Gauging the Effectiveness of Airline Schedule Buffers in Reducing Arrival Delays

by

Jan K. Brueckner
University of California, Irvine

Achim I. Czerny
Hong Kong Polytechnic University

Alberto A. Gaggero
University of Aosta Valley

October 2024

Abstract

In a novel contribution to the literature on airline flight delays, this paper explores the effects of airline schedule buffers on arrival delays, showing that buffers reduce delays and are thus beneficial. The exercise relies on the approach followed in our previous work, which allows us to measure buffer sizes and flight delays by using government data on the flight sequences of individual aircraft. Our empirical results roughly match the predictions of an analytical model regarding the expected sizes of the buffer effects.

Gauging the Effectiveness of Airline Schedule Buffers in Reducing Arrival Delays

by

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

1. Introduction

Airline delays and delay propagation have been the focus of a growing literature in transportation engineering and economics. An arrival delay of an inbound flight inconveniences its own passengers, but lateness may also lead to delay propagation, where the flight's late arrival also leads to a late departure of the aircraft's next flight and then its late arrival. Late arrival of inbound aircraft is in fact the major cause of subsequent arrival delays, based on government data cited by Brueckner, Czerny and Gaggero (2021a). One remedy for arrival delays as well as delay propagation is the use of schedule buffers, which are increments added to aircraft flight and ground times, so that they exceed minimum feasible times. A larger flight buffer for the incoming flight reduces the chance that random flight-time shocks lead to its late arrival as well as a possible subsequent departure delay. A larger ground buffer also reduces the chance of a late departure and subsequent delay by absorbing the impact of a late-arriving inbound flight as well as the impact of random ground-time shocks. A larger flight buffer for the subsequent flight helps to absorb a late departure as well as the effects of unfavorable flight-time shocks.

The goal of the present paper is one that has not been pursued, to our knowledge, so far in the literature. In particular, we investigate the *empirical effectiveness of airline flight and ground buffers in reducing arrival delays.* We ask if buffers are actually beneficial, and if so, how big are the arrival-delay effects? In particular, we are interested in measuring the minutes of reduced arrival delay resulting from an extra flight-buffer minute as well as the delay reduction from an extra ground-buffer minute. Our approach uses the analytical and empirical frameworks from our previous papers, with the measurement of flight delays as well as the lengths of flight and ground buffers relying on US government data on the flight sequences of individual aircraft. These data allow us to compute flight-buffer lengths as the difference

* We thank Mark Hansen and Bo Zou for comments, but the usual disclaimer applies.

between scheduled and minimum feasible flight times, and to compute ground-buffer lengths as the difference between scheduled and minimum feasible ground times. To measure the desired effects, we use a regression equation relating a flight's arrival delay to the size of its flight buffer, the size of the ground buffer preceding the flight, and the incoming flight's arrival delay, along with a host of other control variables. The results show that a minute increase in either the flight or ground buffer leads to *less than a minute's reduction* in arrival delay, but that the flight buffer has a stronger effect, with the latter result matching the predictions of our model. Our results thus show that airline schedule buffers are productive, reducing flight delays, so that an airline's investment of resources in generating them is worthwhile.

A key aspect of our empirical approach is recognition of the likely endogeneity of both the flight and ground buffers, which would lead to biased estimates of the buffers' effects on arrival delays if not corrected. The problem is that the airline would choose large buffers for flights whose unobservable characteristics make them prone to arrival delays, making the buffers positively correlated with the regression error term and leading to upward biased coefficients. Our correction makes use of instrumental variables, which are equal to the average buffer values for different flights whose route and airport characteristics are similar to those of the observed flight.

Our paper connects to a large existing literature, which is surveyed by Li, Mao and Li (2024). Previous papers that investigate the optimal lengths of schedule buffers include AhmadBeygi, Cohn and Lapp (2010), Sohoni, Lee and Klabjan (2011), Deshpande and Arikan (2012), Arikan, Deshpande and Sohoni (2013), Kafle and Zou (2016), Kang and Hansen (2017), and Brueckner, Czerny and Gaggero (2021a,b). In the two latter papers (referred below to as BCG 2021a,b), we use highly stylized economic models rather than realistic engineering-style frameworks to investigate the optimal buffer choices. Based on their theory, BCG 2021a investigate the empirical determinants of buffers sizes, and Kafle and Zou (2016), Kang and Hansen (2017) also carry out empirical exercises related to their models.

