Transportation Research Part B 162 (2022) 55–71

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Airline delay propagation: A simple method for measuring its extent and determinants[☆]

Jan K. Brueckner [a,*], Achim I. Czerny [b], Alberto A. Gaggero [c]

[a] Department of Economics, University of California, Irvine, Irvine, CA 92697, USA
[b] Department of Logistics and Maritime Studies, Hong Kong Polytechnic University, Hong Kong
[c] Department of Economics and Management, University of Pavia, 27100 Pavia, Italy

ARTICLE INFO

Keywords:
Propagated delay
Arrival delay
Ground buffer

ABSTRACT

This paper offers a simple approach for identifying propagated departure delays and measuring their contribution to arrival delays. Under our approach, a propagated departure delay occurs when the arrival delay of the inbound flight exceeds the subsequent flight's ground buffer. The size (or frequency) of such propagated delays relative to the size (or frequency) of arrival delays then measures the contribution of propagated delays to late arrivals. This approach differs from earlier attempts to quantify the contribution of delay propagation since it focuses on an individual flight and its immediate predecessor, without attempting to trace the sources of delay propagation back through the entire sequence of prior flights. The paper's empirical results show that the contribution of propagated departure delays to arrival delays depends on several key determinants.

1. Introduction

Delay propagation, where late arrival of an inbound flight leads to a late departure and then late arrival of the subsequent outbound flight, is the main cause of flight delays in the US, according to FAA data.[1] A literature has emerged in transportation engineering and economics studying delay propagation as part of a broader inquiry into the stochastic aspects of airline operations. Two papers in particular, Arikan et al. (2013) and Kafle and Zou (2016), propose methods for measuring the extent of delay propagation.[2] These methods are ambitious, with the goal of quantifying the potential cascade of delay propagation through the entire sequence of daily flights operated by a given aircraft. In other words, a late inbound arrival may lead to a departure delay and then late arrival of the next flight, which may in turn lead to late departure of the subsequent flight, and so on. With this broad view, isolating the contribution of delay propagation to the on-time performance of any particular flight is complex analytically, making an empirical implementation correspondingly complex. The goal of the present paper is to offer and make use of a simpler approach to quantifying delay propagation and its effects. Our approach has a limited purpose, which is to identify propagated departure delays for individual flights *without considering their ultimate sources*. In other words, we seek to determine whether a flight's late departure was caused by late arrival of the inbound flight, *without asking whether the inbound flight's lateness was itself*

☆ We thank Hanxiang Zhang for assistance in compiling the weather data and thank (without implicating) two referees for helpful comments.
* Corresponding author.
E-mail address: jkbrueck@uci.edu (J.K. Brueckner).
[1] See Fig. 1 in Brueckner et al. (2021a).
[2] See AhmadBeygi et al. (2008, 2010), and Deshpande and Arikan (2012) for related papers. Kim and Park (2021) apply Kafle and Zou's (2016) methods to Korean airports.

https://doi.org/10.1016/j.trb.2022.05.003
Received 10 August 2021; Received in revised form 4 May 2022; Accepted 4 May 2022
Available online 30 May 2022
0191-2615/

caused by prior delay propagation. This narrower focus, which looks at an individual flight and its immediate predecessor rather than the entire daily flight sequence, yields straightforward measures of the contribution of delay propagation.

Central to our exercise is the ability to identify propagated departure delays in the data by applying the theoretical scheduling models of Brueckner et al. (2021a,b), hereafter BCG. The key theoretical prediction borrowed from BCG's work is that a propagated departure delay occurs when late arrival of the inbound flight, measured in minutes, exceeds the ground buffer for the subsequent flight, equal to minutes of extra scheduled ground time beyond the minimum feasible aircraft turnaround time (quantities that all can be measured in the data). The difference between the inbound delay and the ground buffer, when positive, equals minutes of propagated departure delay. A ratio based on this quantity, minutes of propagated departure delay divided by minutes of subsequent arrival delay, gives our measure of the contribution of propagated departure delays to subsequent late arrivals.

We also compute variants of this propagation measure that rely on flight counts at the route level rather than minutes of delay for individual flights. One variant equals the number of flights on a route having both a propagated departure delay and an arrival delay divided by the number of flights with an arrival delay. This route-level propagation measure is also adjusted to capture the contribution to arrival delays of moderate inbound arrival delays, which do not automatically lead to a late departure.

The ultimate goal of the paper is to explore how the contribution of propagated departure delays to arrival delays varies with route, airport, and airline characteristics. Our first finding is that this contribution is higher when the origin airport is *not* a hub for the airline. Therefore the contribution of propagated delays is larger at spoke endpoints, where the late inbound flight may be arriving from a hub after a confluence of factors delayed its departure.[3] Another finding is that the contribution of propagated departure delays rises over the day, with flights later in the day affected more than earlier flights. This finding presumably reflects the cumulative effect of prior delays during the day, whose causes are a mixture of propagated delay and delays from other sources. A third finding is that propagated departure delays account for a larger share of arrival-delay minutes for low-cost carriers than for other airlines, partly reflecting the quick turnarounds (short ground buffers, thus more propagated delays) that are a feature of their business model.[4]

These findings are presented late in the paper (in Section 5), after we present preliminary regressions that validate the underpinnings of our approach, which is done in Section 4. The first preliminary regression verifies that arrival delays for inbound flights do indeed create subsequent departure delays. This flight-level regression relates minutes of departure delay to minutes of inbound arrival delay, and the estimated positive relationship illustrates the source of propagated departure delays.[5] Next, we verify that an inbound arrival delay can indeed create an arrival delay for the subsequent flight, with the regression relating minutes of arrival delay to minutes of inbound arrival delay and showing a positive connection. We also present spline versions of these regressions, where the impact of the inbound arrival delay is allowed to be larger the greater its size relative to the ground buffer (a pattern that is confirmed).

Additional preliminary regressions explore the main connection of interest, the one relating arrival delays to departure delays and other covariates.[6] We present flight-level regressions showing that the relationship is slightly less than one-to-one, even when departure delays are decomposed into propagated delays and delays from other sources. The near one-to-one connection appears to reflect the difficulty of overcoming a departure delay by faster flying, as illustrated in additional regressions relating airborne time to departure delay, which have very small negative coefficients.

With the underpinnings of our approach verified, we then return in Section 5 to our measures that quantify the contribution of propagated departure delays to late arrivals. We present regressions relating the flight-level and route-level contribution measures to a host of covariates, as described above.

The present paper is an outgrowth of our earlier research on flight delays and schedule buffers BCG (2021a) and BCG (2021b). These papers developed theoretical models of how airlines choose the magnitudes of schedule buffers, which are partly designed to reduce the likelihood of delay propagation. The flight buffer is the amount by which scheduled flight time exceeds the minimum possible flight time, and the ground buffer is the analogous component of ground time, as mentioned above. These models generated predictions relating the sizes of buffers to various determinants, and BCG (2021a) tested these predictions via regression analysis relating buffer sizes to route characteristics, time-of-day and month, airline identities, aircraft type, and previous flight-time variability.[7] They relied on voluminous US Department of Transportation data on the daily scheduled and actual flight and ground times for individual aircraft, and these same data are used in the present paper to generate and exploit measures of actual delay propagation.[8] The present study relies on several of the key delay formulas from BCG's papers as a basis for its empirical analysis.

The plan of the paper is as follows. Section 2 explains the model-based derivations underlying the empirical work, and Section 3 discusses the data. As mentioned above, Section 4 presents preliminary empirical results, while Section 5 presents the regressions showing how the arrival-delay contribution of propagated departure delays varies by route, airport and airline characteristics. Section 6 offers conclusions.

---

3 Using a completely different methodology, Diana (2009) explores whether delay propagation is higher at concentrated airports, which are usually hubs.

4 In addition to showing that low-cost carriers have short ground buffers, some of BCG's (2021a) regressions indicate that their flight buffers are also short, potentially contributing to late inbound arrivals.

5 Tan et al. (2021) carry out a similar exercise for Chinese airports, but are forced because of data limitations to rely on airport averages rather than information for individual aircraft.

6 Wong and Tsai (2012) estimate somewhat similar regressions with a hazard (survival-analysis) approach.

7 Eufrásio et al. (2021) carry out a similar exercise.

8 A collection of papers related to research on buffers focuses on the determinants of scheduled block (flight) times and the issue of schedule "padding," where flight times expand in response to growing airport congestion and other sources of delays. See Sohoni et al. (2011), Hao and Hansen (2014), Kang and Hansen (2017, 2018), Zhang et al. (2018), Fan (2019), Forbes et al. (2019a), Forbes et al. (2019b), Wang et al. (2019), and Yimja and Gorjidooz (2019). Following earlier work, a different line of inquiry is exemplified by Li and Jing (2021), who test for interconnections between delays at different airports, thus characterizing a "delay propagation network." Dou et al. (2020) and Fleurquin et al. (2013) carry out similar exercises.

2. The propagated delay measures

This section of the paper uses the scheduling model of BCG to derive the relationships that underlie the subsequent empirical work. Section 2.1 derives the equation relating arrival delays to departure delays, which underlies several of the regressions reported in Section 4. Section 2.2 derives the formula that relates propagated departure delay to inbound arrival delay and the ground buffer, which is used in quantifying the contribution of propagated departure delays to arrival delays in Section 2.3.

2.1. Delay definitions

For each of the millions of flights in our 2018 data set, we have information on its scheduled and actual times of departure and arrival as well as information on the magnitude of the flight buffer and the ground buffer prior to the flight. Following BCG, let $t_{d,i}$ and $t_{a,i}$ denote the scheduled departure and arrival times for flight i, and let $\hat{t}_{d,i}$ and $\hat{t}_{a,i}$ denote the flight's actual departure and arrival times. The departure and arrival delays for flight i, denoted $D_{d,i}$ and $D_{a,i}$ are then given by

$$D_{d,i} = \hat{t}_{d,i} - t_{d,i} \tag{1}$$

$$D_{a,i} = \hat{t}_{a,i} - t_{a,i}. \tag{2}$$

Note that these delay expressions can be negative (with departures or arrivals being early), as they often are in the data.

