CENTER FOR AIRLINE ECONOMICS

WORKING PAPER SERIES

Are on-time performance statistics worthless? An empirical study of the flight scheduling strategies of Brazilian airlines

Ana Beatriz R. Eufrásio
Rogéria A. G. Eller
Alessandro V. M. Oliveira

December 2019
Aeronautics Institute of Technology
São José dos Campos, Brazil

Are on-time performance statistics worthless?
An empirical study of the flight scheduling strategies of Brazilian airlines

Ana Beatriz R. Eufrásio

Rogéria A. G. Eller

Alessandro V. M. Oliveira[✈]

Abstract

Airlines may manage their on-time performance by lengthening schedules with engineered increases in planned flight times. In Brazil, we suspect that a recent upsurge in the proportion of early flight arrivals may be the outcome of strategic additions of flight buffers by carriers, aiming to improve on-time performance statistics. However, longer times may also be the outcome of changes in the operating conditions of flights, such as the cost indexes. We develop an econometric method of high dimensional sparse (HDS) regression to decompose the extra schedule block times into operational and strategic factors, after accounting for the uncertainty in the scheduling decision-making. We estimate the impact of extra times on flight delays, allowing for moderation effects of runway congestion, slots, and propagated delay. We test and confirm the hypothesis of existence and effectiveness of schedule padding practices. Airlines apparently set longer extra times on denser routes, possibly to minimize the reputational risk contagion. We find that a 2012 on-time disclosure rule may have produced the unintended consequence of inducing the padding behavior by carriers. In contrast, slot regulation may mitigate the formation of extra block times.

Keywords: airlines; schedule padding; on-time performance; econometrics; LASSO regression.

JEL Classification: D22; L11; L93.

[✈] Corresponding author. Email address: alessandro@ita.br.

▪ Affiliations: Center for Airline Economics, Aeronautics Institute of Technology, Brazil (all authors).
▪ Acknowledgements: São Paulo Research Foundation (FAPESP) - grants n. 2013/14914-4 and 2015/19444-1; National Council for Scientific and Technological Development (CNPq) - grants n. 301654/2013-1, n. 301344/2017-5 and PIBIC; and CAPES Institutional Postgraduate degree funding. The authors wish to thank José Guerreiro Fregnani, Cristian Vieira dos Reis, Luiz André Gordo, Vitor Caixeta Santos, Paula Guimaraes, Joao B. T. Szenczuk, and Renan P. Oliveira. All mistakes are ours.

1. Introduction

"Airlines are taking a little more conservative approach to ensure they're going to arrive on time." (...) "It's part of their marketing campaigns, part of their affinity programs to develop consumer loyalty." - Sean Cassidy (vice president of the US Air Line Pilots Association labor union and Alaska Airlines pilot).[1]

Flight scheduling is one of the most important tools of airline network management, as it is a key driver of operating costs. However, carriers may have strong incentives to plan scheduled flight times not only based on cost conditions but also on the status of service quality competition in the market. Setting longer flights confers airlines with more flexibility to deal with unexpected delays and still accomplish the scheduled arrival times, a strategy labeled as "schedule padding" in the industry. By padding their schedules, airlines have often been blamed for artificially improving their on-time performance (OTP).[2]

Scheduling longer travel times may be inevitable for an airline from the flight operations standpoint. Actually, the flight management systems on modern airliners optimally determine the cruise speed of a flight in line with the cost index parameter (CI), a ratio between time-dependent costs and fuel costs. Each flight cruise is typically assigned with a speed that is within the "maximum range cruise speed" - a low speed consistent with a null CI level - and the "maximum permissible cruise speed" - a high speed consistent with the maximum CI level.[3] In the first case, the time-dependent cost is low relative to the unit fuel cost, allowing for longer flight duration and less fuel consumption; in the second case, the time-dependent cost strongly dominates the fuel cost, leading to flights with shorter duration and higher fuel consumption. In this sense, the task of scheduling flight times is strictly dictated by the relative operating costs of the airline and as a result, not all extra times added to scheduled flight times constitute real strategic time buffers.

The objective of this paper is to empirically decompose the extra times incorporated by carriers to their schedules into strategic and operational determinants. We also aim at assessing the efficiency of extra times in enhancing OTP by estimating their impacts on the odds of flight delays. We analyze the Brazilian airline industry from 2001 to 2018. In this period, the country has witnessed relevant variations on its OTP records. In 2008, the São Paulo/Guarulhos (GRU) airport, a key international gateway in the country, was considered one of the most delayed airports in the world.[4] Ten years

---

[1] "Airlines pad flight schedules to boost on-time records", USA Today, available at *www.usatoday.com*, Feb 14, 2013.

[2] See Yimga & Gorjidooz (2019).

[3] Young (2018), and Deo, Silvestre, & Morales (2020).

[4] *''The World's most-delayed airports'',* Forbes, Jan, 14, 2008.

later, however, the situation has completely changed, with GRU now ranking number 10 among the top 20 major airports with respect to OTP.[5] While the occurrence of delays in the Brazilian market has dropped from 27,5% in 2008 to 15,8% in 2018, the episodes of early arrivals have increased from 0,4% to 44,3% in the same period.[6]

We inspect the market incentives of airlines to engage in strategic flight scheduling by means of an econometric of high dimensional sparse (HDS) regression that estimates the drivers of the extra scheduled block times in Brazil. We also investigate the impact of three regulatory reforms: the 2012 on-time disclosure resolution - which made mandatory to airlines to publish their delays and cancellations statistics for each flight on their websites and other sales channel; the 2014 slots reform at major airports in the country; and finally, the introduction of a major ATM innovation - the implementation of Performance-Based Navigation (PBN) procedures in Brazilian airports since the late 2000s.

This paper aims at contributing to the recent econometric literature on airline strategic scheduling - Skaltsas (2011), Forbes, Lederman, &Yuan (2018), Fan (2019), Yimga & Gorjidooz (2019), Brueckner, Czerny, & Gaggero (2019). In particular, we study the decoupling of scheduled block times from the “unimpeded” block times, i.e., the gate-to-gate travel times accomplished under ideal flight circumstances.[7] We then test whether a lengthening in flight duration by carriers is motivated by a set of competition-related drivers that are ceteris paribus to the cost index-related factors. To the best of our knowledge, this is the first study that empirically distinguishes the strategic buffer time from the operational extra time of airlines. We therefore examine if carriers set longer scheduled block times purely as a schedule padding practice, by adding a strategic buffer to artificially improve OTP, or if such extra time is actually an unavoidable consequence of changes in operating conditions.

This paper is organized as follows. Section 2 provides a discussion of the literature on airlines scheduling decisions and the determinants of flight on-time performance. Section 3 presents the empirical model. Section 4 presents the estimation results, which is followed by the conclusions.

---

[5] Source: Official Aviation Guide (OAG), Punctuality League 2019, Jan, 2019.
[6] Source: Active Scheduled Flight Report (VRA), National Civil Aviation Agency.
[7] Yimga & Gorjidooz (2019) and Fan (2019).

2. Airline scheduling and the duration of flights

Although aircraft technology is more advanced than ever, flight duration has actually increased for several routes over the years (Fan, 2019). For example, certain non-stop flights from London Gatwick to JFK and from Heathrow to Bangkok in 2018 were lengthened by 20 minutes, when compared to their 2008 scheduled block times,[8] while other flights from Heathrow to Newark were 35 minutes longer.[9] In parallel, and not coincidentally, airlines have reached notable high OTP records. In 2012, for the first time, there were more early arrivals than disrupted flights in the US, with 20% of the major airlines' domestic flights arriving at least 15 minutes earlier.[10] So far, a limited number of econometric studies have analyzed the determinants of variations in flight times. In what follows we present details of the key findings of the previous literature.

2.1. Operational aspects of the setting of scheduled block times

Although airlines aim at building reliable schedules, some characteristics, as weather and congestion, are not fully manageable. The predictability of flight times is key for the planning of schedules by airlines, as the more uncertain actual flight times are, the longer scheduled block times tend to be (Kang & Hansen, 2017). Adverse weather conditions are one of the main external factors that potentially affect delays and add important uncertainty to travel times. The scheduled block times can vary by season, on account of the prevailing winds, and even by time of day, because of possible congestions occurring during peak hours.[11] Fan (2019) shows that, as airports and air traffic control are increasingly denser, there are more variabilities that can cause disruptions and cascade delays through other flights. Brueckner, Czerny, & Gaggero (2019) find that a higher variability of flight times, measured by the standard deviation of the actual flight times distribution, contributes to increase the additional time of flight schedules.

A longer scheduled block time may be related to infrastructure constraints. Airports' congestions and crowded airspace possibly contribute to airlines increase their actual block times, aiming at managing the risks associated with passengers' connections and flight disruptions (Fan, 2019). The OTP of a flight may be damaged by the propagation of a small flight delay that occurred much earlier (Kafle & Zou, 2016). Adding extra times to the minimum necessary time of flight and of ground turnaround operations can incorporate unexpected delays and absorb their propagation (Kafle &

[8] The scheduled block time represents the difference between the gate arrival time at the destination airport and the gate departure time at the origin airport.

