Research in Transportation Economics 103 (2024) 101405

Contents lists available at ScienceDirect

Research in Transportation Economics

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

Research paper

Estimating tactical surface metering management's effect on aircraft fuel savings at airport

Hideki Fukui [a,*], Chikage Miyoshi [b]

[a] Faculty of Law and Letters, Ehime University, 3 Bunkyo-cho, Matsuyama, 790-8577, Ehime, Japan
[b] Centre for Environment, Cranfield University, College Road, Cranfield, MK43 0AL, United Kingdom

ARTICLE INFO

JEL classification:
L93
Q51
R11
R41

Keywords:
Tactical surface metering
Taxi-out/in fuel consumption
Terminal flight data manager
Difference-in-differences
Propensity score matching

ABSTRACT

We estimated the effect of surface metering technology deployed at Charlotte Douglas International Airport (CLT) by the National Aeronautics and Space Administration (NASA) and the Federal Aviation Administration (FAA) on fuel and CO_2 savings. We employed a difference-in-differences strategy, covariate-balancing methods, and a fixed-effects model. Furthermore, we used a doubly robust strategy that combines matching with parametric regression using datasets constructed by data-driven approaches, which minimizes discretion in selecting comparison units. Our estimates from the daily panel of US carriers from November 2015 to November 2019 suggested that about 39.156 kg (95% confidence interval (CI) 64.017–13.436) of fuel and 120.599 kg (CI 197.174 kg–41.382 kg) of CO_2 could be saved per flight even when the single-engine taxiing rate was as high as 75% during the entire deployment phase. At the same time, our results suggested a concerning pattern: taxi-out time savings generated by the technology diminished over time.

1. Introduction

The more time aircraft taxiing on the surface, the more fuel consumed. The vast majority of fuel consumption occurs while aircraft are flying. However, a surface management strategy can be a practical option for conserving fuel and reducing emissions because it may face fewer implementation barriers than other methods (Hao et al., 2017). Stettler et al. (2011) and Koudis et al. (2018) estimated that typical landing and takeoff operations for aircraft account for approximately 36% of fuel consumption, suggesting taxiing is a significant source of airport emissions. The estimated statistical evidence highlights the importance of decreasing aircraft taxiing times to reduce airport pollutant emissions (Stettler et al., 2011). Emissions from aircraft operations are categorized in Scope 3[1] and constitute the most significant source of emissions at airports, but airport managers do not currently control or manage them effectively (ACI, 2018). Efficient management of traffic on the surface, however, significantly contributes to emission reduction and fuel conservation.

Indeed, without coordination between the air traffic control tower (ATCT) and ramp controllers, the latter tend to assign release times to aircraft potentially earlier than the optimal times. If an aircraft is pushed back too early, surface congestion and queue times increase, which results in extra fuel consumption and CO_2 emissions. To address this issue, NASA developed new software that enables integrated arrival, departure, and surface (IADS) operations and deployed a tactical surface metering tool provided by the software on November 29, 2017, at CLT. This paper examines how and to what extent the surface tactical surface metering reduced taxi times, fuel consumption, and CO_2 emissions at CLT.

Since November 29, 2017, the FAA and NASA have been testing the new software at CLT to reduce fuel burned by minimizing taxi delays and ramp congestion. According to the FAA and NASA, tactical surface metering enabled by the software has saved more than 275,000 gallons of fuel annually and reduced CO_2 emissions by 8 tons daily by reducing the time aircraft spend taxiing (as of September 2021). The number of performed departures at CLT in 2018 was about 255,000[2] (US

* Corresponding author.
E-mail addresses: fukui.hideki.hz@ehime-u.ac.jp (H. Fukui), c.miyoshi@cranfield.ac.uk (C. Miyoshi).

[1] According to the US Environmental Protection Agency (EPA), Scope 1 refers to GHGs emitted directly from sources under the control or ownership of an organization. Scope 2 is indirect GHGs attributable to the consumption of electricity, steam, heat, or cooling purchases. Scope 3 is GHGs attributed to operations from resources that the reporting organization does not own or control but that have an indirect influence in its value chain. (US EPA, GHG Inventory Development Process and Guidance <https://www.epa.gov/climateleadership/ghg-inventory-development-process-and-guidance> (Accessed 10 June 2022).).

[2] The actual number is 243,401 for domestic flights and 11,974 for international flights, respectively.

https://doi.org/10.1016/j.retrec.2023.101405
Received 9 July 2023; Received in revised form 19 December 2023; Accepted 20 December 2023
Available online 31 January 2024

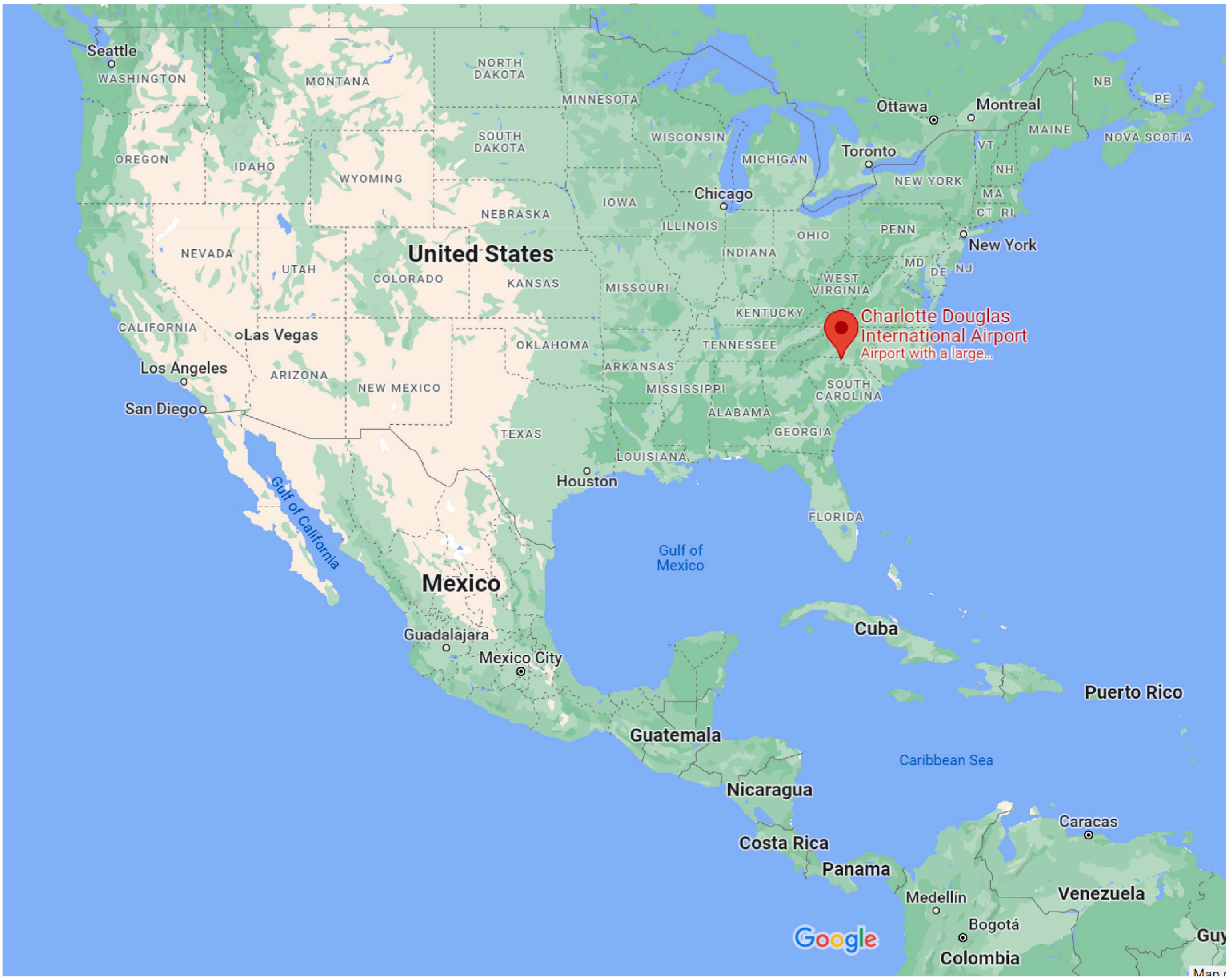

Fig. 1. Charlotte Douglas International Airport. Source: Google Maps (2023).

Department of Transportation and US Department of Transportation). As 1 gallon of jet fuel at 59 °F (i.e., 15 °C) has a mass of 3.103 kg, the estimate by FAA and NASA suggests that surface metering technology saved about 3.3 kg (=275,000 × 3.103/255,000) of fuel and 10.16 kg (=3.3 × 3.08) of CO_2 per flight.[3]

Based on the results, the FAA plans to deploy surface metering technology via the Terminal Flight Data Manager (TFDM) program in 27 airports (FAA, 2021). The TFDM supports surface scheduling and metering by implementing electronic flight strips and data-sharing functionalities. By interfacing the TFDM data with the flight operator, airport data, and other FAA systems, flight operators, air traffic controllers, and airports can more efficiently manage aircraft flow through all flight phases from the departure gate to the arrival gate. Deployment of surface metering technology will help minimize taxi delays, fuel burn, and CO_2 emissions (FAA, 2022).

The FAA and NASA reported promising effects in the impact assessment. However, the fuel savings calculation during the test period is based on actual gate-holding times and taxi-time savings, defined as the difference between the initial and revised runway release times for approval request (APREQ) flights (NASA ATD-2 team, 2020). The calculation does not control for time-varying (but not varying across units) confounding factors, such as traffic demand changes over time. For example, suppose a downward trend in air traffic occurs at an airport, with average taxi times decreasing over time. In that case, this downward trend may not be caused by surface metering technology but by other factors, such as declining passenger demand due to the recession. If we do not control for the effects of these time-varying (but not varying across units) confounding factors, we may wrongly attribute the reduction in taxi times to the deployment of surface metering technology. This predicament necessitates a counterfactual comparison group in which the average taxi times show a trend parallel to those in the treatment group before deploying the technology. Such a comparison group would mitigate the effects of time-varying (but not varying across units) confounding factors and ensure estimates more precisely reflect the impact of the tactical surface metering on fuel savings.

In addition, the FAA and NASA did not fully assess the potential adverse side effects of gate-holding: airborne and taxi-in delays for arriving flights. The fuel savings from reducing aircraft taxi-out times could be canceled out if arriving flights spend extra airborne and taxi-in time due to gate-holding departures prompted by surface metering technology. However, the FAA and NASA only compared the standard on-time performance metrics of outbound/inbound flights from/to CLT

[3] The NASA ATD-2 team (2020) set the CO_2 emissions coefficient for jet fuel at 3.08 [kg CO_2/kg jet fuel] (=9.57 [kg CO_2/gallon]).

Table 1
Descriptions of variables.

Variables	Descriptions
Taxi-out	Average taxi-out times (minutes) for outbound flights from origin calculated on a carrier–route–departure time block basis (logged) (Source: US Department of Transportation)
Airborne	Average airborne times per mile (minutes) for inbound flights to destination calculated on a carrier–route–arrival time block basis (logged) (Source: US Department of Transportation)
Taxi-in	Average taxi-in times (minutes) for inbound flights to destination calculated on a carrier–route–arrival time block basis (logged) (Source: US Department of Transportation)
Post	Post dummy (1 on and after November 29, 2017, 0 otherwise)
O_CLT	CLT as origin dummy (1 if origin is CLT, 0 otherwise)
D_CLT	CLT as destination dummy (1 if destination is CLT, 0 otherwise)
Post1	Post1 dummy (1 from November 29, 2017 to February 18, 2018, 0 otherwise)
Post2	Post2 dummy (1 from February 19, 2018 to September 30, 2018, 0 otherwise)
Post3	Post3 dummy (1 from October 1, 2018 to August 31, 2019, 0 otherwise)
Post4	Post4 dummy (1 from September 1, 2019 to November 28, 2019, 0 otherwise)
Post*O_CLT	Post * CLT as origin dummy (interaction term between Post and CLT as origin dummies)
Post1*O_CLT	Post1 * CLT as origin dummy (interaction term between Post1 and CLT as origin dummies)
Post2*O_CLT	Post2 * CLT as origin dummy (interaction term between Post2 and CLT as origin dummies)
Post3*O_CLT	Post3 * CLT as origin dummy (interaction term between Post3 and CLT as origin dummies)
Post4*O_CLT	Post4 * CLT as origin dummy (interaction term between Post4 and CLT as origin dummies)
Post*D_CLT	Post * CLT as destination dummy (interaction term between Post and CLT as destination dummies)
Post1*D_CLT	Post1 * CLT as destination dummy (interaction term between Post1 and CLT as destination dummies)
Post2*D_CLT	Post2 * CLT as destination dummy (interaction term between Post2 and CLT as destination dummies)
Post3*D_CLT	Post3 * CLT as destination dummy (interaction term between Post3 and CLT as destination dummies)
Post4*D_CLT	Post4 * CLT as destination dummy (interaction term between Post4 and CLT as destination dummies)
Runway_Origin	Number of runways at origin airport (Source: US DOT, National Flight Data Center (NFDC) Airport Database's Airport Facilities and Runways Tables)
Runway_Dest	Number of runways at destination airport (Source: US DOT, National Flight Data Center (NFDC) Airport Database's Airport Facilities and Runways Tables)
MonthDepShare_Origin_Carrier	Monthly share of performed departures at origin calculated on a carrier basis (Source: US Department of Transportation and US Department of Transportation)
MonthArrShare_Dest_Carrier	Monthly share of performed arrivals at destination calculated on a carrier basis (Source: US Department of Transportation and US Department of Transportation)
DaySchDep_Origin_Carrier_Route_TimeBlock	Daily number of scheduled departures at origin aggregated on a carrier–route–departure time block basis at origin (Source: US Department of Transportation)
DaySchArr_Dest_Carrier_Route_TimeBlock[a]	Daily number of scheduled arrivals at destination aggregated on a carrier–route–arrival time block basis at destination (Source: US Department of Transportation)
DaySchDep_Route_TimeBlock	Daily number of scheduled departures aggregated on a route–departure time block basis at origin (Source: US Department of Transportation)
DaySchArr_Route_TimeBlock	Daily number of scheduled arrivals aggregated on a route–arrival time block basis at destination (Source: US Department of Transportation)
MonthSchDep_Origin_Carrier	Monthly number of scheduled departures (domestic and international) at origin aggregated on a carrier basis (Source: US Department of Transportation and US Department of Transportation)
MonthSchArr_Dest_Carrier[b]	Monthly number of scheduled arrivals (domestic and international) at destination aggregated on a carrier basis (Source: US Department of Transportation and US Department of Transportation)
AvgDayDepDelay_Carrier_Route_TimeBlock[c]	Average daily departure delays (in minutes) calculated on a carrier–route–departure time block basis at origin (Source: US Department of Transportation)
TotalDayDepDelay_Origin_TimeBlock	Total daily departure delays (in minutes) aggregated on an origin–departure time block basis (Source: US Department of Transportation)
TotalDayArrDelay_Dest_TimeBlock[d]	Total daily arrival delays (in minutes) aggregated on a destination–arrival time block basis (t-1) (Source: US Department of Transportation)
DayCancel_Carrier_Route_TimeBlock	Daily number of canceled flights aggregated on a carrier–route–departure time block basis (Source: US Department of Transportation)
DayCancel_Origin_TimeBlock_DueToCarrier[e]	Daily number of flights canceled due to reasons within control of carrier aggregated on an origin–departure time block basis (Source: US Department of Transportation)
DayCancel_Origin_TimeBlock_DueToWeather	Daily number of flights canceled due to severe weather aggregated on an origin–departure time block basis (Source: US Department of Transportation)
DayCancel_Origin_TimeBlock_DueToNAS	Daily number of flights canceled due to the National Airspace System (NAS) aggregated on an origin–departure time block basis (Source: US Department of Transportation)
DayCancel_Origin_TimeBlock_DueToSecurity	Daily number of flights canceled due to security concerns aggregated on an origin–departure time block basis (Source: US Department of Transportation)
DayDivert_Carrier_Route_TimeBlock	Daily number of diverted flights aggregated on a carrier–route–arrival time block basis (Source: US Department of Transportation)
DayDivert_Dest_TimeBlock	Daily number of diverted flights aggregated on a destination–arrival time block basis (Source: US Department of Transportation)
Distance	Route distance (logged) (Source: US Department of Transportation)

[a] An increase or decrease in the number of departures and arrivals might worsen or reduce congestion, which in turn may affect the increase or decrease in taxi times. Therefore, we used the numbers of scheduled departures and arrivals on a carrier–route–departure/arrivals time basis as controls.

[b] The number of international flights could also affect domestic flight operations and taxi times. Unfortunately, T-100D and T-100I, the international segment data, is monthly data and is not available on a daily basis. Therefore, we used the aggregated monthly number of international and domestic flights as a control.

[c] Departure delays could be caused not only by airport or terminal congestion but also by a variety of events such as security problems, severe weather, or carrier issues (maintenance or crew problems, aircraft cleaning, baggage loading, fueling, late-arriving aircraft, etc.) (Richetta & Odoni, 1994; Ryerson et al., 2014). In addition, departure delays might consequently induce carriers to taxi at higher speeds with lower fuel efficiency in order to recover their time. Therefore, departure delay variables should be included to control for the effects of mechanisms that may lead to increased fuel use due to departure delays.