To see the connection between arrival and departure delays, the determinants of flight times must be considered. The actual arrival time of flight i is the actual departure time plus the actual flight time, given by

$$\hat{t}_{a,i} = \hat{t}_{d,i} + m_i + \epsilon_i. \tag{3}$$

In (4), $m_i + \epsilon_i$ is the actual flight time, equal to minimum feasible flight time m_i plus a positive random term ϵ_i that accounts for the randomness of flight durations.

The scheduled arrival time $t_{a,i}$ equals the scheduled departure time $t_{d,i}$ plus the scheduled flight time, which equals the minimum flight time m_i plus the flight buffer $b_{f,i}$. With $t_{a,i}$ thus equal to $t_{d,i} + m_i + b_{f,i}$, arrival delay is equal to

$$\begin{aligned} D_{a,i} = \hat{t}_{a,i} - t_{a,i} &= \hat{t}_{d,i} + m_i + \epsilon_i - (t_{d,i} + m_i + b_{f,i}) \\ &= D_{d,i} + \epsilon_i - b_{f,i}, \end{aligned} \tag{4}$$

using (2). Therefore, arrival delay equals departure delay plus the random flight-duration term minus the flight buffer. With the random term held fixed, arrival delay then rises in one-for-one fashion as departure delay increases. But since the random factor is not fixed in reality, and since the buffer $b_{f,i}$ varies across flights, this one-for-one relationship is likely to hold only approximately.

2.2. Propagated departure delay

The discussion now turns to the determinants of departure delay, $D_{d,i}$ in (4). In BCG (2021a), the only cause of a departure delay is late arrival of the inbound flight. In other words, a propagated departure delay is the only kind of departure delay in their model. In their follow-up paper, BCG (2021b) realistically added random shocks to the ground time prior to a flight, which provide another source of departure delays.

Suppressing such shocks to focus on propagated departure delays, let $t_{a,i-1}$ and $\hat{t}_{a,i-1}$ denote the scheduled and actual arrival times for the inbound flight (flight $i-1$). Then, the inbound arrival delay is

$$D_{a,i-1} = \hat{t}_{a,i-1} - t_{a,i-1}. \tag{5}$$

Furthermore, let the scheduled ground time prior to flight i be denoted $t_{g,i}$ and let the minimum flight turnaround time be denoted $\bar{t}_{g,i}$. Flight i's ground buffer then equals

$$b_{g,i} = t_{g,i} - \bar{t}_{g,i} > 0, \tag{6}$$

being equal to the excess of scheduled ground time above the minimum turnaround time, as noted above. The minimum turnaround is computed from the data and is specific to airports and aircraft types.

To derive the departure delay in (4), note that the actual departure time of flight i equals

$$\hat{t}_{d,i} = \max\{t_{d,i},\ \hat{t}_{a,i-1} + \bar{t}_{g,i}\}. \tag{7}$$

To understand (7), note that $\hat{t}_{a,i-1} + \bar{t}_{g,i}$ is the earliest possible departure time for flight i, equal to the arrival time of the inbound flight plus the minimum aircraft turnaround time. If the inbound flight arrives on time ($t_{a,i-1} = \hat{t}_{a,i-1}$), then this expression equals $t_{a,i-1} + \bar{t}_{g,i} < t_{a,i-1} + t_{g,i} = t_{d,i}$, being less than the scheduled departure time (note $\bar{t}_{g,i} < t_{g,i}$). The max in (8) then equals the first $t_{d,i}$ argument, so that flight i departs on time. But if the reverse inequality holds, so that $\hat{t}_{a,i-1} + \bar{t}_{g,i} > t_{d,i}$, then the flight's earliest possible departure time exceeds the scheduled departure time, and flight i departs late.

To derive the magnitude of the resulting propagated departure delay, the scheduled departure time $t_{d,i}$ is subtracted from both sides of (7), which becomes

$$D_{d,i} = \hat{t}_{d,i} - t_{d,i} = \max\{0,\ \hat{t}_{a,i-1} + \bar{t}_{g,i} - t_{d,i}\}. \tag{8}$$

Adding and subtracting $t_{a,i-1}$ from the second max expression in (8) yields

$$\begin{aligned} D_{d,i} &= \max\{0,\ \hat{t}_{a,i-1} - t_{a,i-1} + t_{a,i-1} + \bar{t}_{g,i} - t_{d,i}\} \\ &= \max\{0,\ D_{a,i-1} + t_{a,i-1} + \bar{t}_{g,i} - t_{d,i}\} \\ &= \max\{0,\ D_{a,i-1} + t_{a,i-1} + t_{g,i} + \bar{t}_{g,i} - t_{g,i} - t_{d,i}\} \\ &= \max\{0,\ D_{a,i-1} - b_{g,i}\}, \end{aligned} \tag{9}$$

where the last equality comes from adding and subtracting $t_{g,i}$ in the previous line and using $t_{d,i} = t_{a,i-1} + t_{g,i}$ and $t_{g,i} - \bar{t}_{g,i} = b_{g,i}$. Therefore, the departure delay for flight i equals the inbound arrival delay minus the ground buffer if this quantity is positive, equaling zero otherwise.[9] While this analysis, following BCG, does not recognize other sources of departure delays beyond inbound arrival delays, such sources are obviously important empirically, and they are considered below.[10]

2.3. Measuring the arrival-delay contribution of propagated departure delays

As just noted, sources beyond delay propagation contribute to departure delays in actual flight operations, and the notation of the model must be adjusted to reflect this reality. Accordingly, let expression $D_{d,i}$ in (9) be renamed $D^{prop}_{d,i}$ to reflect its measurement of propagated departure delay. Also, let $D^{tot}_{d,i}$ equal total minutes of departure delay from all sources, and let $D^{other}_{d,i} = D^{tot}_{d,i} - D^{prop}_{d,i}$ denote minutes of departure delay from sources other than propagation.

Since our data allow measurement of the actual and scheduled arrival times of all flights along with magnitude of the ground buffer for individual flights, $D^{prop}_{d,i}$ in (9) can be computed along with $D_{a,i}$ in (2). To measure the contribution of propagated departure delays to late arrivals, we then form the ratio

$$R^{flight}_i = \frac{D^{prop}_{d,i}}{D_{a,i}} = \frac{\text{minutes of propagated departure delay for flight } i}{\text{minutes of arrival delay for flight } i}, \tag{10}$$

where the $flight$ superscript on R indicates that the ratio is measured at the individual flight level. Note that R^{flight}_i captures the share of just one source, propagated departure delay, in minutes of arrival delay. Additional sources are minutes of departure delay caused by factors other than propagation along with random elements (weather, unanticipated congestion) that affect flight times (note that flight time equals airborne time plus taxi-in and taxi-out times).

While R^{flight}_i is computed for each flight in our data set, it is also possible to compute an R ratio at the route level, where a route consists of two endpoint airports connected by nonstop service. Rather than using minutes of delay, this route-level aggregation simply counts the number of flights on a route with both propagated departure delay and arrival delay along with the number of flights with arrival delay, yielding

$$R^{route}_r = \frac{\#\ \text{route}-r \text{ flights with propagated departure delay and arrival delay}}{\#\ \text{route}-r \text{ flights with arrival delay}}, \tag{11}$$

where r denotes the route and an arrival delay occurs when $D_{a,i} > 15$ minutes, following the FAA definition. Among flights with arrival delay, R^{route}_r thus gives the share that also had a propagated departure delay. In actuality, the route-level aggregation is finer than indicated in (11), with aggregation occurring at the route × week (or month) level.

While R^{route}_r tells whether a late-arriving flight also had a propagated departure delay, it does not net out the other factors that may have contributed to the late arrival. However, an adjusted measure can be constructed that does so. To produce this adjusted R ratio, which better captures the share of late arrivals caused solely by propagated departure delays, we compute the share of flights at the route level with no departure delay that also experienced an arrival delay, equal to A^{route}_r. The adjusted ratio is then $R^{adj}_r = (1 - A^{route}_r) \times R^{route}_r$, where the share of flights with no departure delay that arrived late is netted out.[11] However, because A^{route}_r is on average near zero, this adjustment has little effect. The near-zero value shows that flights without a departure delay are very seldom late (with $D_{a,i} > 15$ minutes), so that arrival delays are almost exclusively due to departure delays. This conclusion can

[9] This analysis ignores the possibility of aircraft substitution, in which another plane is substituted on a route experiencing a substantial propagated departure delay. To explore this issue, suppose that, under its original flight plan, aircraft 1 flies from city A to city B and then to city C. Assuming a zero ground buffer for simplicity, the scheduled arrival time in B and the scheduled departure from B to C are both t^*, Aircraft 1 is late, arriving at B at time $t^{**} > t^*$. Knowing in advance that its arrival will be late, delaying aircraft 1's departure to city C, the airline assigns aircraft 2 (possibly a spare plane) to the B-C route, so that the departure can occur on time. In addition, aircraft 1 is reassigned to fly from city B to city E, with the plane that was originally assigned to that flight used for a different flight. Suppose that the B-E flight was originally scheduled to depart at t**, exactly the time of aircraft 1's late arrival. That flight can thus be operated using aircraft 1 with no departure delay. The upshot is that, because of aircraft substitution, the calculation in (9) will yield a propagated departure delay of $t^{**} - t^*$ even though the B-E flight departs (and may eventually arrive) on time. Thus, the predicted association between propagated departure delays and subsequent arrival delays can be disrupted in the presence of aircraft substitution. Although the extent of substitution in response to scheduling (as opposed to mechanical) disruptions is unknown, the resulting empirical problem will be small if such substitutions are rare, as seems likely.

[10] In an extension of their basic model, BCG (2021b) explore the effect of random shocks to ground time, which capture the other sources of departure delay.

[11] Specifically,

$$A^{route}_r = \frac{\text{number of route}-r \text{ flights without departure delay but with arrival delay}}{\text{number of route}-r \text{ flights without departure delay}}.$$

Table 1
Summary statistics of the delay variables.