A large group of additional papers is purely empirical. Papers that investigate the determinants and delay impacts of scheduled "block" times (gate-to-gate times), which are closely related to the length of flight buffers, include Hao and Hansen (2014), Kang and Hansen (2017)

Wang, Zhou, Hansen and Chin (2019), and Fan (2019). Other papers ask how airport and network characteristics affect delays and delay propagation, and they include AhmadBeygi, Cohn, Guan and Belobaba (2008), Diana (2009), Fleurquin, Wong and Tsai (2012), Ramasco, and Eguiluz (2013), Dou, Kastl and Lazarev (2020), Bubalo and Gaggero (2021), Kim and Park (2021), Li and Jing (2021), and Tan, Jia, Yan, Wang and Bian (2021). Forbes, Lederman and Yuan (2019) (a paper that, like BCG's work, is written by economists) investigates the determinants of "schedule padding" (longer block times), while Forbes, Lederman and Wither (2019) explore how required posting of on-time performance affected this phenomenon. Finally Brueckner, Czerny and Gaggero (2022) expose the factors leading to a high contribution of propagated departure delays to arrival delays.

The plan of the paper is as follows. Section 2 presents the analytical framework used to derive the estimating equation, drawing on our previous work. Section 3 discusses the data and the construction of the instruments, while section 4 presents the regression results. Section 5 offers conclusions.

2. Analytical framework for deriving the estimating equation

As a basis for the subsequent empirical investigation, this section of the paper derives an estimating equation that relates the arrival delay for flight i to the arrival delay for flight $i-1$ (the incoming flight) and the ground and flight buffers, using the analytical framework of our previous papers. The arrival delay for flight i equals

$$D_{a,i} \;=\; \widehat{t}_{a,i} \;-\; t_{a,i}, \tag{1}$$

where $\widehat{t}_{a,i}$ is the actual arrival time and $t_{a,i}$ is the scheduled arrival time. Note that an early arrival generates a negative delay. The two arrival times in (1) are given by

$$\widehat{t}_{a,i} \;=\; \widehat{t}_{d,i} \;+\; m_i \;+\; \epsilon_{f,i}, \qquad t_{a,i} \;=\; t_{d,i} \;+\; m_i \;+\; b_{f,i}, \tag{2}$$

where $\widehat{t}_{d,i}$ and $t_{d,i}$ are the actual and scheduled departure times, m_i is the minimum feasible flight duration, $\epsilon_{f,i}$ is a positive random shock capturing deviation from the minimum duration, and $b_{f,i}$ is the flight buffer, the amount added to the minimum flight duration to generate the scheduled arrival time.

Substituting (2), $D_{a,i}$ in (1) is then given by

$$D_{a,i} \;=\; \widehat{t}_{d,i} \;-\; t_{d,i} \;+\; \epsilon_{f,i} \;-\; b_{f,i} \;=\; D_{d,i} \;+\; \epsilon_{f,i} \;-\; b_{f,i}. \tag{3}$$

where $D_{d,i} = \widehat{t}_{d,i} - t_{d,i}$ gives flight i's departure delay. Eq. (3) then says that the flight's arrival delay equals its departure delay plus the difference between the flight-time shock and the flight buffer.

The goal of the next steps in this analysis is to derive an expression that allows the departure delay $D_{d,i}$ in (3) to be written in terms of other variables. To do so, note that the actual and scheduled departure times in (3) are given by

$$\widehat{t}_{d,i} \;=\; \max\{\widehat{t}_{a,i-1} \;+\; n_i \;+\; \epsilon_{g,i},\; t_{d,i}\}, \qquad t_{d,i} \;=\; t_{a,i-1} \;+\; n_i \;+\; b_{g,i}, \tag{4}$$

where $\widehat{t}_{a,i-1}$ and $t_{a,i-1}$ are the scheduled and actual arrival times for flight $i-1$, n_i is the minimum feasible ground time prior to flight i, $\epsilon_{g,i}$ is a positive shock that determines actual ground time above the minimum, and $b_{g,i}$ is the ground buffer prior to flight i, which is added to n_i to get scheduled ground time and thus the scheduled departure time of flight i.[1]

Note that the first term inside the max expression in (4) gives the earliest time at which flight i can depart, equal to the arrival time of the incoming flight plus the minimum ground time plus the ground-time shock. If this expression is less than $t_{d,i}$, flight i could in principle depart early but does not (departing at $t_{d,i}$), so as to respect the schedule. When this term exceeds $t_{d,i}$, however, the flight cannot depart on time, and a departure delay occurs. The magnitude of the departure delay $D_{d,i}$ in this case equals $\widehat{t}_{a,i-1} + n_i + \epsilon_{g,i} - t_{d,i}$.

