# The Effect of Adopting the Next Generation Air Transportation System on Air Travel Performance

Ziyan Chu
Yale School of the Environment

Yichen Christy Zhou*
Clemson University

May 2023

**Abstract**

The U.S. Federal Aviation Administration (FAA) has implemented a large-scale multi-year infrastructure program called the Next Generation Air Transportation System (NextGen) to improve air transportation efficiency. To assess its efficacy, we estimate how NextGen projects completed between 2014 and 2017 affected air travel time and delays using an event study approach. We find sizable time savings in air travel time and delays from implementing NextGen. The time savings are more substantial for flights that experience unexpected shocks, such as poor weather and prior delays. In contrast, while NextGen seemed to close the performance gap between the hub and non-hub carriers generated by market power, the effect was short-lived and quickly reversed. Although we also find some small social benefits from carbon emission reductions, we cannot rule out that aggregate carbon emissions may have increased due to rebound effect.

Keywords: Air transportation, transportation infrastructure, air travel time and delay, private benefits of infrastructure improvement

JEL codes: L93, R4, O18, Q53

---

*Chu: Alumni of Yale School of the Environment. Zhou (corresponding author): John E. Walker Department of Economics, 320M Wilbur O. and Ann Powers Hall, Clemson, SC, 29634. Contact: yichen2@clemson.edu. Zhou thanks the National Bureau of Economic Research (NBER) and Sloan Foundation for the financial support. We thank Kenneth Gillingham for his generous donation of covering some RA hours under the same NBER grant. We thank Silke Forbes and James People for discussing this paper. We thank Tom Lam and Michael Makowsky for great comments. We thank participants at the NBER transportation workshop, the Southeast Energy & Environmental Economics Workshop, the AERE Summer Meetings, and the ASSA Annual Meetings. We thank the editor and the anonymous reviewers for insightful comments.

# 1 Introduction

The US air transportation system serves a large volume of passengers and continues to grow. In 2017 the number of flights departing from US airports accounted for 27% of all flights worldwide according to the World Bank data.[1] The Federal Aviation Administration (FAA) forecasted air traffic to grow at an average rate of 2.4% per year between 2016 and 2037 (FAA, 2018a). The commercial aviation industry accounted for 5.4% of the US GDP in 2012 (FAA, 2014). Despite its importance, US aviation infrastructure is much less competitive when compared to other developed countries and US air traffic congestion has been rising since the 2000s (FAA and Eurocontrol, 2010; Mayer and Sinai, 2003). In November 2021, Congress passed the Infrastructure Investment and Jobs Act (IIJA) and the Biden Administration signed the $1.2 trillion bill into law. The bipartisan IIJA includes a sizable budget, $25 billion, for aviation infrastructure upgrades aimed at improving air traffic efficiency.[2] Retrospective estimates of recent large-scale air infrastructure investments will be crucial in providing insights into the magnitude and the source of potential benefits.

This study examines the impact of a recent air traffic infrastructure program on a scale similar to that of the aviation budget in the IIJA - the Next Generation Air Transportation System (NextGen). We assess how NextGen projects that were implemented between 2014 and 2017 have led to changes in air travel performance in the US.

NextGen is a multi-year large-scale effort to improve airspace infrastructure that the U.S. government began planning in 2004, during the George W. Bush Administration, with a total budget of $20 billion USD (GAO, 2017). NextGen projects were adopted in 39 airports selected by the Joint Planning and Development Office (JPDO), which represent an important market share of US air traffic (JPDO, 2004). The selected airports were treated over time and some underwent multiple waves of technological upgrades (see Section 2 for details). For example, many treated airports implemented multiple runway operations (MRO), a technology that allows simultaneous takeoff and landing operations, which can increase the usage of runway and airspace and reduce delays and queueing time on the runway. Despite the fact that planning began in 2004, construction did not begin until 2013. In 2014 the first wave of projects were completed in multiple large airports such as the Hartsfield-Jackson Atlanta International Airport (ATL) and San Francisco International Airport (SFO). Since then, NextGen has continued to roll out in other airports, and previously treated airports continued to adopt additional NextGen projects.