Variable	Description	Mean	Std. Dev.	Min	Max	Obs.
$D_{a,i-1}$	arrival delay for flight $i-1$	−2.235	21.688	−49	120	4,205,229
$D_{a,i}$	arrival delay for flight i	−0.609	23.224	−49	120	4,205,229
$D^{tot}_{d,i}$	departure delay for flight i	4.529	19.296	−19	120	4,205,229
$D^{prop}_{d,i}$	$\max\{0, D_{a,i-1} - b_{g,i}\}$	1.733	8.556	0	113	4,205,229
$D^{other}_{d,i}$	$D^{tot}_{d,i} - D^{prop}_{d,i}$	2.796	15.199	−71	120	4,205,229
R^{flight}_{i}	$\frac{D^{prop}_{d,i}}{D_{a,i}}$, defined if $D_{a,i} > 0$	0.127	0.418	0	44	1,432,446
	if $D_{a,i} > 15$	0.202	0.330	0	3	678,244
$R^{route}_{r,month}$	$\frac{\text{\# route-}r\text{ flights with } D^{prop}_{d,i}>0 \text{ and } D_{a,i}>15 \text{ by month}}{\text{\# route-}r\text{ flights with } D_{a,i}>15 \text{ by month}}$	0.393	0.292	0	1	55,717
$R^{route}_{r,week}$	$\frac{\text{\# route-}r\text{ flights with } D^{prop}_{d,i}>0 \text{ and } D_{a,i}>15 \text{ by week}}{\text{\# route-}r\text{ flights with } D_{a,i}>15 \text{ by week}}$	0.388	0.375	0	1	177,959
$\widetilde{R}^{route}_{r,month}$	$\frac{\text{\# route-}r\text{ flights with } 0.5b_{g,i}<D_{a,i-1}\le b_{g,i}\text{, with } D^{tot}_{d,i}>0\text{, and } D_{a,i}>15 \text{ by month}}{\text{\# route-}r\text{ flights with } D_{a,i}>15 \text{ by month}}$	0.123	0.168	0	1	55,717
$\widetilde{R}^{route}_{r,week}$	$\frac{\text{\# route-}r\text{ flights with } 0.5b_{g,i}<D_{a,i-1}\le b_{g,i}\text{, with } D^{tot}_{d,i}>0\text{, and } D_{a,i}>15 \text{ by week}}{\text{\# route-}r\text{ flights with } D_{a,i}>15 \text{ by week}}$	0.124	0.236	0	1	177,959

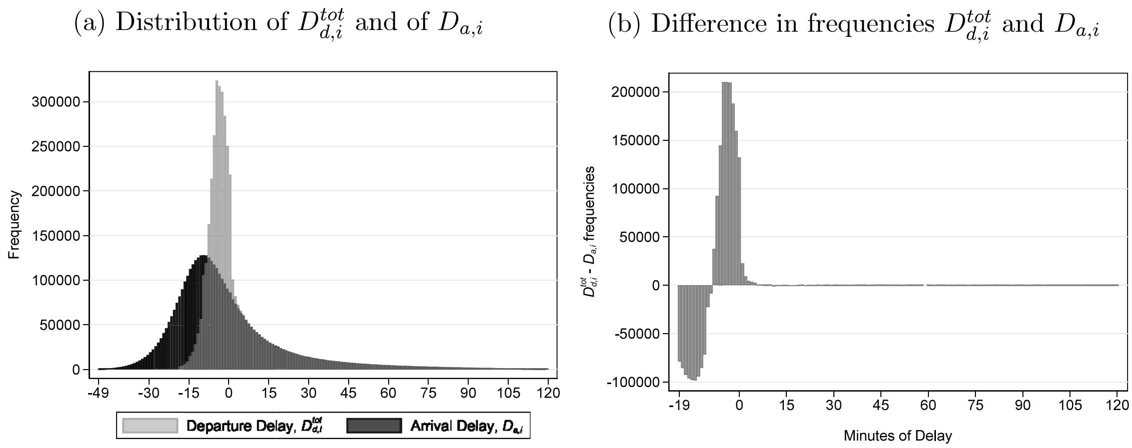

Fig. 1. Arrival and departure delay.

be seen in a different way by replacing $D^{prop}_{d,i}$ in the ratio in (10) with $D^{tot}_{d,i}$, which yields a modified ratio whose average value is close to 1 (equal to 0.85), showing that arrival delay minutes on average approximately equal minutes of total departure delay.

If the inbound flight experiences an arrival delay smaller in magnitude than the ground buffer, then a propagated departure delay as we have defined it does not occur. However, the shortening of the available turnaround time as a result of the inbound delay offers a greater chance for stochastic ground-time factors to make the flight depart late. As a result, it could be argued that a moderate inbound arrival delay could generate an indirect propagated departure delay via this channel. To capture the arrival-delay contribution of departure delays from such indirect propagation, the following variant of R^{route}_{r} is computed, where b_g is again the ground buffer:

$$\widetilde{R}^{route}_{r} = \frac{\text{\# route}-r\text{ flights with } 0.5b_g < D_{a,i-1} \le b_g\text{, with departure delay, and with arrival delay}}{\text{\# route}-r\text{ flights with arrival delay}} \tag{12}$$

This ratio equals the share of flights with an arrival delay that have both a departure delay and a moderate inbound arrival delay (between half and the full ground buffer), which may have contributed to the departure delay.

3. Data description

The sample used in this paper covers US domestic airline operations for the year 2018.[12] It is constructed by combining information from the US Department of Transportation and Weather Underground.[13]

The USDOT data is from the 'Marketing Carrier On-Time Performance' dataset, which at the aircraft level, uniquely identified by the aircraft's tail number, provides information on the carrier operating the flight, the origin and destination airports, the departure date, the scheduled and actual departure and arrival times, the taxi-out and taxi-in times, and the airborne time (flights to and from foreign destinations are not covered). From this initial dataset, we remove canceled or diverted flights as well flights from/to US

[12] The data files for the regressions in the paper are available at https://data.mendeley.com/datasets/72z2mb843k/1.

[13] See https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGK for the former and https://www.wunderground.com for the latter source of data.

Commonwealth areas and Territories. Flights with an arrival or departure delay greater than 120 min are also excluded to prevent outliers from affecting our results.[14]

Our regressions have a panel structure, with the time dimension corresponding to the days of 2018 and the cross sectional dimension corresponding to aircraft with different tail numbers. Accordingly, the regressions include date fixed effects as well as tail-number fixed effects.

To calculate the ground buffer using (6), we first compute the scheduled ground time $t_{g,i}$ that separates flight i and flight $i-1$ in the sequence of flights operated by the observed aircraft during the day. 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 flight i is then obtained by subtracting this minimum feasible ground time from the scheduled ground time, as in (6).[15] Finally, as in BCG (2021a), we exclude from the sample aircraft that operate more than 8 flights during the day.[16]

The weather data are obtained from Weather Underground, a commercial meteorological service that provides weather reports at each airport roughly every half-hour or every hour, depending on the airport. Among various meteorological metrics, each weather report shows the general weather condition (i.e., sunny, cloudy, rainy, etc.) and the temperature, which are two relevant pieces of information used in the regressions. The weather data are included in our sample by matching the actual take-off and landing times of each flight with the latest available weather report at the origin and destination airports, respectively. For example, if a flight leaves at JFK at 9:15 AM and arrives at LAX at 12:15 PM, then JFK hourly weather at 9:00 AM and LAX weather at 12:00 PM are the origin and destination weather observations for the flight.[17] To reduce the number of weather variables, we consolidate the variables for snow and storm conditions, which appear to capture the most severe weather conditions, into a Snow/Storm variable. Storm conditions consist of thunderstorms or other storms with high winds.

Table 1 presents summary statistics for the sample, which contains more than 4.2 million flights. The mean arrival delay, listed in the second row, is -0.61 minutes, reflecting the fact that, despite widespread concerns about delays, most flights arrive early. This fact is also illustrated in the darkly shaded histogram in the first panel of Fig. 1, which shows that the mode of the delay distribution lies in the negative range (the minimum is -49 min). A long positive tail exists, however, which is truncated at 120 min per our sample restriction.

As seen in Table 1, the mean (total) departure delay is positive at 4.53 min. The lightly shaded histogram in the first panel of Fig. 1, which shows departure delays, reveals that negative delays are again common, occurring when all passengers have boarded and an early departure is otherwise feasible. But since aircraft cannot leave too far in advance of their scheduled departure, the minimum departure delay is -19 minutes, with most values above -10 minutes.

The difference between the arrival and departure-delay distributions is shown in the second panel of Fig. 1, which confirms that the distributions have very similar positive tails while showing that, in the negative range, departure delays are more concentrated near zero. Returning to Table 1, the first row shows that the mean arrival delay of -2.24 min for previous flights (flight $i-1$) is even more negative than the flight-i mean, a consequence of the fact that previous flights are earlier in the day on average, when delays are less common.

Propagated departure delays, given by $D_{d,i}^{prop}$ in (9), have a mean of 1.73 min (by definition, they must be positive) and maximum of 113 min. Other departure delays, given by $D_{d,i}^{other}$ and equal to the difference between total and propagated delays, have a mean of 2.80 min and a maximum at the truncated value of 120 min.

The remaining rows of Table 1 show summary statistics for the different versions of the R ratio, which capture the contribution of propagated departure delays to arrival delays. The flight-based R_i^{flight}, which is defined only for positive arrival delay, has a mean of 0.127, showing that minutes of propagated departure delay are on average about 13% of arrival-delay minutes. Thus, for flights that arrive late, the contribution of propagated departure delay is fairly modest in size when delays are measured in minutes.

Note that one would expect R_i^{flight} to be less than 1 on average since other sources beyond delay propagation contribute to departure delay and hence arrival delay. However, favorable conditions (high tailwinds, for example) could mostly eliminate the effect of a departure delay, leading to the flight to arrive nearly on time. With its denominator near zero, the magnitude of R_i^{flight} could then be very large, as seen in the maximum value of 44 from Table 1. However, cases with $R_i^{flight} > 1$ are relatively rare, accounting for just 2.5% of the 1,432,430 observations with a positive arrival delay, and among these observations, the median value is 1.25, not far above 1.[18]

The next row of Table 1 shows the route-based R, computed at the route × month level. The mean value of 0.393 indicates that, among flights with an arrival delay longer than 15 min, 39% had a propagated departure delay. The subsequent row of the table shows R_r^{route} instead computed at the route × week level, which has a very similar mean equal to 0.388.

[14] We also excluded flights below the first percentile of the arrival or departure delay distributions (very early arrivals or very early departures).