To compute the departure delay $D_{d,i} = \widehat{t}_{d,i} - t_{d,i}$ using (4), $t_{d,i}$ is subtracted from both elements of the max expression in (4), yielding

$$D_{d,i} \;=\; \widehat{t}_{d,i} - t_{d,i} \;=\; \max\{\widehat{t}_{a,i-1} - t_{a,i-1} + \epsilon_{g,i} - b_{g,i},\; 0\} \;=\; \max\{D_{a,i-1} + \epsilon_{g,i} - b_{g,i},\; 0\}, \tag{5}$$

[1] The ground-time shock was mostly absent in BCG's previous papers, being added only as an extension in the analysis of the discrete model of BCG (2021b).

where $D_{a,i-1} = \widehat{t}_{a,i-1} - t_{a,i-1}$ is the arrival delay for flight $i-1$ (n_i cancels). From (5), if the incoming arrival delay plus the ground-time shock minus the ground buffer exceeds zero, then the departure delay equals the value of that expression. Otherwise, no departure delay occurs. Note that an early arrival for the incoming flight (a negative value of $D_{a,i-1}$) can help to offset an unfavorable value of $\epsilon_{g,i}$, the ground-time shock, reducing the likelihood of a departure delay.

Substituting (5) in (3) gives the desired expression for flight i's arrival delay, which depends on the incoming arrival delay and both buffers. The expression, which forms the basis for the following empirical work, is

$$D_{a,i} \;=\; \max\{D_{a,i-1} \;+\; \epsilon_{g,i} \;-\; b_{g,i},\; 0\} \;+\; \epsilon_{f,i} \;-\; b_{f,i}. \qquad (6)$$

Converting (6) into a regression equation yields

$$D_{a,i} \;=\; \alpha \;+\; \beta D_{a,i-1} \;+\; \gamma b_{g,i} \;+\; \theta b_{f,i} \;+\; X_i\delta \;+\; u_i, \qquad (7)$$

where X_i is a vector of route, airport and daily weather characteristics that may affect arrival delays and u_i is an error term that captures the influence of the random shocks $\epsilon_{g,i}$ and $\epsilon_{f,i}$ as well as other unmeasured factors.

The likely magnitudes of the estimated coefficients $\widehat{\beta}$, $\widehat{\gamma}$, and $\widehat{\theta}$ from (7) can be predicted from (6). First, the estimated coefficient $\widehat{\theta}$ of $b_{f,i}$ should be close to -1, its coefficient in (6). In other words, an increase in the flight buffer should reduce flight i's arrival delay in one-for-one fashion. However, while an increase in $b_{g,i}$ will decrease $D_{a,i}$ in one-for-one fashion when the max expression in (6) is larger than zero (when flight i's departure is delayed), a higher ground buffer will have no effect when the expression is less than zero, in which case there is no departure delay. As a result, we would expect the estimated $b_{g,i}$ coefficient, which captures an average effect, to be smaller than (instead of equal to) 1 in absolute value, with $-1 < \widehat{\gamma} < 0$. Thus, because the size of $b_{f,i}$ always matters while the size of $b_{g,i}$ only matters in the presence of a departure delay, the flight buffer's beneficial effect on arrival delay should be stronger than that of the ground buffer.

Similarly, since the magnitude of the incoming arrival delay only matters when flight i's departure is delayed, the estimated coefficient $\widehat{\beta}$ of $D_{a,i-1}$ should satisfy $0 < \widehat{\beta} < 1$, again being less than 1 in absolute value. Moreover, since the coefficients of $D_{a,i-1}$ and $b_{g,i}$ in (6) are symmetric (equal to 1 and -1), it follows that the (positive and negative) estimated coefficients $\widehat{\beta}$ and $\widehat{\gamma}$ should be equal in absolute value.

In estimating (7), the ground and flight buffers must be treated as endogenous. The reason is that the buffers are likely to be correlated with the regression error term u_i. To see why, observe that a particular flight will experience unobservable factors that influence arrival delays, which are embodied in u_i. For example, these unobservables could include unmeasured characteristics of the destination airport, such as its location with respect to other (especially nearby) airports, which may affect in-route delays. The airline is likely to set the ground and flight buffers taking these unobservables into account, so that the buffers are then correlated with the error term u_i. Since unobservables that tend to make u_i and hence $D_{a,i}$ large will lead the airline to set large values for the buffers, the direction of this correlation is positive. Such correlation leads to upward bias in the estimated buffer coefficients $\widehat{\gamma}$ and $\widehat{\theta}$, whose likely negative values are then biased toward zero. An instrumental-variables approach is then needed to generate consistent estimates, and the choice of instruments is discussed below.