[d] This variable lagged one period (i.e., one arrival time block) because the arrival delay of a given flight at the destination did not affect the taxi-out time at the origin or the airborne time of the same flight.

[e] US DOT defines the reasons for cancellation as follows: Carrier: Circumstances within the control of carrier (e.g., maintenance or crew problems, aircraft cleaning, baggage loading, fueling); Weather: Significant meteorological conditions (actual or forecasted) that, in the judgment of the carrier, prevent the operation of a flight such as tornado, blizzard, or hurricane; National Aviation System (NAS): A broad set of conditions, such as non-extreme weather conditions, airport operations, heavy

traffic volume, and air traffic control; Security: Evacuation of a terminal or concourse, re-boarding of aircraft because of a security breach, inoperative screening equipment and/or long lines in excess of 29 min at screening areas. Source: US DOT, Understanding the Reporting of Causes of Flight Delays and Cancellations (Monday, January 10, 2022)https://www.bts.gov/topics/airlines-and-airports/understanding-reporting-causes-flight-delays-and-cancellations <Accessed 2 October 2022>.

before and after the deployment of tactical surface metering. Again, their assessment did not control for the effects of time-varying (but not varying across units) factors (NASA ATD-2 team, 2020).

Our contribution is to address the limitations identified in the previous studies, including the NASA ATD-2 team (2020), and present the quantitative evidence for the efficacy of aircraft surface management at airports using new traffic management technologies. We estimate how traffic management technologies for airport surfaces could reduce taxi times, fuel consumption, and CO_2 emissions at CLT. Specifically, we use a difference-in-differences strategy along with covariate-balancing methods and a fixed-effects model using data from US DOT Airline On-Time Performance Data (hereafter, US DOT, OTP), Air Carrier Statistics (Form 41 Traffic) T-100 Domestic/International Segment (All Carriers) (hereafter, US DOT, T100D and US DOT, T100I), and US DOT National Flight Data Center (US DOT, NFDC) Airport Database's Airport Facilities and Runways Tables from 2015 to 2019. The treatment group comprised flights to and from CLT. The comparison group included flights to and from airports other than CLT (see Appendix A). Among all American Airlines hubs (Charlotte, Chicago, Dallas/Fort Worth, Los Angeles, Miami, New York, Philadelphia, Phoenix, and Washington, DC) (American Airlines Group Inc, 2021), CLT has the highest share of flights operated by the airline, which is over 80%, during the analysis period. As such, it is difficult to single out any airport as a suitable comparison unit, even from the American Airlines hub airports. Indeed, it can be an arbitrary choice by the researchers. Therefore, we constructed comparison groups using data-driven approaches, which minimized discretion in selecting the comparison units (see Section 4).

In our study, we used a doubly robust strategy that combines matching with parametric regression (explained in Sections 4 and 5). This two-step approach is doubly robust because even if either the matching or the parametric model is not accurately specified, the causal estimates will still be consistent as long as one of them is correctly specified (Bang & Robins, 2005; Ho et al., 2007). To account for the effects of changes in factors at each airport, for which data are not compiled in a publicly available form and are difficult to obtain (e.g., the number of gates and the taxi distance from a gate to a runway), we controlled them through airport-specific year-fixed effects. This is because it is not practical to identify changes at all airports in our datasets, which comprise more than 200 airports.

We structure this paper as follows. Section 2 provides a literature review on the impact of taxiing, flight delays, and gate-holding on fuel consumption. Section 3 explains the benefit mechanism of tactical surface metering, our data, and the estimation model. Section 4 examines the validity of the model assumption, i.e., the parallel trend assumption and the covariate balance between the treatment and comparison groups. Section 5 presents the estimation results. Section 6 calculates the fuel and CO_2 savings surface metering technology achieved. Finally, Section 7 presents the conclusions and limitations of this study and identifies some future research directions.

Table 2
Parallel-trends test (pre-treatment period).

H_0: Linear trends are parallel	Prob > F
Dataset A (Taxi-out)	0.120
Dataset B (Airborne)	0.336
Dataset C (Taxi-in)	0.269

2. Literature review

Researchers have estimated the impact of taxi-out and flight delays and assessed the fuel consumption benefits of gate-holding (Hao et al., 2017; Koudis et al., 2018; Nikoleris et al., 2011; Ravizza et al., 2013; Ryerson et al., 2014; Stettler et al., 2011). The results of these studies suggested that airports could achieve fuel and emission reductions through surface metering technology that contributes to shorter taxi times.

Ravizza et al. (2013), Nikoleris et al. (2011), and Stettler et al. (2011) provided an overall picture of the relationship between fuel consumption and aircraft movements on airport surfaces. Ravizza et al. (2013) developed a new multi-objective approach to analyze the tradeoff between taxi time and fuel burn when aircraft taxi on airport surfaces. They developed and combined two algorithms: the first found the best routes concerning minimal taxi time, and the second approximated the most time-efficient speed profile for each route. Then, they ran simulations and performed a sensitivity analysis for data from Zurich Airport. They found that the fuel-related objective function significantly influenced the tradeoff. Using actual position data from aircraft operated at Dallas/Fort Worth International Airport, Nikoleris et al. (2011) estimated fuel consumption and emissions for taxiing. They modeled fuel consumption and emissions during taxi operations as a linear function of four taxi phases: stopping, turning, accelerating, and moving at a constant speed or braking (tracked as one function). They found that stop-and-go movement accounted for approximately 18% of the fuel consumed. They identified two significant fuel combustion and emission contributors: idling and taxiing at a constant speed or braking. Stettler et al. (2011) developed and applied a method to calculate emissions from UK airports after quantifying uncertainties and pointed out the potential negative impact of airport emissions on human health and climate. They developed an inventory of emissions from aircraft landing and takeoff operations. Then, they quantified the scientific and operational uncertainty in the emissions based on previous studies and the Monte Carlo simulation to determine the most important sources of uncertainty. According to their estimates based on OAG data in 2005, the 20 busiest UK airports generated 10.2 Gg of NOx (−23 to +29%), 0.73 Gg of SO2 (−29 to +32%), 11.7 Gg of CO (−42 to +77%), 1.8 Gg of HC (−59 to +155%), 2.4 Tg of CO_2 (−13 to +12%), and 0.31 Gg of PM2.5 (−36 to +45%).

Hao et al. (2017) focused on analyzing the effects of eliminating taxi delays on fuel savings. First, they modeled flight fuel burn as a linear function of different components of flight time: taxi times (nominal and delayed times) and airborne times (nominal and delayed times). Then, they estimated the contribution of each flight time component to fuel consumption by a fixed effects model using the fuel and flight data from a major US-based carrier and the flight-level performance data from the Aviation System Performance Metrics (ASPM) database provided by the FAA. They found that gate-holding eliminated taxi delays (i.e., extra taxi time) and reduced flight fuel consumption by approximately 1%. They also found that, for some airports, reducing taxi delays could reduce fuel consumption by as much as 2% when they calculated the savings per airport.

In contrast, Koudis et al. (2018) examined how single-engine taxing (SET) affected fuel consumption and emissions. They aimed to develop a model to quantify the impact of SET on fuel consumption and pollutant emission by using high-resolution trajectory data for aircraft, encompassing SET activities for 3510 flights at London Heathrow Airport in

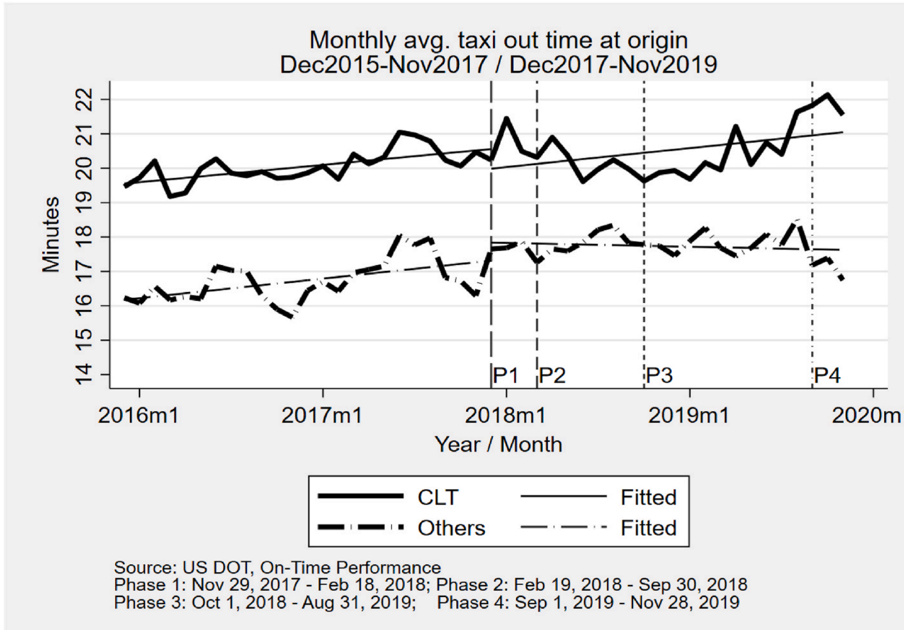

Fig. 2. Dataset A – Monthly average taxi-out time of outbound flights from origin airports.

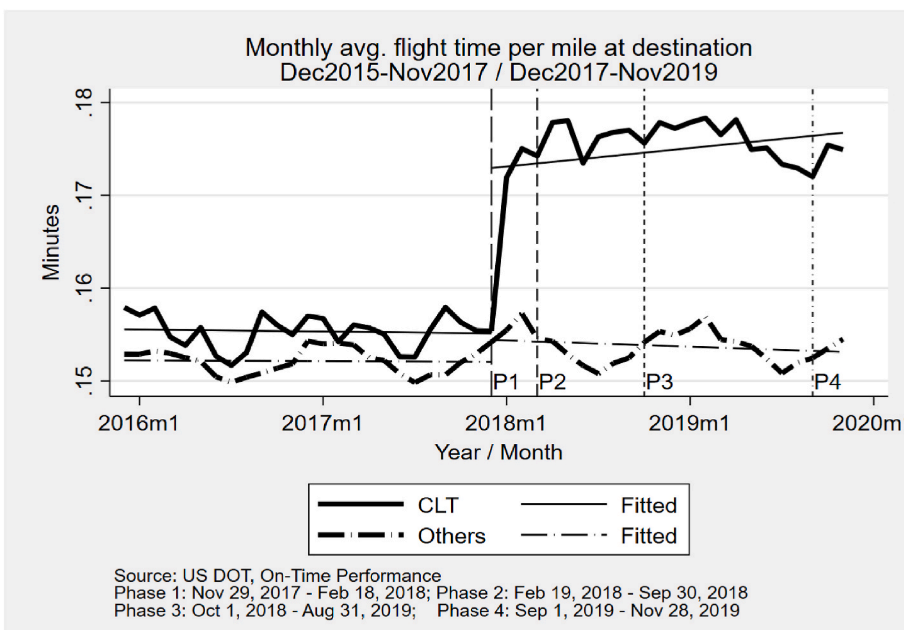

Fig. 3. Dataset B – Monthly average airborne time of inbound flights to destination airports.

November 2012. Based on their SET fuel consumption model, they estimated that fuel consumption and pollutant emissions would increase up to 50% without SET during taxi-in. However, their estimates also suggested that shortening the time to start SET to the 25th percentile of recorded values would result in a 7–14% reduction in fuel consumption and pollutant emissions compared to the current operation.

Ryerson et al. (2014) estimated fuel savings from air traffic management (ATM) improvements by analyzing actual flight-level data on fuel consumption provided by a US carrier, data from the FAA National Airspace System Performance Analysis Capability (NASPAC), and the flight-level performance data from the FAA ASPM database. Using a fixed effects model, they regressed fuel consumption on departure delay, airborne delay, excess planned flight time, and terminal area inefficiencies. Their estimation results suggested that approximately 1.0–1.5% and 1.5–4.5% of the system-wide averages of flight fuel consumption occurred due to ATM delay and terminal inefficiencies, respectively. Their estimation also suggested that the impact of predicted delay on fuel consumption was 80–90% smaller than that of unanticipated delay, underscoring the importance of having reliable ATMs to reduce fuel consumption and CO_2 emissions.

In these studies, the researchers assessed the amount of fuel consumed when aircraft taxi on surfaces and examined the effects of techniques that reduce fuel consumption during taxiing. The results of these studies provided basic methods, information, and data to estimate

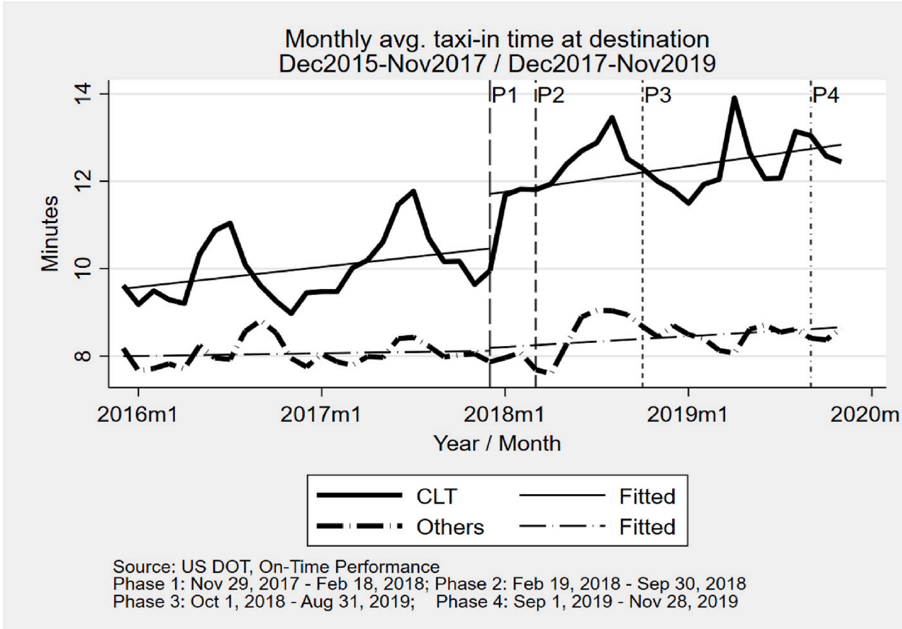

Fig. 4. Dataset C – Monthly average taxi-in time of inbound flights to destination airports.

Table 3
Variable balance across the treatment and comparison groups in Dataset A (Taxi-out): Standardized differences (%) between two groups.

Dataset A		
Sample type	(1) Original	(2) Matched
	Absolute mean standardized differences (%)	
(1) Runway_Origin	19.1	2.4
(2) Runway_Dest	47.9	3.1
(3) MonthDepShare_Origin_Carrier	40.1	5.5
(4) MonthArrShare_Dest_Carrier	39.3	5.5
(5) DaySchDep_Origin_Carrier_Route_TimeBlock	14.5	0.8
(6) DaySchArr_Dest_Carrier_Route_TimeBlock	16.5	0.7
(7) DaySchDep_Route_TimeBlock	162.9	8.0
(8) DaySchArr_Route_TimeBlock	54.3	1.3
(9) MonthSchDep_Origin_Carrier	215.4	5.1
(10) MonthSchArr_Dest_Carrier	42.9	3.9
(11) AvgDayDepDelay_Carrier_Route_TimeBlock	0.1	0.8
(12) TotalDayDepDelay_Origin_TimeBlock	47.5	1.5
(13) TotalDayArrDelay_Dest_TimeBlock	4.4	1.2
(14) DayCancel_Carrier_Route_TimeBlock	2.1	0.2
(15) DayCancel_Origin_TimeBlock_DueToCarrier	17.6	1.3
(16) DayCancel_Origin_TimeBlock_DueToWeather	21.1	0.9
(17) DayCancel_Origin_TimeBlock_DueToNAS	4.5	0.7
(18) DayCancel_Origin_TimeBlock_DueToSecurity	1.6	0.5
(19) DayDivert_Carrier_Route_TimeBlock	0.6	0
(20) DayDivert_Dest_TimeBlock	4.6	0.1
(21) Distance	28.2	4.6
Mean bias (average of absolute mean standardized differences)	37.4	2.3
Observations	5,534,608	239,222

Table 4
Variable balance across the treatment and comparison groups in Dataset B (Airborne): Standardized differences (%) between two groups.