[15] Since the Marketing Carrier On-Time Performance dataset does not include international flights, except those from/to US Commonwealth areas and Territories, we restrict the ground buffer to be not greater than 200 min, to exclude possible incomplete records that may arise when an aircraft flies to a non-domestic destination (e.g., Canada or Mexico, returning to the same US airport) between two domestic flights (the resulting time gap would be incorrectly identified as a ground buffer). Furthermore, in order to have confidence in the accuracy of the observed minimum ground time used in the buffer computation, we require that the number of observations used to generate it is at least equal to 30.

[16] This restriction mostly removes ‘ping-pong’ flights, which are flights that, multiple times in a day, operate a very short-haul route linking a hub to non-hub airports (e.g., Hawaiian flights between Honolulu and the other Hawaiian islands). Although these flights are generally characterized by a short ground buffer, delay propagation is not an issue.

[17] A similar approach is adopted by Bubalo and Gaggero (2015, 2021).

[18] Although flight buffers tend to reduce arrival delay, suggesting that R_i^{flight} should be greater than 1 more often, other random effects that tend to increase flight times (unanticipated congestion at the destination, for example) prevent this outcome.

Table 2
R variables.

	R_i^{flight}	$R_{r,month}^{route}$	$R_{r,week}^{route}$	$\widetilde{R}_{r,month}^{route}$	$\widetilde{R}_{r,week}^{route}$
AIRLINE					
Alaska Airlines	0.071	0.286	0.279	0.136	0.148
Allegiant Air	0.191	0.480	0.492	0.125	0.128
American Airlines	0.113	0.339	0.333	0.124	0.125
Delta Air Lines	0.131	0.373	0.364	0.115	0.117
Frontier Airlines	0.196	0.457	0.455	0.136	0.137
Hawaiian Airlines	0.061	0.179	0.175	0.097	0.110
JetBlue	0.182	0.458	0.465	0.130	0.129
Southwest Airlines	0.133	0.455	0.456	0.140	0.140
Spirit Airlines	0.137	0.392	0.392	0.114	0.114
United Airlines	0.127	0.338	0.334	0.103	0.104
Virgin America	0.106	0.276	0.287	0.127	0.125
CARRIER TYPE					
Legacy carrier	0.118	0.341	0.336	0.116	0.118
Low-cost carrier	0.146	0.453	0.455	0.133	0.134
ORIGIN AIRPORT SIZE					
Large	0.094	0.307	0.306	0.137	0.137
Medium	0.120	0.366	0.367	0.127	0.129
Small	0.208	0.493	0.494	0.103	0.102
ORIGIN HUB STATUS					
Hub	0.074	0.257	0.256	0.133	0.134
Non-hub	0.158	0.439	0.437	0.119	0.120

Notes. The classification of large, medium, and small airports is based on the ranking obtained from the Department of Transportation (DOT), which reports the total number of enplaned passengers on U.S. and foreign airlines for the main U.S. airports. Large airports are top-10 airports in terms of total number of enplaned passengers in year 2018 (the year of our sample), medium airports are those ranked between positions 11 and 50, and small airports are those whose rank is below position 50 or which do not appear in the airport list provided by DOT (for the raking list see https://www.bts.gov/airport-rankings-2018, sheet 'Total').

The final two rows of the table show the route-level *R* variables based on moderate inbound arrival delays, computed at the month and week levels. The values are smaller than those of the R_r^{route} variables, indicating that the combination of a moderate inbound arrival delay and a departure delay is less common than a propagated departure delay as a cause of late arrival.

Summary statistics for our other covariates are shown in Table A.1 in the appendix. The weather and hub variables are self explanatory, while the congestion variables are equal to airport traffic in the hour of a flight's arrival or departure divided by the number of runways (see BCG, 2021a) and the competition variable equals the number of airlines competing nonstop on the route.

Table 2 shows how the mean *R* ratios vary by airline, carrier type, and the origin airport's size as well as its hub status. The flight-based *R* in the first column ranges between 0.061 (for Hawaiian) and 0.196 (for Frontier). Because Hawaiian operates mostly long distance flights where aircraft turnaround is not a large factor in delays, its small *R* may make sense. By contrast, the high *R* values for Allegiant, Frontier, and Jet Blue could reflect the small ground buffers (and thus quick turnarounds) favored by low-cost carriers, which raise the likelihood of propagated departure delays. While Southwest and Spirit have lower *R* values, low-cost carriers on average have higher *R* values than legacy carriers, as seen in the next panel of the table. Consistent with this difference, the three big network carriers (American, Delta and United) have relatively low *R* values, consistent with their less-aggressive turnaround scheduling. The regressions presented in Section 5 confirm that, when controlling for other factors, LCCs continue to have higher *R* values than other carriers.

The magnitudes of the *R* variables also depend on airport characteristics, the most important of which is the hub status of the origin airport. As can be seen in first column, non-hub origins have higher flight-based *R* values than hub origins, a pattern again confirmed by our section-5 regressions. Thus, the contribution of propagated departure delays to late arrivals is greater for flights with non-hub origins. Column 1 of the table also shows that the mean *R* varies directly with origin-airport size, although this variation may be mainly capturing a hub/non-hub difference.

Turning to the other columns of Table 2, the means of the route-level variable R_r^{route} at both the month and the week level show rankings across airlines similar to the ranking in the first column. The effects of airport characteristics are also similar. By contrast, the $\widetilde{R}_r^{route}$ variable, designed to capture the arrival-delay contribution of indirect propagated departure delays, shows different patterns. The range of values in the third column, which shows the month-based measure, is narrower than in column 2, and the ranking across carriers differs. Now, the LCC value is only slightly larger than the legacy value, and hub/non-hub and large/small patterns are reversed, with $\widetilde{R}_r^{route}$ larger at hubs and at large airports. This measure, which attempts to capture the contribution of indirect propagated departure delay, thus varies across airports and airlines differently than the variable R_r^{route}, which is based on a strict definition of propagated departure delay and thus captures the direct effect. Further discussion of these differences is presented in Section 5.

Table 3
Departure and arrival delay.

Dependent variable	(1) $D^{tot}_{d,i}$	(2) $D^{tot}_{d,i}$	(3) $D_{a,i}$	(4) $D_{a,i}$
$D_{a,i-1}$	0.303***		0.112***	
	(0.002)		(0.002)	
$D_{a,i-1} \leq 0.5b_{g,i}$		0.040***		−0.179***
		(0.001)		(0.002)
$0.5b_{g,i} < D_{a,i-1} \leq b_{g,i}$		0.535***		0.446***
		(0.005)		(0.005)
$b_{g,i} < D_{a,i-1}$		0.872***		0.694***
		(0.002)		(0.003)
Distance	0.070***	0.072***	−0.178***	−0.176***
	(0.003)	(0.002)	(0.004)	(0.003)
Non-hub origin	−0.617***	−1.131***	−0.724***	−1.261***
	(0.052)	(0.047)	(0.070)	(0.066)
Non-hub destination	−0.066	0.079*	0.520***	0.673***
	(0.049)	(0.045)	(0.067)	(0.065)
Congestion origin	0.134***	0.114***	0.237***	0.215***
	(0.003)	(0.003)	(0.004)	(0.004)
Congestion destination	0.084***	0.071***	0.256***	0.242***
	(0.003)	(0.003)	(0.004)	(0.004)
Competitors	−0.082***	−0.111***	0.034	0.004
	(0.017)	(0.016)	(0.023)	(0.022)
Morning	1.579***	1.505***	1.834***	1.742***
	(0.036)	(0.034)	(0.050)	(0.048)
Afternoon	3.833***	3.719***	4.778***	4.638***
	(0.042)	(0.041)	(0.057)	(0.056)
Late Afternoon	5.283***	5.062***	7.142***	6.886***
	(0.049)	(0.048)	(0.067)	(0.066)
Evening	4.990***	4.685***	7.010***	6.661***
	(0.050)	(0.047)	(0.070)	(0.068)
Rain origin	1.149***	1.470***	2.895***	3.238***
	(0.042)	(0.039)	(0.054)	(0.052)
Rain destination	0.856***	0.841***	4.018***	4.005***
	(0.041)	(0.039)	(0.053)	(0.051)
Fog origin	0.174***	0.258***	1.613***	1.698***
	(0.062)	(0.057)	(0.080)	(0.076)
Fog destination	1.570***	1.559***	2.957***	2.948***
	(0.072)	(0.068)	(0.091)	(0.088)
Snow/Storm origin	2.803***	3.327***	13.679***	14.243***
	(0.076)	(0.073)	(0.098)	(0.097)
Snow/Storm destination	−0.959***	−0.652***	4.564***	4.901***
	(0.069)	(0.063)	(0.093)	(0.090)
Low temperature orig.	0.547***	0.747***	0.994***	1.213***
	(0.044)	(0.041)	(0.060)	(0.058)
Low temperature dest.	0.183***	0.137***	0.541***	0.496***
	(0.044)	(0.040)	(0.058)	(0.055)
R^2	0.163	0.275	0.063	0.148
Observations	4,205,120	4,205,120	4,205,120	4,205,120

Notes. Ordinary Least Squares estimation with tail number and date fixed-effects, airport of origin fixed effects, airport of destination fixed effects; constant not reported. The estimated coefficients marked with ***, ** and * are statistically significant at, respectively the 1%, 5% and 10% level. The standard errors are clustered by tail number.

4. Preliminary regressions

As explained in the introduction, this section of the paper reports the results of preliminary regressions that have two purposes. The first is to validate the underpinnings of our approach by showing that an inbound arrival delay can lead to a late departure and late arrival for the subsequent flight. The second purpose is to explore the connection between departure delays and arrival delays, showing that the two are closely linked.

4.1. Effects of inbound arrival delays on departure and subsequent arrival delays

The first two columns of Table 3 confirm the basic mechanism of delay propagation by relating departure delay in minutes to minutes of arrival delay for the inbound flight along with a host of control variables. While column 1 shows the expected positive effect in an equation with a single inbound delay variable (whose coefficient is 0.303), column 2 allows the effect of the inbound delay to depend on its size relative to the ground buffer $b_{g,i}$.

Table 4
Arrival delay.