[9] *"Schedule padding: Is this why air travel is getting slower?"* Which, Aug, 27, 2018.

[10] "*Airlines pad flight schedules to boost on-time records"* USA Today, Feb, 14, 2013.

[11] See a discussion in Holloway (2008).

Zou, 2016; Brueckner, Czerny, & Gaggero, 2019). Kafle & Zou (2016) analyze how flight and ground buffers can be implemented to reduce newly formed and propagated delays. They find that flight buffers usually are smaller and vary less than ground buffer, with ground operations having a greater heterogeneity than flight operations.

Mayer & Sinai (2003) and Forbes, Lederman, & Yuan (2018) note that setting longer scheduled block times imply in higher crew cost and in less efficient use of aircraft, by assigning less missions for the existing fleet. In addition, travel duration also impacts aircraft fuel efficiency. Fuel expenses are one of the main cost components of carries, and their importance increases with higher fuel price (Şafak, Atamtürk, & Aktürk, 2019). In situations of a rise in fuel prices, carries may decide to fly at slower cruising speeds to reduce fuel consumption (Fan, 2019). Even though fuel consumption reduces with slower flights, time-dependent costs, as crew costs and maintenance, increase with travel time (Edwards, Dixon-Hardy, & Wadud, 2016). There is an optimal cruise speed that correspond to the lowest flight operating costs, considering fuel, time-dependent and fixed costs, determined by the *cost index* of the flight. The cost index represents the ratio between the time-dependent costs and fuel cost per flight, being unique to the aircraft and to the airline (Young, 2018). The lower the cost index, the lower the cruise speed and also the fuel burnt. If a longer block time is a consequence of flying at slower speeds to reduce fuel consumption, the operating costs are not necessarily lower, as longer flights increase time-dependent costs. Higher fuel prices incentive airlines to improve their fuel efficiency (Zou *et al.*, 2014), and may affect travel time. However, Fan (2019) note an almost imperceptible impact of the fuel price on flight block time.

2.2. Scheduled block times and strategic time buffers

Regarding the passenger point of view, flights with shorter duration may be preferred (Kang & Hansen, 2017; Prince & Simon, 2009). In fact, shorter flights can be seen as a competitive advantage, especially in routes with intense competition (Skaltsas, 2011). But adding an extra scheduled time - a "buffer" to accommodate possible unexpected events - may have a positive impact on airline's service quality as it generates better on-time performance (OTP) statistics. Performance and reputation are important to improve passenger perception about airline service quality. Although such lengthening may attract more customers due to the guarantee of less delay (Kang & Hansen, 2017; Prince & Simon, 2009; Skaltsas, 2011), improving one dimension of airlines service quality may deteriorate others. Adding a planned extra time will increase the reliability of arrival times, but will also increase the total time of travel, not being straightforward evident if service quality will improve (Forbes, Lederman, & Yuan, 2018). Yimga & Gorjidooz (2019) associate the schedule padding practices with negative effects to the consumer welfare. They find that, although this practice generates a better OTP, its demand-increasing effect is offset by the aversion of passengers for longer scheduled flight times.

Kang & Hansen (2017) find that in a highly competitive route, airlines tend to increase their scheduled block times. Prince & Simon (2009) and Fan (2019) find evidence that carries reduce their scheduled travel time in less competitive markets, being more vulnerable to delays. Forbes, Lederman, & Wither (2019) discuss that as airlines engage in price and quality competition, if one carrier is encouraged to improve its reported OTP statistics, this would possible induce its competitors to adopt the same strategy. Miranda & Oliveira (2018) find that schedule padding allows reducing the likelihood of flight disruptions without provoking cost-related price increases.

2.3. Impacts of changes in policies and regulations on flight scheduling

Changes in policies and regulation may provide incentives for airlines to adjust their scheduled block times. Quality disclosure programs set rules to make the OTP statistics publicly available. For example, the punctuality statistics of US large airlines have been published since 1995, following the On-Time Disclosure Rule (OTDR), set by the US Department of Transportation in 1987. Shumsky (1993) observed that the scheduled block times of some domestic routes of the US market had been lengthened during the years after the OTDR implementation. Forbes, Lederman, & Wither (2019) find evidence that airlines lengthened their scheduled times as a response to the OTDR.

Air traffic management (ATM) innovations, and the resulting regulations, constitute other sources of incentives for airline scheduling adjustments. The Single European Sky ATM Research (SESAR) in Europe, and the Next Generation Air Transportation System (NextGen) in the US are examples of initiatives that aim at implementing a more efficient navigation system. They aim at implementing a paradigm shift from ground-based to a fully satellite-based ATM system, with Performance Based Navigation (PBN) being one of the tools for targeting a more precise information about the position of airplanes in the air space. Diana (2017) finds evidence that the NextGen programs and the airspace redesign in the US have possibly improved OTP since their implementation.

Airline on-time management is also influenced by the enforcement of airport slot regimes. Santos & Robin (2010) analyze the possible causes of delays in European airports, including an investigation of the impacts of slot coordination. They find evidence that delays are higher at fully coordinated airports, lower at schedules-facilitated airports, and the lowest at slot constrained ones.

3. Research design

3.1. Application

We consider the domestic Brazilian airline industry from 2001 to 2018 to develop empirical models of extra scheduled block time and flight delays. Deregulation and low-cost carrier entry have produced notable intensification of price competition in this industry in the period. As a result, the market has expanded significantly, from 29.9 million of domestic passengers transported in 2001, to 93.7 million in 2018.[12] The rapid growth put pressure on the existing airport and air traffic management (ATM) infrastructure, with higher congestion and many episodes of cascading delays in the late 2000s. As a way to address the problems associated with congestion, the authorities have engaged in a sequence of important regulatory reforms to mitigate flight disruptions in the country. First, new management procedures and regulations of the air space, with the enforcement of Performance Based Navigation (PBN), have been gradually introduced in some Brazilian airports since 2010.[13] PBN allows a satellite-based air navigation with more accurate three-dimensional flight paths and, as a result, a more optimized airspace with more direct routes. Some of the expected operational gains of PBN are the reduction of flight times and flight delays.[14]

A second governmental measure to alleviate flight disruptions in the country was the launch of a new on-time disclosure rule in 2012.[15] The main objectives of that regulation are to allow the disclosure of the characteristics of the services provided by the airlines and to enhance the transparency in the consumer relations in the industry. The new regulation then enforced the obligation that airlines have to make their delay and cancellations statistics publicly available. More specifically, carriers must disclose the percentages of delays and cancellations of each flight not only on the airlines' website, but in all their ticket distribution channels - either direct or third party. The disclosed figures must be the same calculated by the regulator, namely the National Civil Aviation Agency (ANAC), for each airline-flight stage-period combination.[16] The information must be visible since the beginning of sales process, i.e. as soon as the passenger informs the desired flight itineraries and dates.

---

[12] Source: National Civil Aviation Agency and Department of Civil Aviation's Air Transport Yearbooks (2001, 2018)

[13] According to the Aeronautical Information Circulars (AICs) released by the Department of Air Traffic Control from Brazil (DECEA).

[14] "Fact Sheet: NextGen and Performance-Based Navigation", Federal Aviation Administration, 2014. Available at: *www.faa.gov*.

[15] ANAC Resolution nº 218, Feb 28, 2012.

[16] In the face-to-face and telephone offer of the service, the information shall be presented to the consumer upon request.

The third regulatory reform was implemented during the country's arrangements to host the 2014 FIFA World Cup and the 2016 Summer Olympics, and was concerned with airport slots.[17] In the country, ANAC is the authority in charge to regulate the operation of slots at coordinated airports. The first airport declared as coordinated in the country was São Paulo/Congonhas Airport (CGH), in 1996. The 2014 regulation incorporated the concept of seasons from IATA's Worldwide Slot Management Standards. A number of major airports in the country were declared coordinated besides CGH, such as São Paulo/Guarulhos (GRU), Rio de Janeiro/Santos Dumont (SDU) and Belo Horizonte/Pampulha (PLU), among others. The new rules stipulate fines to the carrier that intentionally keeps allocated slot that it does not intend to operate, among other fines and penalties. The regulation utilizes the concept of airport "incumbent" and "entrant" airlines to define the target participation of each of these groups in the total amount of slots - at least 50% for entrants. Minimum slot usage percentages are enforced in a "use-it-or-lose-it" rule with targeted flight regularity and punctuality of at least 80% and 75%, respectively - 90% and 80% at CGH -, and 15 minutes of tolerated flight time deviation for computation of flight delays. In the rule, the agency stipulated that the concepts of "entrant" could be flexibly redefined on occasions when the expected benefit from a more intense use of airport infrastructure could be technically proven. The new rule aimed to reduce the entry barriers to competition on slot-constrained airports.