Dataset B		
Sample type	(1) Original	(2) Matched
	Absolute mean standardized differences (%)	
(1) Runway_Origin	46.9	0.9
(2) Runway_Dest	33.8	2
(3) MonthDepShare_Origin_Carrier	43.1	3.5
(4) MonthArrShare_Dest_Carrier	22.7	2.9
(5) DaySchDep_Origin_Carrier_Route_TimeBlock	4.6	2.1
(6) DaySchArr_Dest_Carrier_Route_TimeBlock	7.8	1.6
(7) DaySchDep_Route_TimeBlock	54.1	2.4
(8) DaySchArr_Route_TimeBlock	136.4	0
(9) MonthSchDep_Origin_Carrier	40	6
(10) MonthSchArr_Dest_Carrier	179.1	2.3
(11) AvgDayDepDelay_Carrier_Route_TimeBlock	2.7	0.5
(12) TotalDayDepDelay_Origin_TimeBlock	24.3	0.2
(13) TotalDayArrDelay_Dest_TimeBlock	3.4	0.5
(14) DayCancel_Carrier_Route_TimeBlock	0.1	0.1
(15) DayCancel_Origin_TimeBlock_DueToCarrier	8.3	0.1
(16) DayCancel_Origin_TimeBlock_DueToWeather	6.2	0.1
(17) DayCancel_Origin_TimeBlock_DueToNAS	3.8	0.5
(18) DayCancel_Origin_TimeBlock_DueToSecurity	0.2	0.4
(19) DayDivert_Carrier_Route_TimeBlock	0.4	0.2
(20) DayDivert_Dest_TimeBlock	7.7	0.4
(21) Distance	31.2	3
Mean bias (average of absolute mean standardized differences)	31.3	1.4
Observations	9,890,246	360,442

the fuel and emission reduction achieved by decreasing taxi and airborne delays and using SET. Unfortunately, they shared the same limitation as the assessment by the FAA and NASA (NASA ATD-2 team, 2020): their assessment did not control for the effects of time-varying (but not varying across units) factors. Also, none of these researchers assessed these techniques' potential adverse side effects. For example, Hao et al. (2017) suggested that eliminating extra taxi time through gate-holding would reduce flight fuel consumption by about 1%, an important finding. However, it is equally important to note that gate-holding could cause adverse side effects. The fuel savings from reducing aircraft taxi-out times could be canceled out if arriving flights spend extra airborne or taxi-in times due to gate-holding for departing flights.

This paper endeavors to fill the gaps identified in the previous studies

Table 5
Variable balance across the treatment and comparison groups in Dataset C (Taxi-in): Standardized differences (%) between two groups.

Dataset C		
Sample type	(1) Original	(2) Matched
	Absolute mean standardized differences (%)	
(1) Runway_Origin	39.9	3.3
(2) Runway_Dest	33.6	11.7
(3) MonthDepShare_Origin_Carrier	25.7	5.9
(4) MonthArrShare_Dest_Carrier	24.6	12.4
(5) DaySchDep_Origin_Carrier_Route_TimeBlock	3.5	2.5
(6) DaySchArr_Dest_Carrier_Route_TimeBlock	7.2	2
(7) DaySchDep_Route_TimeBlock	48.3	3
(8) DaySchArr_Route_TimeBlock	69.8	7.4
(9) MonthSchDep_Origin_Carrier	29.2	5.2
(10) MonthSchArr_Dest_Carrier	67.2	13.9
(11) AvgDayDepDelay_Carrier_Route_TimeBlock	6.4	0
(12) TotalDayDepDelay_Origin_TimeBlock	22.6	1.2
(13) TotalDayArrDelay_Dest_TimeBlock	8.9	1.4
(14) DayCancel_Carrier_Route_TimeBlock	0.6	0.6
(15) DayCancel_Origin_TimeBlock_DueToCarrier	7.1	0.4
(16) DayCancel_Origin_TimeBlock_DueToWeather	5.7	0.5
(17) DayCancel_Origin_TimeBlock_DueToNAS	3.5	0.3
(18) DayCancel_Origin_TimeBlock_DueToSecurity	0.2	0.2
(19) DayDivert_Carrier_Route_TimeBlock	0.8	0.2
(20) DayDivert_Dest_TimeBlock	0.7	1.1
(21) Distance	28.7	3.3
Mean bias (average of absolute mean standardized differences)	20.7	3.6
Observations	2,828,787	544,160

and examines how and to what extent traffic management technologies for airport surfaces could reduce taxi times, fuel consumption, and CO_2 emissions.

3. Benefit mechanism, data, and estimation model

In this section, we describe (1) the benefit mechanism of tactical surface metering, (2) the data used for estimation, and (3) the estimation model.

3.1. Benefit mechanism of tactical surface metering

"First-come, first-served" (FCFS) operations were one of the primary sources of surface traffic congestion at CLT before implementing tactical surface metering. Many airports currently employ the FCFS approach for departure management, based on the sequence in which aircraft arrive at their spots and request taxi clearance. Pilots strive to push back aircraft from their gates as early as possible to achieve on-time performance metrics during this operation. With multiple flights scheduled at similar times, congestion on the airport's surface will worsen unless measures are taken to balance the demand for runway use. Moreover, nominal traffic demand, as determined by the operating carriers' schedules, has regularly exceeded the airspace capacity available to CLT during multiple periods of the day. Due to inefficient stop-and-go taxi-out operations brought on by the mismatch between capacity and demand for both airport surface and airspace, actual takeoff times are delayed, and the airport's departure throughput is restricted (Jung et al., 2018; Talebi et al., 2019).

Based on data from 2014, NASA conducted an operational management analysis of CLT and found that during taxiing from the gate to the runway, aircraft stopped an average of 4.5 times, with an average duration of 4 min per stop. Any aircraft speed below one knot for longer than a minute is considered a stop here, as is any delay in the queue's progress. According to NASA's calculation, the average excess taxi-out time on the airport surface for all flights from January to March 2017 was approximately 6.3 min per flight when actual taxi-out times measured between pushback and takeoff were compared to undisturbed times on the same taxi route. NASA estimates that this additional time equates to 2040 tons of excess fuel consumed over the same period (Jung et al., 2018; Talebi et al., 2019).

Tactical surface metering developed by NASA improves the situation mentioned above. Specifically, it enables the user to adjust surface traffic demand by shifting excess queue time from the runway queue to the gates based on the capacity prediction, reducing fuel consumption during taxi-out. Then, by what mechanism can the benefits of tactical surface metering be achieved?

Let us denote unimpeded taxi-out times as *UniOut*, unimpeded gate holding times as *UniGate*, and unimpeded taxiing times as *UniTaxi*, respectively. Then, unimpeded taxi-out times can be expressed by the following equation:

$$UniOut = UniGate + UniTaxi \quad (1)$$

If we assume that unimpeded gate holding times, *UniGate*, can be regarded as practically zero, Eq. (1) can be rewritten as follows:

$$UniOut = UniTaxi \quad (2)$$

The actual taxi-out times in real operations involve gate holding times greater than zero. Therefore, by denoting the actual taxi-out times as *Out*, the actual gate holding times as *Gate*, and the actual additional taxiing times as *AddTaxi*, the actual taxi-out times (*Out*) can be expressed by the following equation:

$$Out = UniOut + Gate + AddTaxi \quad (3)$$

The excess taxi-out times (*Exc*) sum the latter two terms in Eq. (3), i.e., *Gate* and *AddTaxi*.

$$Exc = Gate + AddTaxi \quad (4)$$

If airport surface congestion is constant, the left term in Eq. (4) will be approximately constant regardless of the introduction of tactical surface metering. Let us denote the excess taxi-out times before the introduction of tactical surface metering as *ExcBef* and the excess taxi-out times after the introduction as *ExcAft*, and if airport surface congestion remains constant regardless of the introduction of tactical surface metering, then

$$ExcBef = ExcAft \quad (5)$$

However, even if the excess taxi-out times do not change after the introduction of tactical surface metering, the distribution of the gate holding times (*Gate*) and the additional taxiing times (*AddTaxi*) varies due to the metering, which results in different fuel consumption for *ExcBef* and *ExcAft*. For example, if *Gate* increases and *AddTaxi* decreases, fuel consumption for *ExcAft* will decrease compared to that for *ExcBef*. This is because fuel consumption per unit of time in gate holding is less than in taxiing.

NASA's analysis, for example, shows that the average excess airport surface taxi-out time, i.e., *ExcBef* was 6.3 min per flight for all flights operated from January to March 2017 (Jung et al., 2018). To illustrate, let's assume that before introducing tactical surface metering, *Gate* was 2 min, and *AddTaxi* was 4.3 min.

$$ExcBef = Gate\ (2\ \text{min}) + AddTaxi\ (4.3\ \text{min})$$

Next, assume that *Gate* is 3 min and *AddTaxi* is 3.3 min due to the introduction of tactical surface metering.

$$ExcAft = Gate\ (3\ \text{min}) + AddTaxi\ (3.3\ \text{min})$$

Table 6
Descriptive statistics of dataset A1 (Taxi-out): All observations.

Dataset A1 (all observations)	Mean	SD	Min	Max
Taxi-out	2.758	0.467	0	5.226
Post	0.590	0.492	0	1
O_CLT	0.120	0.325	0	1
Post*O_CLT	0.0804	0.272	0	1
Post1	0.0534	0.225	0	1
Post2	0.183	0.387	0	1
Post3	0.277	0.448	0	1
Post4	0.0757	0.265	0	1
Post1*O_CLT	0.00679	0.0821	0	1
Post2*O_CLT	0.0256	0.158	0	1
Post3*O_CLT	0.0375	0.190	0	1
Post4*O_CLT	0.0104	0.102	0	1
Runway_Origin	3.835	1.304	1	6
Runway_Dest	4.312	1.752	1	8
MonthDepShare_Origin_Carrier	26.574	21.605	0.00474	100
MonthArrShare_Dest_Carrier	25.237	21.803	0.00289	100
DaySchDep_Origin_Carrier_Route_TimeBlock	1.0130	0.119	1	4
DaySchArr_Dest_Carrier_Route_TimeBlock	1.0207	0.149	1	4
DaySchDep_Route_TimeBlock	18.408	16.389	1	79
DaySchArr_Route_TimeBlock	25.0965	21.756	1	110
MonthSchDep_Origin_Carrier	2585.156	2761.185	0	9734
MonthSchArr_Dest_Carrier	3839.626	5247.121	0	24,662
AvgDayDepDelay_Carrier_Route_TimeBlock	9.722	46.316	−114	2710
TotalDayDepDelay_Origin_TimeBlock	181.223	425.192	−397	9701
TotalDayArrDelay_Dest_TimeBlock	81.559	520.555	−1564	20,356
DayCancel_Carrier_Route_TimeBlock	0.000249	0.0160	0	2
DayCancel_Origin_TimeBlock_DueToCarrier	0.0751	0.501	0	38
DayCancel_Origin_TimeBlock_DueToWeather	0.124	0.843	0	64
DayCancel_Origin_TimeBlock_DueToNAS	0.0450	0.355	0	20
DayCancel_Origin_TimeBlock_DueToSecurity	0.000187	0.0150	0	3
DayDivert_Carrier_Route_TimeBlock	0.0000322	0.00570	0	2
DayDivert_Dest_TimeBlock	0.0504	0.582	0	51
Distance	6.377	0.826	3.332	8.536
Observations	5,534,608			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

Note that this does not change the average excess taxi-out times. Eq. (5) shows that if airport surface congestion remains constant, the average excess taxi-out times will remain constant regardless of introducing tactical surface metering. However, fuel consumption can be expected to change. Indeed, the aircraft can keep its engines off while waiting at the gate, which implies that fuel consumption can be lower in gate holding times than in taxiing times. Therefore, if tactical surface metering increases gate holding times and decreases additional taxiing times, fuel consumption will naturally decrease. If we denote the fuel consumption in excess taxi-out times before the introduction of tactical surface metering as *ExcFuelConsBef* and that after the introduction as *ExcFuelConsAft*, then

$$ExcFuelConsBef > ExcFuelConsAft \tag{6}$$

Sometimes, a gate conflict can occur when a departing aircraft is held at the gate where the arriving aircraft needs to be. Then, the arriving aircraft must wait to run its engines, which leads to more fuel burning. The gate conflict can cancel out the fuel savings achieved by reducing taxi-out times. However, in this situation, the ramp controller can assign the departing aircraft to the hardstand/holding area to absorb the hold time with engines off elsewhere in the ramp area. According to NASA (Jung et al., 2018), so long as this type of ramp-controlling measure is feasible, even if gate conflicts occur, we can expect that the negative impact of increased fuel consumption due to surface metering will be negligible.

3.2. Data

To estimate the effects of tactical surface metering on taxi-out, airborne, and taxi-in times, we constructed daily unbalanced panel data on a carrier–route–departure/arrival time block[4] basis using the US DOT, OTP, US DOT, T100D, US DOT, T100I, and US DOT, NFDC Airport Database's Airport Facilities and Runways Tables. Tactical surface metering at CLT, which is shown in Fig. 1, has been enabled since November 29, 2017. Therefore, our analysis spans a four-year period from November 29, 2015 to November 28, 2019. The period includes two years before and after implementing tactical surface metering at CLT. Of course, relevant data are available after November 28, 2019. However, since December 2019, the novel coronavirus spread around the world. The airline industry was hit particularly hard by the coronavirus pandemic and was virtually grounded for several months. Therefore, we have limited the dataset to the period before the pandemic to eliminate the unprecedented exceptional effects the coronavirus pandemic caused.

The period after the introduction of tactical surface metering at CLT was divided into four phases: (1) November 29, 2017, to February 18, 2018; (2) February 19, 2018, to September 30, 2018; (3) October 1, 2018, to August 31, 2019; and (4) September 1, 2019, to November 28, 2019. During the first phase, tactical surface metering was applied exclusively to the operations in the second bank (from around 9 a.m.–11 a.m.). In the second phase, the application of tactical surface metering

[4] The departure/arrival time blocks in the US DOT OTP data are defined as follows: (1) 12:00 a.m. to 5:59 a.m., (2) 6:00 a.m. to 6:59 a.m., (3) 7:00 a.m. to 7:59 a.m., (4) 8:00 a.m. to 8:59 a.m., (5) 9:00 a.m. to 9:59 a.m., (6) 10:00 a.m. to 10:59 a.m., (7) 11:00 a.m. to 11:59 a.m., (8) 12:00 p.m. to 12:59 p.m., (9) 1:00 p.m. to 1:59 p.m., (10) 2:00 p.m. to 2:59 p.m., (11) 3:00 p.m. to 3:59 p.m., (12) 4:00 p.m. to 4:59 p.m., (13) 5:00 p.m. to 5:59 p.m., (14) 6:00 p.m. to 6:59 p.m., (15) 7:00 p.m. to 7:59 p.m., (16) 8:00 p.m. to 8:59 p.m., (17) 9:00 p.m. to 9:59 p.m., (18) 10:00 p.m. to 10:59 p.m., and (19) 11:00 p.m. to 11:59 p.m..

Table 7
Descriptive statistics of dataset A2 (Taxi-out): Matched observations.

Dataset A2 (matched observations)	Mean	SD	Min	Max
Taxi-out	2.877	0.403	0	5.136
Post	0.589	0.492	0	1
O_CLT	0.500	0.500	0	1
Post*O_CLT	0.295	0.456	0	1
Post1	0.0710	0.257	0	1
Post2	0.184	0.387	0	1
Post3	0.283	0.451	0	1
Post4	0.0510	0.220	0	1
Post1*O_CLT	0.0546	0.227	0	1
Post2*O_CLT	0.0752	0.264	0	1
Post3*O_CLT	0.142	0.349	0	1
Post4*O_CLT	0.0227	0.149	0	1
Runway_Origin	3.989	0.959	1	6
Runway_Dest	4.267	1.729	1	8
MonthDepShare_Origin_Carrier	18.432	17.00132	0.00474	43.0652
MonthArrShare_Dest_Carrier	23.117	21.713	0.0257	100
DaySchDep_Origin_Carrier_Route_TimeBlock	1.00126	0.0355	1	2
DaySchArr_Dest_Carrier_Route_TimeBlock	1.00374	0.0610	1	2
DaySchDep_Route_TimeBlock	27.625	17.819	1	79
DaySchArr_Route_TimeBlock	22.851	19.768	1	109
MonthSchDep_Origin_Carrier	3772.972	3621.620	1	9368
MonthSchArr_Dest_Carrier	3486.260	5265.785	1	24,662
AvgDayDepDelay_Carrier_Route_TimeBlock	10.695	47.357	−44	1445
TotalDayDepDelay_Origin_TimeBlock	259.769	492.0006	−381	9701
TotalDayArrDelay_Dest_TimeBlock	81.487	498.733	−1520	20,356
DayCancel_Carrier_Route_TimeBlock	0.0000125	0.00354	0	1
DayCancel_Origin_TimeBlock_DueToCarrier	0.0911	0.385	0	19
DayCancel_Origin_TimeBlock_DueToWeather	0.178	0.982	0	47
DayCancel_Origin_TimeBlock_DueToNAS	0.0341	0.250	0	16
DayCancel_Origin_TimeBlock_DueToSecurity	0.000276	0.0166	0	1
DayDivert_Carrier_Route_TimeBlock	0	0	0	0
DayDivert_Dest_TimeBlock	0.0461	0.537	0	46
Distance	6.351	0.716	4.00733	8.514
Observations	239,222			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

was extended to include the second and third banks (from around 9 a.m. to 1 p.m.). In the third phase, tactical surface metering was applied to all banks. In the fourth phase, i.e., after September 2019, tactical surface metering was used for all banks. However, due to changes in airline schedules in September 2019, more flights were assigned metering holds in the fourth phase (NASA ATD-2 team, 2020). As pointed out above, we excluded the period after the beginning of the coronavirus pandemic to prevent outliers caused by the pandemic from unduly affecting our estimates. Therefore, the fourth phase started on September 1, 2019, and ended on November 28, 2019.

We constructed our datasets on a directional trip basis because the tactical surface metering at CLT might have exerted different effects on inbound flights from those on outbound flights. Therefore, Dataset A included outbound flights from CLT but not inbound flights to CLT, whereas Datasets B and C included inbound flights to CLT but not outbound flights from CLT.

