 \;<\; b_1 + t_g - \overline{t}_g \;=\; b_1 + b_g. \quad (8)$$

Note that satisfaction of (8) is ensured if flight 1 is early, or if $\epsilon_1 < b_1$. But the inequality can also be satisfied when flight 1 is late provided that it is not too late. Using (8), the probability of an on-time departure for flight 2 is $F_1(b_1 + b_g)$, with late departure (which reverses the inequality in (8)) having probability $1 - F_1(b_1 + b_g)$.

The formula for $\widehat{t}_{d2}$ allows derivation of conditions for late arrival of flight 2. Using (6), flight 2 is late in arriving when

$$\widehat{t}_{a2} \;=\; \widehat{t}_{d2} + m_2 + \epsilon_2 \;>\; t_{a2} \;=\; m_1 + b_1 + t_g + m_2 + b_2. \tag{9}$$

If flight 2 departs on time, so that $\widehat{t}_{d2} = m_1 + b_1 + t_g$ from (7), then (substituting in (9)), it arrives late when

$$\epsilon_2 \;>\; b_2, \tag{10}$$

an outcome that has probability $1 - F_2(b_2)$ (early arrival occurs when (10) is reversed and has probability $F_2(b_2)$). If flight 2 departs late, so that $\widehat{t}_{d2} = m_1 + \epsilon_1 + \overline{t}_g$, then (substituting in (9)), it arrives late when

$$\begin{aligned} \epsilon_1 + \epsilon_2 \;&>\; b_1 + b_2 + b_g, \quad \text{or} \\ \epsilon_2 \;&>\; b_1 + b_2 + b_g - \epsilon_1. \end{aligned} \tag{11}$$

With late departure, flight 2 arrives early when (11) is reversed.

Table 1 summarizes the preceding information, while showing flight 2's arrival delay. A key observation from the table is that when flight 2's departure is delayed, its arrival delay depends on the sum $b_1 + b_2 + b_g$. Therefore, the ground buffer affects delay by altering this sum, conditional on a late departure for flight 2. However, b_g also affects whether a late departure occurs via the direction of the inequality in the second column of the table. As will be seen, the overall impact of b_g on the airline's objective function operates through both these channels.

Using (11) and the reverse of inequality (8), the probability of late departure and late arrival for flight 2 is given by

$$\Pr(\widehat{t}_{a2} > t_{a2} \ \cap \ \widehat{t}_{d2} > t_{d2}) \quad = \quad \int_{\epsilon_1 = b_1 + b_g}^{\overline{\epsilon}_1} \int_{\epsilon_2 = b_1 + b_2 + b_g - \epsilon_1}^{\overline{\epsilon}_2} f_1(\epsilon_1) f_2(\epsilon_2) d\epsilon_2 d\epsilon_1. \tag{12}$$

Similarly, using the reverse of inequality (11) and the reverse of (8), the probability of late departure and early arrival for flight 2 is

$$\Pr(\widehat{t}_{a2} < t_{a2} \ \cap \ \widehat{t}_{d2} > t_{d2}) \quad = \quad \int_{\epsilon_1 = b_1 + b_g}^{\overline{\epsilon}_1} \int_{\epsilon_2 = \underline{\epsilon}_2}^{b_1 + b_2 + b_g - \epsilon_1} f_1(\epsilon_1) f_2(\epsilon_2) d\epsilon_2 d\epsilon_1. \tag{13}$$

Combining all this information, the probability of a late arrival for flight 2 is given by

$$\begin{aligned} &\Pr(\widehat{t}_{a2} > t_{a2}) \quad = \\ &\Pr(\widehat{t}_{a2} > t_{a2} \mid \widehat{t}_{d2} = t_{d2}) \ \Pr(\widehat{t}_{d2} = t_{d2}) \ + \ \Pr(\widehat{t}_{a2} > t_{a2} \ \cap \ \widehat{t}_{d2} > t_{d2}) \quad = \\ &[1 - F_2(b_2)] F_1(b_1 + b_g) \ + \ \int_{\epsilon_1 = b_1 + b_g}^{\overline{\epsilon}_1} \int_{\epsilon_2 = b_1 + b_2 + b_g - \epsilon_1}^{\overline{\epsilon}_2} f_1(\epsilon_1) f_2(\epsilon_2) d\epsilon_2 d\epsilon_1. \end{aligned} \tag{14}$$

Note that a conditional probability can be used in the first half of (14) because whether flight 2 is late conditional on an on-time departure (which depends only on ϵ_2) is independent of whether the departure is on-time (which depends only on ϵ_1). The absence of this kind of independence when flight 2 is late requires the different kind of expression in last half of (14).

Similarly, the probability of early arrival for flight 2 is given by

$$\begin{aligned} &\Pr(\widehat{t}_{a2} < t_{a2}) \quad = \\ &\Pr(\widehat{t}_{a2} < t_{a2} \mid \widehat{t}_{d2} = t_{d2}) \ \Pr(\widehat{t}_{d2} = t_{d2}) \ + \ \Pr(\widehat{t}_{a2} < t_{a2} \ \cap \ \widehat{t}_{d2} > t_{d2}) \quad = \\ &F_2(b_2)[1 - F_1(b_1 + b_g)] \ + \ \int_{\epsilon_1 = b_1 + b_g}^{\overline{\epsilon}_1} \int_{\epsilon_2 = \underline{\epsilon}_2}^{b_1 + b_2 + b_g - \epsilon_1} f_1(\epsilon_1) f_2(\epsilon_2) d\epsilon_2 d\epsilon_1. \end{aligned} \tag{15}$$

3.2. The airline's objective function

The disutility for passengers of flight 1 is again given by in Ω_1 in (2). The disutility from late arrival of flight 2 is given by

$$\Omega_{2,late} \;=\; xF_1(b_1+b_g)\int_{b_2}^{\bar{\epsilon}_2}(\epsilon_2-b_2)f_2(\epsilon_2)d\epsilon_2 \;+$$
$$x\int_{\epsilon_1=b_1+b_g}^{\bar{\epsilon}_1}\int_{\epsilon_2=b_1+b_2+b_g-\epsilon_1}^{\bar{\epsilon}_2}[\epsilon_1+\epsilon_2-(b_1+b_2+b_g)]f_1(\epsilon_1)f_2(\epsilon_2)d\epsilon_2 d\epsilon_1. \tag{16}$$

The first half of (16) captures disutility from a late arrival when flight 2 departs on time (note that $1-F_2$ in (14) is replaced by the integral). The rest of (16) captures late disutility when the departure is late (note that the bracketed term is added inside the integrand in (14)).

Similarly, the disutility from early arrival of flight 2 is

$$\Omega_{2,early} \;=\; yF_1(b_1+b_g)\int_{\underline{\epsilon}_2}^{b_2}(b_2-\epsilon_2)f_2(\epsilon_2)d\epsilon_2$$
$$+\; y\int_{\epsilon_1=b_1+b_g}^{\bar{\epsilon}_1}\int_{\epsilon_2=\underline{\epsilon}_2}^{b_1+b_2+b_g-\epsilon_1}[b_1+b_2+b_g-(\epsilon_1+\epsilon_2)]f_1(\epsilon_1)f_2(\epsilon_2)d\epsilon_2 d\epsilon_1. \tag{17}$$

The passenger-disutility portion of the airline's objective function, denoted by Ω, is given by the sum of (2), (16), and (17):

$$\Omega \;=\; \Omega_1 \;+\; \Omega_{2,late} \;+\; \Omega_{2,early}. \tag{18}$$

The flight buffer costs must be added to Ω along with the cost of ground time, with the airline seeking to minimize Ω plus $c_f b_1 + c_f b_2 + c_g b_g$. The derivatives of Ω with respect to b_g, b_1 and b_2 are computed in the appendix. Adding the relevant buffer cost to these derivatives and setting the resulting expressions equal to zero yields the following first-order conditions:

$$\frac{\partial\Omega}{\partial b_g} \;+\; c_g \;=\;$$

$$(x+y)\int_{\epsilon_1=b_1+b_g}^{\bar{\epsilon}_1}\left[F_2(b_1+b_2+b_g-\epsilon_1)) - \frac{x}{x+y}\right] f_1(\epsilon_1)d\epsilon_1 \;+\; c_g \;=\; 0 \qquad (19)$$

$$\frac{\partial\Omega}{\partial b_1} \;+\; c_f \;=\; \frac{\partial\Omega}{\partial b_g} \;+\; (x+y)\left[F_1(b_1) \;-\; \frac{x}{x+y}\right] \;+\; c_f \;=\; 0 \qquad (20)$$

$$\frac{\partial\Omega}{\partial b_2} \;+\; c_f \;=\; \frac{\partial\Omega}{\partial b_g} \;+\; (x+y)F_1(b_1+b_g)\left[F_2(b_2) \;-\; \frac{x}{x+y}\right] \;+\; c_f \;=\; 0. \qquad (21)$$

Note in (20) and (21) that $\partial\Omega/\partial b_1$ and $\partial\Omega/\partial b_2$ are equal to $\partial\Omega/\partial b_g$ plus the remaining terms in the relevant equation.

3.3. Characterizing the optimum

The immediate implication of the first-order conditions is that the flight buffer b_1 has the same value as in the single-flight model. This conclusion can seen by using (19) to substitute $-c_g$ in place of $\partial\Omega/\partial b_g$ in (20), which yields a condition that matches (3) from the single-flight model. With flight 1's buffer set as if the flight were operating in isolation, the buffer therefore plays no role in mitigating delay propagation. Thus,

Proposition 1. *Responsibility for mitigation of delay propagation falls only on the ground buffer and the flight buffer for flight 2.*

The lack of a delay-propagation role for flight 1's buffer makes sense. Even though b_1 and the ground buffer are, in effect, perfect substitutes in addressing delay propagation, use of the ground buffer instead of the flight buffer does not distort the balance between late and early disutilities for flight 1, making it the preferred tool. However, suppose the ground buffer were somehow constrained below its optimal value, due to a shortage of airport gate capacity, for example, which could be caused by hoarding of gates by a dominant airline (see Ciliberto and Williams (2010) for empirical evidence). In this case, b_1 would help to address delay propagation along with the other buffers. Such a constraint would make $\partial\Omega/\partial b_g$ in (20) less than $-c_g$, causing b_1 to rise above its single-flight value, thus addressing delay propagation.

To see how the delay-propagation responsibility is apportioned between flight 2's buffer and the ground buffer, it is useful to first consider the the unrealistic case where $c_g = 0$, with ground time being costless. When $c_g = 0$, $F_1(b_1 + b_g) = 1$ must hold at the optimum, so that the probability of late departure for flight 2 (which requires $\epsilon_1 > b_1 + b_g$) equals zero. Suppose to the contrary that $F_1(b_1 + b_g) < 1$ is satisfied along with the first-order conditions. Then, from (20), $F_2(b_2) - x/(x + y) = 0$ must hold. But since $F_2(b_1 + b_2 + b_g - \epsilon) \leq F_2(b_2)$ holds over the range of integration in (19), which is nonempty given the maintained assumption, it follows that

$$\begin{aligned} \frac{\partial \Omega}{\partial b_g} &< (x+y)\left[F_2(b_2) - \frac{x}{x+y}\right]\int_{\epsilon_1=b_1+b_g}^{\overline{\epsilon}_1} f_1(\epsilon_1)d\epsilon_1 \\ &= (x+y)\left[F_2(b_2) - \frac{x}{x+y}\right][1 - F_1(b_1+b_g)] = 0. \end{aligned} \tag{22}$$

This inequality contradicts the assumption that (19) equals zero, ruling out the premise that $F_1(b_1 + b_g) < 1$. Therefore, the ground buffer b_g is set at a value large enough to eliminate the chance of late departure for flight 2, so that $F_1(b_1 + b_g) = 1$.[4]

Then, setting $F_1(b_1 + b_g)$ in (21) equal to 1 and replacing $\partial\Omega/\partial b_g$ with $-c_g$, (21) matches the optimality condition (3) for the single-flight model, implying that the optimal b_2 equals the single-flight value. Thus, when $c_g = 0$, the flight buffer for flight 2 is set as if the flight were operating in isolation, with delay propagation not an issue. Since the same conclusion has already been established for flight 1, we can state

Proposition 2. *When ground time is costless, the ground buffer does all the work in mitigating delay propagation, fully eliminating it, with no contribution from flight 2's buffer.*

Now consider the realistic case where the cost of ground time is positive, with $c_g > 0$. Setting (19) equal to zero now implies that the integral must be negative, which means that

[4] It should be noted that $F_1(b_1 + b_g) = 1$ does not yield a unique solution for b_g given that any b_g value satisfying $b_1 + b_g \geq \overline{\epsilon}_1$ makes the equality true. However, replacing b_1 by b_1^*, the airline might be assumed to choose the smallest b_g satisfying the equality, so that $b_1^* + b_g = \overline{\epsilon}_1$, yielding a unique solution given by $b_g^* = \overline{\epsilon}_1 - b_1^*$.

the lower limit of integration cannot exceed the upper limit as before, leading to a zero value for the integral. In other words, $F(b_1 + b_g)$ must now be less than rather than equal to 1, indicating that there is a chance of late departure for flight 2.