Table 1 displays the recent evolution of early and delayed flight arrivals in the country. In that table, it is possible to see a significant increase in early flight arrivals, concomitant with a decrease in delayed flights, especially since 2015 - when the proportion of early arrivals exceeds the proportion of delayed flights for the first time. We suspect that part of the explanation for such dynamics may be a consequence of on-time performance improvement accomplished by strategic scheduling practices - a hypothesis that we will formally test in the empirical modelling.

[17] ANAC Resolution nº 338, Jul 22, 2014.

Table 1 - Proportion of early and delayed flight arrivals in Brazil[18]

year	Early Arrivals	Delayed Arrivals
2007	0,3%	38,8%
2008	0,4%	27,5%
2009	0,4%	20,0%
2010	1,3%	24,0%
2011	2,6%	24,7%
2012	4,7%	21,7%
2013	5,6%	16,2%
2014	5,2%	16,4%
2015	13,9%	13,0%
2016	20,4%	12,9%
2017	19,4%	15,2%
2018	44,3%	15,8%

Notes: (i) Source: ANAC's VRA Report - domestic flights, with own calculations; (ii) A flight is computed as delayed if it arrives 15 or more minutes later than the scheduled arrival time; (iii) An arrival is set as early if the flight's arrival time is earlier than the scheduled arrival time.

Figure 1 presents the evolution of the mean actual and scheduled flight times of on some of the densest routes in Brazil.[19] In line with Table 1, we observe that the actual flight times have decreased in parallel with an increase in the scheduled flight times.

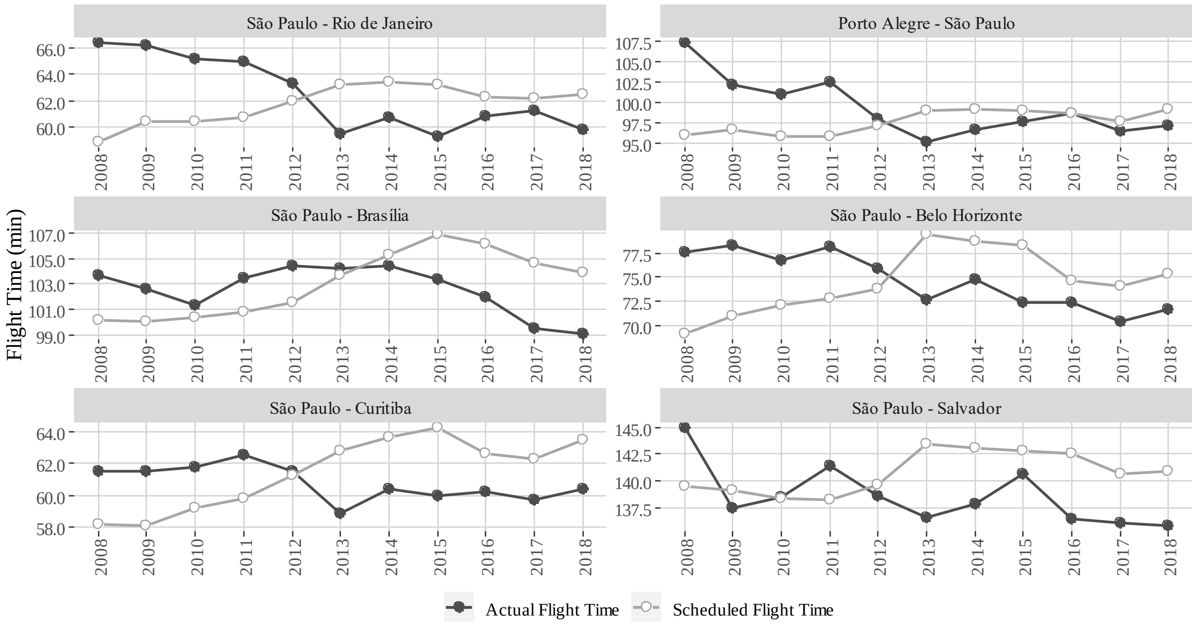

Figure 1 - Evolution of mean actual scheduled flight times of a selection of top-densest routes in Brazil

***Source:* National Civil Aviation Agency – VRA Report, with own calculations**

---

[18] We consider a flight delayed if it arrives 15 or more minutes later than the scheduled. Also, we consider an early arrival when the real arrival time is earlier than the scheduled time.

[19] The selected routes are 6 among the top 12 densest city-pairs of Brazilian domestic market in 2018.

3.2. Data

Our dataset consists of a panel of 322 routes of the domestic airline industry in Brazil, with monthly observations between January 2001 to December 2018.[20] We consider only passenger flights and define route as a directional city-pair, grouping multiple airports belonging to the same catchment area. We restrict our analysis to routes that involve two state capitals. Most data utilized are publicly available from the National Civil Aviation Agency (ANAC) in an online database. ANAC provides information on all the scheduled flights on the Active Scheduled Flight Report (VRA). We also utilized ANAC's Air Transportation Market Statistics Database, available online as well, which contain flights operational data aggregate by month of each city-pair/airline. Jet fuel price information is obtained from the National Agency for Petroleum, Natural Gas and Biofuels (ANP)'s website.

3.3. Strategic and operational components of the extra scheduled block time

We now clarify the terminology utilized in our modelling of strategic flight buffer time formation. Consider that the *actual block time* (ACTBT) of a flight depends on its operating conditions, on the status of the infrastructure of the airports, and on the intensity of competition of the route, as displayed by Equation (1).

$$\text{ACTBT} = \text{O}(\text{DIST}, \text{AIRC}, \text{CI}, \text{WEA}) + \text{I}(\text{CONG}) + \text{C}(\text{OTP}), \tag{1}$$

where O(.) are the flight operations-related drivers of block time, namely flight distance (DIST), the assigned aircraft, including its technology, aerodynamics, engine and weight (AIRC), the cruise speed chosen considering the cost index of the operation (CI) and the weather conditions, including the intensity and direction of winds (WEA). I(.) are the infrastructure-related drivers, namely the scarcity stemming from airports and ATM, which provokes their congestion effects (CONG) on flight duration. C(.) are the market incentives that impact flight times, namely the satisfaction levels and loyalty of passengers, the carrier's reputational risk and its goals related to on-time performance (OTP), assuming service-quality competition in the market. Note that we allow I(.) to be decoupled from C(.), and thus assume that although the shortage of terminal, runway and air space infrastructure may increase block times, we do not regard them as sources of strategic behavior by carriers to control their OTP.

[20] The number of sample periods is 212. Due to the special procedures for the 2014 World Cup, some months of that year were not available in the data sample.

We conceive the scheduled block time (SCHBT) as a function of the expected actual block time ($ACTBT^e$) on the occasion in which the airline takes its flight scheduling decisions. The superscript *e* means "expected at the time of flight scheduling". We model SCHBT as:

$$SCHBT = ACTBT^e + \mu, \tag{2}$$

where μ is a random error term associated with the uncertainty over the flight at the scheduling decision making. By substituting (1) into (2) and considering the airline's expectations with respect to the operating conditions, congestion and its OTP goals we then have:

$$SCHBT = O(DIST, AIRC, CI^e, WEA^e) + I(CONG^e) + C(OTP^e) + \mu, \tag{3}$$

in which we assume that the flight distance and the assigned aircraft are known by the airline at the scheduling time. The other factors, namely CI, WEA, CONG and OTP, are subject to the formation of expectations by the carrier. Let us consider the possibility that the formation of expectations about the competition in the market creates an incentive for the airline to insert a *flight time buffer* when planning its schedule. Such padding is exclusively designed to take better control of the airline's OTP. We therefore substitute C(.) by a buffer function $BUFF(.) = C(OTP^e) + \zeta$, where ζ is an error term associated with the carrier's misinterpretations of the future status of competition:

$$SCHBT = O(DIST, AIRC, CI^e, WEA^e) + I(CONG^e) + BUFF(OTP^e) + u, \tag{4}$$

where $u = \mu + \zeta$ is a component error term. Now consider that the flight operating conditions and the infrastructure-related determinants of the block time O(.) and I(.) may be redefined and decomposed into the following parts: an unimpeded block time function UNIBT(.),[21] and an optimal addition to that unimpeded time, OPTADD(.). We then have:

$$\begin{aligned} SCHBT = {} & UNIBT(DIST, AIRC, CI^{max}) \\ & + OPTADD(CI^e, WEA^e, CONG^e) \\ & + BUFF(OTP^e) + u, \end{aligned} \tag{5}$$

where CI^{max} is the maximum cost index employed on route *k* at time *t*. In the unimpeded block time (minimum feasible block time), the speed is set close to the maximum permissible cruise speed associated with the maximum CI level, i.e., a situation in which the fuel costs on that route are

[21] See Zou & Hansen (2012). Fan (2019) utilizes the term "uncongested baseline block time".

minimum related to the time-dependent costs. In this hypothetical case, the carrier minimizes the flight time, but at the cost of consuming more fuel (Young, 2018).