3.3. Model

The basic estimation model is given by Eq. (7):

$$Y_{iodbt} = \theta_t + \alpha POST + \beta CLT + \gamma POST * CLT + x_{iodbt}\delta + c_{iodb} + u_{iodbt} \tag{7}$$

The subscripts i, o, d, b, and t, denote carrier, origin, destination, departure/arrival time block, and day, respectively. The dependent variable, Y_{iodbt}, represents either (a) average taxi-out times (minutes)[5] for outbound flights from the origin (logged), (b) average airborne times per mile (minutes) for inbound flights to the destination (logged), or (c) average taxi-in times (minutes) for inbound flights to the destination (logged). We used averaged values as the dependent variables because the US DOT OTP data implied that some carriers scheduled multiple flights to the same destination during the same departure/arrival time block.

In this study, the treatment was the implementation of tactical surface metering at CLT, which started on November 29, 2017. The treatment group consisted of flights to and from CLT, and the comparison group consisted of flights to and from airports other than CLT. The estimation model included *POST* and *CLT* dummies and an interaction term between these dummies, $POST * CLT$. The *POST* dummy indicated the periods after tactical surface metering was enabled at CLT; its value was one if the flight date occurred between November 29, 2017 and November 28, 2019, and 0 otherwise. The *CLT* dummy represented the treatment group; its value was one if the origin or destination were CLT and 0 otherwise. The interaction term, $POST * CLT$, estimated the effects of tactical surface metering at CLT. We also included controls (x_{iodbt}) and time-varying errors (u_{iodbt}) in Eq. (7). The variable c_{iodb} represented fixed effects defined on a carrier–route–departure/arrival time block basis,

[5] The taxi-out time was defined as the time that passes between an actual off-block time (AOBT) and actual take-off time (ATOT) (Eurocontrol, 2021). The taxi-in time refers to the period between the actual landing time (ALDT) and the actual in-block time (AIBT).

Table 8
Descriptive statistics of dataset B1 (Airborne): All observations.

Dataset B1 (all observations)	Mean	SD	Min	Max
Airborne	−1.892	0.186	−3.332	0.997
Post	0.563	0.496	0	1
D_CLT	0.0669	0.250	0	1
Post*D_CLT	0.0448	0.207	0	1
Post1	0.0529	0.224	0	1
Post2	0.176	0.381	0	1
Post3	0.263	0.440	0	1
Post4	0.0713	0.257	0	1
Post1*D_CLT	0.0669	0.250	0	1
Post2*D_CLT	0.00379	0.0614	0	1
Post3*D_CLT	0.0143	0.119	0	1
Post4*D_CLT	0.0209	0.143	0	1
Runway_Origin	0.00581	0.0760	0	1
Runway_Dest	4.350	1.783	1	8
MonthDepShare_Origin_Carrier	3.660	1.475	1	6
MonthArrShare_Dest_Carrier	26.490	22.379	0.00498	100
DaySchDep_Origin_Carrier_Route_TimeBlock	28.866	23.619	0.00475	100
DaySchArr_Dest_Carrier_Route_TimeBlock	1.0155	0.142	1	7
DaySchDep_Route_TimeBlock	1.0116	0.118	1	7
DaySchArr_Route_TimeBlock	25.592	21.861	1	109
MonthSchDep_Origin_Carrier	18.466	17.283	1	84
MonthSchArr_Dest_Carrier	3790.587	5161.297	0	24,649
AvgDayDepDelay_Carrier_Route_TimeBlock	2785.322	2793.284	0	9734
TotalDayDepDelay_Origin_TimeBlock	8.419	41.854	−204	2468
TotalDayArrDelay_Dest_TimeBlock	257.728	567.395	−411	15,213
DayCancel_Carrier_Route_TimeBlock	33.473	281.597	−1580	11,620
DayCancel_Origin_TimeBlock_DueToCarrier	0.000165	0.0129	0	2
DayCancel_Origin_TimeBlock_DueToWeather	0.0820	0.432	0	44
DayCancel_Origin_TimeBlock_DueToNAS	0.139	0.935	0	68
DayCancel_Origin_TimeBlock_DueToSecurity	0.0532	0.441	0	38
DayDivert_Carrier_Route_TimeBlock	0.000181	0.0162	0	7
DayDivert_Dest_TimeBlock	0.0000203	0.00451	0	1
Distance	0.0342	0.427	0	34
Observations	9,890,246			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

differenced-out in the fixed-effects model. Therefore, we can control unobserved heterogeneity or confounding factors specific to the carrier–route–departure/arrival time block.

The controls included 21 variables, such as the numbers of runways and scheduled departures/arrivals at origin/destination airports, constructed from the following data sources: US DOT, NFDC Airport Database's Airport Facilities and Runways Tables; US Department of Transportation, and US Department of Transportation; and US Department of Transportation. Our estimation model also included aggregated time-fixed effects (year, month, week, and day dummies) as controls. Table 1 presents detailed descriptions and abbreviations of the variables.

4. Validity of the model assumption

4.1. Parallel-trend assumption checks

Using a difference-in-differences approach with covariate-balancing methods, we estimated the effects of tactical surface metering at CLT on fuel and CO_2 reduction. Therefore, we constructed our datasets in which the dependent variables exhibited a parallel trend before the deployment of tactical surface metering at CLT (before November 29, 2017).

There is considerable uncertainty regarding the selection of comparison groups in difference-in-differences analyses. The choice of comparison groups often depends on the affinities between treated and untreated units perceived subjectively by the researchers. To address the uncertainty, we used data-driven approaches in constructing comparison groups to minimize discretion in selecting the comparison units. Indeed, it is difficult to single out any airport as a suitable comparison unit, even from the American Airlines hub airports, because CLT has the highest share of flights operated by the airline during the analysis period. The idea behind our data-driven approaches is equivalent to the synthetic control approach: a combination of airports can provide a better comparison for the airport exposed to the intervention than any single airport alone (Abadie et al., 2010; Abadie & Gardeazabal, 2003).

First, we selected and combined observations of flights from/to CLT and only one other airport that appeared in the US DOT OTP data. Then, we examined whether the trends of the dependent variables were parallel in the treatment (i.e., CLT) and comparison (i.e., one other airport) groups. We performed this parallel-trends test by adding two three-way interactions between a pre-treatment period variable (*PRE*), the *CLT* dummy, and the date variable (*PRE * CLT * DATE*), and a post-treatment period variable (*POST*), the *CLT* dummy, and the date variable (*POST * CLT * DATE*) in Eq. (7). The coefficient of *PRE * CLT * DATE* captured the differences in slopes between the treatment and comparison groups during the period before tactical surface metering was enabled at CLT. If the coefficient of *PRE * CLT * DATE* was not statistically significant (i.e., 0), we regarded it as statistical evidence of a parallel linear trend between the two groups during the pre-treatment period.

We simplify the notation by rewriting Eq. (7) as follows:

$$Y_{iodbt} = DID + u_{iodbt} \tag{8}$$

Then, we can write Eq. (9) for the parallel-trends test as follows:

Table 9
Descriptive statistics: Dataset B2 (Airborne): Matched observations.

Dataset B2 (matched observations)	Mean	SD	Min	Max
Airborne	−1.853	0.221	−2.567	0.0531
Post	0.709	0.454	0	1
D_CLT	0.500	0.500	0	1
Post*D_CLT	0.354	0.478	0	1
Post1	0.0623	0.242	0	1
Post2	0.219	0.414	0	1
Post3	0.346	0.476	0	1
Post4	0.0811	0.273	0	1
Post1*D_CLT	0.0465	0.211	0	1
Post2*D_CLT	0.0912	0.288	0	1
Post3*D_CLT	0.177	0.382	0	1
Post4*D_CLT	0.0392	0.194	0	1
Runway_Origin	3.973	1.739	1	8
Runway_Dest	4.0130	1.279	1	6
MonthDepShare_Origin_Carrier	24.956	22.312	0.00952	100
MonthArrShare_Dest_Carrier	18.390	15.176	0.00475	53.822
DaySchDep_Origin_Carrier_Route_TimeBlock	1.00822	0.0956	1	6
DaySchArr_Dest_Carrier_Route_TimeBlock	1.00358	0.0600	1	3
DaySchDep_Route_TimeBlock	20.787	20.134	1	109
DaySchArr_Route_TimeBlock	33.492	20.878	1	84
MonthSchDep_Origin_Carrier	3394.779	5534.990	0	24,649
MonthSchArr_Dest_Carrier	3863.0590	3339.948	0	9734
AvgDayDepDelay_Carrier_Route_TimeBlock	8.655	44.0655	−32	1594
TotalDayDepDelay_Origin_TimeBlock	203.372	498.0658	−411	14,685
TotalDayArrDelay_Dest_TimeBlock	67.698	428.845	−1580	9026
DayCancel_Carrier_Route_TimeBlock	0.0000916	0.00957	0	1
DayCancel_Origin_TimeBlock_DueToCarrier	0.0639	0.354	0	44
DayCancel_Origin_TimeBlock_DueToWeather	0.123	0.911	0	66
DayCancel_Origin_TimeBlock_DueToNAS	0.0504	0.431	0	32
DayCancel_Origin_TimeBlock_DueToSecurity	0.0000943	0.0137	0	6
DayDivert_Carrier_Route_TimeBlock	0.00000832	0.00288	0	1
DayDivert_Dest_TimeBlock	0.0687	0.645	0	28
Distance	6.282	0.756	4.205	8.287
Observations	360,442			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

$$Y_{iodbt} = DID + DATE + PRE * CLT * DATE + POST * CLT * DATE + \epsilon_{iodbt} \tag{9}$$

We repeated the parallel-trends test using Eq. (9) for all datasets, combining observations of flights from/to CLT and each of the other airports that appeared in the US DOT OTP data. We then pooled the datasets showing parallel trends ($p \geq 0.1$) and tested them again to confirm that the combined datasets showed parallel trends. After the test, we assessed the balance of variables between the treatment and comparison groups using the absolute standardized difference (the details are given in Section 4.2). If the standardized difference for each variable after one-to-one propensity score matching was larger than 15% between the treatment and comparison groups, the covariate balance check was repeated with datasets combining observations of flights from/to CLT and other airports with higher p-values for the coefficient of *PRE * CLT * DATE* until the standardized difference of each variable became less than 15%. The test statistics for the coefficients of *PRE * CLT * DATE* in the datasets that passed the covariate balance check are shown in Table 2. The airports in the datasets and their three-letter codes are shown in Appendix A. These airports provide commercial services according to the FAA Airport Master Record (5010).

In Fig. 2 through 4, using the data selected above, we plotted the monthly average values of the dependent variables for the treatment and comparison groups (taxi-out, airborne, and taxi-in times) to examine whether they exhibited a parallel trend during the analysis period. The bold solid lines represent the trends of the treatment groups, and the dashed bold lines represent the trends of the comparison groups. Thin, straight lines show fitted value trends. In all figures, the above variables exhibited roughly parallel trends before the deployment of tactical surface metering at CLT. The parallel trends implied that our comparison group is a reasonable control for our difference-in-differences approach.

4.2. Mitigation of potential selection bias

The above observation regarding the validity of the comparison group may be too strong. Indeed, CLT was not randomly selected as the deployment site for tactical surface metering. Consequently, the distributions of covariates might be significantly different between the treatment and comparison groups in our datasets. Table 3 shows the absolute mean standardized differences of each variable between the two groups in the original and matched samples from Dataset A (including outbound flights from CLT but not inbound flights to CLT). Column 1 of Table 3 shows that 14 of the 21 variables had more than 10% standardized differences in the original sample. Although less than 10–25% standardized differences are generally considered an acceptable imbalance (Austin, 2009; Garrido et al., 2014; Stuart et al., 2013), nine variables had more than 25% standardized differences. Column 1 suggests that many variables in the original sample would likely be unacceptable for the difference-in-differences analysis without matching. The comparison group in the original sample may not be regarded as a valid comparison group. We observed a similar tendency for samples from Datasets B and C (including inbound flights to CLT but not outbound flights from CLT), as Tables 4 and 5 show.

We employed propensity score matching (PSM) to balance observed covariates across the treatment and comparison groups (Guo & Fraser, 2014). Furthermore, we tried to employ other balancing methods: coarsened exact matching (CEM; Blackwell et al., 2009; Iacus et al., 2012) and entropy balancing (EB; Hainmueller, 2012; Hainmueller &

Table 10
Descriptive statistics of dataset C1 (Taxi-in): All observations.

Dataset C1 (all observations)	Mean	SD	Min	Max
Taxi-in	2.0134	0.605	0	5.598
Post	0.625	0.484	0	1
D_CLT	0.234	0.423	0	1
Post*D_CLT	0.156	0.363	0	1
Post1	0.0541	0.226	0	1
Post2	0.196	0.397	0	1
Post3	0.294	0.456	0	1
Post4	0.0811	0.273	0	1
Post1*D_CLT	0.0132	0.114	0	1
Post2*D_CLT	0.0499	0.218	0	1
Post3*D_CLT	0.0730	0.260	0	1
Post4*D_CLT	0.0203	0.141	0	1
Runway_Origin	4.133	1.843	1	8
Runway_Dest	4.406	1.967	1	7
MonthDepShare_Origin_Carrier	22.281	20.785	0.00900	100
MonthArrShare_Dest_Carrier	29.466	21.199	0.00475	100
DaySchDep_Origin_Carrier_Route_TimeBlock	1.0132	0.116	1	4
DaySchArr_Dest_Carrier_Route_TimeBlock	1.00945	0.0992	1	4
DaySchDep_Route_TimeBlock	23.0711	21.967	1	109
DaySchArr_Route_TimeBlock	29.336	26.615	1	103
MonthSchDep_Origin_Carrier	3090.754	4940.899	0	24,649
MonthSchArr_Dest_Carrier	5122.143	5165.496	1	15,921
AvgDayDepDelay_Carrier_Route_TimeBlock	9.588	47.696	−80	2149
TotalDayDepDelay_Origin_TimeBlock	231.314	545.424	−411	15,213
TotalDayArrDelay_Dest_TimeBlock	72.282	443.571	−1580	15,040
DayCancel_Carrier_Route_TimeBlock	0.000217	0.0148	0	2
DayCancel_Origin_TimeBlock_DueToCarrier	0.0730	0.401	0	44
DayCancel_Origin_TimeBlock_DueToWeather	0.127	0.898	0	68
DayCancel_Origin_TimeBlock_DueToNAS	0.0501	0.435	0	32
DayCancel_Origin_TimeBlock_DueToSecurity	0.000165	0.0159	0	7
DayDivert_Carrier_Route_TimeBlock	0.0000389	0.00624	0	1
DayDivert_Dest_TimeBlock	0.0694	0.809	0	51
Distance	6.339	0.713	3.807	8.255
Observations	2,828,787			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

Xu, 2013). Unfortunately, we did not achieve balance in our datasets using CEM and EB. Therefore, we used PSM in our analysis.

We present the procedure details for obtaining the balancing scores (i.e., propensity scores) in Appendix B. This section reports the balance of covariates across the treatment and comparison groups after PSM. We used the nearest neighbor one-to-one matching process to match one observation in the treatment group to one in the comparison group with a similar probability of being treated. We matched the observation *i* in the treatment group with the observation *j* in the comparison group using the absolute difference of propensity scores between *i* and *j*. The caliper was defined as the standard deviation of the propensity score multiplied by 0.2 (Austin, 2011). We selected the pair of *i* and *j* with the smallest absolute difference in propensity scores among all pairs as a matched pair if the absolute difference in propensity scores was within the above caliper. We assessed the balance of variables between the treatment and comparison groups using the absolute standardized difference, which is the absolute value of the difference in average values of a variable in two groups divided by the pooled standard deviation of that variable (Austin, 2019).

Column 2 of Table 3 shows each variable's absolute mean standardized differences in the matched samples from Dataset A. All variables showed less than 10% standardized differences in the matched samples presented in Column 2. We observed a similar result for Datasets B and C. Column 2 of Table 4 indicates that no variable shows a standardized difference of more than 10% after matching. In Column 2 of Table 5, three variables showed more than 10% standardized differences. However, they are less than 15%, and the mean bias (i.e., the average of the absolute mean standardized differences of all variables) is 3.6%. Therefore, most variables seem balanced, suggesting that PSM reduced the observable differences between the treatment and comparison groups.

In the following analyses, we combine matching with parametric regression called "doubly robust estimation." The two-step procedure is doubly robust because if either the matching or the parametric model is misspecified but one of them is correctly specified, causal estimates will still be consistent. A simple difference in means does not have this double robustness (Bang & Robins, 2005; Ho et al., 2007). However, one crucial point is that these covariate-balancing methods can only control for observed confounders. Unobserved selection bias cannot be fully controlled, even with PSM. This is "always a limitation of non-randomized studies compared with randomized studies" (Rubin, 1997).

5. Estimation results

5.1. The average effect over the entire post-deployment period

Table 6 through 11 present descriptive statistics for the samples we obtained from the raw and matched Datasets A, B, and C. A negative sign of variables representing departure/arrival delays[6] indicates that the actual departure or arrival of the flight was earlier than the original scheduled departure or arrival time, and the value indicates the difference between the actual departure or arrival time and the original departure or arrival time. Table 12 through 14 report the estimation results using each dataset. Columns 1 through 4 in each table show the estimation results from the original and matched samples with and without origin- and destination-specific year-fixed effects, respectively.