It perhaps could be argued that $D_{a,i-1}$, the arrival delay of the incoming flight, is also correlated with the regression error term. For example, unmeasured characteristics of flight i's origin airport (the destination airport for the incoming flight $i-1$) could affect both the arrival delay for that flight as well as the arrival delay for flight i, which would lead to such correlation. This correlation could bias the coefficients of both $D_{a,i-1}$ and the buffers. However, such an issue, if it exists, seems to be of second-order importance compared to the endogeneity of the buffers. As a result, we do not seek instruments for $D_{a,i-1}$.

3. Data and generation of the instrumental variables

The sample used in this paper covers US domestic airline operations for the year 2018 and is constructed by combining information from two main data sources: the US Department of

Transportation (DOT) and Weather Underground.[2] This section provides details, replicating as needed the data descriptions in BCG (2021a, 2022).

3.1. DOT data

The US DOT data are from the 'Marketing Carrier On-Time Performance' dataset, which reports information on the sequence of flights operated by individual aircraft, uniquely identified by the aircraft's tail number, on a given date. For each flight in the sequence, the dataset provides information on the carrier operating the flight, the origin and destination airports, 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 Commonwealth areas and Territories.

To calculate ground buffers, we first compute $t_{d,i} - t_{a,i-1}$, the scheduled ground time 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. For flight i, the actual ground time is equal to $\widehat{t}_{d,i} - \widehat{t}_{a,i-1}$, where the hats denote actual values, and the minimum value is computed across all aircraft j of the same type with turnarounds at the same airport. The ground buffer $b_{g,i}$ for flight i is then obtained by subtracting this minimum feasible ground time from the scheduled ground time $t_{d,i} - t_{a,i-1}$.[3]

Similarly, the flight buffer is obtained from the actual flight time, computed as the difference between flight i's actual arrival and departure times, $\widehat{t}_{a,i} - \widehat{t}_{d,i}$. Among all aircraft of flight i's type flying the same route, we pick the smallest value of the actual flight time to obtain the minimum flight time by route and aircraft type. In analogy to the ground buffer, the flight buffer $b_{f,i}$ for flight i is then obtained by subtracting this minimum flight time from the

[2] 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.

[3] Recalling that 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 minutes so as 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).

scheduled flight time, which equals $t_{a,i} - t_{d,i}$.

Note that in order to have confidence in the accuracy of the observed minimum flight and ground times used in the buffer computations, we require that the number of observations used to generate them is at least equal to 30. Finally, as in BCG (2021a, 2022), we exclude from the sample those aircraft that operate more than 8 flights during the day.[4]

3.2. Generation of the ground- and flight-buffer instruments

The instruments for the ground and flight buffers are the average buffer values for 'similar' airports and 'similar' routes, respectively. Airport similarity is based passenger volume, using the DOT's assignment of each US commercial airport to one of four major categories (large hub, medium hub, small hub, and nonhub) based on its passenger volume. Route similarity is gauged with respect to distance. The Marketing Carrier On-Time Performance database divides routes into eleven distance groups: up to 249 miles, 250-499 miles, 500-749 miles, 750-999 miles, 1,000-1,249 miles, 1,250-1,499 miles, 1,500-1,749 miles, 1,750-1,999 miles, 2,000-2,249 miles, 2,250-2,499 miles, and above 2,500 miles.

To generate the value of the ground-buffer instrument for flight i, we calculate, for same airline and flight date, the weighted average of the ground buffers at other airports belonging to the same passenger-volume category as the origin airport of flight i, with flight i's buffer value thus excluded. The result, $\widetilde{b}_{g,i}$, is the value of the ground buffer instrument for flight i. Similarly, $\widetilde{b}_{f,i}$, the value of the flight-buffer instrument for flight i, is the weighted average of the flight buffers for the same airline and date on other routes in the same distance category as the route flown by flight i, with the buffer for flight i again dropped.[5]

Since ground (flight) buffers for similar airports (routes) involve different unobservables than those for flight i, the average values of these buffers will be uncorrelated with the flight i's unobservables and thus uncorrelated with the regression error term u_i. However, the in-

[4] 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 the minimum flight and ground times. Moreover, this 8-flight 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 short flight and ground buffer, delay propagation is not an issue.