[1] Source: https://data.worldbank.org/indicator/IS.AIR.PSGR?year_high_desc=true.

[2] Description from the White House https://www.whitehouse.gov/bipartisan-infrastructure-law/ and the FAA https://www.faa.gov/bil.

We collected historical flight data for all domestic flights from 2010 to 2017 using the On-Time Performance dataset by the U.S. Department of Transportation (DOT) for our analysis. We observe detailed air travel performance measures such as departure delay, taxi time, airborne time, and arrival delay for each flight. We also collected detailed NextGen completion history data from the FAA – specifically, the year and quarter each treated airport completed a NextGen project and the nature of each NextGen project.

To study the effect of implementing NextGen technologies, we need to have a comparable control group to help us control for the temporal-varying unobservables. Among 39 eventually treated airports (hereafter *NextGen* airports), most of them are top 40 airports with a few outliers from the top 60 airports. Therefore, we use the nine top 40 airports that are non-NextGen airports as the control group. In particular, we restrict our sample to include domestic flights that (i) depart from either a NextGen airport or a top 40 US airport, and (ii) land at a top 75 US airport. Given the control group and that some large airports are treated later in the panel (e.g., the Dulles International Airport (IAD) was not treated until 2017), we end up with a reasonable-sized control group over the time horizon.

To guide the empirical strategy, we begin with an event study exercise as in Dobkin et al. (2018) (details in Section 3). We use the time of the initial treatment at each airport to define the event time. We regress a set of air travel performance measures on event-time dummies for flights that depart from both a NextGen and a control-group airport. The rich historical flight dataset allows us to construct fixed effects and compare similar flights that depart from the same origin airport, are operated by the same carrier, and depart at the same time block (measured by the day of the week and the hour of the day). We find no evidence of a systematic pre-trend. For almost all air travel performance measures, the point estimates have a negative slope after the initial event, suggesting that implementing NextGen technologies have reduced air travel time and delays, and these time savings continue to increase post the event. We find no evidence of level shift before and after the initial event, suggesting that we should not deploy a typical difference-in-differences estimator in this context.

We design the baseline empirical model and specify two factors to explain the pattern we observe in the event study results. First, we allow the effect of NextGen to intensify over time by specifying a parametric quadratic post-trend, similar to Dobkin et al. (2018), to approximate the changes in slope after the completion of the initial project. Second, as many treated airports are repeatedly treated, we measure this continuous treatment using the cumulative number of NextGen categories completed at an airport and interact it with the quadratic post-trend variables. To be consistent with the event study, we include the same set of fixed effects and controls.

The identification rests upon the assumption that the *specific quarter* that we observe a treatment completed at an airport is exogenous. While certain airports can be potentially endogenously chosen to be early adopters, at which particular quarter the project can be completed is subject to various implementation difficulties. In section 4 we show that the FAA website frequently reschedules the planned milestones and that the actual completion can differ from the planned milestones; we also present FAA implementation reports documenting similar changes. The randomness in the specific timing is related to the uncertainty and variation in the duration of multiple milestones prior to the implementation stage. This holds even though early adopters and late adopters are not necessarily identical. Moreover, the uncertainty and variation in flight (re)scheduling further remove the potential endogeneity concern.

We find that implementing NextGen has led to an improvement in air travel performance, with diminishing marginal returns over time. Completing one more category of NextGen at the origin airport would reduce the *total air travel time* (measured by the elapsed time plus the departure delay) by 1.3 minutes in year one post the initial implementation, 2.3 minutes in year two, and 3.2 minutes in year three. The time savings initially come from all parts of the total air travel time. A year after the initial implementation, flights would experience 0.7-minute time savings in airborne time (54% of the overall air travel time 1.3 minutes), 0.2-minute taxi-out time (13%), and 0.4-minute time savings in departure delay (33%). By years two and three, more time savings come from airborne time and departure delay.