[6] AvgDayDepDelay_Carrier_Route_TimeBlock, TotalDayDepDelay_Origin_TimeBlock, and TotalDayArrDelay_Dest_TimeBlock.

Table 11
Descriptive statistics: Dataset C2 (Taxi-in): Matched observations.

Dataset C2 (matched observations)	Mean	SD	Min	Max
Taxi-in	2.0801	0.608	0	5.46
Post	0.691	0.462	0	1
D_CLT	0.500	0.500	0	1
Post*D_CLT	0.346	0.476	0	1
Post1	0.0557	0.229	0	1
Post2	0.191	0.393	0	1
Post3	0.349	0.477	0	1
Post4	0.0953	0.294	0	1
Post1*D_CLT	0.0346	0.183	0	1
Post2*D_CLT	0.0860	0.280	0	1
Post3*D_CLT	0.180	0.384	0	1
Post4*D_CLT	0.0461	0.210	0	1
Runway_Origin	4.147	1.756	1	8
Runway_Dest	3.900	1.714	1	7
MonthDepShare_Origin_Carrier	22.656	20.260	0.0125	100
MonthArrShare_Dest_Carrier	24.830	18.329	0.00475	73.529
DaySchDep_Origin_Carrier_Route_TimeBlock	1.0105	0.105	1	4
DaySchArr_Dest_Carrier_Route_TimeBlock	1.00556	0.0746	1	3
DaySchDep_Route_TimeBlock	22.969	20.512	1	109
DaySchArr_Route_TimeBlock	30.152	27.416	1	103
MonthSchDep_Origin_Carrier	3466.272	5229.0500	0	24,649
MonthSchArr_Dest_Carrier	5392.140	5336.715	1	15,921
AvgDayDepDelay_Carrier_Route_TimeBlock	9.339	44.212	−63	1934
TotalDayDepDelay_Origin_TimeBlock	228.686	523.832	−411	14,920
TotalDayArrDelay_Dest_TimeBlock	69.546	423.353	−1580	15,040
DayCancel_Carrier_Route_TimeBlock	0.000182	0.0135	0	1
DayCancel_Origin_TimeBlock_DueToCarrier	0.0726	0.401	0	44
DayCancel_Origin_TimeBlock_DueToWeather	0.132	0.938	0	68
DayCancel_Origin_TimeBlock_DueToNAS	0.0558	0.458	0	32
DayCancel_Origin_TimeBlock_DueToSecurity	0.000176	0.0156	0	4
DayDivert_Carrier_Route_TimeBlock	0.0000110	0.00332	0	1
DayDivert_Dest_TimeBlock	0.0699	0.786	0	51
Distance	6.343	0.711	4.00733	8.243
Observations	544,160			

Note: Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. Negative values of departure/arrival delay variables mean early departures or arrivals.

Table 12 shows the estimation results for Dataset A. Even columns show the estimation results controlling for origin- and destination-specific year-fixed effects. Although we are controlling for the effects of the number of runways at each airport and other factors, changes in the number of gates, taxi distance from a given gate to a given runway, and other additional factors can affect the dependent variables in this study. However, identifying these changes at all airports in our datasets, which include more than 200 airports, is not feasible because data on these changes are not publicly available. Therefore, we controlled for the effects of changes other than the number of runways at each airport through airport-specific year-fixed effects. If we construct interaction terms of origin/destination airport and year dummies, around 1000 dummies must be included in our estimation equations. It is computationally challenging to perform estimations with such many dummies. Thus, we performed estimations with origin- and destination-specific year-fixed effects using reghdfe, a user-written Stata program (Constantine & Correia, 2021; Correia, 2017; Correia et al., 2020). It absorbs high-dimensional fixed effects by recalculating dependent and independent variables to have a mean 0 within the groups specified and adding back the overall mean of each variable. Then, it regresses the adjusted dependent variable on the adjusted independent variables. As a result, the distance variable is omitted in some of the estimations performed by reghdfe because it is highly collinear with the fixed effects.

In Table 12, the dependent variable is the daily average taxi-out times (minutes) for outbound flights from the origin airport calculated on a carrier–route–departure time block basis (logged). All columns of Table 12 show that the coefficients of the interaction term of the *Post* and *CLT* dummies, *POST * CLT*, are negative and statistically significant. The estimates controlling for origin/destination-specific year fixed effects suggest that, after the deployment of tactical surface metering at CLT, average taxi-out times of outbound flights from CLT decreased by about 2.1%[7] (Column 2) or 5.4%[8] (Column 4) compared to outbound flights from the comparison group. However, as described in the previous section, the covariate balance in the original datasets cannot be regarded as a satisfactory comparison group. Therefore, hereafter, we consider the results obtained using matched samples (Column 4) the most reliable. The results support the conclusion reached by the FAA and NASA: tactical surface metering technology can reduce taxi-out times.

The coefficients of other variables are statistically significant except for the following six variables: the number of runways at the destination, the monthly share of performed arrivals at destination calculated on a carrier basis, the daily number of scheduled departures/arrivals aggregated on a carrier–route–departure/arrival time block basis, the monthly number of scheduled departures/arrivals at an origin/destination aggregated on a carrier basis, the daily number of canceled flights aggregated on a carrier–route–departure/arrival time block basis, and the daily number of flights canceled due to reasons within the control of carrier and security concerns aggregated on an origin–departure time block basis. (The number of runways at origin, the daily number of diverted flights aggregated on a carrier–route–departure/arrival time block basis, and the distance variable are omitted because it is highly collinear with the fixed effects.)

The sign of the statistically significant coefficients of other control variables is positive except for one variable: the average daily departure delays (minutes) calculated on a carrier–route–departure time block basis at the origin. The negative and statistically significant coefficient

[7] $(\exp(-0.0216)-1)*100 \approx -2.1$
[8] $(\exp(-0.0552)-1)*100 \approx -5.4$

Table 12
Estimation results – Taxi-out: Datasets A1 and A2.

Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post	0.0354***	0.0250***	0.0509***	0.0455**
	(0.00227)	(0.00219)	(0.0133)	(0.0160)
Post*O_CLT	−0.145***	−0.0216***	−0.0654***	−0.0552**
	(0.00564)	(0.00643)	(0.0131)	(0.0187)
Runway_Origin	−0.00517	0.00865	0.165***	–
	(0.00379)	(0.00516)	(0.0119)	–
Runway_Dest	−0.00314	−0.00637	−0.00594	−0.00156
	(0.00391)	(0.00390)	(0.0112)	(0.0102)
MonthDepShare_Origin_Carrier	−0.00134***	−0.000927***	0.00371*	0.00859***
	(0.000165)	(0.000148)	(0.00185)	(0.00210)
MonthArrShare_Dest_Carrier	−0.000594**	−0.000297	−0.000158	−0.000829
	(0.000212)	(0.000188)	(0.000600)	(0.000647)
DaySchDep_Origin_Carrier_Route_TimeBlock	0.0149***	0.0143***	0.00985	0.0155
	(0.00341)	(0.00320)	(0.0235)	(0.0236)
DaySchArr_Dest_Carrier_Route_TimeBlock	0.00296	0.00321	−0.0158	−0.0134
	(0.00305)	(0.00294)	(0.0201)	(0.0194)
DaySchDep_Route_TimeBlock	0.00310***	0.00369***	0.00228***	0.00222***
	(0.000111)	(0.000124)	(0.000293)	(0.000272)
DaySchArr_Route_TimeBlock	0.00152***	0.00143***	0.00128***	0.00147***
	(0.0000792)	(0.0000771)	(0.000302)	(0.000269)
MonthSchDep_Origin_Carrier	0.0000207***	0.0000309***	0.00000220	−0.00000205
	(0.00000198)	(0.00000202)	(0.00000734)	(0.00000733)
MonthSchArr_Dest_Carrier	−0.00000103	−0.00000154	−0.00000169	−0.00000317
	(0.00000113)	(0.00000116)	(0.00000433)	(0.00000405)
AvgDayDepDelay_Carrier_Route_TimeBlock	−0.000106***	−0.000107***	−0.000128***	−0.000129***
	(0.00000807)	(0.00000808)	(0.0000294)	(0.0000294)
TotalDayDepDelay_Origin_TimeBlock	0.0000732***	0.0000722***	0.0000501***	0.0000498***
	(0.00000137)	(0.00000137)	(0.00000305)	(0.00000305)
TotalDayArrDelay_Dest_TimeBlock	0.0000577***	0.0000573***	0.0000507***	0.0000509***
	(0.000000896)	(0.000000895)	(0.00000296)	(0.00000308)
DayCancel_Carrier_Route_TimeBlock	0.0216	0.0216	−0.110	−0.0957
	(0.0140)	(0.0140)	(0.191)	(0.188)
DayCancel_Origin_TimeBlock_DueToCarrier	−0.0108***	−0.0104***	−0.00184	−0.00204
	(0.000396)	(0.000392)	(0.00217)	(0.00218)
DayCancel_Origin_TimeBlock_DueToWeather	0.00920***	0.00954***	0.0103***	0.0104***
	(0.000371)	(0.000372)	(0.00129)	(0.00130)
DayCancel_Origin_TimeBlock_DueToNAS	0.0226***	0.0233***	0.0106*	0.0106*
	(0.000932)	(0.000927)	(0.00510)	(0.00508)
DayCancel_Origin_TimeBlock_DueToSecurity	0.0226	0.0190	−0.0432	−0.0370
	(0.0158)	(0.0158)	(0.0381)	(0.0375)
DayDivert_Carrier_Route_TimeBlock	0.0152	0.0163	–	–
	(0.0292)	(0.0291)	–	–
DayDivert_Dest_TimeBlock	0.00772***	0.00781***	0.00764***	0.00768***
	(0.000434)	(0.000433)	(0.00212)	(0.00213)
Distance	−1.623	–	27.38*	–
	(2.260)	–	(13.85)	–
Observations	5,534,608	5,534,608	239,222	239,222
Adjusted R^2	0.018	0.422	−0.007	0.285
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓

Note: Standard errors in parentheses are clustered by carrier, route, and departure-arrival time block. The data set is an unbalanced panel. Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. In the estimations performed by reghdfe (a user-written Stata program for absorbing high-dimensional fixed effects), which are shown in columns 2 and 4, several variables (e.g., the distance variable) are omitted because they are highly collinear with the fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

of this variable suggests that carriers experiencing departure delays in a given time block were trying to recover from departure delays during the taxi-out phase.

On the other hand, the positive and statistically significant coefficients of other control variables suggest that increases in the following variables correlate to increased taxi-out times: the monthly share of performed departures at origin, the number of scheduled departures/arrivals, the length (in minutes) of total daily departure/arrival delays for each time block, the number of flights canceled due to reasons beyond the control of the carrier, and the number of diverted flights.

First, it is straightforward to understand why the number of canceled and diverted flights due to reasons beyond the carriers' control would lead to increased taxi-out times. Factors such as weather conditions in the origin airport regions, congestion at NAS, and incidences at arrival airports can be expected to delay aircraft takeoffs at the origin airports and directly cause an increase in taxi-out times. Second, it is also understandable that carriers with a higher monthly share of departures performed at origin airports and carriers with a greater number of scheduled departures/arrivals on a given route would experience increased taxi-out times because they have a greater opportunity to be affected by factors that would directly cause an increase in taxi-out times, such as those described above. Finally, the increase in the length (in minutes) of total daily departure/arrival delays for each time block means that the nominal traffic demand based on carriers' published schedules at a given airport increased during a given time period. It is straightforward to understand that surface congestion would cause inefficient stop-and-go taxiing operations, increasing taxi-out times (Jung et al., 2018).

Table 13 reports the estimation results for Dataset B. The dependent

Table 13
Estimation results – Airborne: Datasets B1 and B2.

Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post	0.00164***	0.00190***	0.00174	0.00630*
	(0.000397)	(0.000412)	(0.00280)	(0.00306)
Post*D_CLT	−0.00822***	−0.00894**	−0.00985***	−0.0195***
	(0.00102)	(0.00284)	(0.00183)	(0.00585)
Runway_Origin	0.00209***	0.00818***	−0.00144	0.00250
	(0.000549)	(0.000894)	(0.00246)	(0.00384)
Runway_Dest	0.000353	0.00207**	0.00883	0.0206
	(0.000455)	(0.000682)	(0.00818)	(0.0136)
MonthDepShare_Origin_Carrier	−0.0000416	−0.00000337	0.0000280	0.0000330
	(0.0000308)	(0.0000350)	(0.000149)	(0.000162)
MonthArrShare_Dest_Carrier	−0.000000593	0.0000155	0.000786*	0.000370
	(0.0000242)	(0.0000262)	(0.000333)	(0.000399)
DaySchDep_Origin_Carrier_Route_TimeBlock	0.00279***	0.00249***	−0.00591*	−0.00503
	(0.000594)	(0.000581)	(0.00291)	(0.00285)
DaySchArr_Dest_Carrier_Route_TimeBlock	0.00174**	0.00158**	−0.00178	−0.00330
	(0.000582)	(0.000565)	(0.00346)	(0.00322)
DaySchDep_Route_TimeBlock	−0.0000410***	−0.00000631	0.000101	0.000121*
	(0.0000119)	(0.0000123)	(0.0000541)	(0.0000580)
DaySchArr_Route_TimeBlock	0.000331***	0.000318***	0.000325***	0.000231***
	(0.0000148)	(0.0000148)	(0.0000467)	(0.0000467)
MonthSchDep_Origin_Carrier	0.00000123***	0.00000114***	0.00000695***	0.00000806***
	(0.000000285)	(0.000000313)	(0.00000121)	(0.00000123)
MonthSchArr_Dest_Carrier	−0.000000569	−0.00000121**	−0.000000560	0.00000104
	(0.000000353)	(0.000000400)	(0.00000164)	(0.00000190)
AvgDayDepDelay_Carrier_Route_TimeBlock	−0.0000595***	−0.0000596***	−0.0000937***	−0.0000940***
	(0.00000113)	(0.00000112)	(0.00000553)	(0.00000553)
TotalDayDepDelay_Origin_TimeBlock	0.00000263***	0.00000262***	0.00000413***	0.00000424***
	(0.000000135)	(0.000000134)	(0.000000569)	(0.000000566)
TotalDayArrDelay_Dest_TimeBlock	0.0000274***	0.0000273***	0.0000249***	0.0000249***
	(0.000000252)	(0.000000250)	(0.000000727)	(0.000000729)
DayCancel_Carrier_Route_TimeBlock	0.00845**	0.00804*	0.0180	0.0172
	(0.00315)	(0.00314)	(0.0232)	(0.0237)
DayCancel_Origin_TimeBlock_DueToCarrier	−0.000583***	−0.000590***	−0.00101*	−0.00103*
	(0.0000675)	(0.0000670)	(0.000428)	(0.000429)
DayCancel_Origin_TimeBlock_DueToWeather	0.000492***	0.000506***	0.000482*	0.000487**
	(0.0000453)	(0.0000455)	(0.000188)	(0.000188)
DayCancel_Origin_TimeBlock_DueToNAS	0.0000563	0.0000800	0.000296	0.000301
	(0.0000946)	(0.0000941)	(0.000470)	(0.000462)
DayCancel_Origin_TimeBlock_DueToSecurity	0.00752***	0.00734***	0.0139	0.0126
	(0.00147)	(0.00146)	(0.00936)	(0.00929)
DayDivert_Carrier_Route_TimeBlock	0.0290**	0.0295**	0.0578	0.0568
	(0.0103)	(0.0103)	(0.0334)	(0.0312)
DayDivert_Dest_TimeBlock	0.0147***	0.0147***	0.0150***	0.0150***
	(0.000192)	(0.000192)	(0.000640)	(0.000641)
Distance	−0.986***	–	0.915	–
	(0.257)	–	(2.672)	–
Observations	9,890,246	9,890,246	360,442	360,442
Adjusted R^2	0.022	0.852	−0.003	0.875
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓

Note: Standard errors in parentheses are clustered by carrier, route, and departure-arrival time block. The data set is an unbalanced panel. Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. In the estimations performed by reghdfe (a user-written Stata program for absorbing high-dimensional fixed effects), which are shown in columns 2 and 4, the distance variable is omitted because it is highly collinear with the fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

variable is daily average airborne times (minutes) per mile for inbound flights to the destination, calculated on a carrier–route–arrival time block basis (logged). The coefficients of the interaction term *POST* * *CLT* are again negative regardless of sample type (original or matched) and statistically significant. The estimation result obtained from the matched sample, which controls for origin- and destination-specific year fixed effects (Column 4), suggests that, after the deployment of tactical surface metering at CLT, average airborne times per mile of inbound flights to CLT decreased by about 1.9%[9] compared to inbound flights in the comparison group. The result suggests that surface metering technology did not yield the expected adverse side effects, i.e., increased airborne time due to the gate-holding resulting from the technology.