The optimal addition of time, OPTADD(.), represents the adjustment made to the unimpeded time as an anticipation of a lower cost index, and of possible adversities that create non-ideal operating conditions. For example, if the departure or the landing of a flight is scheduled for a busy hour at an airport, in which multiple flights operate, airlines may decide to add extra minutes to their schedules to guarantee the reliability of their operations. In contrast, the BUFF(.) part of the SCHBT equation represents the schedule padding practice, i.e., the buffer time added exclusively as a strategy to artificially improve OTP levels. Note that we consider that any flight may have an extra time added to the unimpeded block time that is “optimum” with respect to typical flight operations - the OPTADD(.) portion in (5). In this sense, OPTADD(.) is a necessary addition to minimize operating costs, given the non-ideal operating conditions that are anticipated by carriers. Without such optimal addition, the carrier would probably incur in higher than optimal cruise speed, or suboptimal climb or descent procedures. In other words, whereas BUFF(.) is an extra time mainly related to strategic factors based on market pressures, OPTADD(.) is a necessary extra time, exclusively associated with the pursuit of efficient operations by the carrier in the very common cost index situation of fuel costs not being negligible with respect to time-related costs.

Our empirical specification requires a proxy for UNIBT(.) and then moving it to the left-hand side of the equation. We then finally reach:

$$\text{EXTBT} = \text{OPTADD}(\text{CI}^{e}, \text{WEA}^{e}, \text{CONG}^{e}) + \text{BUFF}(\text{OTP}^{e}) + u, \qquad (6)$$

where EXTBT is the extra scheduled block time inserted by the airlines to their unimpeded flight time, i.e. EXTBT = SCHBT - UNIBT*, with UNIBT* being a proxy for UNIBT(DIST, AIRC, CI^{max}). Our empirical model considers EXTBT as the regressand, and a set of shifters as explanatory variables associated with the OPTADD(.) and BUFF(.) functions. Note that the actual flight distance component of Equation (5), i. e. the DIST factor - is not needed in our empirical model, as UNIBT* eventually becomes part of the regressand in (6).

3.4. Empirical model

Equation (7) presents the empirical counterpart of the extra scheduled block time model of (6), in our application to the Brazilian airline industry. Equation (8) presents our flight delays model.

$$
\begin{aligned}
EXTBT_{k,t} = {} & \beta_1\, FUELP_{k,t\text{-}h} + \beta_2\, DENS_{k,t\text{-}h} + \beta_3\, FREQ_{k,t\text{-}h} + \beta_4\, ASIZE_{k,t\text{-}h} \\
& + \beta_5\, RWYCONG_{k,t\text{-}h} + \beta_6\, SLOTPR_{k,t\text{-}h} + \beta_7\, HHI_{k,t\text{-}h} + \beta_8\, T_t \\
& + \beta_9\, T_t \times QUALREG_{k,t} + \beta_{10}\, T_t \times SLOTREG_{k,t} + \beta_{11}\, T_t \times ATMREG_{k,t} \\
& + \beta_{12}\, CASCDEL_{k,t\text{-}h} + u_{k,t},
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
ODDSDEL_{k,t} = {} & \varphi_1\, DENS_{k,t} + \varphi_2\, FREQ_{k,t} + \varphi_3\, ASIZE_{k,t} + \varphi_4\, RWYCONG_{k,t} \\
& + \varphi_5\, SLOTPR_{k,t} + \varphi_6\, HHI_{k,t} + \varphi_7\, T_t + \varphi_8\, T_t \times QUALREG_{k,t} \\
& + \varphi_9\, T_t \times SLOTREG_{k,t} + \varphi_{10}\, T_t \times ATMREG_{k,t} + \varphi_{11}\, CASCDEL_{k,t\text{-}h} \\
& + \varphi_{12}\, EXTBT_{k,t} + \varphi_{13}\, EXTBT_{k,t} \times RWYCONG_{k,t} \\
& + \varphi_{14}\, EXTBT_{k,t} \times SLOTPR_{k,t} + \varphi_{15}\, EXTBT_{k,t} \times CASCDEL_{k,t} + v_{k,t},
\end{aligned} \tag{8}
$$

where k denotes the route ($k = 1, \ldots, 322$ directional city-pairs), t the periods ($t = 1, \ldots, 212$ months) and h is a time lag to denote the horizon for flight scheduling. Below we discuss the components of (7) and (8). On the sequence, Table 2 presents the descriptive statistics and the sources of each of the main variables.

- $EXTBT_{k,t}$ is the route mean extra time added to flights in the sample period, in minutes. In other words, we compute $EXTBT_{k,t} = SCHBT_{k,t} - UNIBT^*_{k,t}$ - where $UNIBT^*_{k,t}$ is a proxy for the unimpeded block time for each flight -, and then extract the route mean for each time period. With respect to the configuration of $UNIBT^*_{k,t}$, we follow the approach of the literature - Yimga & Gorjidooz (2019), and Fan (2019). For each combination of route, airline, month and aircraft, we set the unimpeded block time as a low percentile of the actual flight times distribution.[22] In the ODDSDEL equation, this variable is included as a regressor, aiming at allowing a formal test of the effectiveness of the extra times on the management of flight delays by carriers;

- $FUELP_{k,t}$ is a proxy for the jet A1 fuel price in deflated local currency. We add to that metric a proxy for the Brazilian state tax burden on the jet fuel burned on domestic flights, with rates ranging from 3% to 25% depending on the state and on the period in the sample. An increase in the price of jet fuel may induce airlines to adjust their cost index and consequently their scheduled 4flight time, to achieve fuel consumption reductions targeting energy efficiency. It is computed as the minimum average jet fuel price observed at the endpoint cities of a route;

- $DENS_{k,t}$ is the total number of route passengers (in ten thousand). This variable is intended to capture the adjustments made to the scheduled travel times according to the route density, as

[22] We utilize the the 5th percentile but also experiment with the 10th percentile in a robustness check.

airlines are possibly more concerned with passenger satisfaction on the densest routes. The motivation for that behavior may be related to the reputational risk associated with poor OTP, which can be amplified in these markets and be quickly spread to passengers of other markets. Beyond this strategic motivation, denser routes usually have more complex flight trajectories stemming from specific ATM procedures, and also have more connecting passengers making airport terminal operations more complicated. These factors may lead airlines to schedule more extra block time;

- $FREQ_{k,t}$ is the total number of nonstop flights (in hundreds). This variable aims at capturing the fact that on city-pairs with high flight frequency, carriers possibly have to accomplish better operational efficiency with lower aircraft turnaround times. As a result, there is fewer space for flight time lengthening. In contrast, routes with higher flight frequency are more vulnerable to flight delays;
- $ASIZE_{k,t}$ is the average number of seats of aircrafts on the route. Ceteris paribus, larger jet airliners may present higher cruise speed than smaller regional airliners. Additionally, passenger and baggage handling associated with larger aircraft may be more complex and costly. Finally, although total flight costs are usually higher in large airplanes, economies of density may emerge in such way to decrease the unit costs of the non-fuel-related inputs; as a consequence, the cost index may be an indirect function of the seating capacity of the airliner;
- $RWYCONG_{k,t}$ is the percentage of flights operated at congested hours on the route. An hour is computed as "congested" when any of the endpoint airports operates more flights that its declared runway capacity, considering arrivals and departures. This variable is expected to control how airlines set their extra scheduled time according to past airport congestion;
- $SLOTPR_{k,t}$ is the percentage of flights operated at slot-constrained airports. It allows to investigate the effect of a slot regime on flight scheduling, as airlines may set shorter flight times when operating at coordinated airports;
- $HHI_{k,t}$ is the Herfindahl-Hirschman index of city-pair concentration (multiplied by 100), based on revenue passengers. It is related to the variations in the intensity airline rivalry and its effects on the incentives for scheduling aiming at strategic controlling the on-time performance (OTP) levels by carriers. As it is clearly associated to the market conduct of airlines, we believe that a formal hypothesis tests of the nullity of the coefficient of this variable may allow us to empirically assess the existence of schedule padding practices in the market;
- T_t is a time trend variable, T = 1, 2, ..., TP, where TP is the total number of sample time periods;

- $T_t \times QUALREG_t$ is an interaction variable to capture a possible structural break in the trend from the launch of the airline on-time disclosure regulation in Brazil. $QUALREG_t$ is a dummy variable, assigned with 1 from May 2012, since when new regulation has been in force. As the regulation establishes that all carriers must report their flight delays and cancelations statistics to consumers at the moment of sales, the rule may have strengthened the incentives for padding their schedules targeting better OTP;

- $T_t \times SLOTREG_{k,t}$ is an interaction term to capture a possible break in the trend after the regulatory reform of airports slots of 2014, in which stricter "use-it-or-lose-it" slot rules, at a broader set of major airports, were introduced. $SLOTREG_{k,t}$ is a dummy variable set with 1 after July 2014;