Our estimates imply that in 2017 alone, implementing NextGen has led to an overall reduction in air travel time by 4 minutes per flight, compared to the scenario where the treated airports were never treated. Using these estimates, we evaluate the private gains in passenger time savings and fuel cost savings by relying on both the FAA's Cost and Benefit Guideline (FAA, 2016a) and the FAA's operation-specific fuel use data as modeled with the Aviation Environmental Design Tool (AEDT). Our most conservative calculation suggests that the 4-minute air travel time reduction is equivalent to per-flight private gains of $368 (2017 USD), with 84% resulting from passenger time savings and 16% from fuel cost savings. The overall private gains from all treated flights in 2017 amount to a total of $0.9 billion (2017 USD) which is very sizable. We showed that our calculations are on the conservative side under alternative fuel consumption assumptions, and the private gains could be even greater with additional considerations. Lastly, we find some social co-benefits from carbon reductions but the magnitude is negligible compared to the private gains.

However, we cannot rule out that the increase in flight traffic plausibly enabled by NextGen improvements may have increased the aggregate carbon emissions.

NextGen programs include four categories of projects which we explain in detail in Section 2. To assess the potential outcome from adopting all technologies, we consider the hypothetical scenario in which all NextGen technologies were fully implemented in 2014. We simulate the air travel performance in event year 3 in 2017 after four categories of NextGen technologies were adopted in 2014, and compare that to the scenario in which none of the NextGen airports were treated. On average, we expect private gains of \$1,119 per flight from passenger time savings and fuel cost savings. This benefit would amount to a total of \$2.8 billion from all flights that depart from a treated airport in year 3 alone. Without accounting for other benefits which are excluded in the calculations, the air travel time savings alone may justify the \$20 billion budget within 7-8 years.

Infrastructure bills are usually proposed to improve the performance of flights and the well-being of passengers that suffer from the worst air travel experiences. Therefore, we also study the heterogeneous effects of NextGen across flights. For flights that end up on the right tail of air travel performance and delays due to unexpected shocks, we find the benefits do fall on those most in need. For example, we find greater time savings for flights that experienced severe weather than the baseline estimates, with most reductions resulting from the departure delay - suggesting efficiency improvement in terminal tower logistics. Also, for flights that experienced a severe prior delay from previous operations, they experienced much greater gains from NextGen than an average flight. Their gains come from almost all dimensions - a shorter airborne time, taxi-out time, and in particular departure delay.

However, we do not find the same result for flights that experience poor air travel performance due to plausible pre-existing distortions. Flights operated by non-hub airlines in a hub are likely pre-distorted by congestion created by hub airlines (Mayer and Sinai, 2003). While we find non-hub carriers benefitted more than hub carriers at the hub airports initially, this effect only lasted two years after the initial implementation. Similarly, while we find some initial benefits for low-cost carriers in comparison to legacy carriers, this effect is fully reversed in the third year after the initial implementation. These findings suggest lifting the infrastructure constraints has no long-term co-benefit in reducing pre-existing distortion from market power, and may exacerbate it and create greater inequality among carriers.

Lastly, our baseline results are robust to (i) alternative functional form specifications, (ii) the consideration of the difference between earlier versus later adopters, (iii) the consideration of the size of the treated airports, (iv) the consideration of market structure changes by investigating carriers going through mergers, (v) richer fixed effects and controls, and alternative samples.

This paper contributes to the literature studying the effect of transportation infrastructure investments. Zhang and Zhang (2003) develop an applied theory model and suggest the potential benefit of expanding airport capacity. As for empirical evidence, Morrison and Winston (2008) find that FAA expenditures on airports reduced the number of air traffic delays using cross-sectional aggregated airport-level data from 2000. For large-scale efforts such as the NextGen or the IIJA, most evaluations are done using ex-ante projections (e.g., FAA, 2016b). Our study provides micro-level evidence of the benefits and documents the sources of the benefits and the distributional effects. The NextGen project has been extended through 2030 and another round of aviation infrastructure funding has passed with the IIJA. Given these current large investments, our findings deliver a timely evaluation of the potential benefits of these programs - midway through NextGen, and at the initial planning stage of IIJA investments. Also, our findings provide information regarding unintended consequences.

This paper also contributes to the literature regarding the impact of improving infrastructure and increasing capacity on congestion.[3] Theories that study highway congestion predict ambiguous results - it is possible that improving capacity (e.g., highway capacity) could increase the supply of traffic during peak hours, and in the equilibrium the congestion would be the same in the absence of the capacity improvement. Given that our estimates are local, our results suggest at least short- to medium-run benefits of air travel time savings.