Furthermore, the estimation results reported in Table 14, which we obtained from Dataset C, suggest that the other expected adverse side effect, i.e., increased taxi-in time, was also not observed during the taxi-in phase of operation. In Column 4 of Table 14, the coefficient of the interaction term *POST* * *CLT* is −0.00501 but is statistically insignificant. This estimate suggests that, after the deployment of tactical surface metering at CLT, average taxi-in times of inbound flights to CLT decreased compared to inbound flights in the comparison group but were not statistically significant. Gate-holding of departures from CLT prompted by surface metering technology did not seem to have resulted in a situation where a flight arriving at CLT found its initially assigned gate was occupied by another aircraft.

[9] $(\exp(-0.0195) - 1) * 100 \approx -1.9$

Table 14
Estimation results – Taxi-in: Datasets C1 and C2.

Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post	−0.0192***	−0.0195***	−0.00553	−0.0122
	(0.00403)	(0.00402)	(0.00892)	(0.0102)
Post*D_CLT	−0.00989	−0.0111	−0.0148	−0.00501
	(0.00537)	(0.00870)	(0.00761)	(0.0133)
Runway_Origin	0.00253	−0.0112	0.0282*	0.00853
	(0.00859)	(0.0104)	(0.0129)	(0.0187)
Runway_Dest	−0.0541***	−0.0636***	–	–
	(0.0119)	(0.0145)	–	–
MonthDepShare_Origin_Carrier	0.000884**	0.000254	−0.000300	0.0000386
	(0.000275)	(0.000221)	(0.000674)	(0.000640)
MonthArrShare_Dest_Carrier	−0.000311	−0.000129	0.000493	−0.00114
	(0.000251)	(0.000217)	(0.000590)	(0.000616)
DaySchDep_Origin_Carrier_Route_TimeBlock	0.00958	0.0101	0.0136	0.00993
	(0.00730)	(0.00704)	(0.0172)	(0.0161)
DaySchArr_Dest_Carrier_Route_TimeBlock	0.0297***	0.0292***	0.0395	0.0327
	(0.00802)	(0.00777)	(0.0211)	(0.0199)
DaySchDep_Route_TimeBlock	0.000182	−0.0000927	0.000454	0.000353
	(0.000140)	(0.000127)	(0.000260)	(0.000267)
DaySchArr_Route_TimeBlock	0.00227***	0.00200***	0.00239***	0.00235***
	(0.000115)	(0.000121)	(0.000206)	(0.000241)
MonthSchDep_Origin_Carrier	−0.00000541**	−0.00000420*	0.00000269	0.00000679*
	(0.00000184)	(0.00000185)	(0.00000315)	(0.00000293)
MonthSchArr_Dest_Carrier	−0.00000524**	0.00000149	−0.00000556	0.0000178***
	(0.00000201)	(0.00000188)	(0.00000315)	(0.00000332)
AvgDayDepDelay_Carrier_Route_TimeBlock	−0.000313***	−0.000314***	−0.000516***	−0.000521***
	(0.0000126)	(0.0000126)	(0.0000293)	(0.0000294)
TotalDayDepDelay_Origin_TimeBlock	0.00000751***	0.00000791***	0.0000114***	0.0000119***
	(0.000000992)	(0.000000979)	(0.00000219)	(0.00000213)
TotalDayArrDelay_Dest_TimeBlock	0.0000673***	0.0000673***	0.0000804***	0.0000806***
	(0.00000153)	(0.00000153)	(0.00000297)	(0.00000299)
DayCancel_Carrier_Route_TimeBlock	0.00601	0.00654	−0.00227	−0.00706
	(0.0193)	(0.0191)	(0.0432)	(0.0433)
DayCancel_Origin_TimeBlock_DueToCarrier	0.00148*	0.00160*	−0.000197	−0.000157
	(0.000716)	(0.000697)	(0.00146)	(0.00145)
DayCancel_Origin_TimeBlock_DueToWeather	0.00147***	0.00129***	0.000503	0.000357
	(0.000371)	(0.000367)	(0.000905)	(0.000899)
DayCancel_Origin_TimeBlock_DueToNAS	0.000815	0.000542	0.00279	0.00258
	(0.000838)	(0.000824)	(0.00167)	(0.00165)
DayCancel_Origin_TimeBlock_DueToSecurity	0.00988	0.0132	0.0542	0.0581
	(0.0158)	(0.0156)	(0.0404)	(0.0394)
DayDivert_Carrier_Route_TimeBlock	0.00351	0.00489	0.0875	0.113
	(0.0553)	(0.0543)	(0.279)	(0.277)
DayDivert_Dest_TimeBlock	0.0142***	0.0141***	0.0168***	0.0168***
	(0.000582)	(0.000580)	(0.00131)	(0.00131)
Distance	9.909***	–	10.48	–
	(2.889)	–	(9.831)	–
Observations	2,828,787	2,828,787	544,160	544,160
Adjusted R^2	0.005	0.517	−0.009	0.513
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓

Note: Standard errors in parentheses are clustered by carrier, route, and departure-arrival time block. The data set is an unbalanced panel. Aggregate time fixed effects (year, month, week, and day) are omitted for brevity. In the estimations performed by reghdfe (a user-written Stata program for absorbing high-dimensional fixed effects), which are shown in columns 2 and 4, several variables (e.g., the distance variable) are omitted because they are highly collinear with the fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5.2. Average effects during each phase of the post-deployment period

As described in Section 4.1, the FAA and NASA implemented tactical surface metering stepwise. Indeed, the implementation phase was divided into four stages: (1) November 29, 2017 to February 18, 2018; (2) February 19 2018, to September 30, 2018; (3) October 1, 2018 to August 31, 2019; and (4) September 1, 2019 to November 28, 2019. Tactical surface metering was applied only to limited operations banks during the first and second phases: the second bank (from about 9 a.m. to 11 a.m.) during the first phase and the second and third banks (from about 9 a.m. to 1 p.m.) during the second phase. Tactical surface metering was applied to all banks starting in the third phase. The stepwise implementation of the new technology at CLT also impacted the effects of tactical surface metering. Therefore, we divided the *POST* dummy into four dummies representing each implementation phase of tactical surface metering, which take a value of one if the flight date is within each phase and 0 otherwise. We provide the modified estimation model in Eq. (10):

$$Y_{iodbt} = \theta_t + \sum_{p=1}^{4} \alpha_p POST_p + \beta CLT + \sum_{p=1}^{4} \gamma_p POST_p * CLT + x_{iodbt}\delta + c_{iodb} + u_{iodbt} \tag{10}$$

Again, the subscripts i, o, d, b, and t denote carrier, origin, destination, departure/arrival time block, and day, respectively. The added subscript, p, represents the four implementation phases of tactical surface metering at CLT.

Table 15
Estimates for each phase.

Panel 1: Datasets A1 and A 2				
Dependent variable: Taxi-out				
Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post1	0.0350***	0.0285***	0.0540***	0.0567***
	(0.00223)	(0.00221)	(0.0143)	(0.0162)
Post2	0.0296***	0.0189***	0.0360**	0.0372
	(0.00283)	(0.00283)	(0.0139)	(0.0191)
Post3	0.0369***	0.0272***	0.0665***	0.0670**
	(0.00334)	(0.00335)	(0.0169)	(0.0215)
Post4	0.0228***	0.0127**	0.0475*	0.0559*
	(0.00399)	(0.00401)	(0.0207)	(0.0248)
Post1*O_CLT	−0.103***	−0.0195**	−0.0580***	−0.0584**
	(0.00643)	(0.00645)	(0.0143)	(0.0188)
Post2*O_CLT	−0.182***	−0.0742***	−0.0563***	−0.0584**
	(0.00604)	(0.00762)	(0.0138)	(0.0221)
Post3*O_CLT	−0.156***	−0.0537***	−0.0895***	−0.0787**
	(0.00650)	(0.00876)	(0.0161)	(0.0253)
Post4*O_CLT	−0.0668***	0.0310**	−0.0312	−0.0209
	(0.00783)	(0.0100)	(0.0196)	(0.0295)
Observations	5,534,608	5,534,608	239,222	239,222
Adjusted R^2	0.019	0.422	−0.007	0.285
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓
Panel 2: Datasets B1 and B 2				
Dependent variable: Airborne				
Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post1	0.00352***	0.00347***	0.0118***	0.00876**
	(0.000420)	(0.000430)	(0.00289)	(0.00318)
Post2	0.000650	0.000609	−0.00753*	−0.0108**
	(0.000472)	(0.000475)	(0.00326)	(0.00375)
Post3	0.00496***	0.00495***	−0.00158	−0.00542
	(0.000595)	(0.000607)	(0.00357)	(0.00425)
Post4	0.00570***	0.00559***	−0.00445	−0.00827
	(0.000692)	(0.000706)	(0.00425)	(0.00512)
Post1*D_CLT	−0.0145***	−0.00923**	−0.0248***	−0.0206***
	(0.00185)	(0.00287)	(0.00330)	(0.00590)
Post2*D_CLT	−0.00468***	0.00261	−0.00615**	−0.00101
	(0.000974)	(0.00221)	(0.00221)	(0.00596)
Post3*D_CLT	−0.00804***	−0.000829	−0.00826***	−0.00260
	(0.00121)	(0.00304)	(0.00212)	(0.00652)
Post4*D_CLT	−0.00985***	−0.00245	−0.0138***	−0.00714
	(0.00140)	(0.00330)	(0.00284)	(0.00713)
Observations	9,890,246	9,890,246	360,442	360,442
Adjusted R^2	0.023	0.852	−0.003	0.875
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓
Panel 3: Datasets C1 and C 2				
Dependent variable: Taxi-in				
Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Post1	−0.0134***	−0.0143***	−0.00343	−0.00663
	(0.00400)	(0.00402)	(0.00907)	(0.0103)
Post2	0.00122	0.000136	0.0213*	0.00619
	(0.00478)	(0.00486)	(0.0105)	(0.0124)
Post3	0.0105	0.0178**	0.0241	0.0459**
	(0.00568)	(0.00588)	(0.0126)	(0.0141)
Post4	0.00156	0.0107	0.0375*	0.0674***
	(0.00664)	(0.00679)	(0.0153)	(0.0165)
Post1*D_CLT	−0.0121	−0.0111	−0.0116	−0.00553
	(0.00669)	(0.00874)	(0.00957)	(0.0133)
Post2*D_CLT	−0.0182**	−0.0150	−0.0228*	−0.00823
	(0.00600)	(0.00938)	(0.00893)	(0.0154)
Post3*D_CLT	−0.0131*	−0.0446***	−0.0195*	−0.0656***
	(0.00633)	(0.0106)	(0.00915)	(0.0169)
Post4*D_CLT	0.0278***	−0.0167	0.0125	−0.0548**
	(0.00804)	(0.0121)	(0.0115)	(0.0190)
Observations	2,828,787	2,828,787	544,160	544,160
Adjusted R^2	0.005	0.517	−0.009	0.513

(*continued on next page*)

Table 15 (*continued*)

Panel 3: Datasets C1 and C 2				
Dependent variable: Taxi-in				
Sample type	(1) Original	(2) Original	(3) Matched	(4) Matched
Year, month, week, and day fixed effects	✓	✓	✓	✓
Origin- and destination-specific year fixed effects		✓		✓

Note: Standard errors in parentheses are clustered by carrier, route, and departure-arrival time block. The data set is an unbalanced panel. Control variables and aggregate time fixed effects (year, month, week, and day) are omitted for brevity.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 16
Average fuel and CO_2 savings (kg) – Datasets A1, A2, C1, and C2.

Panel A1 - Using matched data: Average fuel savings (kg) from reducing taxi-out time per outbound flight at CLT							
		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1-4: Taxi-out	−54.818	[−89.624, −18.810]	−46.987	[−76.821, −16.123]	−39.156	[−64.017, −13.436]
Panel A2 - Using matched data: Average CO_2 savings (kg) from reducing taxi-out time per outbound flight at CLT							
		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1-4: Taxi-out	−168.839	[−276.043, −57.935]	−144.719	[−236.608, −49.659]	−120.599	[−197.174, −41.382]
Panel B1 - Using matched data: Average fuel savings (kg) from reducing taxi-in time per inbound flight to CLT							
		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1-4: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]
Panel B2 - Using matched data: Average CO_2 savings (kg) from reducing taxi-in time per inbound flight to CLT							
		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1-4: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]

Phase 1: Nov 29, 2017–Feb 18, 2018; Phase 2: Feb 19, 2018–Sep 30, 2018; Phase 3: Oct 1, 2018–Aug 31, 2019; Phase 4: Sep 1, 2019–Nov 28, 2019

Panels 1 through 3 of Table 15 report estimates from the modified model. The estimates of the changes in taxi-out times obtained from the matched sample by controlling for origin- and destination-specific year fixed effects, shown in Column 4 of Panel 1, suggest that a reduction of taxi-out times of outbound flights from CLT occurred in all phases. The coefficients of interest ($\sum_{p=1}^{4}\gamma_p$) are statistically significant except for the fourth phase. The expected adverse side effects were, again, not observed. Indeed, the airborne times also decreased in all phases, though the coefficients of interest are statistically significant only in the first phase, as Column 4 of Panel 2 shows. Panel 3 also suggests the taxi-in times decreased in all phases, and the coefficients of interest are statistically significant in the third and fourth phases, as shown in Column 4.

The estimates in Panel 1 suggest that the reduction of taxi-out times becomes statistically insignificant in the fourth phase. It would be desirable if we could examine this point further. Still, unfortunately, we were unable to comprehensively analyze this point in this paper because the period after the fourth phase is when the coronavirus pandemic hit the market. The analysis of this point is a subject for future work.

6. Fuel and emissions savings

This section estimates the fuel and CO_2 savings surface metering technology achieved. The estimation results suggest that airborne times can be affected by surface metering technology. Indeed, any change in surface traffic management can reduce airborne times if it allows greater arrival capacity or greater predictability about when an arrival slot is available. However, calculating fuel and CO_2 savings achieved through reducing airborne time is highly complicated (see Seymour et al., 2020). Therefore, applying our estimation results regarding airborne time to calculating fuel and CO_2 savings is premature. In this study, we limit the calculation of fuel and CO_2 savings to those achieved by reducing taxi-out and taxi-in times.

To calculate fuel and CO_2 savings and compare them to the NASA ATD-2 team (2020) estimates, we made the following assumptions, replicating the assumptions made by the NASA ATD-2 team (2020) as closely as possible:

(1) Fuel savings depend on the aircraft engine model and taxi-out/in-time savings.
(2) We calculated the average taxi-out/in time savings per flight (seconds), *AvgTimeSaving*, using the following Eq. (11):

$$AvgTimeSaving\ [sec] = (\exp(EstCoef) - 1) \times TaxiMeanTime_c \tag{11}$$

where *EstCoef* represents the estimated coefficient(s) of the interaction term(s) of the *Post* and *CLT* dummies, *POST * CLT* (or $\sum_{p=1}^{4} POST_p * CLT$). *TaxiMeanTime* is the mean level value (seconds) of taxi-out/in times in each implementation phase of tactical surface metering. The subscript *c* represents the comparison group. Aircraft may burn more fuel during taxi-out than taxi-in because they are heavier during the former operations (Hao et al., 2017).

(3) Based on NASA's assumption, we assumed that the engine thrust level during taxiing is the same as in the idle condition, which means that the level does not vary during taxi operations.
(4) Data on the percentage of single-engine taxi (SET) operations are not publicly available. Therefore, when calculating the effective

Table 17
Average fuel and CO_2 savings (kg) – Datasets A1, A2, C1, and C2.