In this case, the mitigation of delay propagation is apportioned between the ground buffer and flight 2's buffer, with the exact apportionment depending on the relationship between c_f and c_g. Suppose first that $c_f < c_g$, so that the cost of the flight buffer is less than that of the ground buffer. Letting ** denote optimal values in the two flight model, it follows from (20) that $F_2(b_2^{**}) - x/(x+y) > 0$ must hold. Since $F_1(b_1^{**} + b_g^{**}) < 1$, (21) and $F_2(b_2^{**}) - x/(x+y) > 0$ imply

$$
\begin{aligned}
0 &= (x+y)F_1(b_1^{**} + b_g^{**}) \left[F_2(b_2^{**}) \; - \; \frac{x}{x+y} \right] \; + \; c_f \; - \; c_g \\
&< (x+y) \left[F_2(b_2^{**}) \; - \; \frac{x}{x+y} \right] \; + \; c_f \; - \; c_g \qquad (23)
\end{aligned}
$$

(recall $b_1^{**} = b_1^*$). With the last line of (23) thus positive, it follows that b_2^{**} is larger than b_2^*, the single-flight value of b_2, which makes the second line equal to zero. Since flight 2's buffer is thus larger than the value it would assume if the flight were operating in isolation, it follows that flight 2's buffer assists the ground buffer in addressing delay propagation. This conclusion is due to the relative cheapness of the flight buffer, which encourages its use in addressing propagation.

In the reverse case where $c_f > c_g$, reversal of the above argument yields $b_2^{**} < b_2^*$, so that flight 2's buffer is *less than* its single-flight value. Now, relative cheapness of the ground buffer means that it takes extra responsibility in addressing delay propagation, causing flight 2's buffer to be reduced below its single-flight value. This adjustment means that the flight buffer now actually contributes to delay propagation, but this effect is offset by the greater role of the ground buffer. Finally, when $c_f = c_g$, flight 2's buffer equals its single-flight value ($b_2^{**} = b_2^*$), so that it neither helps nor offsets the ground buffer in mitigating delay propagation. With the buffers equally costly, the nondistorting ground buffer is thus set to do all the work in addressing delay propagation, although the chance of propagation is not reduced to zero given the costliness of the buffer. Summarizing yields

Proposition 3. *When $c_g > 0$, mitigation of delay propagation is apportioned between the ground buffer and flight 2's buffer. When $c_f = c_g$, the ground buffer alone addresses delay propagation (while not fully eliminating it), with flight 2's buffer set at its single-flight value. When $c_f < (>)\ c_g$ flight 2's buffer contributes to (partly offsets) the ground buffer's mitigation of delay propagation, taking a value above (below) its single-flight value.*

The results yield a further implication in the case where the distributions of the random flight-duration terms are equal, allowing the 1 and 2 subscripts to be removed from the F functions. With common F's, the single-flight values of b_1 and b_2 are the same. Then, the fact that b_2 is greater than (less than) the common single-flight value as $c_f < (>)\ c_g$ means that flight 2's buffer is greater than (less than) flight 1's buffer, which always equals the common single-flight value, when $c_f < (>)\ c_g$. In other words, $b_2^{**} > (<)\ b_1^{**}$ holds as $c_f < (>)\ c_g$. The reason is that flight 1's buffer plays no role in addressing delay propagation, while flight 2's buffer contributes to (borrows from) the ground buffer's delay-mitigation effect as $c_f < (>)\ c_g$.

The second-order conditions for the optimization problem have not been mentioned so far. In the appendix, it is shown that, that if $c_f \geq c_g$, then Ω is strictly convex in b_1 and b_2 at the optimum, so that conditional on b_g, flight buffers satisfying the first-order conditions based on (20) and (21) yield a local minimum of Ω. However, it is not possible to establish convexity of Ω in all three buffers, a condition that must be assumed to hold.

With b_1 equal to its single-flight value generally and b_2 taking this value in the $c_g = 0$ and $c_f = c_g$ cases, comparative statics for the buffers are then given by (5). Otherwise, comparative-static results are not available except in the case where $c_f = c_g$. Then, (20) and (21) yield $F_i(b_i^*) = x/(x+y),\ i = 1, 2$, so that the flight buffers are increasing in x and decreasing in y (a result that holds generally for b_1). Counterintuitively, however, the b_i^* values are completely independent of the common level of the buffer costs when $c_f = c_g$. It can be shown, though, that b_g^* is decreasing in the common buffer cost, so that the ground buffer falls when both buffer costs increase in parallel fashion. In addition, it can be shown that b_g^* is increasing in x and decreasing and y.[5]

[5] The first conclusion follows because (19) is increasing in b_g and $\partial^2\Omega/\partial b_g^2$ is easily shown to be positive when $c_f = c_g$. The second conclusion follows because (19) can be shown to be decreasing in x and increasing in y.

4. Numerical examples

Figures 2–5 present numerical examples for the case of costly buffers. Since some of the comparative-static effects of x and y and of c_f and c_g have been derived analytically, the analysis focuses on the effects of greater variability in the random terms ϵ_1 and ϵ_2. To generate the results, ϵ_1 and ϵ_2 are assumed to follow independent normal distributions with mean zero. Their standard deviations start out equal, satisfying $\sigma_1 = \sigma_2 = 0.0$. Then, each of the σ's increases up to 1.5 holding the other σ fixed, allowing the effect of greater flight-time variability for flights 1 and 2 to be appraised separately. Next, the σ's are set at a common value and increased simultaneously from 0.0 up to 1.5, allowing the effects of greater overall flight-time variability to be appraised.

Among the other parameters, the minimum turnaround time $\bar{t}_g$ is set at 0.5 hours, and the lateness and earliness disutilities are set at $x = 1.0$ and $y = 0.1$, with y realistically much smaller than x. The buffer costs are initially set at $c_f = 0.05$ and $c_g = 0.01$. For computational reasons, the infinite upper and lower limits of the normal distributions are replaced by 10 and -10 respectively. With σ's taking the values mentioned above, the probability that ϵ_1 or ϵ_2 lies outside this range is virtually zero, making this restriction inconsequential. It should be noted that, given the stylized nature of the model, realism in the choice of parameter values is not possible, and the qualitative (rather than quantitative) effects of parameters changes are of interest.

Figure 2 shows the effect of increasing σ_1 from 0.0 to 1.5 with σ_2 fixed at 0.5. As can be seen, the flight buffer b_1 rises as σ_1 increases, a natural finding, while the ground buffer b_g also rises. The buffer b_2 for flight 2 appears to be constant in the figure, but it increases very slightly with σ_1. The conclusion, therefore, is that b_1 and b_g alone do almost all the work in absorbing the greater chance of an arrival delay and subsequent delay propagation that follows from an increase in flight-time variability for flight 1.

Figure 3 shows the effect of increasing σ_2 from 0.0 to 1.5 with σ_1 fixed at 0.5. Now b_2 rises, while b_1 is constant (note that the b_1 solution from (25) is independent of σ_2). However, the ground buffer b_g is decreasing in σ_2, apparently because greater flight-time variability for flight 2 makes the ground buffer less effective at preventing a late arrival.[6]

Figure 4 shows the effect of simultaneously increasing σ_1 and σ_2 from 0.0 to 1.5. Both flight buffers naturally increase with the common σ value, and although the figure makes the buffers look equal in size, b_1 is slightly larger than b_2, as predicted. In addition, b_g is increasing in the common σ value. However, Figure 5 shows that the behavior of the ground buffer is reversed when $c_f = 0.08$ and $c_g = 0.03$, falling as the common σ value increases (matching the outcome in Figure 2). Figure 5 shows an additional point that does not arise in the other cases. In particular, b_g becomes negative as σ increases, showing that the ground time is set *below* the minimum feasible turnaround time $\bar{t}_g$. Nothing in the model prevents this outcome, which need not lead to late departure for flight 2 if flight 1's buffer is sufficiently large. In the data discussed below, however, the outcome is exceedingly rare, accounting for only 0.2% of the observations.

The implication is that the effect of flight-time variability on the ground buffer could be positive or negative depending on the magnitudes of the other parameters. If the ground buffer is sufficiently cheap compared to the flight buffers ($c_g = 0.01$ vs. $c_f = 0.05$), it is used along with the flight buffers to address the greater threat of delay propagation resulting from higher flight-time variability (Figure 4). But when the ground-buffer cost is larger as a proportion of c_f ($c_g = 0.03$ vs. $c_f = 0.08$), then the flight buffers partly supplant the ground buffer as the threat of delay propagation rises, with b_g falling (Figure 5).

5. Adding passenger connections

So far the analysis has suppressed the possibility that some passengers connect from flight 1 to flight 2. These passengers would travel from the origin city of flight 1 to flight 2's destination city, making a connecting trip in the absence of nonstop service between the two cities. This type of connecting travel, however, has little effect on the model. Assuming that a share α of passengers on both flights are traveling nonstop while $1 - \alpha$ are connecting, the airline's objective function would be altered in a straightforward way. In (18), Ω would be multiplied

[6] Alternatively, recall that, conditional on delay propagation, the second flight's arrival delay depends on the sum of buffer times $b_1 + b_2 + b_g$ (see Table 1). If the second flight's buffer increases because of an increase in σ_1, the result is an increase in the sum of buffer times. The ground buffer is then reduced to moderate the increase in this sum.

by α and then added to the term $(1-\alpha)(\Omega_{2,late} + \Omega_{2,early})$, which represents the late and early disutilities for connecting passengers, who care only about their arrival time at flight 2's destination. Adding the two expressions, the disutility portion of the objective function reduces to (18) with α multiplying Ω_1.

A more interesting and complex connecting scenario arises if two additional flights, again using a single aircraft, are added to the model. The turnaround city for these flights, which are denoted 1B and 2B, is the same as for the original flights, now denoted 1A and 2A. This common turnaround city can thus be viewed as a hub, which is a destination in its own right but also supports passenger connections. Connecting passengers now include those traveling from flight 1A's origin (O_A) to flight 2B's destination (D_B) as well as those traveling from O_B (flight 1B's origin) to D_A (flight 1A's destination).[7]

If flight 1A arrives after the departure of flight 2B, passengers traveling from O_A to D_B miss their connection, suffering disutility V, with same conclusion applying to passengers connecting from flight 1B to flight 2A. Note that since connecting passengers using flights 1A and 2A (or 1B and 2B) do not change planes, a missed connection is not possible for them.

The portion of the airline's objective function applying to nonstop trips and same-plane connecting trips is given by adding the Ω's for the A and B flights, with the previous α modification incorporated.[8] The part of the objective function that applies to the remaining connecting passengers makes use of the probability of a missed connection. For 1A-2B connections, this probability is $P_{AB} \equiv Pr(m_1 + \epsilon_{1A} > \widehat{t}_{d2A})$, using (7) and adding an A subscript, while for 1B-2A connections, the probability is given by the analogous expression P_{BA}. In the previous expression, $m_1 + \epsilon_{1A}$ is the arrival time of flight 1, and a missed connection occurs when it is greater than flight 2's departure time.

Using P_{AB} and P_{BA}, the remaining (disutility) part of the objective function is given by

---

[7] Symmetry holds, with corresponding flight distances equal and with flights 1A and 1B departing at a common time.

[8] The relevant expression is $\alpha\Omega_{1A} + \Omega_{2A,late} + \Omega_{2A,early} + \alpha\Omega_{1B} + \Omega_{2B,late} + \Omega_{2B,early}$. Note that the buffers inside the Ω expressions also acquire A and B subscripts, although their equilibrium values will be the same given symmetry.

$1 - \alpha$ times

$$P_{AB}V \, + \, (1 - P_{AB})(\Omega_{2A,late} \, + \, \Omega_{2A,early}) \, + \, P_{BA}V \, + \, (1 - P_{BA})(\Omega_{2B,late} \, + \, \Omega_{2B,early}). \quad (24)$$

Note that for connecting passengers who make, rather miss, their connection, the early-late disutility is the same as for nonstop passengers on either flight 2A or 2B.