The rest of the paper is organized as follows. We provide the background information for the NextGen program in Section 2 and describe how we assemble the data in Section 3. In Section 4 we present our empirical model. We discuss our estimation results and implications in Section 5. In Section 6, we conduct robustness checks. In section 7, we conclude.

# 2 The NextGen Program

In 2004 the U.S. Congress passed the Integrated National Plan for the Next Generation Air Transportation System (NextGen), which grew out of the Vision 100 – Century of Aviation

[3]Other than underprovision of the infrastructure, a public good problem, there is also inefficiency from the congestion. The latter problem has been well studied and economists would usually prescribe a well-designed congestion price such as peak-time pricing as the first-best solution; or more sophisticated pricing accounting for other inefficiencies that are well-studied such as the carrier market power, the network effect, local airport competition, and other missing market factors (e.g., Brueckner, 2002, 2005; Daniel, 1995, 2001; Morrison and Winston, 2007; Mayer and Sinai, 2003; Lee et al., 2021; Lam and Zhou, 2021; Yan and Winston, 2014). However, when (suboptimally) under-provided infrastructure creates binding constraints for airlines' optimal scheduling problem to maximize their profits, there may be important private and social gains from adopting modern infrastructure. In addition to Zhang and Zhang (2003), the authors also study the capacity expansion consequence when there is market power (Zhang and Zhang, 2006).

Reauthorization Act (CARA). President George W. Bush signed CARA into law, and Congress authorized and created the Joint Planning and Development Office (JDPO) to coordinate NextGen programs with the FAA. According to the FAA, "NextGen is the FAA-led modernization of our nation's air transportation system. Its goal is to increase the safety, efficiency, capacity, predictability, and resiliency of American aviation... Airlines, general aviation operators, pilots, and air traffic controllers gain better information and tools that help passengers and cargo arrive at their destinations more quickly, while aircraft consumes less fuel and produces fewer emissions."[4] The FAA received $7.4 billion from Congress to facilitate NextGen programs from 2004 to 2014 and has projected $20.6 billion in total spending for the NextGen program through 2030 (GAO, 2017).

Policymakers and the airline industry support the NextGen program because investing in modern infrastructure can potentially pay for itself. Potential private benefits include reductions in: the frequency and length of delays which are costly to the airline and the consumer; cruising time by allowing planes to fly closer and to take straighter routes; taxi time on the runways; and jet fuel consumption via the above channels (FAA, 2011). Other benefits may include reduced emissions from improved ground operation efficiency and reduced air travel time (FAA, 2011).[5]

The JPDO selected 39 airports to undergo NextGen upgrades (JPDO, 2011). In Figure 1, we plot these treated airports, as well as the nine non-NextGen airports from the top 40 airports as the control group. NextGen airports tend to be the busiest airports in the US, such as Atlanta (ATL) and Chicago (ORD), representing a sizable share of air traffic; outliers such as Anchorage (ANC) and Louisville (SDF) have lower levels of air traffic. Table 1 Panel 1 shows that 84% of flights in our sample travel from a NextGen airport. Airports in the control groups include Nashville (BNA), Baltimore (BWI), Washington D.C. Reagan (DCA), Fort Lauderdale (FLL), Honolulu (HNL), Orlando (MCO), San Deigo (SAN), Salt Lake City (SLC), and Tempa (TPA). They are primarily midsize airports.[6]

[4] Grandfathered FAA webpage: https://www.faa.gov/nextgen/what_is_nextgen/. See Online Appendix page 3 for the webpage archive: https://zhouyc.files.wordpress.com/2023/03/nextgenfaa_20230301_onlineappendix_a_grandfatherweb.pdf.

[5] NextGen may also provide safer flights and improve national security. These benefits also align with the goals of the Vision 100 – The CAR Act, which was signed after the events of September 11, 2001.