Panel A1 - Using matched data: Average fuel savings (kg) from reducing taxi-out time per outbound flight from CLT

		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1: Taxi-out	−22.269	[−35.648, −8.388]	−19.087	[−30.555, −7.190]	−15.906	[−25.463, −5.992]
	2: Taxi-out	−70.442	[−120.079, −18.732]	−60.379	[−102.924, −16.056]	−50.316	[−85.770, −13.380]
	3: Taxi-out	−76.109	[−120.823, −28.842]	−65.236	[−103.562, −24.722]	−54.363	[−86.302, −20.601]
	4: Taxi-out	0	[0, 0]	0	[0, 0]	0	[0, 0]

Panel A2 - Using matched data: Average CO_2 savings (kg) from reducing taxi-out time per outbound flight from CLT

		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1: Taxi-out	−68.588	[−109.795, −25.836]	−58.789	[−94.110, −22.145]	−48.991	[−78.425, −18.454]
	2: Taxi-out	−216.963	[−369.842, −57.695]	−185.968	[−317.007, −49.453]	−154.973	[−264.173, −41.211]
	3: Taxi-out	−234.415	[−372.134, −88.834]	−200.927	[−318.972, −76.143]	−167.439	[−265.810, −63.453]
	4: Taxi-out	0	[0, 0]	0	[0, 0]	0	[0, 0]

Panel B1 - Using matched data: Average fuel savings (kg) from reducing taxi-in time per inbound flight to CLT

		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]
	2: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]
	3: Taxi-in	−9.900	[−14.668, −2.330]	−8.486	[−12.573, −1.997]	−7.071	[−10.477, −1.664]
	4: Taxi-in	−9.038	[−14.912, −0.442]	−7.747	[−12.782, −0.379]	−6.455	[−10.651, −0.316]

Panel B2 - Using matched data: Average CO_2 savings (kg) from reducing taxi-in time per inbound flight to CLT

		Single Engine Taxi rate					
Covariate balancing	Phase	25%	[95% Conf. interval]	50%	[95% Conf. interval]	75%	[95% Conf. interval]
PSM	1: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]
	2: Taxi-in	0	[0, 0]	0	[0, 0]	0	[0, 0]
	3: Taxi-in	−30.491	[−45.177, −15.310]	−26.136	[−38.723, −13.123]	−21.780	[−32.270, −10.936]
	4: Taxi-in	−27.836	[−45.929, −9.056]	−23.859	[−39.368, −7.762]	−19.883	[−32.806, −6.468]

Phase 1: Nov 29, 2017–Feb 18, 2018; Phase 2: Feb 19, 2018–Sep 30, 2018; Phase 3: Oct 1, 2018–Aug 31, 2019; Phase 4: Sep 1, 2019–Nov 28, 2019.

number of engines for each flight, we used multiple assumed rates of SET operations (25%, 50%, and 75%).

$$Effective\ engine\ count = SET\ rate \times 1 + (1 - SET\ rate) \times Engine\ count \tag{12}$$

(5) We obtained aircraft types, engine models, and the number of engines from the FAA Aircraft Registration Database and merged them with US DOT OTP data using the aircraft registration number (N-Number).
(6) We obtained values for fuel flow (kg/sec) at idle conditions from the International Civil Aviation Organization's (ICAO) Aircraft Engine Emissions Databank. Then, we merged them with the FAA Aircraft Registration Database and US DOT OTP data by using the name of the engine model.
(7) Based on a calculation by NASA, we computed fuel and emissions savings using Eq. (13) (NASA ATD-2 team, 2020):

$$Average\ Fuel\ saving\ per\ flight\ [kg] = \sum_{j=1}^{n} AvgTimeSaving\ [sec] \times FuelFlow_j[kg/sec] \times EffEngCount_j \Big/ FlightFreq_j \tag{13}$$

where *AvgTimeSaving* is either the average taxi-out time saving per flight or the average taxi-in time savings per flight. *FlightFreq* represents flight frequencies and $j = 1 \cdots n$. The subscript j represents the aircraft that appeared in our dataset.

$$Average\ CO_2\ saving\ per\ flight\ [kg] = AvgFuelSaving\ per\ flight\ [kg] \times CO_2EmissionCoef\ [3.08\ kg\ CO_2\ per\ kg\ jet\ fuel] \tag{14}$$

We set the carbon dioxide emission coefficient, $CO_2EmissionCoef$, in Eq. (14) to the same value (i.e., 3.08 *kg* CO_2 *per kg jet fuel*) that the NASA ATD-2 team (2020) used[10].

Panel A in Table 16 presents estimated fuel and CO_2 savings from reducing taxi-out times that surface metering technology achieved at CLT. The lower the SET rate, the higher the fuel and CO_2 savings will be. If we compute the estimated amounts of fuel and CO_2 savings by using the coefficient estimate obtained from PSM and setting the SET rate at 75%, about 39.156 kg (95% confidence interval (CI) 64.017 kg–13.436 kg) of fuel and 120.599 kg (CI 197.174 kg–41.382 kg) of CO_2 can be saved per flight. As described in Section 1.2, the estimation by the FAA and NASA suggests that surface metering technology saved about 3.3 kg of fuel and 10.16 kg of CO_2 per flight. Our estimation results suggest that fuel and CO_2 savings from the reduction of taxi-out times the technology achieved may be more significant than the estimates of the FAA and

[10] The US Energy Information Administration reports that the CO_2 emission coefficient for jet fuel as of February 9, 2022, is 9.75 kg CO_2/gallon. Using this coefficient and a density of jet fuel (Jet A type for US) of 6.84 [lb/gallon] (=3.103 [kg/gallon]) at 59 °F (=15 °C), which is assumed by NASA, the CO_2 emissions for an aircraft consuming 1 kg of jet fuel is 3.14 [kg CO_2/kg jet fuel] (=9.75 [kg CO_2/gallon]/3.103 [kg jet fuel/gallon]). However, for the purpose of comparison, we set the carbon dioxide emission coefficient for jet fuel at 3.08 [kg CO_2/kg jet fuel] (=9.57 [kg CO_2/gallon]), as NASA did. See US Energy Information Administration, Carbon Dioxide Emissions Coefficients <https://www.eia.gov/environment/emissions/co2_vol_mass.php> (Accessed 7 June 2022).

NASA, even when the SET rate is as high as 75%.

The difference between the estimates of the FAA/NASA and our estimates may be due to different estimation assumptions and models. Specifically, using counterfactuals (i.e., comparison groups) in our analysis may contribute to the significant difference between the estimates of the FAA/NASA and ours. Column 4 in Table 12 shows that the coefficient of the *Post* variable is 0.0455, which is statistically significant. The value suggests that taxi-out times in the comparison group increased by about 4.7%[11] after introducing tactical surface metering at CLT. As a result of the difference between the comparison group, which is inefficient in terms of reduction in taxi-out times, and the treatment group, i.e., CLT, our analysis may have resulted in a larger estimate of decrease in fuel consumption and CO_2 emissions than the FAA/NASA's estimate.

Panel B in Table 16 reports estimated fuel savings and CO_2 reduction achieved by reducing taxi-in times after the deployment of tactical surface metering at CLT. The estimated coefficient of the interaction term, *POST * CLT*, is not statistically significant, so the estimated fuel and CO_2 savings are zero.

Panel B in Table 17 suggests that the estimated savings in fuel consumption and CO_2 emissions due to reduced taxi-in times were not zero throughout all phases. Indeed, fuel and CO_2 emissions reductions seem to have been achieved in the third and fourth phases. In contrast, Panel A in Table 17 shows a slightly alarming trend. Indeed, Panel A suggests that the estimated fuel and CO_2 savings from reducing taxi-out times become zero in the fourth phase.

The results suggest that the fuel and CO_2 savings achieved by surface metering technology may not be stable and could be quickly canceled out or reversed by some minor shocks. Indeed, the scale of savings per flight achieved through surface metering technology is modest. In our datasets, one of the most popular aircraft operated at CLT was the Bombardier CL-600-2D24 (maximum number of passenger seats is 90). The fuel flow at the idle condition of its engine, CF34–8C5, is 0.0643 kg/s, according to the ICAO Aircraft Engine Emissions Databank. Thus, the 39.156 kg (CI 64.017 kg–13.436 kg) of fuel savings will be canceled out while the aircraft sits on the tarmac and waits for take-off for around 10 min (CI 17 min.–3.5 min) during single-engine taxiing. Although the cumulative impact of the savings is not negligible, this possibility is an important caveat that should be considered and investigated further in future research, considering that the average taxi-out time at CLT is still around 20 min during the analysis period after the introduction of tactical surface metering.

7. Conclusion

We estimated the effects of the surface metering technology developed and deployed at CLT by NASA and the FAA on fuel and CO_2 savings. Using the coefficient estimates obtained from a doubly robust strategy that combines matching with parametric regression using data sets constructed by data-driven approaches, which minimize discretion in the selection of the comparison units, our results suggested that tactical surface metering technology can achieve higher fuel and CO_2 savings than FAA and NASA reported. Our results also suggest that surface metering technology did not yield the expected adverse side effects, i.e., increased airborne and taxi-in times due to the gate-holding resulting from the technology.

At the same time, we must consider the possible limitations of surface metering technology. Our estimation results revealed a concerning finding: the fuel- and CO_2-saving effects of tactical surface metering may not be stable and be easily canceled out or reversed by some minor shocks. Our results suggest that the estimated fuel and CO_2 savings from reducing taxi-out times become zero in the fourth phase, even after controlling for the effects of changes in factors such as congestion level at CLT.

This finding has an important policy implication regarding the effectiveness of surface metering technology measures: departure metering may function more effectively with coordinated arrival metering. To achieve a net zero in aviation, novel and revolutionary technologies are required, such as electrified and hydrogen-powered aircraft, and the industry plans to have them in commercial in 2035–40. However, it takes time, and we urgently need effective measures. The currently available technology can still contribute to our target goals to some extent. We hope this paper can contribute to the policy implication for this reason.

One limitation of our study is relevant to this implication; we did not analyze how increased or decreased taxi-out/in times affect taxi-in/out times. Therefore, future research should examine the impact of taxi-out/in times on changes in taxi-in/out times and explore practical ways to coordinate departure and arrival metering.

This study was also limited because we needed to analyze the possible externalities of airport surface metering technology. Changing taxi-out/in times at one airport could result in changes in taxi-out/in times at other airports that have flights to and from the former airport. In other words, surface metering technology can have external effects on other airports. Future research should also examine the size and scope of externalities from airport surface metering technology.

In light of the possible limitations of surface metering technology, airport congestion management mechanisms such as congestion pricing and market-based slot allocation systems (Avenali et al., 2015; Basso & Zhang, 2010; Brueckner, 2009; Castelli et al., 2011; Fukui, 2010, 2012, 2014; Gillen & Starkie, 2016; Guiomard, 2018; Pellegrini et al., 2012; Pertuiset & Santos, 2014; Sheng et al., 2015; Starkie, 1998; Verhoef, 2010; Zhang & Czerny, 2012) and possible sustainable-fuel flight incentives (International Air Transport Association IATA, 2022) warrant further investigation. Combining surface metering technology with airport congestion management mechanisms may reduce taxi times, fuel consumption, and CO_2 emissions more effectively.

CRediT authorship contribution statement

Hideki Fukui: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Chikage Miyoshi:** Writing – review & editing, Validation, Software.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank two anonymous referees and participants at the 2022 Air Transport Research Society Conference, the 2022 Applied Regional Science Conference, the 2023 International Transportation Economics Association Conference, and the 2023 World Conference on Transport Research for their valuable and insightful comments that improved the paper. We also acknowledge financial support from the Japan Society for the Promotion of Science (20K01448; 21KK0028). All remaining errors are the authors' responsibility.

[11] $(\exp(0.0455) - 1) * 100 \approx 4.7$

Appendix A. Three-letter codes for the airports that appear in the resulting datasets

Table A1
Three-letter codes for the airports

	Origin	Destination
Dataset A (Taxi-out) Origin: 150 Destination: 255	ABY, ACK, ACT, ADK, AEX, AMA, APN, AUS, AVL, AZA, BET, BFL, BFM, BGR, BHM, BKG, BOS, BPT, BQK, BRO, BRW, BTR, CAE, CDC, CGI, CHO, CHS, CLL, CLT, CNY, COD, COS, COU, CPR, CRP, CSG, DAB, DCA, DHN, DIK, DRO, DVL, EFD, EGE, ELM, EUG, EWN, FCA, FLG, GCC, GFK, GGG, GPT, GRI, GSP, GST, GUC, HDN, HGR, HHH, HNL, HOB, HOU, HYS, IAD, IAG, IAH, IFP, ISN, JFK, JLN, KOA, LAW, LBB, LBF, LCH, LFT, LGB, LIH, LIT, LRD, MAF, MEI, MFE, MFR, MGM, MKG, MLB, MLU, MMH, MQT, MRY, MSY, MTJ, MVY, MYR, OAJ, OGD, OME, ONT, ORH, PAH, PBG, PGD, PHF, PHL, PIR, PNS, PRC, PSE, PSG, PUB, PVU, RDU, RHI, RKS, ROW, SAF, SAV, SBP, SCE, SCK, SGU, SHV, SIT, SLN, SMX, SNA, SPN, SPS, SRQ, STC, SUX, SWO, TLH, TOL, TUS, TWF, TXK, TYR, TYS, UIN, USA, UST, VEL, VLD, VPS, XWA, YNG, YUM	ABE, ABQ, ACK, ACY, AEX, AGS, ALB, AMA, ANC, ART, ASE, ATL, ATW, ATY, AUS, AVL, AVP, AZA, BDL, BFL, BGR, BHM, BIL, BIS, BLI, BLV, BNA, BOI, BOS, BQN, BRO, BTR, BTV, BUF, BUR, BWI, BZN, CAE, CAK, CGI, CHA, CHO, CHS, CID, CKB, CLE, CLL, CMH, CMI, COS, CRP, CRW, CVG, DAB, DAL, DAY, DCA, DEN, DFW, DRO, DSM, DTW, ECP, EGE, ELM, ELP, EUG, EVV, EWN, EWR, EYW, FAI, FAR, FAY, FCA, FLL, FLO, FNT, FSD, FWA, GFK, GJT, GNV, GPT, GRI, GRK, GRR, GSO, GSP, GTF, GUC, GUM, HDN, HGR, HHH, HNL, HOB, HOU, HPN, HRL, HSV, HTS, HVN, HYA, HYS, IAD, IAG, IAH, ICT, IDA, ILM, IMT, IND, ISN, ISP, ITH, ITO, JAC, JAN, JAX, JFK, JMS, JNU, KOA, KTN, LAN, LAS, LAX, LBB, LBE, LBL, LCH, LCK, LEX, LFT, LGA, LGB, LIH, LIT, LRD, LWB, LYH, MAF, MCI, MCO, MDT, MDW, MEM, MFE, MFR, MGM, MHT, MIA, MKE, MLB, MLI, MLU, MOB, MOT, MSN, MSO, MSP, MSY, MTJ, MVY, MYR, OAJ, OAK, OGD, OGG, OGS, OKC, OMA, ONT, ORD, ORF, ORH, OTZ, PAH, PBG, PBI, PDX, PGD, PGV, PHF, PHL, PHX, PIA, PIB, PIE, PIT, PLN, PNS, PPG, PSC, PSE, PSM, PSP, PVD, PVU, PWM, RAP, RDM, RDU, RFD, RIC, RNO, ROA, ROC, RSW, SAN, SAT, SAV, SBN, SCC, SCE, SCK, SDF, SEA, SFB, SFO, SGF, SGU, SHD, SHV, SJC, SJU, SLC, SLN, SMF, SNA, SPI, SRQ, STC, STL, STT, STX, SWF, SYR, TLH, TOL, TPA, TRI, TTN, TUL, TUS, TVC, TYR, TYS, UIN, USA, UST, VPS, WRG, XNA
Dataset B (Airborne) Origin: 328 Destination: 235	ABE, ABQ, ABR, ACK, ACT, ACV, ACY, ADK, ADQ, AEX, AGS, AKN, ALB, AMA, ANC, APN, ASE, ATL, ATW, ATY, AUS, AVL, AVP, AZA, AZO, BDL, BET, BFF, BFL, BFM, BGR, BHM, BIL, BIS, BJI, BKG, BLI, BLV, BMI, BNA, BOI, BOS, BRD, BRO, BRW, BTM, BTR, BTV, BUF, BUR, BWI, BZN, CAE, CAK, CDC, CDV, CGI, CHA, CHO, CHS, CID, CIU, CKB, CLE, CLL, CMH, CMI, CNY, COD, COS, COU, CPR, CRP, CRW, CVG, CWA, DAB, DAL, DAY, DCA, DEN, DFW, DIK, DLG, DLH, DRO, DSM, DTW, DVL, EAR, ECP, EFD, EGE, EKO, ELM, ELP, ERI, EUG, EVV, EWN, EWR, EYW, FAI, FAR, FAT, FAY, FCA, FLG, FLL, FLO, FNT, FSD, FWA, GCC, GEG, GFK, GJT, GNV, GPT, GRB, GRI, GRK, GRR, GSO, GSP, GST, GTF, GUC, GUM, HDN, HGR, HHH, HIB, HLN, HNL, HOB, HOU, HPN, HRL, HSV, HTS, HVN, HYS, IAD, IAG, IAH, ICT, IDA, IFP, ILM, IMT, IND, INL, ISN, ISP, ITH, ITO, JAC, JAN, JAX, JFK, JMS, JNU, KOA, KTN, LAN, LAR, LAS, LAX, LBB, LBE, LBF, LBL, LCH, LCK, LEX, LFT, LGA, LGB, LIH, LIT, LNK, LRD, LSE, LWB, LWS, LYH, MAF, MBS, MCI, MCO, MDT, MDW, MEI, MEM, MFE, MFR, MGM, MHK, MHT, MIA, MKE, MLB, MLI, MLU, MMH, MOB, MOT, MQT, MRY, MSN, MSO, MSP, MSY, MTJ, MYR, OAJ, OAK, OGD, OGG, OGS, OKC, OMA, OME, ONT, ORD, ORF, OTH, OTZ, OWB, PAE, PAH, PBG, PBI, PDX, PGD, PGV, PHF, PHL, PHX, PIA, PIB, PIE, PIH, PIR, PIT, PLN, PNS, PPG, PRC, PSC, PSG, PSM, PSP, PUB, PVD, PVU, PWM, RAP, RDM, RDU, RFD, RHI, RIC, RKS, RNO, ROA, ROC, ROW, RST, RSW, SAF, SAN, SAT, SAV, SBA, SBN, SBP, SCC, SCE, SCK, SDF, SEA, SFB, SFO, SGF, SGU, SHD, SHV, SIT, SJC, SJU, SLC, SLN, SMF, SNA, SPI, SRQ, STC, STL, STS, STT, STX, SUN, SWF, SYR, TLH, TOL, TPA, TRI, TTN, TUL, TUS, TVC, TWF, TYR, TYS, UIN, USA, VEL, VPS, WRG, WYS, XNA, XWA, YAK, YNG, YUM	ABE, ABI, ABQ, ABR, ABY, ACV, AEX, AGS, ALO, ANC, ART, ASE, ATW, AVL, AVP, AZA, AZO, BFL, BGM, BHM, BIL, BIS, BJI, BKG, BLI, BMI, BNA, BOI, BPT, BQK, BQN, BRD, BRO, BTM, BTR, BTV, BUF, CAE, CDV, CGI, CHO, CHS, CKB, CLE, CLL, CLT, CMH, CMI, CMX, CNY, COD, COU, CPR, CRW, CSG, CWA, DAB, DAL, DBQ, DEN, DIK, DLG, DLH, DRO, DSM, ECP, EGE, EKO, ERI, ESC, EUG, EVV, EWN, EYW, FAT, FCA, FLG, FNL, FNT, FSD, FSM, GFK, GGG, GNV, GPT, GRB, GRI, GRR, GSO, GSP, GTF, HGR, HHH, HIB, HLN, HNL, HOB, HSV, HTS, HYS, IAD, IAH, ICT, IDA, IFP, IMT, INL, ISN, ITH, ITO, JAN, JAX, JLN, JNU, KOA, KTN, LAN, LAR, LAW, LBE, LBF, LBL, LCH, LCK, LFT, LGB, LIT, LNK, LSE, LWB, LWS, MDT, MDW, MEI, MFR, MGM, MHK, MIA, MKE, MLB, MLI, MLU, MMH, MOB, MOT, MQT, MRY, MSN, MSO, MSP, MTJ, MVY, MYR, OGD, OGG, OGS, OKC, OMA, ORF, ORH, OWB, PAH, PBG, PGD, PHX, PIA, PIB, PIH, PIR, PIT, PLN, PNS, PSC, PSE, PSG, PSM, PSP, PUB, PVU, PWM, RAP, RDD, RDM, RFD, RHI, RIC, RKS, RST, SAN, SAV, SBA, SBP, SCC, SCE, SEA, SFB, SGF, SHV, SLC, SLN, SMX, SNA, SPN, SPS, STC, STL, STS, STT, STX, SUN, SUX, SWF, SWO, SYR, TLH, TOL, TRI, TUL, TVC, TXK, TYR, TYS, UIN, USA, UST, VEL, VLD, VPS, WRG, WYS, XNA, XWA, YAK, YNG, YUM
Dataset C (Taxi-in) Origin: 249 Destination: 70	ABE, ABI, ABQ, ACK, ACT, AEX, AGS, ALB, AMA, ANC, ART, ASE, ATL, ATW, AUS, AVL, AVP, AZA, BDL, BFL, BGR, BHM, BIL, BIS, BKG, BLV, BMI, BNA, BOI, BOS, BPT, BRO, BRW, BTR, BTV, BUF, BUR, BWI, BZN, CAE, CAK, CGI, CHA, CHO, CHS, CID, CLE, CLL, CMH, CMI, COS, COU, CRP, CRW, CVG, CYS, DAB, DAL, DAY, DCA, DEN, DFW, DRO, DRT, DSM, DTW, ECP, EGE, ELM, ELP, EUG, EVV, EWN, EWR, EYW, FAR, FAT, FAY, FCA, FLG, FLL, FLO, FNT, FSD, FSM, FWA, GCK, GEG, GGG, GJT, GNV, GPT, GRI, GRK, GRR, GSO, GSP, GUC, HDN, HGR, HHH, HNL, HOU, HPN, HRL, HSV, HTS, HVN, IAD, IAG, IAH, ICT, ILM, IND, ITH, JAC, JAN, JAX, JFK, JLN, JNU, KOA, LAS, LAW, LAX, LBB, LBL, LCH, LCK, LEX, LFT, LGA, LGB, LIH, LIT, LRD, LYH, MAF, MCI, MCO, MDT, MDW, MEI, MEM, MFE, MGM, MHK, MHT, MIA, MKE, MLB, MLI, MLU, MOB, MRY, MSN, MSO, MSP, MSY, MTJ, MYR, OAJ, OAK, OGG, OGS, OKC, OMA, OME, ONT, ORD, ORF, PBG, PBI, PDX, PGD, PGV, PHF, PHL, PHX, PIA, PIB, PIE, PIT, PNS, PSM, PSP, PVD, PWM, RAP, RDU, RFD, RIC, RNO, ROA, ROC, ROW, RSW, SAF, SAN, SAT, SAV, SBA, SBN, SBP, SCE, SDF, SEA, SFB, SFO, SGF, SGU, SHD, SHV, SJC, SJT, SJU, SLC, SLN, SMF, SNA, SPI, SPN, SPS, SRQ, STL, STS, STT, STX, SUX, SWF, SWO, SYR, TLH, TOL, TPA, TRI, TTN, TUL, TUS, TVC, TXK, TYR, TYS, USA, UST, VPS, XNA, YNG, YUM	ABE, ABR, ACK, ACV, ACY, AKN, ART, ASE, BET, BLI, BLV, BMI, BOI, BTV, CHA, CHO, CHS, CID, CLT, CMX, CNY, COU, CRW, CSG, CVG, DBQ, DFW, EAR, FAI, GCK, GRB, GRK, GSO, GUC, GUM, HTS, HYS, IFP, JAC, JAN, LBF, LRD, LWB, MDT, MEI, MHK, MMH, MSO, MTJ, OAJ, OGS, OWB, PAH, PBG, PHL, PIB, PIE, PUB, PVU, SAF, SAV, SGF, SMF, SUX, SWO, TLH, TOL, TTN, UIN, VEL