[5] The weights used to calculate $\widetilde{b}_g$ and $\widetilde{b}_f$ are given by the daily number of operated flights by the airline.

struments will be correlated, respectively, with the ground and flight buffers of flight i because of airport or route similarity. In other words, $\widetilde{b}_g$ and $\widetilde{b}_f$ are likely to fulfill the instrument requirements of relevance (correlation with the endogenous variable) and exogeneity (lack of correlation with the error term), thus being valid instruments. The strength of the instruments is also confirmed empirically by tests shown in the upcoming econometric analysis.

3.3. Weather data and more

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. The weather data are used to indicate whether a flight is affected by different types of bad weather (rain, fog, snow, a low temperature) when it operates. As in Bubalo and Gaggero (2015, 2021) and BCG (2022), the weather observations match the actual take-off and landing times of each flight. For instance, if a flight leaves at JFK at 9:15 AM and arrives at LAX at 12:15 PM, then JFK weather at 9:00 AM and LAX weather at 12:00 PM are the origin and destination weather observations for the flight. 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.

In addition to the weather indicators, the vector X of variables in (7) includes time-of-day dummies (morning, afternoon, late afternoon, and evening, with early morning being the default), route distance, dummy variables indicating whether the origin and destination airports are hubs, and the number of competing airlines flying on the route. In addition, X includes congestion measures for the origin and destination airports, equal to the number of landing/departing flights at the airport in the same hour that the flight is scheduled to take off or land divided by the number of runways of the airport.

Table 1 shows summary statistics for the sample. Observe that, while the mean arrival delay is only 2.8 minutes, the mean delay rises to 11.2 when early arrivals are assigned a delay of zero, reflecting the offsetting effect of early arrivals in computing the first average value. Note also that on average, inbound flights are early by less than a minute. The average flight and ground buffers are respectively 30.7 and 37.6 minutes (with the average instrument values being similar), and between 30 and 40 percent of flights have either a hub origin or destinations.

4. Results

The first two columns of Table 2 show the regression results based on equations (6) and (7), with column 1 showing the OLS results and column 2 showing the two-stage least squares (2SLS) results, which make use of the instruments. The OLS and 2SLS results are qualitatively similar, but we focus on the preferred 2SLS results in column 2. As expected, both buffer coefficients are negative and significant, while the incoming arrival delay coefficient is significantly positive. Thus, longer ground and flight buffers reduce arrival delay, while a greater incoming arrival delay increases the arrival delay of the next flight, other things equal.

Despite this qualitative confirmation of the model's predictions, the coefficient magnitudes in the table do not fully align with expectations. First, while (6) implies that the coefficient of the flight buffer should equal -1, the $b_{f,i}$ coefficient in column 2 equals -0.571, showing a somewhat smaller effect on flight delays than expected. The model also predicts that the incoming-delay and ground-buffer coefficients should be smaller in absolute value than the flight-buffer coefficient, and since the coefficients of $D_{a,i-1}$ and $b_{g,i}$ are equal to 0.269 and -0.180 respectively, this prediction is confirmed. While the model also predicts that these two coefficients should be equal in absolute value, so that their sum equals zero, a significance test rejects this hypothesis even though the coefficient magnitudes are similar.

Overall, the model's predictions thus fare reasonably well in light of the empirical results. The signs of the key coefficients all match predictions. In addition, the flight buffer has a larger absolute effect on arrival delay than either the incoming arrival delay or the ground buffer, and the two latter opposing effects are close in absolute value. A further confirmation of expectations comes from comparing the OLS results to those of 2SLS. Recall that we expect the OLS buffer coefficients to be upward biased (biased toward zero) as a result of positive correlation between the buffers and the regression error term. Comparing columns 1 and 2 of Table 2, this is exactly the outcome we see, with the OLS buffer coefficients smaller in absolute value than the 2SLS coefficients. This finding increases confidence in the appropriateness of our analytical framework.

Table 3 shows the first-stage results underlying the 2SLS regressions of Table 2. With two endogenous variables (the flight and the ground buffers), two instruments are used, with both appearing in each of the first-stage equations. As can be seen, each instrument has a bigger effect on the variable it is meant to predict, with the ground-buffer instrument $\widetilde{b}_{g,i}$ having a larger effect than the flight-buffer instrument $\widetilde{b}_{f,i}$ in the ground-buffer regression (column 1) and vice versa in the flight-buffer regression (column 2). In addition, the instruments are strong, with the Kleibergen-Paap Wald F statistic for the combination of the two instruments taking a large value, as seen at the bottom of column 2 in Table 2.