Dependent variable	(1) $D_{a,i}$	(2) $D_{a,i}$	(3) $Pr(D_{a,i} \geq 15) = 1$
$D^{tot}_{d,i}$	0.875***		
	(0.001)		
$D^{prop}_{d,i}$		0.763***	0.086***
		(0.002)	(0.001)
$D^{other}_{d,i}$		0.907***	0.103***
		(0.001)	(0.000)
Distance	−0.244***	−0.246***	−0.004***
	(0.003)	(0.003)	(0.001)
Non-hub origin	−0.437***	−0.289***	−0.195***
	(0.050)	(0.050)	(0.013)
Non-hub destination	0.577***	0.555***	−0.030**
	(0.048)	(0.048)	(0.013)
Congestion origin	0.110***	0.110***	0.014***
	(0.003)	(0.003)	(0.001)
Congestion destination	0.202***	0.199***	0.026***
	(0.003)	(0.003)	(0.001)
Competitors	0.110***	0.117***	0.005
	(0.015)	(0.015)	(0.004)
Morning	0.170***	0.163***	0.015
	(0.035)	(0.035)	(0.013)
Afternoon	0.857***	0.821***	0.188***
	(0.035)	(0.035)	(0.013)
Late Afternoon	1.620***	1.603***	0.412***
	(0.039)	(0.039)	(0.013)
Evening	1.391***	1.447***	0.279***
	(0.038)	(0.038)	(0.013)
Rain origin	1.619***	1.572***	0.355***
	(0.038)	(0.038)	(0.011)
Rain destination	3.552***	3.488***	0.606***
	(0.037)	(0.037)	(0.011)
Fog origin	1.133***	1.160***	0.278***
	(0.055)	(0.055)	(0.021)
Fog destination	1.721***	1.654***	0.342***
	(0.060)	(0.060)	(0.022)
Snow/Storm origin	11.705***	11.464***	1.821***
	(0.080)	(0.080)	(0.016)
Snow/Storm destination	6.090***	5.978***	1.036***
	(0.073)	(0.072)	(0.015)
Low temperature orig.	0.384***	0.359***	0.298***
	(0.043)	(0.043)	(0.015)
Low temperature dest.	0.316***	0.330***	−0.112***
	(0.040)	(0.040)	(0.015)
R^2 or pseudo R^2	0.544	0.546	0.500
Observations	4,205,120	4,205,120	1,297,732

Notes. Columns (1) and (2): Ordinary Least Squares estimation with tail number and date fixed-effects, airport of origin fixed effects, airport of destination fixed effects; constant not reported; standard errors are clustered by tail number. Column (3): Logit estimation with tail number and date fixed-effects; constant not reported. The estimated coefficients marked with ***, ** and * are statistically significant at, respectively the 1%, 5% and 10% level.

This second regression is a continuous spline specification that allows the inbound delay coefficient to be different for the cases where $D_{a,i-1} \leq 0.5b_{g,i}$ ($D_{a,i-1}$ smaller than half the buffer), $0.5b_{g,i} < D_{a,i-1} \leq b_{g,i}$ (between half the buffer and the full buffer), and $b_{g,i} < D_{a,i-1}$ (greater than the buffer).[19] Since this latter case is the one where a propagated departure delay exists according to our definition, we expect the third coefficient to be the largest of the three and near 1. This is exactly what the regression shows, with an extra minute of inbound arrival delay leading to 0.872 additional minutes of departure delay when the inbound delay is already larger than the ground buffer. By contrast, when the inbound arrival delay is less than half the size of the buffer, an additional minute

[19] If the three spline regimes are given by $D^{tot}_{d,i}=\alpha+\beta D_{a,i-1},\ \theta+\gamma D_{a,i-1},\ \mu+\tau D_{a,i-1}$, with continuity required at the break points $G_1=0.5b_{g,i}$ and $G_2=b_{g,i}$, then the departure delay regression is

$$\begin{aligned} D^{tot}_{d,i} &= \alpha+\beta[(D_{a,i-1}-G_1)I_1+G_1] \\ &\quad +\gamma[(D_{a,i-1}-G_2)I_2+(G_2-G_1)(1-I_1)] \\ &\quad +\tau(D_{a,i-1}-G_2)(1-I_1-I_2), \end{aligned}$$

where I_1=1 if $D_{a,i-1} \leq b_{g,i}$ and zero otherwise and I_2=1 if $0.5b_{g,i} < D_{a,i-1} \leq b_{g,i}$ and zero otherwise. The estimated β, γ, and τ coefficients are shown in column 2.

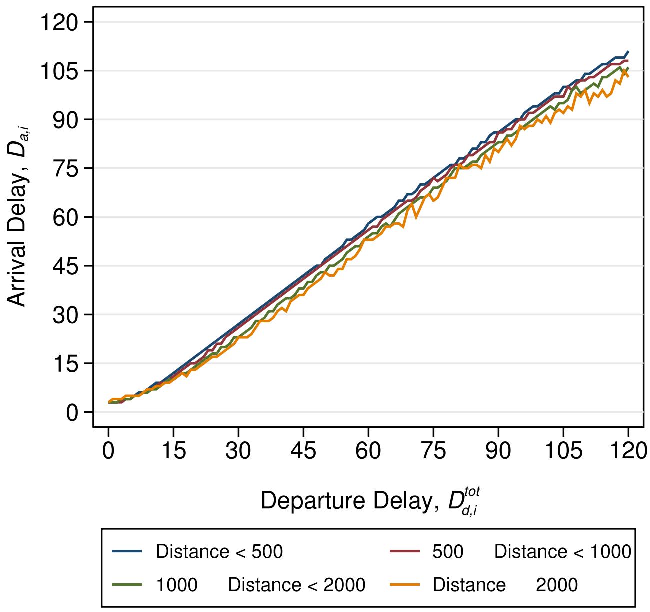

Fig. 2. Minutes of arrival delay vs. departure delay by route distance (in miles).

of inbound delay leads to a negligible increase in departure delay (0.040 min), as expected. But in the intermediate case, where the inbound arrival delay is moderate, lying between half the ground buffer and the full buffer, an extra minute raises departure delay by 0.535 min. While the inbound delay in this case is not large enough to not automatically create a departure delay, it narrows the window for avoiding a delay, making it more likely that other factors will tip the flight into a situation where it departs late. This effect is seen in the moderate magnitude of the estimated coefficient, and its existence provides the motivation for the $\widetilde{R}_r^{route}$ ratio discussed above. Table A.2 in the appendix shows that the adjacent spline coefficients are significantly different from one another in pairwise tests.

The coefficients of the control variables in columns 1 and 2 show that departure delays are higher for longer flights and when there are more competitors on the route, and that delays are mostly higher later in the day (early morning is the default). Among the variables distinguished by origin and destination values, Table A.2 shows that the origin and destination coefficients are significantly different from one another in each case. Congestion at the origin and destination both increase departure delays, with the origin effect being larger. Note that congestion at the destination could increase departure delays via the imposition of a "ground hold" on departing flights. Rain and fog at either endpoint raise departure delays, with the origin effect larger for rain and the destination effect larger for fog (which may impair landings to a greater extent than takeoffs). Again, the destination effects may arise through imposition of a ground hold. Snow or storm conditions at the origin raise departure delays, although these conditions at the destination reduce departure delays, which might reflect efforts to depart early so as to arrive prior to additional expected storms. Low temperatures at the origin or destination raise departure delays, and the larger size of the origin effect may arise because of de-icing requirements.

While these regressions confirm the mechanism leading to a propagated departure delay, the last two columns of Table 3 study a more-indirect linkage by relating the arrival delay of flight i ($D_{a,i}$) to the inbound arrival delay of the previous flight ($D_{a,i-1}$). The first regression has a single delay coefficient (equal to 0.112), while the second regression again is a spline specification. The pattern of coefficients is similar to that in column 2, with the magnitude of the third coefficient showing that an increase in $D_{a,i-1}$ when it is already larger than $b_{g,i}$ leads to an additional 0.694 min of arrival delay. Since arrival delay is less directly linked to $D_{a,i-1}$ than is departure delay, it is natural that this coefficient is smaller than the 0.872 coefficient in column 2. The first coefficient in column 4 is unexpectedly negative, while the second is positive and moderate in size but smaller than the corresponding coefficient in column 2. Again, Table A.2 in the appendix shows that the adjacent coefficients are significantly different from one another. The coefficients of the control variables in columns 3 and 4 are discussed in the next subsection, where closely related regressions are presented.

4.2. Effect of departure delays on subsequent arrival delays

Table 4 serves the second purpose mentioned above by presenting regressions that show the connection between arrival delays and departure delays. These regressions, which follow Eq. (4) in Section 2, provide quantitative evidence on the arrival-delay impact of propagated and other departure delays. This impact underlies the propagated-delay contribution measures R_i^{flight} and R_r^{route}, whose determinants are explored in next section of the paper.

Table 5
Flight-time regressions.

	(1)	(2)	(3)
Dependent variable	*AirborneTime*	*AirborneTime*	*AirborneTime*
Sample restrictions	None	$D^{tot}_{d,i} > 0$	$D^{tot}_{d,i} \geq 15$
$D^{tot}_{d,i}$	−0.049***	−0.059***	−0.070***
	(0.001)	(0.001)	(0.001)
Distance	11.684***	11.656***	11.665***
	(0.006)	(0.006)	(0.008)
Westbound	6.952***	7.615***	6.769***
	(0.056)	(0.062)	(0.070)
Non-hub destination	−0.073	0.072	0.099
	(0.045)	(0.062)	(0.088)
Congestion destination	0.171***	0.175***	0.191***
	(0.003)	(0.005)	(0.007)
Competitors	−0.402***	−0.615***	−0.542***
	(0.017)	(0.021)	(0.031)
Morning	−0.343***	−0.944***	−1.392***
	(0.028)	(0.065)	(0.118)
Afternoon	0.351***	−0.263***	−0.977***
	(0.030)	(0.065)	(0.120)
Late Afternoon	1.198***	0.465***	−0.564***
	(0.035)	(0.070)	(0.124)
Evening	−0.458***	−1.378***	−2.800***
	(0.031)	(0.065)	(0.121)
Rain destination	2.518***	2.414***	2.197***
	(0.031)	(0.058)	(0.087)
Fog destination	0.923***	0.904***	0.723***
	(0.050)	(0.099)	(0.151)
Snow/Storm destination	4.078***	4.228***	3.711***
	(0.051)	(0.091)	(0.129)
Low temperature dest.	−0.497***	−0.890***	−0.956***
	(0.041)	(0.076)	(0.116)
R^2	0.956	0.955	0.950
Observations	4,205,166	1,456,809	663,051

Notes. Ordinary least squares estimation with tail number and date fixed-effects, airport of destination fixed effects; constant not reported. The estimated coefficients marked with ***, ** and * are statistically significant at, respectively the 1%, 5% and 10% level. The standard errors are clustered by tail number.