[6] We choose the top 40 airports as the cutoff to both (i) allow treat and control airports to be reasonably comparable, and (ii) exactly include FAA's core 30 airports for NextGen that FAA deems as important and plausibly intended to treat. See webpage archives in Online Appendix pages 28-33 for details. Also, our treatment and control airports are comparable in terms of runway capacities, weather conditions, and level of competition at the route level (see Appendix Table A.2).

Initial planning of NextGen started in 2004, but the implementation is a decade behind schedule (GAO, 2016).[7] The first wave of projects did not complete the "pre-implementation stages" until around 2012 and 2013, the "implementation stage" until 2013, and the "completion stage" until 2014. We further indicate the year of the initial completed treatment for NextGen airports in Figure 2. The earliest adopters are likely the busiest airports. By 2014 this group, including selected airports such as Atlanta (ATL) and San Francisco (SFO) had completed their first NextGen project. We observe other high-traffic airports treated in subsequent years, including John F. Kennedy Airport (JFK), Dallas/Fort Worth (DFW), and Washington Dallas (IAD), with initial projects completed in years 2015, 2016, and 2017, respectively.

The FAA has implemented four main categories of projects: Multiple Runway Operations (MRO), Performance-based Navigation (PBN), Surface Operation and Data Sharing (SO), and Data Communication (DC).[8] MRO can increase runway capacity and improve runway accessibility by various means, such as reducing the separation between aircraft. For example, ATL implemented the wake recategorization (a type of MRO project) in July 2014, which allows aircraft to safely take off and land closer to each other, increasing capacity and flight efficiency. PBN can improve the flight path during cruising time and increase the predictability of arrival time. SO projects primarily improve logistical efficiency on the ground at the gate, which indirectly affects departure and arrival delays, and operations between the gate and the runway (for takeoff or landing). As a result, SO would mostly benefit surface efficiency. Lastly, DC can reduce communication errors between the pilots and the controller at the air traffic control (ATC) tower and airport terminal towers, which would benefit both safety and surface operations and may reduce delays. To explain the overall effect of adopting NextGen, we later produce estimates of the effect of adopting individual categories of NextGen projects.

We collected the completion history of NextGen projects published by the FAA.[9] For each treated airport, we observe the year and the quarter that each NextGen project was

[7] In 2007 the JPDO was established and FAA started the testing bed in Florida. After congressional effort and logistics, the midterm task force was formed in 2010. In 2012, the initial 35 treated airports were selected and in 2014 the first project was completed. After our sample period, the NextGen projects continued to roll in, especially projects under the category of Data Communication (DC). See details at https://www.faa.gov/nextgen/background

[8] These four NextGen priorities were listed on the grandfathered webpage: https://www.faa.gov/nextgen/snapshots/priorities/. See Online Appendix pages 4–5 for the webpage archive: https://zhouyc.files.wordpress.com/2023/03/nextgenfaa_20230301_onlineappendix_a_grandfatherweb.pdf.

[9] Grandfathered completion history were on https://www.faa.gov/nextgen/snapshots/priorities/completion_history/ and is included in the Online Appendix, pages 6–9. We cross-checked the priority records published by the FAA, e.g., the record of MRO https://www.faa.gov/nextgen/snapshots/

completed. Figure A.2 shows that 36% of airports have more than one category of project adopted in our sample period. We therefore construct a continuous variable to measure the intensity of the treatment as the key treatment variable of interest - the cumulative number of categories of projects completed at an airport in a given quarter. In Figure A.3 we plot the total categories of projects completed from 2010 to 2017. Consistent with Figure 2, most variation takes place between the first half of 2014 and early 2016. Similarly, Figure A.4 shows most treated airports underwent their first project in 2015, then in 2016 and 2014, and lastly a few in 2017. These numbers and patterns suggest that our estimates will be mainly driven by adopters from 2014 to early 2016. Consistent with this timing, an ex-ante cost and benefit analysis projected almost zero benefits from 2010 to 2014, and the majority of the simulated benefits began in 2015 (e.g., see Table 1 of FAA, 2016d).