- $T_t \times ATMREG_{k,t}$ is an interaction variable to capture the effect of the introduction of new air navigation procedures and regulations of Performance Based Navigation operations from the air traffic management (ATM) authorities in a set of airports since 2010. $ATMREG_{k,t}$ is a dummy variable that indicates if a given city-pair is under PBN operations at the endpoint cities;

- $CASCDEL_{k,t}$ is a proxy for cascading delays that may affect all the flights of a given route. It is equal to the maximum proportion of delayed flights between the origin and the destination airports of the route. A flight is considered delayed if the difference between the actual and the scheduled arrival time is, at least, of 15 minutes. This variable is aimed to capture the impact of overall delays on the expectations of the schedule planners, which may add extra time to routes that historically involve more delayed airports. In the ODDSDEL equation this variable helps in the controlling of the time-varying unobservables at the level of the terminal control area, e.g. a propagation of flight delay following an adverse weather condition;

- $ODDSDEL_{k,t}$ is the log odds of flight delays, i.e., $ODDSDEL_{k,t} = \ln [DEL_{k,t}/(1-DEL_{k,t})]$, where the $DEL_{k,t}$ is the proportion of scheduled non-stop flights reported with arrival delays, divided by the total scheduled non-stop flights on the route.[23] Again, we utilize a period of 15 or more minutes do identify if a flight is delayed.

- $EXTBT_{k,t} \times RWYCONG_{k,t}$ is an interaction variable that allows us to test the moderation effects of runway congestion on the possible relationship between EXTBT and ODDSDEL in Equation (8). We therefore aim to inspect if the impact of the inclusion of extra scheduled block time on flight delays, is stronger or weaker depending on the intensity of runway congestion on the route;

[23] To simplify the visualization of the scale of the estimated coefficients, we multiply this variable by 10.

this and the next two variables aim to extend the model of Miranda & Oliveira (2018), which does not allow for moderation effects of EXTBT in their flight delays model.

- $EXTBT_{k,t} \times SLOTPR_{k,t}$ is an interaction variable to inspect existence of possible moderation effects of the slots regime on the relationship between ODDSDEL and EXTBT; here we aim to investigate if the effectiveness of the extra block times in affecting the odds of flight delays is altered on routes of slot-constrained airports;
- $EXTBT_{k,t} \times CASCDEL_{k,t}$ is an interaction variable to inspect the existence of moderation effects of cascading delays on the relationship between ODDSDEL and EXTBT; this variable aims to analyze if the effect of the extra block times on flight delays is weaker or stronger with the occurrence of propagated delays at the airport level;
- u and v are additive functions of the composite error terms;[24] and the β's and the φ's are the parameters to be estimated.

Table 2 – Descriptive statistics of the model variables

Variable	Description	Metric	Mean	Std Dev	Min	Max
$EXTBT_{k,t}$	Extra scheduled block time (i)	minutes	14.255	14.989	0	220
$FUELP_{k,t}$	fuel price (v), (viii)	BRL deflated	2.482	0.573	1.168	4.187
$DENS_{k,t}$	passengers (ii)	count, 10,000s	2.090	3.676	0.001	44.059
$FREQ_{k,t}$	flight frequencies (i), (ii)	count, 100s	2.215	3.799	0.010	41.120
$ASIZE_{k,t}$	nr of seats (i), (ii)	count	141.844	33.461	9.643	207.455
$RWYCONG_{k,t}$	runway congestion (iv)	%	12.679	19.709	0	100
$SLOTPR_{k,t}$	slot flights proportion (i), (iii)	%	14.008	27.053	0	100
$HHI_{k,t}$	market concentration (ii)	index x 100	54.800	22.431	20.817	100
T_t	time trend	discrete sequence	115.669	60.320	1	212
$QUALREG_t$	quality regulatory reform (iii), (viii)	dummy	0.417	0.493	0	1
$SLOTREG_{k,t}$	slot regulatory reform (iii), (viii)	dummy	0.104	0.305	0	1
$ATMREG_{k,t}$	ATM regulatory reform (iii), (viii)	dummy	0.044	0.206	0	1
$CASCDEL_{k,t}$	propagated delays (i)	%	19.016	8.296	0	79.339
$DEL_{k,t}$	proportion of delayed flights (i)	fraction	0.183	0.139	0	1

Sources: National Civil Aviation Agency (ANAC)'s Active Scheduled Flight Report - VRA (i) and Air Transportation Market Statistics Database (ii); governmental regulations available online (iii); "Study of the Air Transport Sector in Brazil" - Brazilian Development Bank, 2010 (iv); National Agency for Petroleum, Natural Gas and Biofuels - ANP (v); Brazilian Institute of Geography and Statistics - IBGE (vi), Central Bank of Brazil (vii) digital media press websites (viii). All figures computed with own calculations.

[24] See further discussion below.

3.5. Estimation strategy

Our empirical framework requires the specification of the unobserved component related to the addition of the extra block time (EXTBT) by carriers, namely the error term u of Equation (6). We have already defined $u = \mu + \zeta$, with u therefore being a sum of random terms associated with the uncertainties when taking scheduling decisions with respect to future flight operations (μ), and with respect to quality-service competition in the market (ζ). To account for such factors, our econometric approach treats u as a component error term related to route-specific idiosyncrasies and a set of time-varying controls. Using indices *k* (route) and *t* (time), we then define:

$$\begin{aligned} u_{k,t} &= \mu_{k,t} + \zeta_{k,t} \\ &= \Phi(ROUTE_k, UNC_{k,t}, RWYCONG_{k,t} \times UNC_{k,t}, \\ &\qquad DIR^{o}_{k,s}, DIR^{d}_{k,s}, AIRL^{i}_{k,t}, AIRC^{j}_{k,t}) + \varepsilon_{k,t}, \end{aligned} \tag{9}$$

where:

- $ROUTE_k$ is the fixed effect of route *k*, designed to control for the route-specific operating and time invariant market idiosyncrasies not observed by the econometrician.
- $UNC_{k,t}$ is a proxy for the level of uncertainty on the scheduling decision-making of carriers. It accounts for both the operational and strategic uncertainty. It is calculated as the coefficient of variation - i.e. the ratio of the standard deviation to the mean - of the actual block time distribution of route *k* and time *t*. Ceteris paribus, airlines may need to add more extra time to flights on routes associated with more uncertain times;
- $RWYCONG_{k,t} \times UNC_{k,t}$ is an interaction variable to capture the moderating effects of airport runway congestion (RWYCONG) on the relationship between uncertainty and the unobserved error;
- $DIR^{o}_{k,s}$ and $DIR^{d}_{k,s}$ are a set of dummy variables that account for the unobserved effects associated with the direction (DIR), the region of the endpoint city (origin or destination, namely *o* and *d*), and the season (*s*), on route *k*. With such terms we aim at controlling the impact of weather and prevailing winds on flight times, which are likely to be region- and season- specific. Including a combination of route direction-endpoint airports regions-seasons controls is essential when analyzing scheduled and actual flight times, especially to account for the potential effect of prevailing winds, because airlines can take advantage of the tailwinds to reduce their flight times (Irvine, Shine and Stringer, 2016). We believe that our proposed framework may be also utilized in different contexts, being especially useful in case of strong jet streams - a phenomenon that is

very important in some regions of the globe but typically neglected in econometric airline studies. In Brazil, jet streams typically occur in the South region. The dummies are associated to the origin and destination regions of a city-pair, considering the five Brazilian geographic regions (North, Northeast, Center-West, Southeast and South). We interact these endpoint regions dummies with other dummies that relate each directional-route with the cardinal and intercardinal directions of the 8-point compass rose, using the azimuth angle. We therefore associate the orientation of each flight with the geographic locations of its origin and destination. Finally, we interact these combinations with a set of dummy variables representing the seasons (summer, winter, fall and spring). As an illustration, a São Paulo - Rio de Janeiro flight has both endpoint cities located in the Southeast of Brazil, and a predominant orientation to the East; we therefore have a dummy of Southeast (origin)/East (route direction) and a dummy of Southeast (destination)/East (route direction), for each season, and thus totaling eight dummies for each directional route.[25]

- $AIRL^{i}_{k,t}$ is a set of dummy variables that control for the presence of airline *i* on route *k* at time *t*. With these dummy variables, we take into consideration the impacts of carrier/time-specific scheduling strategies on the overall flight operations on the route.
- $AIRC^{j}_{k,t}$ is a set of continuous variables that measure the route share of flights of aircraft model *j* on route *k* at time *t*. With these proportions, we aim at accounting for the possible (unobserved) effects of changes in the aircraft mix and technology evolution on the flight operations of the route.
- $\Phi(.)$ is an additive function, and $\varepsilon_{k,t}$ is the random error.