Column 1 of Table 4 shows that total departure delay translates almost one-for-one into arrival delay, with an extra minute of total departure delay leading to 0.875 extra minutes of arrival delay.[20] Column 2 decomposes the total departure delay effect by allowing propagated departure delay ($D^{prop}_{d,i}$) and other departure delay ($D^{other}_{d,i}$) to have different coefficients. As can be seen, the $D^{prop}_{d,i}$ coefficient is the smaller of the two, indicating that an extra minute of propagated departure delay leads to 0.763 additional minutes of arrival delay, while an extra minute of other departure delay leads to 0.907 additional minutes of arrival delay. These coefficients are statistically different from one another, as seen in Table A.2 in the appendix. This difference evidently shows that airlines work harder to overcome the arrival-delay impact of a departure delay when the departure delay is propagated rather than caused by other factors. Carriers perhaps view lateness due to propagated departure delays as more damaging to their reputations, although the coefficient difference is not large.

While columns 1 and 2 use delays measured in minutes, it is interesting to convert arrival delay into a dichotomous variable based on the US government definition of a late flight, having a delay greater than 15 min. Using this dependent variable, column 3 of Table 4 shows a logit version of the regression in column 2, with a flight counted as late if $D_{a,i}$ is greater than 15 min. As can be seen, the $D^{prop}_{d,i}$ coefficient is once again smaller than the $D^{other}_{d,i}$ coefficient, again suggesting that airlines try harder to offset propagated departure delays than delays from other causes.

The results from columns 1 and 2 of Table 4 suggest that it is hard to overcome departure delays, which translate almost one-for-one into arrival delays. The most obvious way of overcoming a departure delay is for the aircraft to fly faster en route, and the one-for-one relationship suggests that this potential remedy is not able to help much.[21] The regressions in Table 5 provide explicit evidence for this conclusion by relating airborne time at the individual flight level to total departure delay. Regardless of whether the magnitude of $D^{tot}_{d,i}$ is unrestricted, constrained to be positive, or constrained to be greater than 15 min, the regressions show that

[20] An unreported spline regression like those in Table 3 shows that this one-for-one relationship is present over various ranges of departure delay.

[21] Aktürk et al. (2014) analyze increases in aircraft speed as a means of offsetting departure delays, recognizing the penalty of higher fuel consumption.

an extra minute of departure delay leads to a very small reduction in airborne time (between 0.049 and 0.070 min). Thus, flying faster does not provide much scope for overcoming departure delays.[22]

It should be noted that this regression includes a new dummy variable indicating a westbound flight (with a positive coefficient due to prevailing winds),[23] but it does not include origin-airport weather variables, hub status and congestion. The reason is that these variables should have no effect on flight time once an aircraft is airborne. As seen in the table, a greater distance naturally increases airborne time, while time is longer when the destination is congested or has bad weather (although low temperatures reduce it). The presence of more competitors on the route reduces flying time, and it is lower early and late in the day. Columns 2 and 3 of Table 5 show that imposing restrictions on the magnitude of departure delay has little effect on the results.

Returning to Table 4, the control-variable coefficients almost always have the same signs and significance as the coefficients in columns 3 and 4 of Table 3. Arrival delays are lower for longer flights (see Fig. 2), higher for flights with more competitors, and higher for flights later in the day, although delays are slightly lower in the evening than in the late afternoon. In addition, arrival delays are lower (higher) when the origin (destination) is a non-hub airport. In other words, arrival delays are higher for flights out of hubs, and lower for flights into hubs, with the latter effect perhaps reflecting efforts to avoid network disruption.

Among the other variables with origin and destination values, higher congestion at either endpoint raises arrival delays, with the destination effect stronger (origin and destination coefficients for all variables are almost always statistically different, as seen in appendix Table A.2). Bad weather at either endpoint raises arrival delays, with the destination effects stronger for rain and fog but the origin effect stronger for snow or storm conditions. Low temperatures at either endpoint raise arrival delays, with the origin and destination effects statistically indistinguishable in Table 4 but significantly different in Table 3, where the origin effect is larger.

Given that departure delay is being controlled for in Table 4, one might wonder how congestion and weather at the origin can still exert an effect on arrival delays. The reason is these factors can raise an aircraft's origin ground time after its gate departure, thus having an impact on eventually arrival delay.

Finally, it is useful to consider an econometric issue that may influence the magnitude of the departure-delay coefficients in Table 4. Recall from (4) that arrival delay equals departure delay plus the random flight-duration term minus the flight buffer $b_{f,i}$, with the latter factors constituting part of the regression error term in Table 4's regressions. It is possible, however, that carriers adjust the flight buffer upward in response to repeated departure delays, so as to reduce arrival delays. If so, the $-b_{f,i}$ component of the regression error term is negatively correlated with the departure-delay covariate, biasing its coefficient downward. Therefore, the fact that the 0.876 coefficient in column 1 is less than 1 may reflect this downward bias, rather than revealing a true arrival-delay impact slightly less than one-for-one. However, since this estimated coefficient is very close to one, the issue may be mostly moot. One might argue that the bias could be eliminated by including the flight buffer (which we can measure) as an additional covariate in the arrival-delay regressions. But since the buffer is endogenous (being chosen by the airline schedulers) and suitable instruments are not apparent, this remedy would simple replace one econometric issue with a new one.

5. Determinants of the arrival-delay contribution of propagated departure delays

5.1. Flight-based results

Having explored the interconnections among inbound arrival delays, departure delays, and subsequent arrival days in Sections 3 and 4, this section presents regressions showing the determinants of the R ratios, which capture the contribution of propagated departure delays to arrival delays. Table 6 presents the regression results when the flight-based R ratio is the dependent variable. Column 1 shows the basic regression, while columns 2–4 show the results under several sample restrictions. Referring to column 1, the results show that the contribution of propagated departure delays increases over the day, with the time-of-day coefficients rising monotonically across the time intervals. Recalling that its mean in Table 1 is 0.13, the increase in R_i^{flight} moving from morning to evening, equal to $0.180 - 0.042 = 0.138$, is roughly equal to the mean value. Thus, time-of-day has a large effect on the contribution of propagated departure delays, apparently reflecting the cumulative effect of delays over the day, which make a propagated delay more likely near the day's end.

In addition, a non-hub origin has a large positive effect on the contribution of propagated departure delays, increasing R by 0.104, an amount just below the variable's mean (matching the pattern in Table 2). This result evidently indicates that flights from non-hub origins have few causes of arrival delays *other than* a late inbound flight. By contrast, a hub origin offers many different sources of departure delays other than propagation, which then raise subsequent arrival delays and thus the denominator of R_i^{flight}, which in turn reduces its value. These sources include airport congestion effects, which may be incompletely captured by our congestion variables, late arrivals of incoming crews who are switching aircraft at the hub, which often create departure delays, and the holding of flights for late connecting passengers, another source of delays. With all these possible sources of departure delays, none of which involves propagation, it follows that the contribution of propagated departure delays at hubs tends to be smaller than at non-hub origins, where these sources are mostly absent, thus explaining the positive non-hub-origin coefficient.

A non-hub destination, by contrast, reduces the contribution of propagated departure delays. This result evidently reflects lower airline effort to curb arrival delays when the destination is a non-hub airport, where network disruption is less of a factor than at

[22] When departure delay and distance are interacted in regressions like those in Table 5, the resulting variable has a significantly negative coefficient, indicating that more of a departure delay can be made up by faster flying on longer flights.

[23] This variable takes the value one if the destination–origin longitude difference is greater than the latitude difference.

Table 6
Flight-based R regressions.

	(1)	(2)	(3)	(4)
Dependent variable	R_i^{flight}	R_i^{flight}	R_i^{flight}	R_i^{flight}
Sample restrictions	$D_{a,i} > 0$	$D_{a,i} > 15$	$D_{a,i} > 0,\ D_{d,i}^{tot} > 0$	$D_{a,i} > 0,\ R_i^{flight} < 1$
Distance	0.003***	0.003***	0.004***	0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)
Non-hub origin	0.104***	0.127***	0.139***	0.056***
	(0.003)	(0.004)	(0.005)	(0.002)
Non-hub destination	−0.029***	−0.031***	−0.026***	−0.017***
	(0.003)	(0.004)	(0.005)	(0.001)
Congestion origin	0.0005**	−0.001***	−0.0001	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Congestion destination	−0.005***	−0.006***	−0.005***	−0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Competitors	−0.001	0.001	−0.001	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Morning	0.042***	0.094***	0.067***	0.027***
	(0.002)	(0.003)	(0.004)	(0.001)
Afternoon	0.060***	0.152***	0.091***	0.048***
	(0.002)	(0.003)	(0.004)	(0.001)
Late Afternoon	0.107***	0.240***	0.153***	0.084***
	(0.003)	(0.004)	(0.004)	(0.001)
Evening	0.180***	0.368***	0.247***	0.131***
	(0.003)	(0.004)	(0.004)	(0.002)
Rain origin	−0.015***	−0.021***	−0.017***	−0.008***
	(0.002)	(0.002)	(0.002)	(0.001)
Rain destination	−0.047***	−0.064***	−0.055***	−0.030***
	(0.002)	(0.002)	(0.002)	(0.001)
Fog origin	0.005	0.007*	0.007	0.010***
	(0.003)	(0.004)	(0.005)	(0.002)
Fog destination	−0.035***	−0.059***	−0.042***	−0.023***
	(0.004)	(0.004)	(0.005)	(0.002)
Snow/Storm origin	−0.116***	−0.142***	−0.135***	−0.073***
	(0.003)	(0.003)	(0.004)	(0.001)
Snow/Storm destination	−0.083***	−0.112***	−0.084***	−0.064***
	(0.003)	(0.003)	(0.003)	(0.001)
Low temperature orig.	−0.024***	−0.040***	−0.027***	−0.015***
	(0.003)	(0.003)	(0.005)	(0.001)
Low temperature dest.	0.002	0.001	0.001	0.002
	(0.003)	(0.003)	(0.004)	(0.002)
R^2	0.064	0.225	0.086	0.095
Observations	1,432,411	678,230	1,010,689	1,396,135

Notes. Ordinary least squares estimation with tail number and date fixed-effects, airport of origin fixed effects, airport of destination fixed effects; constant not reported. The estimated coefficients marked with ***, ** and * are statistically significant at, respectively the 1%, 5% and 10% level. The standard errors are clustered by tail number.

hubs. More sources of arrival delay then raise the denominator of R_i^{flight}, reducing contribution of propagated departure delays. The effect is smaller in absolute value, however, than that of a non-hub origin.