Returning to Table 2, the arrival-delay effects of the X covariates deserve discussion before turning to the results in columns 3 and 4. Column 2 shows that arrival delay is higher when route is long, when the origin is a hub, when either the origin or destination is congested, and when the time of operation is after the early morning (the default category). Moroever, since the time-of-day coefficients increase monotically over the day, arrival delays rise as the day progresses, other things equal. In addition, bad weather of any type raises delays, as can be seen from the coefficients in the bottom third of the table.

The regressions in columns 3 and 4 explore the effect of replacing the arrival delay variable $D_{a,i}$ with $max\{D_{a,i}, 0\}$, which converts negative delays into zeros. In (6), this change converts the RHS of (6) into $max\{$previous RHS expression, $0\}$. The predicted effect of this change is to reduce the effect of the flight buffer, for which a larger value now only matters when the flight is late, having no effect for early arrivals. This prediction is upheld in Table 2, with the absolute value of the flight-buffer coefficient in column 4 reduced by half relative to column 2 (the new coefficient is -0.236 compared to -0.571). By comparison, the changes in the coefficients of $D_{a,i-1}$ and $b_{g,i}$ are relatively small, which appears consistent with the model.

The main results on the arrival-delay effects of airline schedule buffers can be summarized as follows. When negative delays are recognized, an extra ground-buffer minute reduces arrival delay by 0.180 minutes while an extra flight-buffer minute reduces arrival delay by 0.571 minutes. When negative delays are not recognized, as might be appropriate when passengers only suffer from being late, an extra ground-buffer minute reduces arrival delay by 0.170 minutes while an extra flight-buffer minute reduces arrival delay by 0.236 minutes. Schedule buffers are thus productive, so that an airline's devotion of resources to creating them is worthwhile.

5. Conclusion

In a novel contribution to the literature on airline flight delays, this paper has explored the effects of airline schedule buffers on arrival delays, showing that buffers reduce delays and are thus beneficial. The empirical exercise relies on the approach followed in our previous work, which allows us to measure buffer sizes and flight delays by using government data on the flight sequences of individual aircraft. Our empirical results roughly match the predictions of our analytical model regarding the expected sizes of the buffer effects.

While some previous papers have analyzed the optimal choice of buffer sizes, we cannot judge the optimality of the observed buffers based on the current empirical results. Doing so would require additional data on losses from the higher expenses and reduced aircraft utilization due to buffer-generated increases in scheduled flight times, as well as data on the costs of the additional gate space necessitated by longer ground buffers (and hence ground times). While such information would allow the losses from longer buffers to be compared to the gains from reduced arrival delays (which could be computed from our results), the required data are not readily available.

Table 1: Description and main statistics of the variables included in the analysis

Variable	Description	Mean (Std. Dev.)
$D_{a,i}$	Arrival delay of flight i, in minutes	2.827 (38.406)
$\max\{D_{a,i}, 0\}$	Arrival delay of flight i with early arrivals set to 0, in minutes	11.203 (34.654)
$D_{a,i-1}$	Arrival delay of flight $i-1$, in minutes	-0.129 (29.456)
$b_{g,i}$	Ground buffer for flight i, in minutes	37.643 (25.762)
$b_{f,i}$	Flight buffer for flight i, in minutes	30.667 (11.810)
$\widetilde{b}_{g,i}$	Mean ground buffer of flight i's airline across airports of similar dimension of the airport of origin of flight i, in minutes	37.343 (11.121)
$\widetilde{b}_{f,i}$	Mean flight buffer of flight i's airline across routes of similar length of route flown by flight i, in minutes	30.754 (8.482)
Distance	Route distance, in 100-mile units	7.600 (5.302)
Hub origin	Dummy variable = 1 if airport of origin is a hub of the airline	0.380 (0.485)
Hub destination	Dummy variable = 1 if airport of destination is a hub of the airline	0.314 (0.464)
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	12.348 (7.665)
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	10.623 (7.565)
Competitors	Number of nonstop competitors on the route	1.032 (1.099)
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	0.060 (0.238)
Low temperature dest.	Dummy variable = 1 if the temperature at the airport of destination is below 32 degrees Fahrenheit	0.061 (0.239)