# 3 Data and Suggestive Event-Study Evidence

**Air travel time data.** The primary dataset for this analysis is the On-Time Performance Dataset from the U.S. Department of Transportation (DOT) from 2010 to 2017. US airports and carriers have been mandated to report detailed air travel and delay information for all non-stop segments since 1987.[10] Our sample includes all domestic flights that flew from (i) a treated airport or a non-treated top 40 airport, and (ii) flew to a top 75 airport (see Appendix Figure A.1 for the destination airports in the sample). For each flight (defined as a non-stop segment), we observe the actual and scheduled departure (gate-out) and arrival (gate-in) time, which allows us to compute the minutes of departure and arrival delays. The gate-in and gate-out time also allow us to calculate the minutes of the elapsed time, hereafter *total air travel time*. We also observe the actual wheel-off (take-off) and wheel-on (landing) time which allow us to further compute airborne time, taxi-out time, and taxi-in time. We have the information on the carrier, origin and destination airports (i.e., the route). Also, we observe other characteristics such as the tail number which allows us to track the aircraft's previous operation flown into the departing airport and the self-reported delay code.

Table 1 Panel A reports the summary statistics of the on-time performance measures. On average, non-airborne time accounts for an important share of total air travel time, with 15% from taxi time and 6% from departure delay.

---

`priorities/?area=mro` (grandfathered webpages stored in the Online Appendix, pages 10–24). We also checked projects at various metroplex areas, `http://metroplexenvironmental.com/oapm.html`.

[10]Rule 14 CFR § 234 1987 requires airlines with at least 1 percent of domestic flights to report on-time performance metrics. See `https://www.govinfo.gov/content/pkg/FR-1987-09-09/pdf/FR-1987-09-09.pdf#page=165` and `https://www.law.cornell.edu/cfr/text/14/part-234`.

Air travel performances worsened over our sample period. Figure 3 shows that the total air travel time has increased by 15 minutes on average from 2010 to 2017. It is not just because passengers took longer flights. While we observe airborne time rose by 10 minutes from 2010 to 2017, we also observe an increasing trend in departure delay (by 2 minutes) and taxi time (by 3 minutes).

**Suggestive event-study evidence.** To guide our empirical strategy and establish some visual evidence, we begin with an event-study exercise. For flight $i$ operated by airline carrier $j$ departing from origin airport $o$ at calendar time $t$, we estimate the effect of implementing NextGen technology at the origin airport:

$$y_{ijot} = \sum_{s=\underline{S}}^{\overline{S}} \beta_s \cdot d_{os} + \phi_{ojb} + \eta_{yq} + \gamma X_{ijot} + u_{ijot} \tag{1}$$

where $y_{iodt}$ represents an air travel performance outcome. The main outcome we study is the *total air travel time* measured by the minutes of elapsed time plus departure delay. The departure delay is the difference between actual and scheduled departure time, and it measures the extra time a passenger spends waiting for the flight to depart from the gate. The elapsed time is the actual time from the departure gate to the arrival gate. We examine the main outcome as well as these two components. Furthermore, we examine four more refined components of the elapsed time: airborne time, taxi time, taxi-in time, and taxi-out time. Lastly, we also examine the arrival delay which also informs air travel performance.

The parameters are the event-time coefficient $\beta_s$. By the starting day of event time $s$, the dummy variable $d_{os}$ equals 1 if there has been a NextGen project completed at the origin airport $o$, i.e., it is set to one after the initial NextGen project and remains one afterward. In this exercise, event time $s$ is measured in years relative to the initial adoption.

The main identifying cross-sectional fixed effect is $\phi_{ojb}$. It interacts origin airport $o$ by carrier $j$ and by time block $b$ (defined as the day of the week and the hour of the day). This fixed effect allows us to compare flights that depart from the same airport, are operated by the same carrier and depart at a similar time in a week and in a day, but only differ in the status of whether a NextGen project has been completed at the origin airport at the time of departure. We include quarterly time fixed effects (calendar year by quarter) $\eta_{yq}$ as the main temporal fixed effect to remove temporal confounders such as market structural changes that are common to all flights and seasonal factors.