The second and fourth (multiplicative) terms in (24) make the objective function considerably more complex than for a model with only two flights. Given the challenging nature of resulting analysis, persuing it is beyond the scope of the paper. Intuitively, however, avoidance of missed flight connections provides an additional reason beyond mitigation of delay propagation to increase the flight buffers for flights 1A and 1B as well as the corresponding ground buffers. Despite the absence of concrete conclusions beyond this simple intuition, it is interesting nevetheless to see the logic under which flight connections can be added to the model.[9]

6. Empirical Analysis

This section aims to test the predictions of the theoretical model using US data that tracks the flights of individual commercial aircraft. The data allow computation of the flight buffer for a particular flight, which is set equal to scheduled flight time minus the average flight time on the route. The data also facilitate computation of the ground buffer, which equals scheduled ground time minus the minimum observed ground time at the airport (details on both calculations are presented below).

One of the model's predictions, which comes from the simulation analysis, is that greater variability in flight time should lead to increases in both flight buffers, while the effect on the ground buffer is ambiguous. We start by measuring this variability indirectly via several sets of variables. The first set is a collection of month dummy variables, which control for weather

[9] A different model exploring missed connections would proceed as follows, returning to the two-flight framework but assuming the flights use different aircraft (so that delay propagation is absent). If flight 1 arrives after the departure of flight 2, then connecting passengers miss their connection, again generating disutility V. Now, there is no ground buffer per se, but the airline chooses flight 2's scheduled departure with an eye toward avoiding missed connections, assuming that a later departure imposes the same costs as a longer ground buffer.

conditions. The expectation is that weather variability and thus the variability of flight times are greatest in the winter months, when snowstorms often disrupt airport operations. In addition, flights that operate between congested endpoints may experience greater flight-time variability due to the randomness of congestion's impact. Since the same variability may arise in operations to or from hub airports, which are often congested, the regressions include congestion and hub measures. The flight buffer regressions have separate congestion and hub variables for the origin and destination, while the ground buffer regressions have single values for the "turnaround" airport (they also include a dummy for slot-controlled airports, which are prone to congestion). Under a second approach that is more explicity linked to the model, flight time variability is measured by the standard deviation of flight times, with the month dummies and other congestion proxies dropped from the regressions.

A second prediction is that the position of a flight in the day's flight sequence affects the flight buffer. When buffer costs are positive, flight 2 has a longer buffer than flight 1 when $c_f < c_g$, with the relationship reversed when $c_f > c_g$. To test for such a sequencing effect we include time-of-day departure variables (morning, evening, etc.) along with variables that measure a flight's exact position in the sequence of an aircraft's daily flights.

A third prediction (which comes from the $c_g = 0$ case) is that a constraint that keeps the ground buffer below its optimal value will lead to longer flight buffers, thus generating a negative correlation between the magnitudes of flight and ground buffers. This prediction is tested via simple correlation analysis.

A fourth prediction (which comes from the $c_g = 0$ and $c_f = c_g$ cases) is that routes where the lateness disutility x is large will tend to have long flight buffers.

6.1. Data collection and empirical variables

The sample is obtained from the U.S. Department of Transportation (DOT) and covers the US domestic airline market for the year 2018. We mainly rely on the 'Marketing Carrier On-Time Performance' dataset, which for each aircraft, uniquely identified by its tail number, contains information on the carrier operating the flight, the origin and destination, the departure date, the scheduled departure and arrival times, the actual departure and arrival times, and the taxi-in and taxi-out times.[10] We exclude flights that are canceled or diverted and flights from/to US Commonwealth areas and Territories.

The aim of our empirical analysis is to investigate the determinants of the two choice variables described in our theoretical model: the flight buffer and the ground buffer. The flight buffer is obtained as follows. First, for any flight i, where i defines the sequence of daily flights operated by a given aircraft during the day, we measure the actual flight time, which is the sum of the taxi-out, airborne and taxi-in times (equivalent to the so-called 'block-time'). Using the model notation, this actual flight time is given by $\widehat{m}_i = \widehat{t}_{ai} - \widehat{t}_{di}$, the difference between actual arrival and departure times. We then average $\widehat{m}_i$ across all aircraft of flight i's type flying the same route to obtain $\overline{m}$, the aircraft/route average flight time (the route subscript is suppressed for simplicity). The flight buffer is given by the difference between flight i's scheduled flight time m_i and the average actual flight time of the same aircraft type flying the same route: $b_i = m_i - \overline{m}$.

To calculate the ground buffer, we first compute the actual ground time that separates flight k and flight $k+1$ in the sequence of flights operated by the observed aircraft during the day: $\widehat{t}_{kg} = \widehat{t}_{d,k+1} - \widehat{t}_{ak}$. Then we calculate the minimum turnaround time as the shortest actual ground time observed across all aircraft of a given type using the turnaround airport. The ground buffer for an aircraft turnaround is obtained by subtracting this minimum feasible ground time from the scheduled ground time.

We initially restrict the analysis to those aircraft that only operate two flights a day, as in the model. This restriction predominantly limits the focus to coast-to-coast flights or flights to/from Alaska. However, we also present results when the sample is expanded to include aircraft flying more than two flights in a day.

Negative values of the flight and ground buffers are technically possible. For the ground buffer, negativity occurs in the rare cases when the shortest actual ground time is larger than

[10] The novelty of this dataset relative to the 'Reporting Carrier On-Time Performance' dataset, previously used in Forbes (2008) and other studies, is that it distinguishes whether the flight is operated by the reporting carrier or by its regional codeshare partners. Both datasets are downloadable at `www.transtats.bts.gov/Tables.asp?DB_ID=120&DB_Name=Airline%20On-Time%%20Performance%20Data&DB_Short_Name=On-Time`. However the Marketing Carrier On-Time Performance dataset only starts with January 2018, while the Reporting Carrier On-Time Performance dates back to 1987.

the scheduled ground time. Negative flight buffers are more frequent because they are obtained by subtracting an average value. In order to limit this phenomenon, we require that the number of observations used to obtain the flight and ground buffers must be at least equal to 30.

6.2. Empirical approach

Our empirical analysis is based on two sets of regressions: one for flight buffers and another for ground buffers. The standard errors are clustered by route and month in order to allow the residuals of different aircraft (possibly of different carriers) flying on the same route, during the same month, to be correlated.

The general equation to be estimated is:

$$Buffer_{jcodt} \;=\; X_{jcodt}\beta + \eta_{cod} + u_{jcodt} \tag{25}$$

where *Buffer* is either the flight buffer or the ground buffer. The subscript j identifies the aircraft tail number, c the carrier, o the origin airport, d the destination airport (or the turnaround airport in the ground-buffer regression), and t the time. The term η_{cod} represents an airline-airport fixed effect,[11] and u_{jcodt} is the regression error, assumed i.i.d. with zero mean. The independent variables are denoted by X, and they control for different aspects of the buffer decision.

These variables, which are described in Table 2, are mainly constructed from the 'Marketing Carrier On-Time Performance' dataset and can be broadly classified as weather, airport, airline and flight-specific variables. In addition to the month dummies, the congestion, hub, and slot-control variables, and the time-of-day and flight sequencing variables discussed above, Table 2 includes a number of other covariates. We include airline dummies, with American Airlines serving as the default carrier, as well as dummies for regional and low-cost carriers. Since we

[11] The airport *od* fixed effects are not only relevant in the flight-buffer regression, but also in the ground-buffer regression. Consider for example two flights that share the same turnaround airport and aircraft type, but that originate from different airports. These flights have the same minimum actual ground time, but they may have different load factors depending on the strength of demand and other forces. The potential difference in the load factor may affect the turnaround time and hence the ground buffer set by the airline. The inclusion of originating airport fixed effects (the origin airport of the previous flight) aims to control for this unobservable characteristic.

observe both the operating and marketing carrier (i.e., the ticketing carrier), we code flights according to the ticketing carrier but set the dummy variable indicating a regional carrier equal to 1 if the flight is operated by such a carrier under a major carrier's code.

Low-cost carriers are well-known for their faster turnarounds, which should be reflected in shorter ground buffers, and other carrier-specific differences may emerge in the estimation. The flight-buffer regressions also include the distance of the flight and the number of competitors on the route. Long buffers are expected for long-distance flights due possible amplification of factors such as weather that lead to flight-time variability on long flights, while the effect of competition is unclear a priori. Finally, both the flight- and ground-buffer regressions include a weekend dummy.

6.3 Flight-buffer regressions

The results of the flight-buffer regression are presented in Table 3. In order to match the setup of our theoretical model, the regressions in columns (1)–(4) are based on a sample of aircraft that operate only two flights a day (mainly flying coast-to-coast or to Alaska). We refer to this sample as 'two-flight sub-sample'. More specifically, column (1) shows the regression with the flight buffer of Flight 1 set as dependent variable, column (2) has Flight 2's flight buffer as dependent variable, and columns (3) and (4) pool the two flights, with (4) adding the dummy variable *Flight 2* to identify the second flight. The regressions in columns (5)–(7) remove the restriction to aircraft operating just two flights per day. The corresponding sample, which we refer to as the 'unrestricted sample', consists of aircraft that operate at most 8 flights in a day.[12] Other differences between the regression specifications are discussed below.

Where statistically significant, the monthly dummy coefficients have negative signs, indicating that compared to January, the reference month, flight buffers tend to be shorter in other months. These results point to a role for weather uncertainty in influencing the flight-buffer decision. The estimated coefficients suggest that the size of the effect varies with expectations of bad whether: the flight buffer falls monotonically from March/April until September, which

[12] The maximum number of flights operated by a single aircraft observed in the 'Marketing Carrier On-Time Performance' database is 16. However, since operation of 9 or more flights is seldom observed, representing less than 0.40% of the initial database, these cases are not likely to meet our criterion of at least 30 observations in the calculation of $\overline{m}_i$ and are therefore excluded from the unrestricted sample.

is a period of the year generally characterized by good weather conditions and therefore by less flight-time uncertainty. The flight buffer then starts increasing monotonically, while still remaining below the January value, until the end of the year. The statistically insignificant coefficients on *February* dummy and sometimes the *March* dummy mean that there is no difference in flight-buffer choices across the months of January, February and (partially) March. This result makes sense because January, February and March are winter months that bear the same bad weather expectations and therefore the same flight-time uncertainty.

In the two-flight sub-sample, the effect of weather uncertainty on the flight buffer decision quantifies to about 4 to 7 minutes of reduced flight-buffer length during Spring/Summer; see columns (3) and (4). In the unrestricted sample, this effect is still present, but with a slightly smaller impact, at about 5 minutes maximum.

The flight-buffer choice is generally different depending on whether the carrier flies from or to its hub, although the estimated coefficients are not always statistically significant. The positive effect of *Hub destination* may be due to the risk that the aircraft is forced to fly a holding pattern, not being permitted to land because of congestion at the hub. In other words, because of the greater chance of a holding pattern compared to a flight that is not hub-bound, extra minutes are added to the flight buffer of a hub-bound flight. Conversely, flights departing from a hub have shorter buffers because they are likely bound to an uncongested nonhub destination, as indicated by the negative sign on *Hub origin*.

When the flight is scheduled to depart or land during a busy time of day, identified by the number of aircraft departing and landing during the hour of the flight's takeoff or landing (variables *Congestion origin* and *Congestion destination*, respectively), airlines add extra minutes to the flight buffer, irrespective of the hub or non-hub status of the airport. In addition to the possibility of flying a holding pattern, a flight arriving during a congested period may need to wait for a gate once it has landed.[13]

The coefficients of the carrier dummies are quite similar across carriers and regressions. The only notable difference is the shorter flight buffers of Hawaiian Airlines relative to the other

[13] Recall that our measure of flight time is block time, which includes the taxi time along with the airborne time.

carriers. This flight-buffer similarity is expected because flying a route with the same aircraft in practice takes the same time irrespective of the carrier operating the aircraft. Different results will be observed, however, in the ground-buffer regression, where some carriers have notably quick scheduled turnarounds.

Regional airlines in the dataset schedule lower flight buffers relative to other carriers, indicating that these airlines may be keen to maximize aircraft utilization. In addition, while the negative coefficient for *Competitors* suggests an airline may want to increase the number of flights, thus seeking shorter flight buffers, when facing competitors, the estimated coefficient is insignificant in all but one column of the table.

Disruptions to flight time may be proportional to distance, and the regressions thus show longer buffers for longer flights. This effect, however, is quantitatively small, on average about 12 seconds per hundred miles flown. As for weekend flights, the buffers for such flights in the two-flight sub-sample are not statistically different from those of non-weekend flights, although we observe a positive weekend effect in the restricted sample. Because weekends account for more than one quarter of total flights, the higher weekend flight buffers might be a reaction to possible disruptions from higher traffic. However, the magnitude of the weekend effect is very small: the estimated coefficient shows that the buffer difference between weekend and non-weekend flights is only a few seconds.