Table 2: The effect of buffers on arrival delay

	(1)	(2)	(3)	(4)
Estimator	OLS	2SLS	OLS	2SLS
Dependent variable	$D_{a,i}$	$D_{a,i}$	$\max\{D_{a,i},0\}$	$\max\{D_{a,i},0\}$
$D_{a,i-1}$	0.260***	0.269***	0.254***	0.263***
	(0.003)	(0.004)	(0.003)	(0.004)
$b_{g,i}$	-0.086***	-0.180***	-0.069***	-0.170***
	(0.001)	(0.031)	(0.001)	(0.027)
$b_{f,i}$	-0.252***	-0.571***	-0.055***	-0.236***
	(0.003)	(0.041)	(0.002)	(0.036)
Distance	0.197***	0.667***	0.086***	0.357***
	(0.007)	(0.059)	(0.006)	(0.052)
Hub origin	1.763***	3.031***	0.730***	2.069***
	(0.103)	(0.416)	(0.090)	(0.373)
Hub destination	-0.052	0.088	-0.064	-0.116
	(0.099)	(0.143)	(0.088)	(0.126)
Congestion origin	0.292***	0.323***	0.160***	0.160***
	(0.007)	(0.017)	(0.006)	(0.015)
Congestion destination	0.352***	0.379***	0.230***	0.242***
	(0.007)	(0.009)	(0.006)	(0.008)
Competitors	0.004	-0.044	0.022	-0.031
	(0.033)	(0.038)	(0.030)	(0.033)
Morning	1.975***	1.874***	1.368***	1.349***
	(0.068)	(0.078)	(0.059)	(0.068)
Afternoon	6.036***	5.811***	4.593***	4.494***
	(0.073)	(0.086)	(0.064)	(0.076)
Late Afternoon	9.977***	10.154***	7.829***	8.003***
	(0.087)	(0.104)	(0.077)	(0.093)
Evening	10.174***	10.470***	8.288***	8.682***
	(0.088)	(0.174)	(0.078)	(0.156)
Rain origin	3.428***	3.454***	2.511***	2.524***
	(0.090)	(0.091)	(0.084)	(0.084)
Rain destination	5.547***	5.586***	4.059***	4.089***
	(0.093)	(0.093)	(0.086)	(0.087)
Fog origin	2.211***	2.215***	1.495***	1.493***
	(0.147)	(0.147)	(0.137)	(0.137)
Fog destination	4.741***	4.831***	3.804***	3.867***
	(0.164)	(0.164)	(0.152)	(0.152)
Snow/Storm origin	15.270***	15.329***	11.688***	11.708***
	(0.163)	(0.164)	(0.155)	(0.155)
Snow/Storm destination	4.995***	5.190***	3.376***	3.496***
	(0.161)	(0.163)	(0.153)	(0.154)
Low temperature orig.	2.474***	2.935***	2.241***	2.504***
	(0.107)	(0.123)	(0.098)	(0.111)
Low temperature dest.	0.993***	0.815***	0.676***	0.593***
	(0.100)	(0.106)	(0.091)	(0.096)
Test $D_{a,i-1} + b_{g,i} = 0$, F Stat.	2835.980***	9.815***	3743.674***	13.090***
Kleibergen-Paap χ^2 Stat.		463.422***		463.422***
Kleibergen-Paap Wald F Stat.		270.066***		270.066***
R^2	0.608	0.289	0.602	0.281
Observations	4,176,605	4,175,430	4,176,605	4,175,430

Notes. Estimation with tail number and date fixed-effects, airport of origin fixed effects, airport of destination fixed effects; constant not reported. Ground buffer $b_{g,i}$ and flight buffer $b_{f,i}$ are treated as endogenous variables and, in columns (2) and (4), instrumented for using, on the same departure date of flight i, $\widetilde{b}_{g,i}$ and $\widetilde{b}_{f,i}$, which are, respectively, the airline's mean ground buffer across airports of similar dimension of the airport of origin of flight i, and the airline's mean flight buffer across routes of similar length of route flown by flight i. 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