Another important setting of the EXTBT model in Equation (7) is related to the setup of the scheduling decision-making horizon. Considering that airlines schedules are designed some months in advance to actual flight operations, and thus utilize a three months lag to examine the impact of past realizations of key variables on the scheduling decisions of carriers - i. e. we set $h = 3$. We also examine the robustness of the results to changes in this setting.[26]

We perform some diagnosis tests of the the residuals, namely, the Pagan-Hall, White/Koenker and Breusch-Pagan/Godfrey/Cook-Weisberg heteroscedasticity tests, and the Cumby-Huizinga test of autocorrelation. In the tests, we confirm the presence of both problems. We employ the procedure

[25] As many routes share the same origin and destination regions, the total number of dummies inserted in the specification is much less than the number of routes times the eight compass points utilized. We actually computed 288 regions/directions/seasons dummies.

[26] See a discussion in 4.2.

of Newey-West to adjust the standard error estimates. Additionally, we examine the problem of multicollinearity in the estimation. We calculate mean and maximum VIF statistics of, respectively, 4.69 and 17.49 (EXTBT model), and of 5.27 and 17.53 (ODDSDEL model) in our preferred specifications.[27] These results reveal relevant multicollinearity issues indicating that we should cautiously interpret the results of the non-significant variables in our empirical models.

With respect to the ODDSDEL model - Equation (8) -, we utilize a similar specification of the error term of Equation (7) discussed above.[28] However, motivated by the previous literature - Greenfield (2014), Bendinelli, Bettini, & Oliveira (2016) -, we assume that the unobserved components of the flight delays on city-pair market *k* at time *t* are correlated with the status of competition in that market. As a result, we must treat the market concentration level - the HHI variable - as endogenous in the ODDSDEL equation. We then employ the Two-Step Feasible Efficient Generalized Method of Moments (2SGMM) with standard errors robust to arbitrary heteroscedasticity and autocorrelation.

Our identification strategy employs demand shifters associated with the size of the air travel market as instrumental variables. In particular, we utilize key socioeconomic metrics such as the gross domestic product (GDP), the population size, the GDP per capita, the Gini index of income inequality, and the number of active bank agencies, all extracted at the endpoint cities level.[29] For all these demand shifters, we also utilize the maximum, the minimum, the simple and the geometric means between the origin and the destination of a route. The data sources are the Brazilian Institute of Geography and Statistics (IBGE) and the Brazilian Central Bank. Other instrumental variables utilized for identification of the ODDSDEL equation is a set of lagged regressors of the EXTBT equation - namely, RWYCONG, SLOTPR and FUELP. Finally, we utilize Hausman-type instrumental variables, considering concentration levels of other routes to instrument the concentration level of a given route.

Apart from the fixed effects related to the 322 routes, our approach has in principle to deal with the estimation of additional 364 parameters of control variables (288 from $DIR^{o}_{k,s}$ and $DIR^{d}_{k,s}$, 37 from $AIRL^{i}_{k,t}$, and 39 from $AIRC^{j}_{k,t}$ controls), besides the set of 61 instrumental variables. We utilize the econometric method of high dimensional sparse (HDS) regression models of Belloni et al (2012), Belloni, Chernozhukov, & Hansen (2014a,b), and Chernozhukov, Hansen, & Spindler (2015). The HDS approach is flexible in allowing a large set of regressors and/or instrumental variables under

[27] To be more conservative, we do not include fixed effects and controls in the regressions for these VIF extractions.
[28] The only difference is that in the ODDSDEL model we do not include the terms related to scheduling uncertainty.
[29] Some of these metrics have yearly periodicity and therefore required interpolation to produce monthly series.

the assumption that the model is sparse - i.e. only a smaller subset of the initial regressors/instruments are important for capturing the main features of the regression in the data generation process. The approach utilizes the Least Absolute Shrinkage and Selection Operator (LASSO) of Tibshirani (1996). We apply the Instrumental Variables Post-Double Selection (PDS) estimation, which has a final step that estimates a traditional (non-LASSO) regression utilizing the controls and instruments selected in the LASSO estimation step. In our case, the final step of the PDS estimation of EXTBT and ODDSDEL utilized the fixed effects implemented with, respectively, the Ordinary Least Squares (OLS) estimator, and the 2SGMM estimator.

4. Estimation results

Tables 3 and 4 present the results of our empirical models of extra scheduled block time (EXTBT) and arrival flight delays (ODDSDEL) in Brazil. To simplify the exposition, we omit indexes *k* and *t*. In both tables, the respective Column (4) contains our preferred model; Columns (1) to (3) present a set of subspecified versions of the main model, in which we drop some key variables; and finally, Columns (5) to (8) display the results of some robustness checks. In both tables, it is possible to observe that most results are in line with our ex-ante expectations regarding signs and statistical significance of the estimated coefficients. Additionally, and regarding the motivations airlines have when planning their flight schedules, the results of Table 3 suggest that block time lengthening is clearly a result of a combination of both operational and strategic behaviors in the market. Furthermore, the results of Table 4 suggest that adding extra times to flights is effective in reducing the chances of delays.

Let us focus on the main results of Table 3 first - the EXTBT model. Considering the operational side of the problem, we find evidence that carriers tend to add extra time to flights in response to jet fuel price increases: the coefficient of the FUELP variable is statistically significant and positively associated with EXTBT in all specifications. This result support that jet fuel price hikes provoke drops in the expected cost indexes of future flights, leading airlines to set slower planned cruise speeds, and consequently longer flight durations, targeting fuel consumption reductions as a response to the expected higher cost pressure. In contrast, on the strategic side, the results of Table 3 suggest that carriers schedule longer block times on routes with higher competition - i.e. with lower market concentration. The negative and statistically significant coefficient of HHI is in line with the findings of Prince & Simon (2009), which found that in more competitive routes, a more intense service-quality competition is observed.[30] We therefore find evidence of actual schedule padding by carriers - a strategic behavior that is ceteris paribus to operational, cost-index-related factors that may also provoke block time lengthening.

The positive and statistically significant coefficient of the DENS variable in Table 3 provides evidence that carriers have dual motivations to strengthen their EXTBT, stemming from operational and market incentives. As discussed, denser routes are usually associated with increased operational complexity stemming from flight connections, but are also the markets in which carriers are more concerned with the satisfaction of passengers, such as to minimize damages in their overall institutional reputation. The reputational risk contagion - i. e. the quick spread of negative consumer assessments, either from online reviews and/or through word-of-mouth -, from one route to the entire system may be higher on those routes due to the higher size of the travelling public.

Let us now turn to Table 4. Our key research interest here is to investigate the impact of the extra block time addition (EXTBT) on the odds of flight delays (ODDSDEL). In line with Miranda & Oliveira (2018), our preferred model (Column 4) provides evidence that one additional minute of extra block time is associated with a reduction of approximately thirteen percent in the odds of flight delays,[31] as shown by the estimated coefficient of -0.1338 of the variable EXTBT. However, given the positive and statistically significant coefficients of the interaction variables EXTBT × SLOTPR and EXTBT × CASCDEL, we have that such effect is clearly moderated by the status of the operating conditions at the endpoint airports. More specifically, the reduction in the odds of flight delays allowed by extra block times is less effective on routes operated at airports under airport slot regime, and on routes subject to more frequent episodes of propagated delays. In contrast, we find no evidence that the intensity of runway utilization (RWYCONG) moderates the relationship between ODDSDEL and EXTBT.

[30] Kang & Hansen (2017) find that airlines set longer SCHBT are longer in highly competitive markets. Here we further explore the issue by examining the relationship between competition and EXTBT.

[31] With a 95% confidence interval of [-15.41, -11.35].

Table 3 - Estimation results: extra scheduled block time (EXTBT)