As suggested by the model, extra minutes of flight buffer may be added to later flights in the day to mitigate delay propagation. We investigate the effects of a flight's position in the day's flight sequence using two alternate approaches. The first approach uses the time-of-day departure dummies, with early morning being the default period. The second method uses either the flight-sequence dummies or the *Aircraft rotation* variable.[14]

Turning to time-of-day and flight-sequencing effects, the regressions from two-flight sub-sample show that a flight operating later in the day (afternoon, late afternoon, evening) has a

[14] The time-of-day dummies are not used in conjunction with either of the other two variables because they are almost functionally related. For example, while inclusion of both the time-of-day dummies and flight-sequence dummies would imply that independent variation is possible for these covariates, the fact that, say, the eighth flight of the day could never be a morning flight means that such independence is not present. For the same reason, we do not include the time-of-day dummies in the regressions in columns (1) and (2), given that flights 1 and 2 are necessarily in the first and latter parts of the day, respectively.

longer flight buffer than earlier flights, with the difference as large as one minute. While these results are shown in column (3), column (4) replaces the time-of-day dummies with a *Flight 2* dummy. The positive and significant coefficient *Flight 2* recapitulates the results of column (3) by showing that the later of the two flights has a longer buffer. These results conform to some of the predictions of the model.

Shifting to the unrestricted sample, column (5) contains the time-of-day dummies, and the results partly match the later-flight premiums seen in column (3). Flights departing in the late afternoon have longer buffers than flights earlier in the day, although the buffer for evening flights is no different from the buffer for early-morning departures (the default period). Therefore, with more flights per day captured in the data, the time-of-day buffer pattern is somewhat more complex than in the two-flight sub-sample.

Column (6) shows results using flight-sequence dummies in place of the time-of-day dummies, and different implications emerge. With the exception of the negative *Flight 3* coefficient, the pattern is for flights 2 through 5 to have longer buffers than flight 1 (the default), while flights 7 and 8 have shorter buffers than flight 1. Because of its focus on just two flights, the model does not generate predictions for the cases covered by the unrestricted sample. But the empirical pattern appears to show a somewhat different logic than in the two-flight situation. In particular, delay propagation is apparently addressed through longer buffers for *earlier*, rather than later flights. The shortest buffers are for last few flights of the day, given that the aircraft will soon terminate its daily operations, removing any concern about delay propagation. Although this pattern is the reverse of the one seen in the two-flight case, it could well be optimal in a more complex model with a longer flight sequence.

The regression in column (7) shows that this buffer pattern also emerges when the flight dummy variables are replaced by the continous *Aircraft rotation* variable, which appears in quadratic form. The positive sign of the *Aircraft rotation* coefficient together with the negative coefficient on the quadratic term point towards an inverse U-shape relationship between the flight buffer and the rotation variable. In fact, the pattern of the parabola based on the estimated coefficients in column (7) is perfectly in line with the results of column (6), in which the set of flight dummies changes sign at *Flight 6*, as seen in Figure 6. The non-linear effect of aircraft rotations on the flight buffer appears to reconcile two opposite forces: the first force pushes towards longer flight buffers to lessen the risk of delay propagation; the second force pushes towards shorter flight buffers to maximize aircraft utilization.

6.4. Flight buffer regressions with flight-time variability

As explained above, an alternative approach to investigating the effects of flight-time variability is to replace the various proxy variables representing variability (the month dummies, for example) with actual variability measures. The measure we use is the standard deviation of the actual flight times, computed by route, month and flight number. To match the setup of the simulation, we only rely on the two-flight sub-sample.

In addition to the month dummies, we exclude the weekend dummy, distance, the airport congestion variables, and origin and destination airport fixed effects. Recalling that our flight-time measure is calculated as the sum of the airborne time and taxi time, it follows that airport congestion will be a determinant of flight-time variability and thus a candidate for exclusion.[15]

Table 4 reports the results under the new specification, using the same setup as the first part of Table 3. Note that, because the two-flight sample is used, we include two covariates to capture flight-time variability: one for the first flight and one for the second flight. These two variables have a very similar distribution, with mean values of 14.31 and 14.54 and standard deviations of 6.53 and 7.01, respectively for flight 1 and flight 2. Except for the insignificant coefficient in column (1), the variability coefficients are positive and significant in all four regressions. Thus, an increase in flight-time variability typically leads an airline to expand its flight buffers.

The estimates column (1) exactly match the simulation results in Figure 1, with flight 1's buffer increasing in its own variability but insensitive to flight 2's variability. This perfect match is absent in column (2), however, since increases in both variability measures raise flight 2's buffer, even though Figure 2 suggests that the buffer should not respond to flight 1's variability. Columns (3) and (4) show that, when the flights are pooled, their buffers

[15] If the month dummies are included in the flight buffer regression, their coefficients are statistically significant with a sign pattern similar to that in Table 3 and with the flight-time variability coefficients unaltered. This finding indicates that the month dummies have their own separate influence beyond the influence that operates through flight-time variability.

are increasing in each flight's variability, an outcome that could be viewed as matching the pattern in Figure 4, where variability is constrained to be equal across flights. Regardless of these nuances, the regressions results are broadly consistent with the hypothesis that greater flight-time variability leads to longer flight buffers. Note finally that the estimate coefficients of the remaining covariates do not show great differences relative to Table 3.

Table 5 shows how flight-time variability is related to the proxy variables that appeared in Table 3 but were deleted in Table 4. With t statistics five to ten times those of other covariates, distance has by far the greatest explanatory power, with longer flights naturally having greater time variability. The congestion coefficients are also mostly positive and significant, as expected, while weekend flights show low time variability. The month dummies somewhat surprisingly indicate that July and August have higher flight-time variability than the Winter months, an outcome that may be due to high summer travel volumes, which can lead to delay from various sources.[16] This conclusion is hard to reconcile with the weather-based interpretation of the month-dummy coefficients in Table 3, but it is possible that weather effects are not adequately captured by our variability measure.

6.5. Ground-buffer regressions

The ground buffer is calculated at a flight's departing airport. In the two-flight sub-sample, the observed airport corresponds to the departing airport of Flight 2; in the unrestricted sample, the observed airport is the departing airport of Flight i, with $i = 2, ..., 8$. Thus, the flight specific variables in the regression, such as departure time, refer to the following flight. In this way, the first flight of the day is not included.

The results for the ground-buffer regressions are presented in Table 6. As before, the table reports the estimates for the two-flight sub-sample, columns (1) to (3), and for the unrestricted sample, columns (4) to (6). A notable finding in Table 6 is that the choice of ground buffer is highly carrier-specific. Looking at column (1), the top-5 airlines with the shortest ground buffers are Allegiant Air, Spirit Airlines, Southwest Airlines, Hawaiian Airlines and Frontier

[16] It is worth pointing out that if flight-time variability is computed only by route and flight number, then the month dummies become statistically insignificant. This result is mainly technical: because the dependent variable is then invariant throughout the year, the month dummies do not explain much more than the constant of the regression would explain, making their coefficients insignificant.

Airlines. Four of them are low-cost carriers and Hawaiian Airlines, the sole legacy carrier on this list, has a partnership with Jet Blue, which is in 6^{th} place on the ranking.

The finding that low-cost carriers set shorter ground buffers than legacy carriers is also confirmed when we remove the airline fixed effects and include a dummy variable for low-cost carriers. The estimated coefficient of the *Low-cost carrier* variable in columns (2) and (5) shows that low-cost carriers have ground buffers 5 to 7 minutes shorter than those of legacy carriers on average.

Recall that we did not observe such a clear difference in the flight buffer regressions. The divergence, however, makes perfect sense: the flying time of the same aircraft on the same route does not depend on the carrier but instead depends on route characteristics. While there should thus be little difference in the flight buffers of low-cost and legacy carriers, the business model of low-cost carriers, which deemphasizes hub-and-spoke operations and seeks high aircraft utilization, yields faster turnarounds and thus shorter ground buffers. The results show that regional carriers also set short ground buffers, a finding that seems at variance with their role in providing flight connections to major carriers.

The effect of weather uncertainty on ground buffers is not as clear as in the flight-buffer regressions, a finding that is consistent with the simulation results. The results show that in both samples, ground buffers are longest (relative to January) in the Fall months of September, October and November, a pattern that does not have a clear weather-based interpretation. However, in the unrestricted sample, the buffers over the March-August period are significantly shorter than in January, which is partly consistent with short buffers being scheduled in months with better weather.

When the flight departs from the carrier's hub, 10 to 14 minutes are added to the ground buffer. This effect is quantitatively large (the average ground buffer is around 39 minutes) and highly statistically significant across all the columns of the table. The desire to accommodate flight connections at hub airports is clearly the source of these longer buffers.

A congested turnaround airport has longer taxi times, which may prompt the airline to cut ground time. While this expectation is confirmed by the negative coefficients on *Congestion turnaround*, the effect is very small in magnitude: if an extra flight lands/departs when the observed flight lands, the airline reduces the ground buffer by about 5 seconds.[17]

However, when a dummy indicating a slot controlled airport is added to the regression, the congestion coefficient becomes insignicant in the two-flight sub-sample and significantly positive in the unrestricted sample. While this latter coefficient is still very small, the slot-control effect in the two-flight subsample is large: ground buffers are 4 minutes shorter on average at such airports, which tend to be congested.[18]

Time of the day plays a significant role in the choice of ground buffers. As seen from the estimated coefficients on the time-of-day departure dummies, airlines keep adding to ground buffers during the day. This effect is strictly monotonic in each regression, rising from 2-7 minutes extra buffer in the morning (relative to early morning) to 4-19 minutes extra buffer in the evening. With the model portraying operation of only two flights, it cannot predict this pattern. But the pattern suggests that, when an aircraft operates many flights per day, ground buffers may play a more prominent role than flight buffers in mitigating delay propagation late in the day.

Airlines may operate with spare capacity during weekends, since most business travel is mid-week, thus explaining why they set slightly longer ground buffers during weekends, about two minutes longer on average. Our data show that fights scheduled during weekends are more punctual than non-weekend fights, and these longer ground buffers may be part of the reason.[19]

6.6. Ground buffer regressions with flight-time variability

When substituting the variability measures, we match the approach for the flight-buffer regressions by excluding those covariates that can directly influence flight-time variability: month dummies, weekend dummy, congestion variables, airport fixed effects. The results are presented in Table 7.

Since the results are somewhat sensitive to how variability is computed, we use two approaches. In columns (1) and (2), variability is the standard deviation of flight times computed

[17] Results are obtained taking the max and the min estimated coefficients of *Congestion turnaround* and then multiplying by 60, re-scaling the marginal effect of one minute into seconds.

[18] In order to obtain these estimates, we remove the turnaround-airport fixed effects, which are perfectly collinear with the *Slot-controlled airport* variable.

[19] While the average arrival delay of a weekend flight is 6 minutes in the unrestricted sample and 0 minutes in the two-flight sub-sample, the delays increase to 10 and 3 minutes, respectively, for non-weekend flights.

by computed by route, month and flight number, as before, whereas variability is computed by route and flight number in columns (3) and (4).

Both flight-time variability variables have significantly negative effects in columns (1) and (2), results that match to some extent the pattern in Figure 5 from the simulation, where common flight-time variability is assumed. But in columns (3) and (4), the sign pattern exactly matches Figures 2 and 3, with the ground buffer rising with an increase in flight 1's variability and falling with an increase in flight 2's variability. The results are thus somewhat mixed, but they are at least partly able to replicate the simulations. In addition, the other covariates behave as in Table 7.

6.7. Additional results

The empirical analysis above has analyzed the flight and ground buffers separately. That is, we have run separate reduced-form regressions for both types of buffers to indirectly test predictions of our theoretical model, with some success. However, the analysis so far has not investigated the model's prediction of a possible negative correlation between the flight and ground buffers. This prediction arose in the zero buffer-cost analysis via possible constraints on the ground buffer, which would reduce it below the optimal value and lead to increases in both flight buffers, creating an inverse association.

This prediction is difficult to test in a regression framework that puts one buffer (flight or ground) on the left-hand side and the other on the right-hand side, looking for a negative coefficient. The reason is that, since the choices of the flight and ground buffers are simultaneous, the right-hand-side buffer must be treated as endogenous. While this endogeneity rules out use of an ordinary least squares regression, an instrumental variable approach using two-stage least squares (2SLS) could be considered. But given that the buffers are generated within the same airline choice problem, thus being dependent on the same set of exogenous factors, such an instrumental variable (which affects one buffer but not the other) may not even exist.

The impossibility of 2SLS, however, does not preclude a correlation analysis, which is not intended to capture causal effects. Accordingly, Table 8 reports the correlation coefficients between the ground buffer and the flight buffer for the following flight. The correlations are almost universally negative, confirming the model's prediction. While the overall correlations are not large in the two-flight sub-sample, restricting attention to ground buffers smaller than either 60 or 30 minutes increases the correlations substantially in absolute value.[20] The same pattern is not present in the unrestricted sample, where the overall correlations are similar in magnitude to the ones under these restrictions.