We include a vector of control variables $X_{ijot}$. It includes a dummy variable that represents whether the destination airport is a NextGen airport, control variables for scheduled buffer time and delay of the prior operation of an aircraft[11], interactions of carriers fixed effects with a linear quarterly trend, and a linear quarterly trend for slot-administrated airports[12]. We include flights that flew from control-group airports to remove the temporal unobservables. We can further control for destination characteristics but results are similar. Following recent developments in the event-study literature, we do not balance the panel around event time and use all the data in our sample for this illustrative exercise as trimming would result in an unbalanced panel by the calendar time (see e.g., Sun and Abraham, 2021; Borusyak et al., 2022). Also, we do not bin the end points and estimate $\hat{\beta}_s$ 7 years prior to the event to 4 years after the initial event, i.e., $\underline{S} = -7$ and $\overline{S} = 4$.[13] We cluster the standard errors at the level of the origin airport by the hour of the day.

Figure 4 shows our estimation results. For most measures, we do not observe a statistically significant pre-trend. While the end point $\hat{\beta}_{-7}$ appears to be lower for the main outcome and some other outcomes, this estimate is driven by the very few airports that were not treated until the second and third quarters of 2017. $\hat{\beta}_{-7}$ is also much less precise for this reason. Given that the majority of airports were initially treated in 2015, followed by 2016 and finally 2014 (see Figure A.4), the coefficients that carry the most weight in determining whether there is a significant pre-trend are estimates $\hat{\beta}_{-4}$ to $\hat{\beta}_{-2}$ (also see Figure A.5 Panel A for the distribution of observations over event time). These estimates are both small in magnitude and statistically insignificant. Nevertheless, we test the joint significance of all pre-event parameters, and the F-statistics are mostly very small.[14]

[11] Mayer and Sinai (2003) suggest the importance of scheduled buffer time as the delay from the previous operation may affect the current flight. We control whether the aircraft of the current flight is scheduled to arrive at the origin airport from its previous operation within a 45-minute or 30-minute window before the scheduled departure. Similarly, we include a linear control for the prior delay measured by the minutes of arrival delay of the previous operation of the current aircraft when it landed at the origin airport.

[12] The dummy of slot-administrated airports equals 1 for seven airports: John F. Kennedy International Airport (JFK), LaGuardia Airport (LGA), and Ronald Reagan Washington National Airport (DCA) are under slot controls; and Chicago O'Hare International Airport (ORD), Los Angeles International Airport (LAX), Newark Liberty International Airport (EWR), and San Francisco International Airport (SFO) are under schedule restriction.

[13] The literature has not settled on their recommendation regarding the practice of binning distant periods. While Sun and Abraham (2021) show that binning does not solve the contamination issue, Schimidheiny and Siegloch (2020) suggest it is important to overcome the underidentification problem. It is still widely used in practice (see all examples in Freyaldenhoven et al. 2019 and past studies surveyed in Sun and Abraham 2021). We choose not to bin for transparency. Instead, we provide the distribution of observations by event time in Figure A.5 so readers can "visually bin" the end points if they choose to. Our results are similar if we expand our sample period to 2010—2018 (See Appendix Figure A.7).

[14] The F-stat of the joint significance of six parameters all of six lags $H_0 : \beta_{-7} = \beta_{-6} = \cdots = \beta_{-2} = 0$ for the eight measures are: 0.6, 0.6, 0.6, 2.0, 3.1, 4.8, 4.2, and 0.9. Despite the small magnitude, these test statistics can still over-exaggerate the joint significance as they are not weighted. Much fewer data are used to estimate $\beta_{-6}$ and $\beta_{-7}$ (see Figure A.5). Nevertheless, we acknowledge that recent event-study literature has proposed alternative test procedures (see, e.g., Sun and Abraham, 2021; Roth, 2022).

Following the initial NextGen adoption, we observe a clear downward trajectory and statistically significant estimates for almost all measures that relate to the origin airport. The estimates for $\hat{\beta}_4$ is less precise because we lack observations with a four-year lead. The downward post-trend of the total air travel time appears to come from all dimensions related to the origin airport, including a negative slope for airborne time (despite being less precise), taxi-out time, and departure delay; and all these components show a sizable reduction after the event. In contrast, we do not find clear evidence of a sizable post-trend in taxi-in time, which is expected as it is not directly related to treatment that occurs at the origin airport. As for arrival delay, we observe an initial downward trend and it recovers in year 4. Lastly, we do not observe a clear pattern of level shifts before and after the event. Similar to this exercise, we estimate quarterly event time coefficients $\beta_s$ with $s$ measured in quarters relative to the initial adoption. Figure A.6 shows similar patterns despite that estimates are noisier. The event study results suggest that we should specify a regression model that allows the treatment to intensify following the initial year of the adoption.