	(1) EXTBT	(2) EXTBT	(3) EXTBT	(4) EXTBT	(5) EXTBT	(6) EXTBT	(7) EXTBT	(8) EXTBT
FUELP (*lagged*)	0.7601***	0.6676***	0.6639***	0.6722***	0.6541***	0.6264***	0.3978**	0.5023**
DENS (*lagged*)	0.1983**	0.1747*	0.2200**	0.2227**	0.2065**	0.2317***	0.3568***	0.4014***
FREQ (*lagged*)	-0.2194*	-0.2087*	-0.2751**	-0.3110***	-0.2997***	-0.2989***	-0.5314***	-0.6269***
ASIZE (*lagged*)	-0.0512***	-0.0514***	-0.0508***	-0.0506***	-0.0507***	-0.0475***	-0.0495***	-0.0562***
RWYCONG (*lagged*)	0.0151*	0.0150*	0.0140*	-0.0014	-0.0012	0.0009	0.0186	-0.0116
SLOTPR (*lagged*)	-0.0418***	-0.0426***	-0.0427***	-0.0425***	-0.0444***	-0.0391***	-0.0468***	-0.0246**
HHI (*lagged*)	-0.0235***	-0.0242***	-0.0201**	-0.0199**	-0.0203**	-0.0173**	-0.0229***	-0.0234***
T	-0.0406***	-0.0394***	-0.0392***	-0.0392***	-0.0395***	-0.0381***	-0.0353***	-0.0365***
T × QUALREG	0.0300***	0.0315***	0.0310***	0.0309***	0.0309***	0.0278***	0.0258***	0.0275***
T × SLOTREG	-0.0115***	-0.0115***	-0.0110***	-0.0109***	-0.0105***	-0.0096***	-0.0136***	-0.0169***
T × ATMREG	0.0037	0.0036	0.0038	0.0039*	0.0039*	0.0032	0.0031	0.0022
CASCDEL (*lagged*)		0.0401***	0.0323**	0.0315**	0.0296**	0.0300**	0.0166	-0.0078
UNC (*lagged*)			0.1325***	0.1142***	0.1137***	0.1253***	0.0871***	0.0128
RWYCONG × UNC (*lagged*)				0.0023**	0.0024**	0.0017	0.0017	0.0033**
UNIBT*	5th pctl route-airc-time	5th pctl route-airc-time	5th pctl route-airc-time	5th pctl route-airc-time	5th pctl route-airc-time	10th pctl route-airc-time	5th pctl route-airc	5th pctl route-airc-time
Route fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Nr lags (regressors)	3	3	3	3	3	3	3	12
High-dim controls								
Dir-Reg-Seas controls	19/288	44/288	44/288	41/288	no	41/288	41/288	36/288
Airline controls	29/37	29/37	29/37	29/37	30/37	29/37	29/37	29/37
Aircraft controls	38/39	38/39	38/39	36/39	36/39	36/39	36/39	37/39
R^2 Adj Statistic	0.2930	0.2938	0.2944	0.2942	0.2933	0.2803	0.3171	0.3100
RMSE Statistic	12.6047	12.5975	12.5915	12.5933	12.6013	12.3196	12.2037	12.2186
Nr Observations	41,643	41,643	41,643	41,643	41,643	41,643	41,643	39,443

*Notes: results produced by the post-double selection (PDS), LASSO-based, methodology of Belloni et al. (2012, 2014a,b). Post-LASSO estimation is implemented with a fixed effects procedure with standard errors robust to heteroskedasticity and autocorrelation. Selected control estimates omitted; R^2 adjusted and RMSE statistics extracted from an equivalent least-squares dummy variables estimator (LSDV); p-value representations: ***p<0.01, ** p<0.05, * p<0.10.*

Table 4 - Estimation results: odds of flight delays (ODDSDEL)

	(1) ODDSDEL	(2) ODDSDEL	(3) ODDSDEL	(4) ODDSDEL	(5) ODDSDEL	(6) ODDSDEL	(7) ODDSDEL	(8) ODDSDEL
DENS	0.5449***	0.1923***	0.2214***	0.2219***	0.2192***	0.2204***	0.2365***	0.2472***
FREQ	0.0144	0.1059*	-0.0403	-0.0303	-0.0454	-0.0098	-0.1010	-0.0305
ASIZE	0.0344**	0.0145	0.0140	0.0118	0.0149	0.0120	0.0142	0.0168
RWYCONG	0.0293***	0.0213***	0.0210***	0.0268***	0.0252***	0.0254***	0.0255***	0.0283***
SLOTPR	0.0399***	0.0134***	0.0067	-0.0014	0.0002	0.0007	-0.0036	-0.0036
HHI (*endogenous*)	0.1501***	-0.0061	-0.0142	0.0016	-0.0172	0.0072	-0.0214	0.0313
T	-0.0172***	0.0134***	0.0085*	0.0096**	0.0082*	0.0107**	0.0083*	0.0148***
T × QUALREG	-0.0289***	-0.0109***	-0.0077***	-0.0078***	-0.0074***	-0.0085***	-0.0083***	-0.0089***
T × SLOTREG	-0.0030	-0.0041***	-0.0039***	-0.0039***	-0.0042***	-0.0038**	-0.0048***	-0.0040***
T × ATMREG	-0.0000	-0.0018	-0.0014	-0.0014	-0.0013	-0.0014	-0.0013	-0.0009
CASCDEL		0.6760***	0.6757***	0.6586***	0.6584***	0.6570***	0.6320***	0.6591***
EXTBT			-0.1070***	-0.1338***	-0.1361***	-0.1318***	-0.1785***	-0.1413***
EXTBT × RWYCONG				-0.0002	-0.0002	-0.0001	-0.0000	-0.0003
EXTBT × SLOTPR				0.0005***	0.0004**	0.0004**	0.0006***	0.0004**
EXTBT × CASCDEL				0.0011**	0.0012***	0.0014***	0.0022***	0.0017***
UNIBT*	-	-	5th pctl route-airc-time	5th pctl route-airc-time	5th pctl route-airc-time	10th pctl route-airc-time	5th pctl route-airc	5th pctl route-airc-time
Route fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
Nr lags (regressors)	3	3	3	3	3	3	3	12
High-dim controls								
Dir-Reg-Seas controls	35/288	67/288	67/288	53/288	no	58/288	55/288	44/288
Airline controls	29/37	29/37	29/37	29/37	29/37	29/37	29/37	29/37
Aircraft controls	37/39	37/39	37/39	34/39	34/39	34/39	34/39	33/39
High-dim instruments	7/61	7/61	7/61	7/61	7/61	7/61	7/61	5/61
R^2 Adj Statistic	0.3143	0.5193	0.5357	0.5359	0.5323	0.5333	0.5421	0.5500
RMSE Statistic	7.8194	6.5465	6.4344	6.4328	6.4573	6.4508	6.3899	6.3347
Nr Observations	39,450	39,450	39,450	39,450	39,450	39,450	39,450	37,545

*Notes: results produced by the post-double selection (PDS), LASSO-based, methodology of Belloni et al (2012, 2014a,b). Post-LASSO estimation is implemented with a 2-step generalized method of moments (2SGMM) estimator with a fixed effects procedure and standard errors robust to heteroskedasticity and autocorrelation. Selected control estimates omitted; R^2 adjusted and RMSE statistics extracted from an equivalent least-squares dummy variables (LSDV) estimator; p-value representations: ***p<0.01, ** p<0.05, * p<0.10.*

Most other results in Tables 3 and 4 are in line with our ex-ante expectations regarding signs and statistical significance of the estimated coefficients. For example, in Table 3 we find negative and significant coefficients of the FREQ variable, suggesting that carries have higher operational efficiency - i. e. lower time-related costs with respect to fuel costs in their cost indices - on routes with higher frequency flights;[32] we also find that both the levels of propagated delays (CASCDEL) and uncertainty (UNC) perceived by planners at the scheduling decision-making period increase EXTBT. In Table 4, all models show evidence that denser routes (DENS), and routes operated to/from airports with more congested runways (RWYCONG), are more likely to present delayed flights. The former set of routes, but not the latter, are associated with longer extra scheduled block times, as presented in Table 3. Therefore, even with block time lengthening, flights on denser routes still seem to be the ones that delay the most. Importantly to note, we only find evidence of a direct effect of airline competition on the odds of flight delays in Column (1) of Table 4, with the statistically significant coefficient of HHI. Such effect is fully dissipated when the CASCDEL variable is inserted in the model (Column 2) and persists statistically insignificant in the other columns.[33]

4.1. Estimated impacts of the regulatory reform events

A key topic in our empirical analysis of the extra block time - and schedule padding - formation by carriers in Brazil is related to the events of regulatory reform observed in the sample period. We find the following results in Tables 3 and 4. Before the introduction the on-time disclosure rules of 2012, carriers had an estimated trend to lessen their block time lengthening behavior in the market, apparently tolerating a concomitant trend of increase in the odds of delays in their flights. Indeed, the estimated coefficients of variable T in most columns of Tables 3 and 4 are statistically significant, and negative and positive, respectively. After the regulatory reform, however, both trends are almost fully dissipated: the coefficient of the interaction variable $T \times QUALREG$ is positive in Table 3 and, in most cases, negative in Table 4, with estimated effects of almost the same magnitude of the estimated coefficient of T. This is suggestive that the new regulation has apparently provoked an intensification in the schedule padding behavior by airlines. We argue that such movement can be classified as "schedule padding" as it is ceteris paribus to the other factors related to operations, and may have been motivated by the need to strategically position in the market with the new on-time disclosure requirements. A similar phenomenon is reported in the US market after the

[32] This result may be a consequence of the lower turnaround times of carriers with higher flight frequency.

[33] This is in contrast with Bendinelli, Bettini, & Oliveira (2016), which obtain statistically significant results for both variables. We recommend that further investigation is required into this issue.

implementation of similar regulation, as observed by Shumsky (1993), Forbes, Lederman, & Wither (2019) and Yimga & Gorjidooz (2019).