An additional hypothesis, mentioned above, is that routes with high values of the delay disutility x should have long flight buffers. Viewing routes with high shares of business passengers as satisfying this criterion, we included the managerial share of the origin city's workforce as an explanatory variable in the flight-buffer regressions, following Luttmann (2019). However, this variable did not perform successfully, leaving us without confirmation of the delay-disutility hypothesis.

Another possibility that our empirical analysis has not considered so far is that past on-time performance may affect buffer decisions, since it is very natural that carriers would change their buffers to remedy past late arrivals. To capture this possibility, we included in the flight-buffer regressions a variable measuring the past on-time performance of the airline on the route. The variable is constructed as follows: for year 2017 we downloaded from DOT the 'Reporting Carrier On-Time Performance (1987-present)' dataset, and then, by carrier-route-month, we calculated the proportion of delayed flights.[21] The coefficient on this *Past-year delay* variable is positive and statistically significant, confirming the presumption that carriers lengthen flight buffers if delays on the route in the past year were more frequent. The results are reported in the appendix Table A1. Finally and importantly, the inclusion of *Past-year delay* does not alter qualitatively or quantitatively the estimated coefficients of the other covariates.

7. Conclusion

This paper has presented an extensive theoretical and empirical analysis of the choice of schedule buffers by airlines. With airline delays a continuing problem around the world, such

[20] Concentrating on shorter ground times is useful because the trade-off between flight and ground times is more likely to be evident when the latter are short.

[21] The 'Reporting Carrier On-Time Performance (1987-present)' dataset is essentially the same as the 'Marketing Carrier On-Time Performance (Beginning January 2018)', which we use in our empirical analysis, but it does not report the actual operating carrier. Thus, it does not distinguish between major airlines and affiliated regional carriers. For this reason and also because an airline could stop serving the route from one year to the next, some 2018 observations are not matched with the 2017 data.

an undertaking is valuable, and its lessons extend to other transport sectors such as rail and intercity bus service. One useful lesson from the theoretical analysis of a two-flight model is that the mitigation of delay propagation is done entirely by the ground buffer and the second flight's buffer. The first flight's buffer plays no role because the ground buffer is a perfect, while nondistorting, substitute. In addition, the apportionment of mitigation responsibility between the ground buffer and the flight buffer of flight 2 is shown to depend on the relationship between the costs of ground- and flight-buffer time.

The empirical results show the connection between buffer magnitudes and a host of variables, including the month of operation and distance of a flight, whether the flight operates early or late in the day, and congestion measures at the endpoints. In addition, one version of the empirical model relates buffer magnitudes to the variability of flight times, which simulations of the model identify as an important determining factor (the results show that high variability lengthens flight buffers).

Fruitful extensions to this work would most likely lie in the theoretical area. The model could be extended to include additional sequential flights, and the sketch of connecting traffic provided in the paper could be expanded into a full analysis. In addition, passenger scheduling preferences could be introduced, with a buffer-related extension of scheduled flight times possibly becoming less desirable if it leads to a divergence between a passenger's preferred and actual arrival times. The resulting models would be complex, but additional insights could be generated, with relevance extending beyond the airline industry.

Another avenue of exploration could be in the area of market structure. For example, if a profit-maximizing airport sets the gate rental cost at an excessive level, the resulting decrease in ground buffers will impair airline mitigation of delay propagation, with negative effects on passengers. Alternatively, the entry barrier of gate shortages resulting from airport dominance by a large carrier (analyzed by Ciliberto and Williams (2010)) would also affect the ability of smaller carriers to address delay propagation via adequate ground time. Exploration of these issues could be illuminating.

Appendix

A1. The derivatives of Ω

This appendix section computes the derivatives of the objective function for the continuous case with respect to b_g, b_1, and b_2. Since (2) does not involve b_g, the objective function's derivative with respect to b_g is found by differentiating the sum of (16) and (17). The derivative of the first line of (16) with respect to b_g equals

$$x f_1(b_1 + b_g) \int_{b_2}^{\bar{\epsilon}_2} (\epsilon_2 - b_2) f_2(\epsilon_2) d\epsilon_2. \qquad (a1)$$

The second line of (16) can be written as

$$x \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} \left[\int_{\epsilon_2 = b_1 + b_2 + b_g - \epsilon_1}^{\bar{\epsilon}_2} [\epsilon_1 + \epsilon_2 - (b_1 + b_2 + b_g)] f_2(\epsilon_2) d\epsilon_2 \right] f_1(\epsilon_1) d\epsilon_1$$

$$= \quad x \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} Q(\epsilon_1, b_g) f_1(\epsilon_1) d\epsilon_1, \qquad (a2)$$

where $Q(\epsilon_1, b_g)$ denotes the term in brackets in the first line of $(a2)$. The b_g-derivative of $(a2)$ is

$$x \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} \frac{\partial Q(\epsilon_1, b_g)}{\partial b_g} f_1(\epsilon_1) d\epsilon_1 \; - \; x Q(b_1 + b_g, b_g) f_1(b_1 + b_g). \qquad (a3)$$

Substituting $\epsilon_1 = b_1 + b_g$ in the bracketed term in $(a2)$ to evaluate $Q(b_1 + b_g, b_g)$ in $(a3)$, the second term in $(a3)$ equals the negative of $(a1)$, so that these terms cancel. The b_g-derivative of (16) is then equal to the first term in $(a3)$.

$\partial Q / \partial b_g$ consists of two components, the first of which comes from differentiating the bracketed expression in $(a2)$ with respect to the limit of integration, a derivative that equals zero upon substituting the limit into the integrand. The second component comes from differentiating with respect to b_g under the integral, which yields the bracketed expression in $(a2)$ with the integrand replaced by $-f_2(\epsilon_2)$. Therefore, the first-term in $(a3)$ reduces to

$$-x \int_{\epsilon_1=b_1+b_g}^{\overline{\epsilon}_1} \left[\int_{\epsilon_2=b_1+b_2+b_g-\epsilon_1}^{\overline{\epsilon}_2} f_2(\epsilon_2) d\epsilon_2 \right] f_1(\epsilon_1) d\epsilon_1. \qquad (a4)$$

Applying the same steps to (17), the b_g-derivative of that expression equals

$$y \int_{\epsilon_1=b_1+b_g}^{\overline{\epsilon}_1} \left[\int_{\epsilon_2=\underline{\epsilon}_2}^{b_1+b_2+b_g-\epsilon_1} f_2(\epsilon_2) d\epsilon_2 \right] f_1(\epsilon_1) d\epsilon_1. \qquad (a5)$$

Replacing the bracketed terms in $(a4)$ and $(a5)$ with $1 - F_2(b_1 + b_2 + b_g - \epsilon_1)$ and $F_2(b_1 + b_2 + b_g - \epsilon_1)$, respectively, the sum of $(a4)$ and $(a5)$ can be written as

$$\begin{aligned} &= \int_{\epsilon_1=b_1+b_g}^{\overline{\epsilon}_1} [-x(1 - F_2(b_1 + b_2 + b_g - \epsilon_1)) + yF_2(b_1 + b_2 + b_g - \epsilon_1)] f_1(\epsilon_1) d\epsilon_1 \\ &= (x + y) \int_{\epsilon_1=b_1+b_g}^{\overline{\epsilon}_1} \left[F_2(b_1 + b_2 + b_g - \epsilon_1)) - \frac{x}{x + y} \right] f_1(\epsilon_1) d\epsilon_1 \quad = \quad \frac{\partial \Omega}{\partial b_g}. \qquad (a6) \end{aligned}$$

The derivative of the objective function with respect to b_1 builds on the previous results. The b_1-derivatives of (16) and (17) are identical to the b_g derivatives, given by $(a4)$ and $(a5)$, with their sum equal to $(a6)$. Since (2) also depends on b_1, the b_1-derivative of the objective function is then the expression in (3) (slightly rearranged) plus $(a6)$:

$$\frac{\partial \Omega}{\partial b_1} \quad = \quad (x + y) \left[F_1(b_1) \quad - \quad \frac{x}{x + y} \right] \quad + \quad \frac{\partial \Omega}{\partial b_g}. \qquad (a7)$$

Using similar steps,

$$\frac{\partial \Omega}{\partial b_2} \quad = \quad (x + y) F_1(b_1 + b_g) \left[F_2(b_2) \quad - \quad \frac{x}{x + y} \right] \quad + \quad \frac{\partial \Omega}{\partial b_g}. \qquad (a8)$$

A2. Convexity of Ω

This appendix section first shows that the objective function Ω is strictly convex in b_1 and b_2, with b_g fixed at the optimal value, doing so for both cases of zero and positive buffer costs. The first step is to compute the second derivatives of Ω with respect to b_1 and b_2, assuming zero buffer costs. Using the shorthand Ω_{ij} for $\partial^2\Omega/\partial b_i \partial b_j$, differentiation of (20) and (21) yields (after suppressing the multiplicative factor $x + y$))

$$\Omega_{11} = f_1(b_1) - f_1(b_1 + b_g)\left[F_2(b_2) - \frac{x}{x+y}\right] + \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} f_1(\epsilon_1) f_2(b_1 + b_2 + b_g - \epsilon_1) d\epsilon_1 \qquad (a9)$$

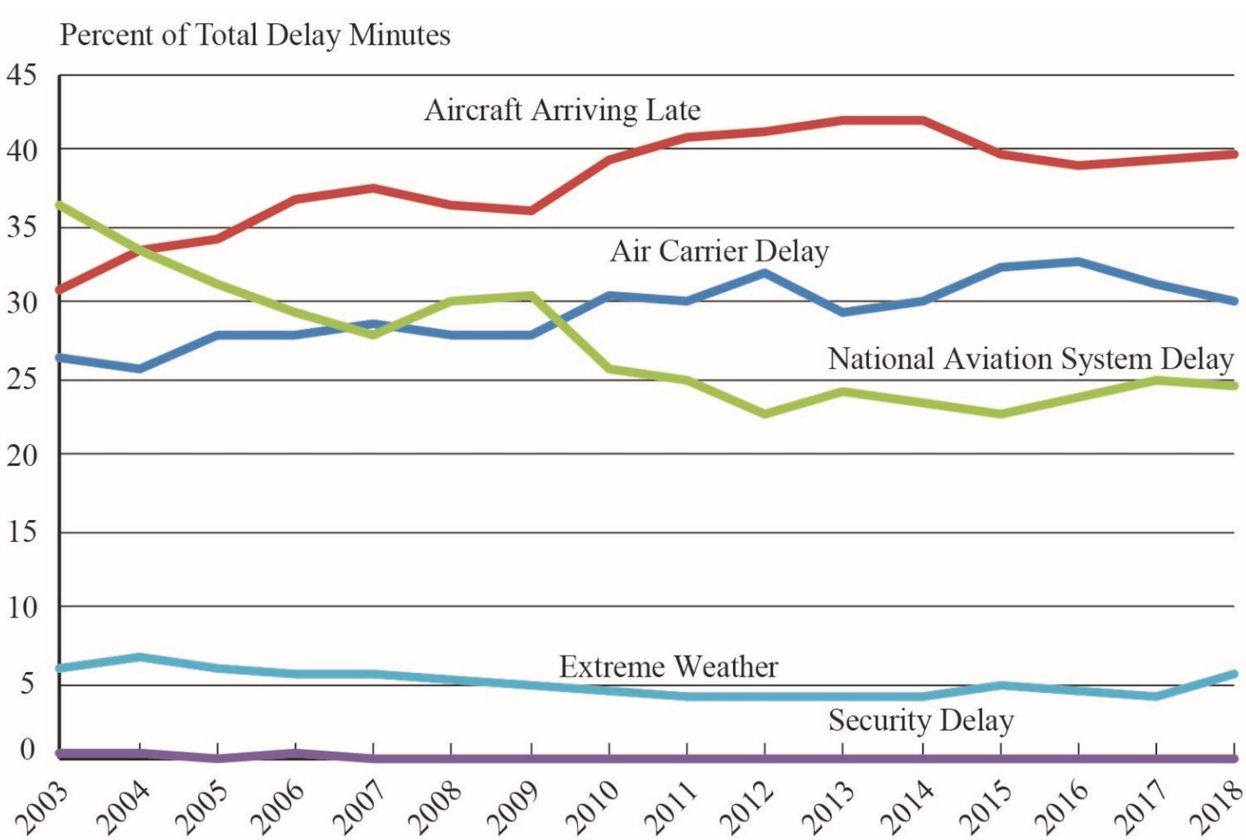
$$\Omega_{22} = F_1(b_1 + b_g) f_2(b_2) + \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} f_1(\epsilon_1) f_2(b_1 + b_2 + b_g - \epsilon_1) d\epsilon_1 \qquad (a10)$$

$$\Omega_{12} = \int_{\epsilon_1 = b_1 + b_g}^{\bar{\epsilon}_1} f_1(\epsilon_1) f_2(b_1 + b_2 + b_g - \epsilon_1) d\epsilon_1 < \Omega_{11},\ \Omega_{22}. \qquad (a11)$$

Observe that the same expressions apply when buffer costs are positive since they vanish in computing the second derivatives.