As for the other covariates in column 1, distance has a positive effect, and while its source is unclear, the magnitude is small, with an extra thousand miles of distance (yielding a 10-fold increase in the variable) raising R_i^{flight} by 0.03. The effects of airport congestion are small and mixed in sign, possibly because congestion is mostly captured by controlling for the hub status of both endpoints, and the level of competition on the route has no effect on the dependent variable. The weather variables have mostly negative and significant coefficients. The apparent reason is that bad weather at either endpoint generates arrival delays for a host of reasons, thereby raising the denominator of R_i^{flight} and thus reducing the contribution of propagated departure delays.

Table A.2 in the appendix shows that, for each control variable in column 1 with origin and destination values, the coefficients differ significantly from one another. This difference is especially noteworthy in the cases of the fog and low-temperature variables, where either the origin or destination coefficient is insignificant. For fog, its presence at the destination may create an arrival delay unrelated to propagation (reducing R_i^{flight}) while its presence at the origin may not interfere as much with takeoffs, leaving the ratio unaffected. As note above, a low temperature at the origin may create a de-icing delay unrelated to propagation, again reducing R_i^{flight}, while a low temperature at the destination has no effect on the ratio.

The results in columns 2–4 of Table 6, which are based on various sample restrictions, are similar to those in column 1, showing the robustness of those findings. While arrival delays are restricted to be positive in column 1, further restricting them to exceed 15 min, as in column 2, has very little effect on the estimated coefficients despite a more than 50% reduction in the sample size. Returning to positive arrival delays but requiring the total departure delay to be positive, as in column 3, again has little effect on the results. Dropping the positive-departure-delay requirement but requiring R_i^{flight} to be less than 1 (which reduces the sample size

Table 7
Route-based R regressions.

Dependent variable	(1) $R^{route}_{r,month}$	(2) $R^{route}_{r,month}$	(3) $R^{route}_{r,week}$	(4) $R^{route}_{r,week}$	(5) $\widetilde{R}^{route}_{r,month}$	(6) $\widetilde{R}^{route}_{r,week}$
Low-cost carrier	0.052	0.084***	0.029	0.059***	0.064***	0.067***
	(0.037)	(0.030)	(0.026)	(0.021)	(0.021)	(0.016)
Non-hub origin	0.131***		0.119***		−0.030	−0.044**
	(0.038)		(0.027)		(0.026)	(0.019)
Non-hub destination	−0.048		−0.037		−0.032	−0.053***
	(0.038)		(0.027)		(0.021)	(0.015)
Congestion origin	0.001	0.001	0.002**	0.002**	0.002**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Congestion destination	−0.005***	−0.005***	−0.005***	−0.005***	0.002**	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Competitors	0.004	0.004	0.009*	0.009*	0.004	0.004
	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.003)
Morning	0.041*	0.043**	0.037**	0.038***	0.018	0.018*
	(0.021)	(0.021)	(0.015)	(0.015)	(0.014)	(0.010)
Afternoon	0.074***	0.077***	0.061***	0.064***	0.006	0.017*
	(0.022)	(0.022)	(0.015)	(0.015)	(0.014)	(0.010)
Late Afternoon	0.114***	0.116***	0.115***	0.117***	0.024	0.028***
	(0.022)	(0.022)	(0.015)	(0.015)	(0.015)	(0.010)
Evening	0.169***	0.170***	0.163***	0.165***	0.020	0.029***
	(0.022)	(0.022)	(0.015)	(0.015)	(0.015)	(0.010)
Rain origin	−0.036	−0.037	0.033***	0.033***	0.050***	0.033***
	(0.025)	(0.025)	(0.010)	(0.010)	(0.016)	(0.007)
Rain destination	−0.091***	−0.091***	−0.080***	−0.080***	0.027	0.008
	(0.025)	(0.025)	(0.010)	(0.010)	(0.017)	(0.007)
Fog origin	0.034	0.033	0.059***	0.058***	−0.003	0.022**
	(0.031)	(0.031)	(0.014)	(0.014)	(0.017)	(0.009)
Fog destination	−0.016	−0.016	−0.009	−0.009	0.017	0.002
	(0.032)	(0.032)	(0.014)	(0.014)	(0.020)	(0.010)
Snow/Storm origin	−0.263***	−0.264***	−0.201***	−0.201***	0.006	0.024***
	(0.034)	(0.035)	(0.013)	(0.013)	(0.025)	(0.009)
Snow/Storm destination	−0.186***	−0.185***	−0.163***	−0.163***	0.016	−0.004
	(0.036)	(0.036)	(0.014)	(0.014)	(0.024)	(0.010)
Low temperature orig.	−0.043***	−0.043***	−0.011*	−0.011*	−0.001	0.008*
	(0.012)	(0.012)	(0.007)	(0.007)	(0.008)	(0.004)
Low temperature dest.	0.102***	0.102***	0.067***	0.067***	0.021**	0.016***
	(0.012)	(0.012)	(0.007)	(0.007)	(0.008)	(0.004)
R^2	0.013	0.012	0.008	0.007	0.002	0.001
Observations	55,717	55,717	177,959	177,959	55,717	177,959

Notes. Ordinary least squares estimation with route fixed-effects; constant not reported. The panel identifier is route r in all columns, whereas the temporal dimension of the panel is month in columns (1), (2) and (5) or week in columns (3), (4) and (6). The estimated coefficients marked with ***, ** and * are statistically significant at, respectively the 1%, 5% and 10% level. The standard errors are clustered by route.

by 2.5%) serves to reduce the absolute magnitudes of the coefficients without changing their qualitative implications, as seen in column 4. This outcome is expected given the smaller average size of the R variable.

5.2. Route-based results

Table 7 presents regression results using the route-based R variables. While carrier dummy variables could not be included in the regressions of Table 6 due to collinearity with the tail-number fixed effects, the route-level regressions circumvent this problem. To see how, note that, since the R variables in the table are measured at the route level, the covariates must also be measured at this level. Therefore, each regression covariate is the route-level average of the corresponding variable, with the average specific to either a month or week. In the case of carriers present on the route, this averaging would yield an average of airline dummy variables, equal to the flight shares of the different carriers. Rather than focusing on specific airlines, we instead chose to only distinguish between low-cost carriers (LCCs) and other airlines, on the belief that the different LCC business model requires a separate treatment. Accordingly, the first variable in Table 7 is the LCC flight share on the route, which is a result of the averaging process applied to LCC dummies.

Turning to the R^{route}_{r} results in Table 7, column 1 shows coefficients often similar in size and significance to those in column 1 of Table 6. The results show a similar sizable positive effect of a non-hub origin on the arrival-delay contribution of propagated departure delays. In addition, their contribution rises over the day, with the coefficients strikingly similar in magnitude to those in column 1 of Table 6, while congestion effects are again mixed. In contrast to Table 6, a non-hub destination has no effect, while greater competition strengthens the contribution of propagated departure delays. The coefficients of the weather variables are insignificant more often than before, but remain mostly negative in column 1 when significant. Table A.2 shows that coefficient equality cannot be rejected for most of the weather variables, in contrast to Table 6. The parallel week-based regression in column 3 yields results very similar to those in column 1.

The effect of the LCC variable depends on the specification of the regression, as seen by comparing columns 1 and 2 (alternatively, columns 3 and 4) of Table 7. While the LCC coefficient is insignificant in column 1, it is significantly positive in column 2, where the two hub variables have been deleted. The apparent reason for this pattern is that the LCC variable is strongly positively correlated with the non-hub origin variable, a consequence of the fact that origins and destinations are never hubs for LCCs given their point-to-point business model. With this strong correlation, the non-hub-origin variable captures both its own positive effect on R_r^{route} along with the positive LCC effect, leaving the latter coefficient insignificant. Dropping both non-hub variables, however, allows the positive LCC effect to be manifested, reflecting Table 2's difference in R_r^{route} means between LCCs and legacy carriers. The same pattern appears in the month-level regressions columns 3 and 4, where the LCC coefficient is insignificant in the presence of the hub variables but significantly positive when they are dropped.

The conclusion, then, is that propagated departure delays contribute more to arrival delays for low-cost carriers than they do for legacy airlines. This finding is apparently tied to one feature of the LCC business model: quick aircraft turnarounds, which make propagated departure delays more likely.

Columns 5 and 6 of Table 7 show regressions using the $\widetilde{R}_r^{route}$ variable. Reflecting the reverse pattern seen in Table 2, the non-hub-origin coefficient is negative in both columns and significant in column 6, instead of being positive. As in the R_r^{route} regressions, $\widetilde{R}_r^{route}$ is mostly higher later in the day, although the evening-morning differential is much smaller than in columns 1 and 2. In contrast to columns 1 and 3, the LCC coefficients are positive and significant in columns 5 and 6, presumably because the positive LCC effect and now negative non-hub-origin effect are in opposite directions, allowing both effects to be measured despite strong correlation between the variables.

To appraise these differences relative to the R_r^{route} regressions, recall that the moderate inbound arrival delay underlying $\widetilde{R}_r^{route}$ is not by itself sufficient to generate a departure delay, with contributions from other random ground-time factors also being required. Since $\widetilde{R}_r^{route}$ then captures a composite of the forces leading to arrival delays, its connections to underlying determinants cannot be expected to exactly mirror the connections between those determinants and R_r^{route}, which solely captures the contribution of propagated departure delays. The upshot is that, while the $\widetilde{R}_r^{route}$ regressions are informative, we focus on the R_r^{route} and R_i^{flight} regressions in drawing conclusions about the contribution of propagated departure delays to late arrivals.