Notes: All airports in Table A1 provide commercial services.

Appendix B. Estimation of the propensity score

This appendix shows the procedure for obtaining the balancing scores (i.e., propensity scores). We use logistic regression to estimate the propensity score using a vector of observed covariates. The regression model is Eq. (B1):

$$CLT = \alpha + X_{iodbt}\beta + u_{iodbt} \tag{B1}$$

The subscripts *i*, *o*, *d*, *b*, and *t*, denote carrier, origin, destination, departure/arrival time block, and time (i.e., day). The treatment group consists of flights to or from CLT. The dependent variable in Eq. (B1), *CLT*, is the dummy variable. *CLT* represents outbound flights from CLT in Dataset A, which we used in our analysis to estimate the effects of tactical surface metering on taxi-out times at CLT. In contrast, *CLT* represents inbound flights to CLT in Datasets B and C, which we used in our analysis to estimate the effects of tactical surface metering on airborne and taxi-in times at CLT. The X_{iodbt} vector is an array of covariates, including variables that are thought to be associated with treatment and those considered to be related to the outcome. Any variables that perfectly predict the treatment status should be excluded from Eq. (B1). Therefore, the X_{iodbt} vector includes all the control variables used in Eq. (7) as components except for aggregate time fixed effects (year, month, week, and day dummies), $POST_p$, and *CLT*, which perfectly predict treatment status in our dataset.

We performed the logistic regression separately for each year to match treated and comparison units within the same year. In other words, we repeated the procedure to obtain the balancing scores 15 times. We omitted the estimation results of the logistic regression for the sample before matching for brevity. The results are available from the authors upon request.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association, 105*(490), 493–505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *The American Economic Review, 93*(1), 113–132.

Airport Council International (ACI). (2018). ACI policy brief: Airports' resilience and adaptation to a changing climate. https://store.aci.aero/wp-content/uploads/2018/10/Policy_brief_airports_adaption_climate_change_V6_WEB.pdf. (Accessed 6 July 2022).

American Airlines Group Inc. (2021). CDP climate change questionnaire 2021. American Airlines. website https://www.aa.com/content/images/customer-service/about-us/corporate-governance/esg/esg-cdp-response-aag-2021.pdf. (Accessed 29 May 2022).

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine, 28*(25), 3083–3107.

Austin, P. C. (2011). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics, 10*(2), 150–161.

Austin, P. C. (2019). Assessing covariate balance when using the generalized propensity score with quantitative or continuous exposures. *Statistical Methods in Medical Research, 28*(5), 1365–1377.

Avenali, A., D'Alfonso, T., Leporelli, C., Matteucci, G., Nastasi, A., & Reverberi, P. (2015). An incentive pricing mechanism for efficient airport slot allocation in Europe. *Journal of Air Transport Management, 42*, 27–36.

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics, 61*(4), 962–973.

Basso, L. J., & Zhang, A. (2010). Pricing vs. slot policies when airport profits matter. *Transportation Research Part B: Methodological, 44*(3), 381–391.

Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). CEM: Coarsened exact matching in Stata. *Stata Journal, 9*(4), 524–546.

Brueckner, J. K. (2009). Price vs. quantity-based approaches to airport congestion management. *Journal of Public Economics, 93*(5–6), 681–690.

Castelli, L., Pesenti, R., & Ranieri, A. (2011). The design of a market mechanism to allocate air traffic flow management slots. *Transportation Research Part C: Emerging Technologies, 19*(5), 931–943.

Constantine, N., & Correia, S. (2021). Reghdfe: Stata module for linear and instrumental-variable/GMM regression absorbing multiple levels of fixed effects. https://ideas.repec.org/c/boc/bocode/s457874.html. (Accessed 1 October 2022).

Correia, S. (2017). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Working Paper http://scorreia.com/research/hdfe.pdf. (Accessed 1 October 2022).

Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *Stata Journal, 20*(1), 95–115.

Eurocontrol. (2021). Additional taxi-out time. https://www.eurocontrol.int/prudata/dashboard/metadata/additional-taxi-out-time/. (Accessed 1 July 2022).

Federal Aviation Administration (FAA) *Airport master Record (5010)*. (). https://www.faa.gov/forms/index.cfm/go/document.information/documentID/185474. (Accessed 1 December 2022).

Federal Aviation Administration (FAA). (2021). New software capability gets planes rolling directly to the runway, reducing fuel burn & taxi time. FAA website https://www.faa.gov/newsroom/new-software-capability-gets-planes-rolling-directly-runway-reducing-fuel-burn-taxi-time. (Accessed 30 September 2021).

Federal Aviation Administration (FAA). (2022). Terminal flight data manager (TFDM). https://www.faa.gov/air_traffic/technology/tfdm/. (Accessed 5 June 2022).

Fukui, H. (2010). An empirical analysis of airport slot trading in the United States. *Transportation Research Part B: Methodological, 44*(3), 330–357.

Fukui, H. (2012). Do carriers abuse the slot system to inhibit airport capacity usage? Evidence from the US experience. *Journal of Air Transport Management, 24*, 1–6.

Fukui, H. (2014). Effect of slot trading on route-level competition: Evidence from experience in the UK. *Transportation Research Part A: Policy and Practice, 69*, 124–141.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research, 49*(5), 1701–1720.

Gillen, D., & Starkie, D. (2016). EU slot policy at congested hubs, and incentives to add capacity. *Journal of Transport Economics and Policy, 50*(2), 151–163.

Google Maps. (2023). Charlotte Douglas international airport. https://www.google.com/maps/@35.2259113,-80.9695278,13z/. (Accessed 1 November 2023).

Guiomard, C. (2018). Airport slots: Can regulation be coordinated with competition? Evidence from Dublin airport. *Transportation Research Part A: Policy and Practice, 114*, 127–138.

Guo, S., & Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications* (2nd ed.). Sage.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20*(1), 25–46.

Hainmueller, J., & Xu, Y. (2013). Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software, 54*(7), 1–18.

Hao, L., Ryerson, M. S., Kang, L., & Hansen, M. (2017). Estimating fuel burn impacts of taxi-out delay with implications for gate-hold benefits. *Transportation Research Part C: Emerging Technologies, 80*, 454–466.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199–236.

Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1–24.

International Civil Aviation Organization (ICAO) *Aircraft engine emissions Databank*. (2021). Ver. 07 https://www.easa.europa.eu/en/domains/environment/icao-aircraft-engine-emissions-databank. (Accessed 30 December 2021).

Jung, Y., Engelland, S., Capps, A., Coppenbarger, R., Hooey, B., Sharma, S., Stevens, L., Verma, S., Lohr, G., Chevalley, E., Dulchinos, V., Malik, W., & Morgan-Ruszkowski, L. (2018). Airspace technology demonstration 2 (ATD-2) phase 1 concept of use (ConUse) (No. NASA/TM-2018-219770). *NASA*.

Koudis, G. S., Hu, S. J., Majumdar, A., Ochieng, W. Y., & Stettler, M. E. J. (2018). The impact of single engine taxiing on aircraft fuel consumption and pollutant emissions. *Aeronautical Journal, 122*(1258), 1967–1984.

National Aeronautics and Space Administration (NASA) ATD-2 team. (2020). *ATD-2 benefits mechanism*. NASA.

Nikoleris, T., Gupta, G., & Kistler, M. (2011). Detailed estimation of fuel consumption and emissions during aircraft taxi operations at Dallas/Fort Worth International Airport. *Transportation Research Part D: Transport and Environment, 16*(4), 302–308.

Pellegrini, P., Castelli, L., & Pesenti, R. (2012). Secondary trading of airport slots as a combinatorial exchange. *Transportation Research Part E: Logistics and Transportation Review, 48*(5), 1009–1022.

Pertuiset, T., & Santos, G. (2014). Primary auction of slots at European airports. *Research in Transportation Economics, 45*, 66–71.

Ravizza, S., Chen, J., Atkin, J. A., Burke, E. K., & Stewart, P. (2013). The trade-off between taxi time and fuel consumption in airport ground movement. *Public Transport, 5*(1–2), 25–40.

Richetta, O., & Odoni, A. R. (1994). Dynamic solution to the ground-holding problem in air traffic control. *Transportation Research Part A: Policy and Practice, 28*(3), 167–185.

Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127*(8S), 757–763.

Ryerson, M. S., Hansen, M., & Bonn, J. (2014). Time to burn: Flight delay, terminal efficiency, and fuel consumption in the National Airspace System. *Transportation Research Part A: Policy and Practice, 69*, 286–298.

Seymour, K., Held, M., Georges, G., & Boulouchos, K. (2020). Fuel Estimation in Air Transportation: Modeling global fuel consumption for commercial aviation. *Transportation Research Part D: Transport and Environment, 88*, Article 102528.

Sheng, D., Li, Z. C., Xiao, Y. B., & Fu, X. (2015). Slot auction in an airport network with demand uncertainty. *Transportation Research Part E: Logistics and Transportation Review, 82*, 79–100.

Starkie, D. (1998). Allocating airport slots: A role for the market? *Journal of Air Transport Management, 4*(2), 111–116.

Stettler, M. E. J., Eastham, S., & Barrett, S. R. H. (2011). Air quality and public health impacts of UK airports. Part I: Emissions. *Atmospheric Environment, 45*(31), 5415–5424.

Stuart, E. A., Lee, B. K., & Leacy, F. P. (2013). Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of Clinical Epidemiology, 66*(SUPPL.8), S84–S90.e1.

Talebi, E., Ging, A., Engelland, S., Capps, A., Eshow, M., Jung, Y., Sharma, S., Downs, M., Freedman, C., Ngo, T., Sielski, H., Wang, E., Burke, J., Gorman, S., Phipps, B., Robeson, I., Wilson, S., & Morgan-Ruszkowski, L. (2019). Airspace technology demonstration 2 (ATD-2) phase 2 technology description document (TDD) (No. NASA/TM-2019-220355). *NASA*.

International Air Transport Association (IATA). (2022). Incentives needed to increase SAF production. Press Release No. 32. 21 June 2022 https://www.iata.org/en/pressroom/2022-releases/2022-06-21-02/. (Accessed 6 July 2022).

Federal Aviation Administration (FAA). Aircraft registration database. https://www.faa.gov/licenses_certificates/aircraft_certification/aircraft_registry/releasable_aircraft_download/. (Accessed 28 December 2021).

US Department of Transportation (US DOT). National Flight Data Center (NFDC) Airport Database. https://www.faa.gov/airports/airport_safety/airportdata_5010/; https://web.archive.org/web/20200303055712/https://www.faa.gov/airports/airport_safety/airportdata_5010/[Defunct] (Accessed on December 5, 2019) .

US Department of Transportation (US DOT) *Airline on-time performance data*. (). https://www.transtats.bts.gov/. (Accessed 2 January 2022).

US Department of Transportation (US DOT) *Air carrier statistics (form 41 traffic) T-100 domestic segment (all carriers) (T100D)*. (). https://www.transtats.bts.gov/. (Accessed 2 January 2022).

US Department of Transportation (US DOT) *Air carrier statistics (form 41 traffic) T-100 international segment (all carriers) (T100I)*. (). https://www.transtats.bts.gov/. (Accessed 2 January 2022).

US Energy Information Administration (US EIA) *Carbon dioxide emissions coefficients*. (). https://www.eia.gov/environment/emissions/co2_vol_mass.php. (Accessed 2 January 2022).

Verhoef, E. T. (2010). Congestion pricing, slot sales and slot trading in aviation. *Transportation Research Part B: Methodological, 44*(3), 320–329.

Zhang, A., & Czerny, A. I. (2012). Airports and airlines economics and policy: An interpretive review of recent research. *Economics of Transportation, 1*(1–2), 15–34.

Estimating tactical surface metering management's effect on aircraft fuel savings at airport

Fukui, Hideki

2024-03-01

Fukui H, Miyoshi C. (2024) Estimating tactical surface metering management's effect on aircraft fuel savings at airport. Research in Transportation Economics, Volume 103, March 2024, Article number 101405

https://doi.org/10.1016/j.retrec.2023.101405