Ω_{22} and Ω_{12} are positive, while the sign of Ω_{11} in $(a9)$ is unclear. However, at the optimum in the zero-cost buffer case, the bracketed term in $(a9)$ is zero, making Ω_{11} positive and also ensuring satisfaction of the inequality in (11). The same conclusion holds in the positive-cost buffer case if $c_f > c_g$, in which case the bracketed term in $(a9)$ is negative at the optimum. With Ω_{11}, $\Omega_{22} > 0$ and $H \equiv \Omega_{11}\Omega_{22} - \Omega_{12}^2 > 0$ (a consequence of the inequalities in $(a11)$), Ω is thus strictly convex in b_1 and b_2 at the optimum. As a result, b_1 and b_2 values satisfying the first-order conditions yield a local minimum for Ω, holding b_g fixed at its optimal value. Convexity of Ω in all three buffers cannot be established analytically and must be assumed.

Figure 1: Incidence of Delay Propagation



Source: US Department of Transportation

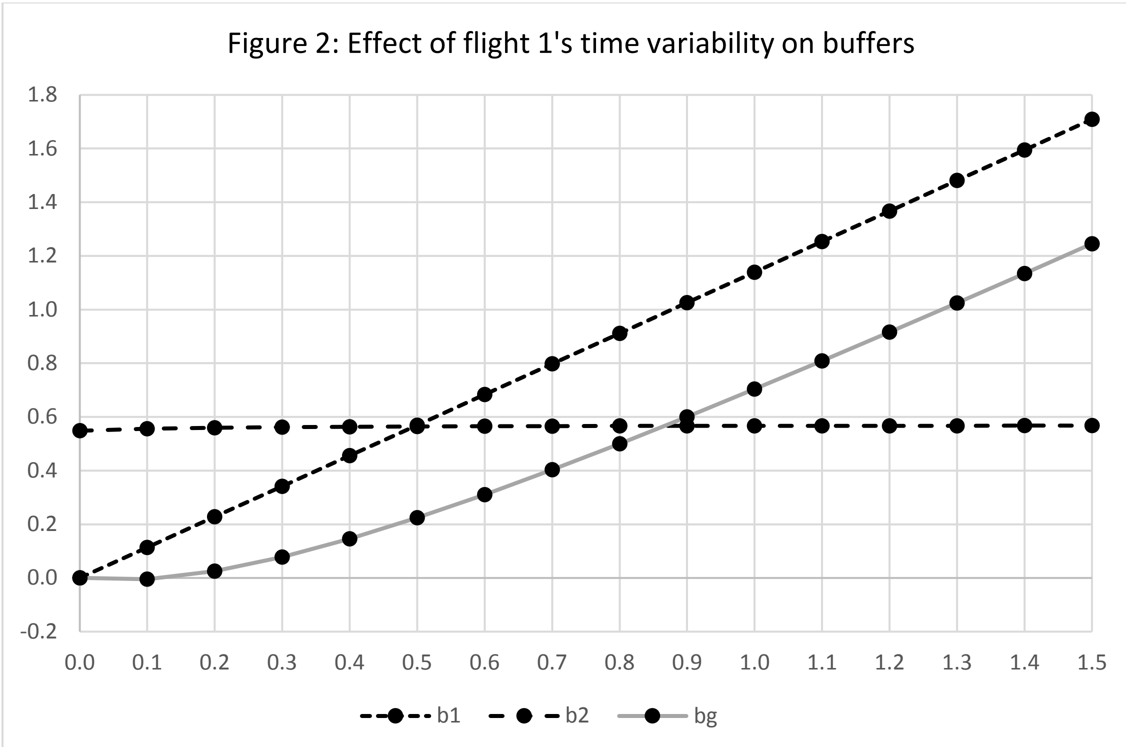


Figure 2: Effect of flight 1's time variability on buffers
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5
b1
b2
bg

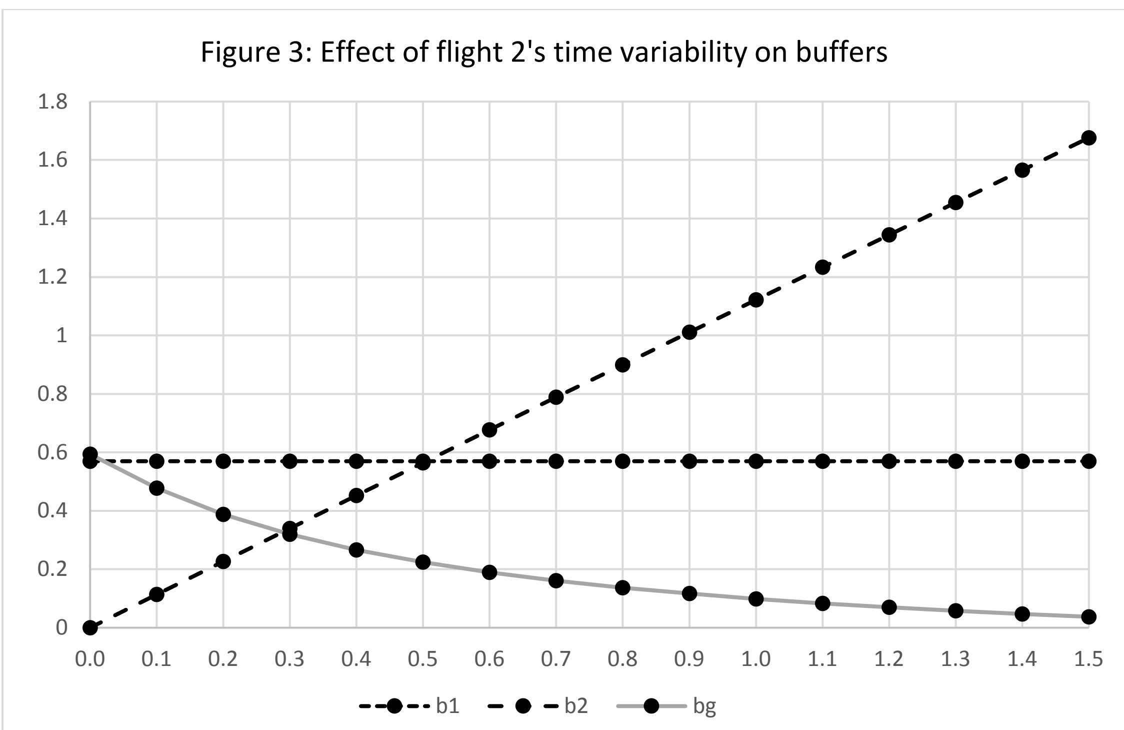


Figure 3: Effect of flight 2's time variability on buffers
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5
b1
b2
bg

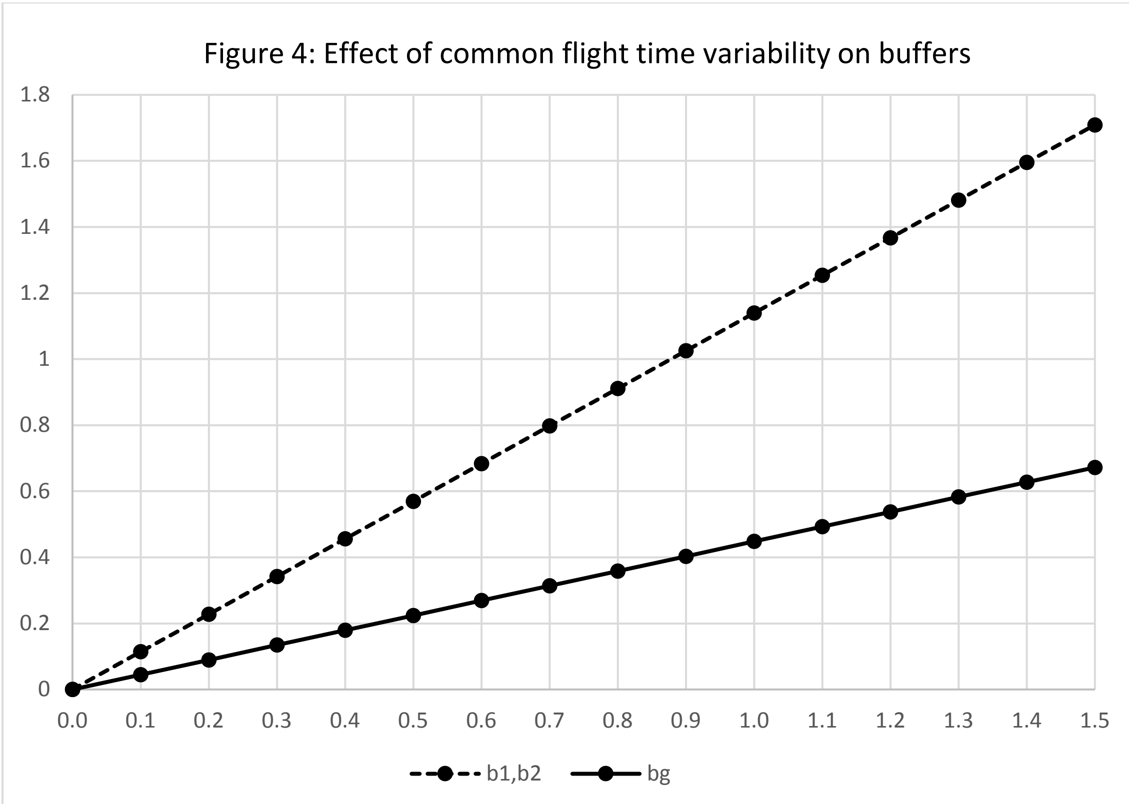


Figure 4: Effect of common flight time variability on buffers
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
b1,b2
bg

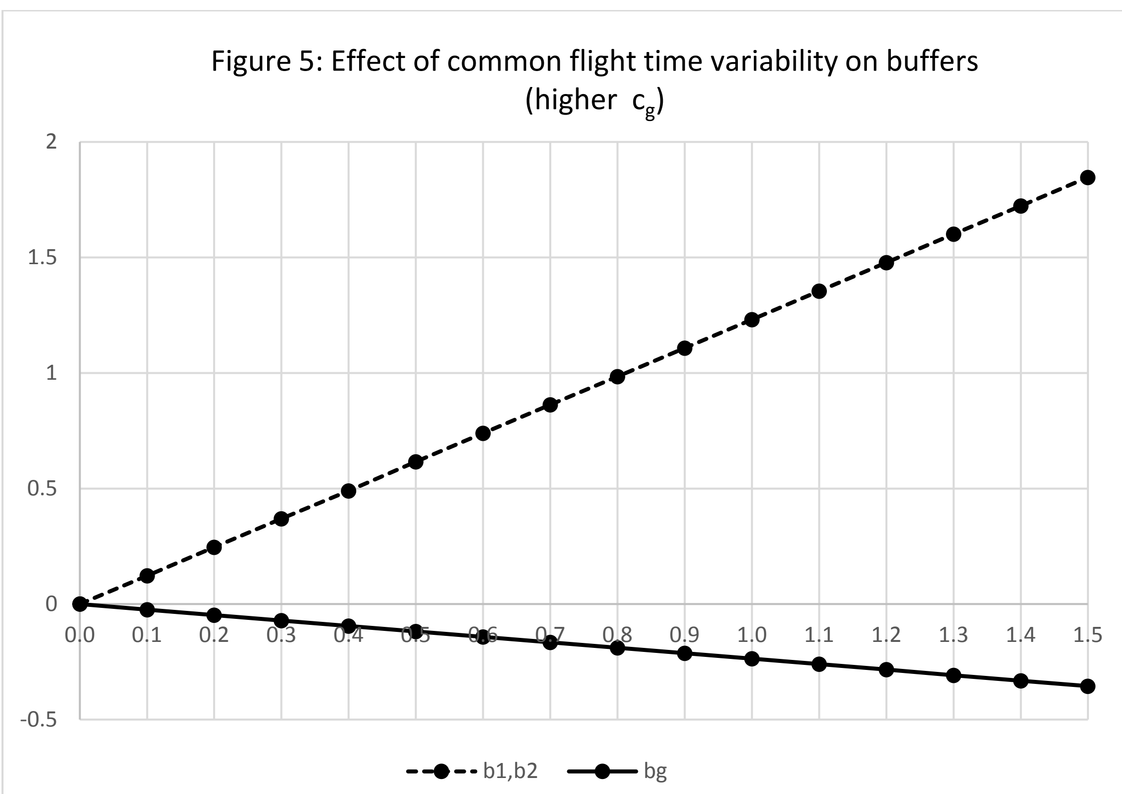


Figure 5: Effect of common flight time variability on buffers
(higher c_g)
2
1.5
1
0.5
0
-0.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
b1,b2
bg