The second regulatory reform observed in the sample period was the 2014 airport slots new regulation. Again, analyzing Tables 3 and 4 jointly, we first note that routes associated with slot-constrained airports present less mean extra scheduled time addition, and higher propensity to delay - as indicated by the estimated coefficients of SLOTPR in those tables. However, the analysis of the results regarding variable T × SLOTREG show that the new regulation has provoked a drop in both EXTBT and ODDSDEL. Indeed, most estimates of those coefficients in both tables are negative and statistically significant. These results indicate that the introduction of a stricter and more comprehensive regulation of airport slots since 2014 may have been effective in reducing flight delays, and thus reducing the need of block time lengthening by carriers - a result that is in line with Fan (2019).

Regarding the implementation of the PBN operations in some Brazilian airports since the early 2010s, the estimated coefficient of the interaction term T × ATMREG suggests no statistically significant effect, as displayed in Tables 3 and 4. Although our results indicate that the new ATM regulatory framework has not produced effects on the extra scheduled block times (EXTBT) of carriers, it may have produced effect on the unimpeded block times (UNIBT), however. Anyway, PBN has apparently did not produced any effect on the likelihood of flight delays (ODDSDEL).

4.2. Robustness checks

We developed a set of robustness checks aiming at studying the sensitivity of our main estimation results displayed in the fourth columns of Tables 3 and 4. We present the results of these experiments in Columns (5), (6), (7), and (8) of those tables. First, in the Columns (5) of Tables 3 and 4 we experiment with an estimation procedure that does not include any region-direction-seasonal dummy, namely the $DIR^{o}_{k,s}$ and $DIR^{d}_{k,s}$ controls discussed in Section 3.4. Second, in the Columns (6) of both tables, we experiment with the 10th percentile of the actual block times distributions as an alternative measure of UNIBT* when computing EXTBT. Third, in Columns (7), we utilize another alternative measure of UNIBT*, this time extracting its values at the route-aircraft level, instead of the route-aircraft-time levels utilized in all other columns. As a final robustness check, in the Columns (8) of Tables 3 and 4 we change the number of lags employed to compute the lagged regressors of the EXTBT model, and the lagged instrumental variables of the ODDSDEL model. We consider a scheduling decision-making window equal to 12 months, and therefore set the number of lags h = 12. As it can be analyzed in the estimates of Columns (5)-(8) in Tables 3 and 4, in none of the robustness checks did we observe significant changes in the main results.

5. Conclusion

This paper investigates the main drivers of the extra times incorporated into the flight schedules of Brazilian airlines. We utilize an econometric method of high dimensional sparse (HDS) regression that employs the Least Absolute Shrinkage and Selection Operator (LASSO) of Tibshirani (1996) to select adequate controls from a vast set of available variables. Our estimation has the methodological contribution of decomposing the extra schedule block times into operational and strategic factors, after controlling for the uncertainty in the scheduling process. We also estimate the effect of extra times on flight delays, allowing for moderation effects of runway congestion, slots, and propagated delay. Finally, our empirical framework allows accounting for a set of key regional and seasonal unobserved effects that are formed at the route level. In particular, we include deeper route direction-specific controls to account for the potential impact of prevailing winds - a procedure that may be even more relevant in other realities characterized by strong jet streams that affect en route performance of flights.

Our results show evidence that the extra time inserted by carriers in their flights can be decomposed into the adjustments following changes in flight operating conditions, and also the strategic buffer time. We find evidence confirming the hypothesis of the existence and effectiveness of schedule padding practices, i.e. extra time additions that are correlated with the status of the competition in the market. Furthermore, the empirical results of on-time performance suggest that one additional minute of extra time is associated with a reduction of roughly thirteen percent in the odds of flight delays, which apparently is an effective strategy to enhance on-time performance.

With respect to the policy implications of our study, the estimation results suggest that on-time disclosure regulation, implemented since 2012, have possibly encouraged airlines to strategically lengthen their flight times. We also estimate a decreasing trend in the odds of flight delays after the introduction of the new regulation. In contrast, the implementation of slot controls at major airports since 2014 have apparently inhibited the addition of extra block time by carriers.

After all, are on-time performance statistics worthless? Our empirical results show evidence that the current official punctuality statistics may not always allow for reliable analyses of on-time performance in the Brazilian airline industry - and possibly in other countries -, as on-time disclosure regulations may create incentives for airlines to engage in schedule padding behavior. The problem associated with the strategic buffers is that they misalign the incentives of carriers in their pursuit of maximum operational efficiency and optimized cruise speeds dictated by cost indexes. The padding behavior induces a suboptimal exploitation of flight resources, and may mask the infrastructure inefficiencies that provoke airport congestion and propagated delays, specially in moments when the air transport industry is growing fast.

References

Belloni, A., Chen, D., Chernozhukov, V., & Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. Econometrica, 80(6), 2369-2429.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014a). Inference on treatment effects after selection among high-dimensional controls. The Review of Economic Studies, 81(2), 608-650.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014b). High-dimensional methods and inference on structural and treatment effects. Journal of Economic Perspectives, 28(2), 29-50.

Bendinelli, W. E. Bettini, H. F., & Oliveira, A. V. M. (2016). Airline delays, congestion internalization and non-price spillover effects of low-cost carrier entry. Transportation Research Part A: Policy and Practice, 85, 39-52.

Borsky, S., & Unterberger, C. (2019). Bad weather and flight delays: The impact of sudden and slow onset weather events. Economics of Transportation, 18, 10-26.

Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Post-selection and post-regularization inference in linear models with many controls and instruments. American Economic Review, 105(5), 486-90.

Edwards, H. A., Dixon-Hardy, D., & Wadud, Z. (2016). Aircraft cost index and the future of carbon emissions from air travel. Applied energy, 164, 553-562.

Errico, A., & Di Vito, V (2017). Performance-based navigation (PBN) with continuous descent operations (CDO) for efficient approach over highly protected zones, 24th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), St. Petersburg, 1-8.

Deo, V. A., Silvestre, F. J., & Morales, M. (2020). The benefits of tankering considering cost index flying and optional refuelling stops. Journal of Air Transport Management, 82.

Diana, T (2017). Has market concentration fostered on-time performance? A case study of seventy-two U.S. airports. Journal of Air Transport Management, 58, 1-8.

Fan, T. P. C. (2019). Schedule creep - In search of an uncongested baseline block time by examining scheduled flight block times worldwide 1986–2016. Transportation Research Part A: Policy and Practice, 121, 192-217.

Forbes, S. J., Lederman, M., & Yuan, Z. (2018). Do Airlines Pad Their Schedules? Review of Industrial Organization, 1-22.

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.

Holloway, S. (2008). Straight and level: practical airline economics. Ashgate Publishing, Ltd.

Irvine, E. A., Shine, K. P. & Stringer, M. A (2016). What are the implications of climate change for trans-Atlantic aircraft routing and flight time? Transportation Research Part D: Transport and Environment, 47, 44-53.

Kafle, N., & Zou, B. (2016). Modeling flight delay propagation: A new analytical-econometric approach. Transportation Research Part B: Methodological, 93, 520-542.

Kang, L., & Hansen, M. (2017). Behavioral analysis of airline scheduled block time adjustment. Transportation Research Part E: Logistics and Transportation Review, 103, 56-68.

Mayer, C., & Sinai, T. (2003). Why do airlines systematically schedule their flights to arrive late? Working paper.

Miranda, V. A., & Oliveira, A. V. M. (2018). Airport slots and the internalization of congestion by airlines: An empirical model of integrated flight disruption management in Brazil. Transportation Research Part A: Policy and Practice, 116, 201-219.

Prince, J., & Simon, D. (2009). Multimarket Contact and Service Quality: Evidence from On-Time Performance in the U.S. Airline Industry. The Academy of Management Journal, 52(2), 336-354.7

Santos, G., & Robin, M. (2010). Determinants of delays at European airports. Transportation Research Part B: Methodological, 44, 3, 392-403.

Şafak, Ö., Atamtürk, A., & Aktürk, M. S. (2019). Accommodating new flights into an existing airline flight schedule. Transportation Research Part C: Emerging Technologies, 104, 265-286.

Shumsky, R. (1993). Response of US air carriers to on-time disclosure rule. Transportation Research Record, 1379, 9–16.

Skaltsas, G. (2011). Analysis of airline schedule padding on US domestic routes (Master's thesis, Massachusetts Institute of Technology).

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

Yimga, J., & Gorjidooz, J. (2019). Airline schedule padding and consumer choice behavior. Journal of Air Transport Management, 78, 71-79.

Young, T. M. (2018). Performance of the Jet Transport Airplane: Analysis Methods, Flight Operations, and Regulations. John Wiley & Sons.

Zou, B., & Hansen, M. (2012). Impact of operational performance on air carrier cost structure: evidence from US airlines. Transportation Research Part E: Logistics and Transportation Review, 48(5), 1032-1048.

Zou, B., Elke, M., Hansen, M., & Kafle, N. (2014). Evaluating air carrier fuel efficiency in the US airline industry. Transportation Research Part A: Policy and Practice, 59, 306-330.