The R_r^{route} and R_i^{flight} regressions yield three important conclusions. First, the contribution of propagated departure delays to arrival delays is higher when the origin is a non-hub airport rather than a hub. Hubs offer many different sources of departure delays unrelated to propagation, diluting its importance, while non-hub origins do not, accounting for this finding. Second, propagated departure delays contribute more to arrival delays later in the day. Thus, late inbound flights, and the propagated departure delays they generate, appear to be a more significant factor toward the end of the day, when the effects of prior delays have accumulated. Third, the arrival-delay contribution of propagated departure delays is greater in the case of LCCs than for other airlines, reflecting their different business model.

6. Conclusion

This paper has offered a simple approach for identifying propagated departure delays and measuring their contribution to arrival delays. Under our approach, a propagated departure delay occurs when the arrival delay of the inbound flight exceeds the subsequent flight's ground buffer. The size (or frequency) of such propagated delays relative to the size (or frequency) of arrival delays then measures the contribution of propagated delays to late arrivals. This approach differs from earlier attempts to quantify the contribution of delay propagation since it focuses on an individual flight and its immediate predecessor, without attempting to trace the sources of delay propagation back through the entire sequence of prior flights. The paper's empirical results show that the contribution of propagated departure delays to arrival delays depends on several key determinants.

We hope that other researchers can make use of our approach or extend it to further analyze the effect of airline scheduling decisions on on-time performance. Given the dislike of flight delays by both airlines and passengers, this subject will always be one of ongoing interest.

CRediT authorship contribution statement

Jan K. Brueckner: Conceptualization, Theoretical analysis, Writing. **Achim I. Czerny:** Conceptualization, Theoretical analysis. **Alberto A. Gaggero:** Conceptualization, Empirical analysis.

Appendix

See Tables A.1 and A.2.

Table A.1
Description and main statistics of the regressors.

Regressors	Description	Mean (Std. Dev.)
Distance	Route distance, in 100-mile units	7.846 (5.610)
Non-hub origin	Dummy variable = 1 if airport of origin is not a hub of the airline	0.618 (0.486)
Non-hub destination	Dummy variable = 1 if airport of destination is not a hub of the airline	0.680 (0.466)
Congestion origin	Number of landing and departing domestic flights at the airport of origin in the same hour when the flight is scheduled to depart divided by the number of runways of the airport of origin	12.366 (7.650)
Congestion destination	Number of landing and departing domestic flights at the airport of destination in the same hour when the flight is scheduled to land divided by the number of runways of the airport of destination	10.640 (7.533)
Competitors	Number of nonstop competitors on the route	1.045 (1.105)
Morning–Evening	Set of departure-time dummy 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)	
Rain–Snow/Storm	Set of dummy variables for the weather condition at the airport of origin or destination. $Sun/Cloud$ is the omitted category	
Low temperature orig.	Dummy variable = 1 if the temperature at the airport of origin is below 32 degrees Fahrenheit	0.059 (0.236)
Low temperature dest.	Dummy variable = 1 if the temperature at the airport of destination is below 32 degrees Fahrenheit	0.060 (0.238)
Low-cost carriers	Share of flights operated by low-cost carriers (i.e. JetBlue, Frontier Airlines, Allegiant Air, Spirit Airlines and Southwest Airlines). The statistics are for $r, month$	0.376 (0.449)
Westbound	Dummy variable = 1 if the flight is westbound	0.335 (0.472)

Table A.2
Are regression coefficients unequal?

Table 3	(1)	(2)	(3)	(4)		
Non-hub origin = Non-hub destination	✓	✓	✓	✓		
Congestion orig. = Congestion dest.	✓	✓	✓	✓		
Rain origin = Rain destination	✓	✓	✓	✓		
Fog origin = Fog destination	✓	✓	✓	✓		
Snow/Storm orig. = Snow/Storm dest.	✓	✓	✓	✓		
Low temp. orig. = Low temp. dest.	✓	✓	✓	✓		
$(D_{a,i-1} \leq 0.5b_g) = (0.5b_g < D_{a,i-1} \leq b_g)$		✓		✓		
$(0.5b_g < D_{a,i-1} \leq b_g) = (b_g < D_{a,i-1})$		✓		✓		
Table 4	(1)	(2)	(3)			
Non-hub origin = Non-hub destination	✓	✓	✓			
Congestion orig. = Congestion dest.	✓	✓	✓			
Rain origin = Rain destination	✓	✓	✓			
Fog origin = Fog destination	✓	✓	✓			
Snow/Storm orig. = Snow/Storm dest.	✓	✓	✓			
Low temp. orig. = Low temp. dest.	✗	✗	✓			
$D^{prop}_{d,i} = D^{other}_{d,i}$		✓	✓			
Table 6	(1)	(2)	(3)	(4)		
Non-hub origin = Non-hub destination	✓	✓	✓	✓		
Congestion orig. = Congestion dest.	✓	✓	✓	✓		
Rain origin = Rain destination	✓	✓	✓	✓		
Fog origin = Fog destination	✓	✓	✓	✓		
Snow/Storm orig. = Snow/Storm dest.	✓	✓	✓	✓		
Low temp. orig. = Low temp. dest.	✓	✓	✓	✓		
Table 7	(1)	(2)	(3)	(4)	(5)	(6)
Non-hub origin = Non-hub destination	✓		✓		✗	✗
Congestion orig. = Congestion dest.	✓	✓	✓	✓	✗	✗
Rain origin = Rain destination	✗	✗	✓	✓	✗	✓
Fog origin = Fog destination	✗	✗	✓	✓	✗	✗
Snow/Storm orig. = Snow/Storm dest.	✗	✗	✗	✗	✗	✓
Low temp. orig. = Low temp. dest.	✓	✓	✓	✓	✗	✗

Notes. F test for inequality of coefficients at 5% level (χ^2 test for column (3) of Table 4). The column number corresponds to the column number of the referring table. The check-mark symbol ✓ indicates that the coefficients are significantly different, whereas the cross-mark symbol ✗ is for the cases where they are not.

References

AhmadBeygi, S., Cohn, A., Guan, Y., Belobaba, P., 2008. Analysis of the potential for delay propagation in passenger airline networks. J. Air Transp. Manag. 14, 221–236.
AhmadBeygi, S., Cohn, A., Lapp, M., 2010. Decreasing airline delay propagation by re-allocating scheduled slack. IIE Trans. 42, 478–489.
Aktürk, M.S., Atamtürk, A., Gürel, S., 2014. Aircraft rescheduling with cruise speed control. Oper. Res. 62, 829–845.
Arikan, M., Deshpande, V., Sohoni, M., 2013. Building reliable air-travel infrastructure using empirical data and stochastic models of airline networks. Oper. Res. 61, 45–64.
Brueckner, J.K., Czerny, A.I., Gaggero, A.A., 2021a. Airline mitigation of propagated delays via schedule buffers: theory and empirics. Transp. Res. E 105, 102333.
Brueckner, J.K., Czerny, A.I., Gaggero, A.A., 2021b. Airline schedule buffers and flight delays: a discrete model. Econ. Transp. 26–27, 100218.
Bubalo, B., Gaggero, A.A., 2015. Low-cost carrier competition and airline service quality in Europe. Transp. Policy 43, 23–31.
Bubalo, B., Gaggero, A.A., 2021. Flight delays in European airline networks. Res. Transp. Bus. Manag. 41, 100631.
Deshpande, V., Arikan, M., 2012. The impact of airline flight schedules on flight delays. Manuf. Serv. Oper. Manag. 14, 423–440.
Diana, T., 2009. Do market-concentrated airports propagate more delays than less concentrated ones? A case study of selected U.S. airports. J. Air Transp. Manag. 15, 280–286.
Dou, L., Kastl, J., Lazarev, J., 2020. Quantifying delay propagation in networks. Unpublished paper.
Eufrásio, A.B.R., Eller, R.A.G., Oliveira, A.V.M., 2021. Are on-time performance statistics worthless? an empirical study of the flight scheduling strategies of Brazilian airlines. Transp. Res. E 145, 102186.
Fan, T., 2019. Schedule creep – in search of an uncongested baseline block time by examining scheduled flight block times worldwide 1986-2016. Transp. Res. A 121, 192–217.
Fleurquin, P., Ramasco, J.J., Eguiluz, V.M., 2013. Systemic delay propagation in the US airport network. Sci. Rep. 3, 1159.
Forbes, S.J., Lederman, M., Wither, M.J., 2019a. Quality disclosure when firms set their own quality targets. Int. J. Ind. Organ. 62, 228–250.
Forbes, S.J., Lederman, M., Yuan, Z., 2019b. Do airlines pad their schedules?. Rev. Ind. Organ. 54, 61–82.
Hao, L., Hansen, M., 2014. Block time reliability and scheduled block time setting. Transp. Res. B 69, 98–111.
Kafle, N., Zou, B., 2016. Modeling flight delay propagation: a new analytical-econometric approach. Transp. Res. B 93, 42–52.
Kang, L., Hansen, M., 2017. Behavioral analysis of airline scheduled block time adjustment. Transp. Res. E 103, 56–68.
Kang, L., Hansen, M., 2018. Assessing the impact of tactical airport surface operations on airline schedule block time setting. Transp. Res. C 89, 133–147.
Kim, M., Park, S., 2021. Airport and route classification by modelling flight delay propagation. J. Air Transp. Manag. 93, 102045.
Li, Q., Jing, R., 2021. Characterization of delay propagation in the air traffic network. J. Air Traffic Manag. 94, 201075.
Sohoni, M., Lee, Y.-C., Klabjan, D., 2011. Robust airline scheduling under block-time uncertainty. Transp. Sci. 45, 451–464.
Tan, X., Jia, R., Yan, J., Wang, K., Bian, L., 2021. An exploratory analysis of flight delay propagation in China. J. Air Transp. Manag. 92, 102025.
Wang, Y., Zhou, Y., Hansen, M., Chin, C., 2019. Scheduled block time setting and on-time performance of U.S. and Chinese airlines: A comparative analysis. Transp. Res. A 130, 825–843.
Wong, J.-T., Tsai, S.-C., 2012. A survival model for flight delay propagation. J. Air Transp. Manag. 23, 5–11.
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.