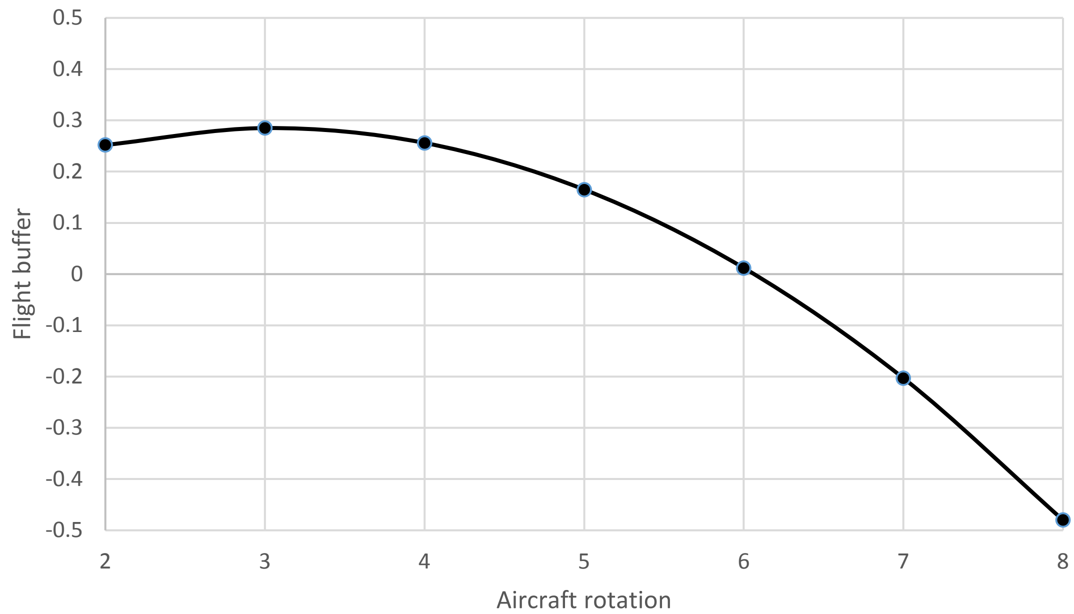


Figure 6: Flight buffer as a function of aircraft rotation

Table 1: Flight 2's arrival time

Flight 2's departure	*Occurs when*	*Flight 2 late (early)*	Arrival delay	*Delay propagation?*
On time	$\epsilon_1 \leq b_1 + b_g$	as $\epsilon_2 > (\leq)\ b_2$	$\max\{0, \epsilon_2 - b_2\}$	NO
Delayed	$\epsilon_1 > b_1 + b_g$	as $\epsilon_2 > (\leq)\ b_1 + b_2 + b_g - \epsilon_1$	$\max\{0, \epsilon_1 + \epsilon_2 - (b_1 + b_2 + b_g)\}$	YES if late

Table 2: Description and main statistics of the empirical variables

b	b_g	Variables	Description	Two-flight sub-sample	Unrestricted sample
✓		Flight buffer (b)	Difference between the scheduled flight time and the average actual flight time, computed by route and aircraft type	6.14 (8.81)	5.01 (6.86)
	✓	Ground buffer (b_g)	Difference between the scheduled ground time and minimum feasible ground time, computed by turnaround airport and aircraft type	40.30 (30.46)	37.88 (25.96)
✓	✓	February-December	Set of monthly dummy variables, *January* is the omitted month		
✓		Hub origin	Dummy variable = 1 if airport of origin is the hub of the airline	0.51 (0.50)	0.36 (0.48)
✓		Hub destination	Dummy variable = 1 if airport of destination is the hub of the airline	0.51 (0.50)	0.36 (0.48)
	✓	Hub turnaround	Dummy variable = 1 if airport of turnaround is the hub of the airline	0.45 (0.49)	0.38 (0.49)
✓		Congestion origin	Number of landing and departing flights at the airport of origin on the same hour when the flight is scheduled to depart	53.08 (33.89)	48.73 (40.14)
✓		Congestion destination	Number of landing/departing flights at the airport of destination on the same hour when the flight is scheduled to land	49.72 (33.13)	47.71 (40.32)
	✓	Congestion turnaround	Number of landing and departing flights at the airport of turnaround on the same hour when incoming flight is scheduled to land	52.80 (41.31)	52.70 (41.31)
	✓	Slot controlled airport	Dummy variable = 1 if airport of turnaround is slot controlled (DCA, JFC and LGA)	0.07 (0.23)	0.05 (0.22)
✓	✓	Alaska-Southwest	Set of airline dummy variables, *American Airlines* is the omitted airline		
✓	✓	Regional carrier	Dummy variable = 1 if the flight is operated by a regional carrier	0.07 (0.26)	0 .31 (0.47)
	✓	Low-cost carrier	Dummy variable = 1 if the flight is operated by a low-cost carrier (i.e. Jet Blue, Frontier Airlines, Allegiant Air, Spirit Airlines and Southwest Airlines)	0.16 (0.46)	0.29 (0.45)
✓		Competitors	Number of competitors on the route	1.58 (1.21)	1.06 (1.11)
✓		Distance	Route distance, in 100-mile units	16.94 (8.66)	8.21 (5.92)
✓	✓	Morning-Evening	Set of departure time variables, *Morning* (9.00-11.59), *Afternoon* (12.00-15.59), *Late afternoon* (16.00-17.59) and *Evening* (18.00-23.59); the omitted category is *Early morning* (0.00-8.59)		
✓	✓	Weekend	Dummy variable = 1 if the flight departs in the weekend	0.31 (0.46)	0.27 (0.44)
✓		Flight i	Dummy = 1 for the i^{th} flight operated in a day by a given aircraft, uniquely identified by its tail number		
✓		Aircraft rotation	Sequence of flights operated in a day by a given aircraft		2.99 (1.69)

(a) The symbol ✓ denotes whether the variable is included in the flight buffer regressions (column b) or in the ground buffer regressions (column b_g).
(b) The second-last and last columns respectively report the mean value of the variable for the two-flight sub-sample and unrestricted sample with the standard errors in parentheses. Empty cells on some variables are due to space reason, since their inclusion would require reporting a set multiple dummies whose mean values would not be very informative.

Table 3: Flight-buffer regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	b_1	b_2	b_1, b_2	b_1, b_2	$b_1, ..., b_8$	$b_1, ..., b_8$	$b_1, ..., b_8$
	Two-flight sub-sample				Unrestricted sample		
February	0.407	-0.162	0.119	0.121	-0.067	-0.065	-0.065
March	-0.786	-0.891**	-0.859*	-0.835*	-1.653***	-1.639***	-1.639***
April	-3.775***	-1.380***	-2.591***	-2.580***	-2.764***	-2.750***	-2.750***
May	-6.369***	-0.901*	-3.677***	-3.671***	-3.468***	-3.454***	-3.455***
June	-7.807***	-1.110**	-4.527***	-4.503***	-3.995***	-3.973***	-3.974***
July	-8.287***	-1.418***	-4.882***	-4.869***	-4.241***	-4.221***	-4.222***
August	-8.564***	-2.068***	-5.336***	-5.330***	-4.298***	-4.281***	-4.281***
September	-9.568***	-4.085***	-6.776***	-6.801***	-4.716***	-4.708***	-4.708***
October	-8.845***	-4.616***	-6.700***	-6.717***	-4.521***	-4.509***	-4.510***
November	-5.548***	-4.092***	-4.757***	-4.784***	-2.607***	-2.599***	-2.600***
December	-2.828***	-3.066***	-2.913***	-2.919***	-1.132***	-1.128***	-1.128***
Hub origin	0.733**	-0.732**	-0.282	-0.259	-0.256***	-0.283***	-0.263***
Hub destination	0.454	0.441	0.461*	0.484*	0.400***	0.407***	0.395***
Congestion origin	0.028***	0.029***	0.028***	0.028***	0.035***	0.033***	0.033***
Congestion destination	0.017***	0.030***	0.029***	0.026***	0.025***	0.025***	0.024***
Alaska Airlines	-1.434***	-0.737*	-0.956***	-0.889***	-0.725***	-0.717***	-0.718***
Allegiant Air	-0.756	-0.407	-0.779	-0.763	-2.805***	-2.900***	-2.889***
Delta Airlines	1.067**	2.280***	1.611***	1.609***	1.812***	1.809***	1.811***
Frontier Airlines	-0.667	1.177	0.071	0.426	1.248***	1.239***	1.254***
Hawaiian Airlines	-6.425***	-6.075***	-6.527***	-6.608***	-6.029***	-6.115***	-6.097***
Jet Blue	-1.555***	-2.509***	-2.531***	-2.241***	-2.049***	-2.038***	-2.027***
Southwest Airlines	-0.411	0.411	-0.025	0.105	1.579***	1.591***	1.600***
Spirit Airlines	-1.541**	-0.309	-1.274***	-0.971**	-0.595***	-0.600***	-0.585***
United Airlines	-2.795***	-1.986***	-2.420***	-2.377***	-0.602***	-0.599***	-0.603***
Virgin America	-5.294***	-1.517	-3.565***	-3.491***	-3.674***	-3.670***	-3.665***
Regional carrier	-1.372***	-1.476***	-1.225***	-1.269***	-0.930***	-0.944***	-0.932***
Competitors	-0.263*	0.115	-0.072	-0.079	-0.064	-0.068	-0.071
Distance	0.157***	0.246***	0.198***	0.195***	0.226***	0.225***	0.225***
Morning			0.011		-0.105***		
Afternoon			0.266*		-0.453***		
Late Afternoon			0.894***		0.338***		
Evening			1.092***		0.001		
Weekend	0.018	-0.029	0.032	0.015	0.094***	0.074***	0.074***
Flight 2				0.658***		0.243***	
Flight 3						-0.075**	
Flight 4						0.098***	
Flight 5						0.248***	
Flight 6						-0.006	
Flight 7						-0.485***	
Flight 8						-1.082***	
Aircraft rotation							0.188***
Aircraft rotation^2							-0.031***
Constant	9.448***	6.090	7.058***	7.548***	4.977***	4.981***	4.819***
R^2	0.261	0.241	0.212	0.211	0.190	0.190	0.189
Observations	114,011	113,195	227,206	227,206	5,643,325	5,643,325	5,643,325

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.
(c) All estimates include airport of origin and airport of destination fixed effects.

Table 4: Flight-buffer regressions with flight-time variability

	(1)	(2)	(3)	(4)
Dependent variable	b_1	b_2	b_1, b_2	b_1, b_2
Flight 1's variability	0.107***	0.052***	0.079***	0.080***
Flight 2's variability	-0.013	0.094***	0.041***	0.042***
Hub origin	0.510*	-0.148	0.016	-0.009
Hub destination	-0.434	0.477*	0.091	0.125
Alaska Airlines	-0.499	-0.211	-0.408	-0.351
Allegiant Air	-1.755*	0.485	-0.808	-0.731
Delta Airlines	2.797***	3.789***	3.255***	3.282***
Frontier Airlines	-0.366	0.828	-0.086	0.174
Hawaiian Airlines	-2.138***	-5.534***	-4.128***	-4.244***
Jet Blue	-0.715	-0.899**	-1.094***	-0.928**
Southwest Airlines	0.381	1.249***	0.585	0.746**
Spirit Airlines	-3.571***	-1.400***	-2.832***	-2.612***
United Airlines	-1.177***	-1.393***	-1.283***	-1.267***
Virgin America	1.130	2.343**	1.502	1.613*
Regional carrier	-2.514***	-2.201***	-2.295***	-2.329***
Competitors	0.309**	0.414***	0.370***	0.362***
Morning			0.209	
Afternoon			-0.459**	
Late Afternoon			0.586***	
Evening			0.355*	
Flight 2				-0.144
Constant	4.451***	3.173***	3.908***	3.981***
R^2	0.046	0.100	0.065	0.064
Observations	114,011	113,195	227,206	227,206

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.