Table 3: First-stage estimates

	(1)	(2)
Estimator	OLS	OLS
Dependent variable	b_g	b_f
$D_{a,i-1}$	0.071***	0.009***
	(0.001)	(0.000)
$\widetilde{b}_{g,i}$	0.215***	0.008***
	(0.009)	(0.003)
$\widetilde{b}_{f,i}$	0.037***	0.229***
	(0.009)	(0.006)
Distance	0.043***	1.095***
	(0.016)	(0.009)
Hub origin	12.803***	0.208***
	(0.134)	(0.062)
Hub destination	-2.604***	1.303***
	(0.099)	(0.052)
Congestion origin	-0.379***	0.208***
	(0.008)	(0.002)
Congestion destination	-0.073***	0.108***
	(0.006)	(0.002)
Competitors	-0.595***	0.077***
	(0.032)	(0.018)
Morning	0.811***	-0.541***
	(0.115)	(0.025)
Afternoon	0.594***	-0.868***
	(0.122)	(0.032)
Late Afternoon	1.541***	0.111***
	(0.132)	(0.040)
Evening	4.712***	-0.445***
	(0.127)	(0.040)
Rain origin	-0.032	0.092***
	(0.058)	(0.018)
Rain destination	0.212***	0.070***
	(0.057)	(0.018)
Fog origin	-0.134	0.048
	(0.096)	(0.032)
Fog destination	0.299***	0.209***
	(0.105)	(0.034)
Snow/Storm origin	-0.174*	0.263***
	(0.089)	(0.030)
Snow/Storm destination	0.185**	0.556***
	(0.090)	(0.028)
Low temperature orig.	0.070	1.451***
	(0.075)	(0.029)
Low temperature dest.	0.359***	-0.680***
	(0.077)	(0.028)
R^2	0.2186	0.624
Observations	4,175,430	4,175,430

Notes. 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.

References

AhmadBeygi, S., Cohn, A., Lapp, M., 2010. Decreasing airline delay propagation by re-allocating scheduled slack. *IIE Transactions* 42, 478-489.

AhmadBeygi, S., Cohn, A., Guan, Y., Belobaba, P., 2008. Analysis of the potential for delay propagation in passenger airline networks. *Journal of Air Transport Management* 14, 221-236.

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.

Brueckner, J.K., Czerny, A.I., Gaggero, A.A., 2021a. Airline mitigation of propagated delays via schedule buffers: Theory and empirics. *Transportation Research Part E* 105, article 102333.

Brueckner, J.K., Czerny, A.I., Gaggero, A.A., 2021b. Airline schedule buffers and flight delays: A discrete model. *Economics of Transportation* 26-27, article 100218.

Brueckner, J.K., Czerny, A.I., Gaggero, A.A., 2022. Airline delay propagation: A simple method for measuring its extent and determinants. *Transportation Research Part B* 162, 55-71.

Bubalo, B., Gaggero, A.A., 2015. Low-cost carrier competition and airline service quality in Europe. *Transport Policy* 43, 23-31.

Bubalo, B., Gaggero, A.A., 2021. Flight delays in European airline networks. *Research in Transportation Business & Management* 41, article 100631.

Deshpande, V., Arikan, M., 2012. The impact of airline flight schedules on flight delays. *Manufacturing & Service Operations Management* 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. *Journal of Air Transport Management* 15, 280-286.

Dou, L., Kastl, J., Lazarev, J., 2020. Quantifying delay propagation in networks. Unpublished paper.

Fan, T., 2019. Schedule creep – In search of an uncongested baseline block time by examining scheduled flight block times worldwide 1986-2016. *Transportation Research Part A* 121, 192-217.

Fleurquin, P., Ramasco, J.J., Eguiluz, V.M., 2013. Systemic delay propagation in the US airport network. *Scientific Reports* 3, article 1159.

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.

Kafle, N., Zou, B., 2016. Modeling flight delay propagation: A new analytical-econometric approach. *Transportation Research Part B* 93, 52-42.

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

Kim, M., Park, S., 2021. Airport and route classification by modelling flight delay propagation. *Journal of Air Transport Management* 93, article 102045

Li, Q. Jing, R., 2021. Characterization of delay propagation in the air traffic network. *Journal of Air Traffic Management* 94, article 201075.

Li, C. , Mao, J. , Li, L., 2024. Flight delay propagation modeling: Data, methods, and future opportunities. *Transportation Research Part E* 185, article 103525.

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

Tan, X., Jia, R., Yan, J., Wang, K., Bian, L., 2021. An exploratory analysis of flight delay propagation in China. *Journal of Air Transport Management* 92, article 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. *Transportation Research Part A* 130, 825-843.

Wong, J.-T., Tsai, S.-C., 2012. A survival model for flight delay propagation. *Journal of Air Transport Management* 23, 5-11.