Table 5: The determinants of flight-time variability

Dependent variable	(1) Flight 1's variability	(2) Flight 2's variability
February	-0.558**	-0.870***
March	-1.079***	-1.198***
April	0.603**	-0.497*
May	-0.300	0.188
June	-0.386	0.135
July	1.130***	0.926***
August	1.097***	1.638***
September	-0.687**	-0.119
October	-0.974***	-0.916***
November	-0.423	0.177
December	0.106	0.023
Congestion origin	0.007**	0.003
Congestion destination	0.009***	0.016***
Distance	0.135***	0.155***
Weekend	-0.104**	-0.123***
Constant	6.051*	-8.365*
R^2	0.283	0.281
Observations	114,011	113,195

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.
(c) All estimates include airport of origin and airport of destination fixed effects.

Table 6: Ground-buffer regressions

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	b_g	b_g	b_g	b_g	b_g	b_g
	Two-flight sub-sample			Unrestricted sample		
February	0.231	0.259	0.051	-0.186	-0.178	-0.305
March	-0.819	-0.854	-1.198	-1.105***	-1.118***	-1.459***
April	-0.626	-0.918	-1.395	-0.874***	-0.885***	-1.236***
May	0.940	0.646	0.492	-0.163	-0.174	-0.523**
June	-0.786	-1.127	-1.276	-1.221***	-1.233***	-1.815***
July	0.336	-0.041	-0.107	-0.912***	-0.917***	-1.522***
August	1.176	0.786	0.696	-0.575***	-0.594***	-1.221***
September	3.108***	2.786***	2.690***	1.281***	1.320***	0.785***
October	2.663***	2.331***	1.970**	0.847***	0.847***	0.302
November	1.712**	1.523**	1.101	0.863***	0.893***	0.409**
December	0.159	-0.074	-0.121	0.350*	0.361*	0.022
Hub turnaround	14.441***	13.130***	10.351***	14.160***	13.423***	13.505***
Congestion turnaround	-0.088***	-0.083***	0.007	-0.085***	-0.085***	0.041***
Slot controlled airport			-4.155***			0.033
Alaska Airlines	0.280		3.198***	5.849***		9.041***
Allegiant Air	-14.563***		-24.813***	-8.478***		-10.362***
Delta Airlines	-3.318***		0.001	-3.210***		-1.098***
Frontier Airlines	-6.511***		-4.135*	-2.996***		-0.666**
Hawaiian Airlines	-6.901***		6.239***	-6.184***		8.705***
Jet Blue	-5.271***		-8.257***	-0.506		-0.151
Southwest Airlines	-7.235***		-8.412***	-7.576***		-5.231***
Spirit Airlines	-8.641***		-9.227***	0.356		2.795***
United Airlines	4.905***		8.001***	5.088***		8.958***
Virgin America	5.673***		4.385**	8.246***		11.150***
Regional carrier	-1.062*		-6.993***	-1.599***		-4.129***
Low-cost carrier		-7.055***			-5.379***	
Morning	7.035***	6.772***	5.917***	4.366***	4.326***	1.989***
Afternoon	14.057***	13.797***	14.142***	4.618***	4.487***	2.811***
Late Afternoon	19.124***	19.148***	16.731***	5.992***	5.985***	3.949***
Evening	19.893***	19.883***	19.064***	9.618***	9.711***	8.060***
Weekend	1.931***	1.962***	2.616***	1.868***	1.930***	2.393***
Constant	32.479***	29.684***	43.968***	26.227***	23.921***	32.446***
R^2	0.204	0.199	0.154	0.235	0.226	0.203
Observations	133,178	133,178	133,178	4,318,421	4,318,421	4318,421

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.
(c) All estimates but columns (16) and (19) include airport of turnaround and airport of origin fixed effects; columns (16) and (19) include only airport of origin fixed effects.

Table 7: Ground-buffer regressions with flight-time variability

	(1)	(2)	(3)	(4)
Dependent variable	b_g	b_g	b_g	b_g
Flight 1's variability	-0.082***	-0.073***	0.157***	0.142***
Flight 2's variability	-0.158***	-0.168***	-0.310***	-0.364***
Morning	3.652***	3.344***	3.246***	2.795***
Afternoon	11.225***	11.628***	10.611***	10.798***
Late Afternoon	12.740***	13.107***	12.324***	12.530***
Evening	15.032***	15.573***	14.466***	14.817***
Hub turnaround	12.883***	12.463***	12.910***	12.503***
Alaska Airlines	3.818***		3.405***	
Allegiant Air	-19.832***		-19.315***	
Delta Airlines	1.375**		0.982	
Frontier Airlines	-0.892		-1.094	
Hawaiian Airlines	5.549***		4.795***	
Jet Blue	-7.982***		-8.827***	
Southwest Airlines	-4.374***		-4.019***	
Spirit Airlines	-7.500***		-8.015***	
United Airlines	9.456***		9.149***	
Virgin America	4.769**		4.837**	
Regional carrier	-0.801		0.004	
Low-cost carrier		-11.455***		-11.792***
Constant	25.203***	28.850***	25.140***	29.927***
R^2	0.115	0.099	0.115	0.099
Observations	133,178	133,178	133,178	133,178

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.
(c) Actual flight-time variability calculated by route, month and flight i in columns (20) and (21), by route and flight i in columns (22) and (23),

Table 8: Correlation of the excess ground buffer with the flight buffers

Two-flight sub-sample

	$\forall\, b_g$	$b_g \leq 60$	$b_g \leq 30$
b_1	-0.0043	-0.0375	-0.0232
b_2	-0.0096	-0.0315	-0.0403

Unrestricted sample

	$\forall\, b_g$	$b_g \leq 60$	$b_g \leq 30$
b_1	-0.0444	-0.0379	-0.0232
b_2	-0.0259	-0.0164	-0.0061
b_3	0.0029	0.0143	0.0097
b_4	-0.0480	-0.0510	-0.0380
b_5	-0.0509	-0.0528	-0.0489
b_6	-0.0469	-0.0372	-0.0469
b_7	-0.0640	-0.0320	-0.0391
b_8	-0.0642	-0.0548	-0.0479

Table A.1: Flight-buffer regressions with past year on-time performance

	(A1)	(A2)	(A3)	(A4)	(A5)
Dependent variable	b_1	b_2	b_1, b_2	$b_1, ..., b_8$	$b_1, ..., b_8$
	Two-flight sub-sample			Unrestricted sample	
February	0.719	0.135	0.456	0.167	0.167
March	-0.597	-0.758*	-0.683	-1.555***	-1.556***
April	-3.944***	-1.203***	-2.581***	-2.838***	-2.839***
May	-6.448***	-0.763	-3.699***	-3.667***	-3.669***
June	-8.210***	-1.052**	-4.707***	-4.168***	-4.170***
July	-8.484***	-1.282**	-4.917***	-4.395***	-4.397***
August	-8.697***	-1.789***	-5.245***	-4.418***	-4.420***
September	-8.745***	-3.413***	-5.992***	-4.647***	-4.648***
October	-8.179***	-4.088***	-6.066***	-4.534***	-4.535***
November	-4.674***	-3.560***	-4.013***	-2.465***	-2.466***
December	-2.462***	-2.977***	-2.702***	-1.277***	-1.278***
Hub origin	0.010	-1.113***	-0.659**	-0.544***	-0.530***
Hub destination	0.565	0.627*	0.597**	0.195*	0.189*
Congestion origin	0.027***	0.026***	0.029***	0.033***	0.033***
Congestion destination	0.021***	0.031***	0.031***	0.028***	0.028***
Alaska Airlines	-0.241	-0.266	0.010	-0.691***	-0.684***
Allegiant Air					
Delta Airlines	1.346***	2.546***	1.873***	1.995***	1.995***
Frontier Airlines	-2.363**	1.360*	-0.038	1.561***	1.566***
Hawaiian Airlines	-6.627***	-5.773***	-6.685***	-6.219***	-6.225***
Jet Blue	-3.791***	-3.373***	-3.592***	-2.773***	-2.766***
Southwest Airlines	-1.113**	0.097	-0.034	1.069***	1.073***
Spirit Airlines	-3.545***	-0.872	-1.767***	-1.142***	-1.137***
United Airlines	-2.979***	-2.064***	-2.500***	-1.220***	-1.221***
Virgin America	-6.629***	-2.590***	-4.765***	-5.213***	-5.205***
Regional carrier	-1.365***	-1.227***	-1.047***	-0.844***	-0.840***
Competitors	-0.238	0.028	-0.121	-0.112**	-0.113**
Distance	0.182***	0.252***	0.196***	0.231***	0.231***
Past-year delay	9.538***	6.228***	8.687***	3.643***	3.650***
Morning	-0.834***	0.174	-0.121	-0.240***	-0.154***
Afternoon	0.310	-0.675**	0.089	-0.408***	-0.399***
Late Afternoon	2.052***	-0.720**	0.736***	0.564***	0.568***
Evening	2.726***	-0.707**	0.696***	0.271***	0.273***
Weekend	0.013	-0.017	0.052	0.059***	0.060***
Flight 2			0.367***	0.402***	
Flight 3				0.218***	
Flight 4				0.126***	
Flight 5				0.017	
Flight 6				-0.313***	
Flight 7				-0.914***	
Flight 8				-1.344***	
Aircraft rotation					0.361***
Aircraft rotation^2					-0.064***
Constant	12.672***	0.974	3.678*	2.158	1.889
R^2	0.280	0.251	0.219	0.198	0.198
Observations	101,920	100,916	202,836	4,126,794	4,126,794

(a) The estimated coefficients marked with ***, ** and * are statistical significance at, respectively the 1%, 5% and 10% level.
(b) The standard errors, not reported to save space, are clustered by route-month.
(c) All estimates include airport of origin and airport of destination fixed effects.

References

ARIKAN, M., DESHPANDE, V., SOHONI, M., 2013. Building reliable air-travel infrastructure using empirical data and stochastic models of airline networks. *Operations Research* 61, 45-64.

BALL, M., BARNHART, C., DRESNER, M., HANSEN, M., NEELS, K., ODONI, A., PETERSON, E., SHERRY, E., TRANI, A., ZOU, B., 2010. *Total Delay Impact Study: A Comprehensive Assessment of the Costs and Impacts of Flight Delay in the United States.* Final Report, National Center of Excellence for Aviation Operations Research (NEXTOR).

CILIBERTO, F., WILLIAMS, J.W., 2010. Limited access to airport facilities and market power in the airline industry. *Journal of Law and Economics* 53, 467-495.

DESHPANDE, V., ARIKAN, M., 2012. The impact of airline flight schedules on flight delays. *Manufacturing & Service Operations Management* 14, 423-440.

FORBES, S.J., 2008. The effect of air traffic delays on airline prices. *International Journal of Industrial Organization* 26, 1218-1232.

FORBES, S.J., LEDERMAN, M., 2010. The performance implications of vertical integration: Evidence from the airline industry. *RAND Journal of Economics* 41, 765-790.

FORBES, S.J., LEDERMAN, M., 2009. Adaptation and vertical integration in the airline industry. *American Economic Review* 99, 1831-49.

FORBES, S.J., LEDERMAN, M., WITHER, M.J., 2019. Quality disclosure when firms set their own quality targets. *International Journal of Industrial Organization* 62, 228-250.

FORBES, S.J., LEDERMAN, M., YUAN, Z., 2019. Do airlines pad their schedules? *Review of Industrial Organization* 54, 61-82.

HAO, L., HANSEN, M., 2014. Block time reliability and scheduled block time setting. *Transportation Research Part B* 69, 98-111.

KANG, L., HANSEN, M., 2017. Behavioral analysis of airline scheduled block time adjustment. *Transportation Research Part E* 103, 56-68.

LUTTMANN, A., 2019. Evidence of directional price discrimination in the U.S. airline industry. *International Journal of Industrial Organization* 62, 291-329.

MAZZEO, M., 2003. Competition and service quality in the U.S. airline industry. *Review of*

Industrial Organization 22, 275-296.

Prince, J.T., Simon, D.H., 2015. Do incumbents improve service quality in response to entry? Evidence from airlines on-time performance. *Management Science* 61, 372-390.

Shumsky, R.A., 1993. Response of U.S. air carriers to on-time disclosure rule. *Transportation Research Record* 1379, 9-16.

Sohoni, M., Lee, Y.-C., Klabjan, D., 2011. Robust airline scheduling under block-time uncertainty. *Transportation Science* 45, 451-464.

Whitin, T.M., 1955. Inventory control and price theory. *Management Science* 2, 61-68.

Wang, X., 2015. Service competition in the airline industry: Schedule robustness and market structure. Unpublished paper.

Zhang, A., Czerny, A.I., 2012. Airports and airlines economics: An interpretive review of recent research. *Economics of Transportation* 1, 15-34.

Zhang, D.J., Salant, Y., Van Mieghem, J.A., 2018. Where did the time go? On the increase in airline schedule padding over 21 years. Unpublished paper.