Lastly, in Figure 4 we focus on treatments at the origin airport. Given congestion is distorted primarily for departing operations (Mayer and Sinai, 2003) and little for landing operations, it is mostly likely that improvements in air travel performance come from changes in departure operations. Nevertheless, we repeat equation (1) for treatment at the destination airports. We find some results for the total air travel time but mixed results for other measures. Both the theoretical plausibility and event-time evidence suggest that we should focus on the treatment at the origin airport.

**Other datasets.** To assess whether the NextGen adoption effect is greater during bad weather, we collected hourly weather information at each airport reported in the National Center for Environmental Information (NCEI) database from the National Oceanic and Atmospheric Administration (NOAA).

Also, to infer fuel consumption and emission reductions from air travel time savings, we acquired the Aviation Environmental Design Tool (AEDT) from the FAA. The AEDT simulates the scenario of fuel consumption and pollution emissions at different stages of air travel such as taxi-out, climb, approach, and taxi-in for each airport and each aircraft model.[15] We collected the fuel use and $CO_2$ emissions data in order to calculate private and social benefits. Since the mapping varies by aircraft, we also collected aircraft model information by linking the On-Time Performance dataset with the DOT Form-B43 using

[15] We applied for the Base of Aircraft Data (BADA) license from the European Organization for the Safety of Air Navigation (EuroControl) which allowed us to obtain the AEDT. The AEDT is based on the Aviation Emission Model (AEM) from EuroControl.

an aircraft's tail number. Table 1 shows the summary statistics of the final dataset after matching with Form-B43.[16] Lastly, we collected jet fuel price data from the U.S. Energy Information Administration (EIA) to evaluate fuel cost savings.

# 4 Empirical Strategy

We design the empirical strategy guided by the patterns of our event studies. There are two possibilities. First, following the initial adoption, the effect of NextGen may intensify over time. We therefore specify a quadratic functional form similar to Dobkin et al. (2018) to estimate the post-trend. In particular, we include $s \cdot \mathbf{1}(s > 0)$ and $s^2 \cdot \mathbf{1}(s > 0)$ where $s$ is the number of years since the initial quarter of treatment.[17]

Second, NextGen airports may continue to be treated after completing the initial project, as shown in Figure A.2. The continuous and repeated treatment may also explain part of the pattern we observe in the event studies. Therefore, we use the cumulative number of categories of project $n_{oq}$ at airport $o$ and quarter $q$ to measure the degree of the treatment. To allow for both factors to explain the post-trend, we choose to use the interaction terms $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ and $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ in our baseline model.

For flight $i$ operated by airline carrier $j$ flying from origin airport $o$ at time $t$, we estimate

$$y_{ijot} = \beta_1 n_{oq} \cdot s \cdot \mathbf{1}(s > 0) + \beta_2 n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0) + \phi_{ojb} + \eta_{yq} + \gamma X_{ijot} + u_{ijot} \tag{2}$$

where $y_{iodt}$ represents the eight air travel performance outcomes previously used in event studies: minutes of elapsed time plus delay departure (i.e., the base measure of the *total air travel time*), elapsed time, airborne time, taxi time, taxi-in time, taxi-out time, departure delay, and arrival delay.[18]

[16] We lose 26% of the sample after matching the On-time Performance data with the Form-B43. Later we show that the baseline results are indistinguishable had we not matched with Form-B43.

[17] There is a slight difference between our specification and Dobkin et al. (2018). As Dobkin et al. (2018) have no control group, they specify a linear event-time variable to control for potential pre-trend and allow for linear extrapolation. Therefore, they use additional quadratic and cubic terms to estimate the post-trend. In addition to the recent development in event-study literature, our estimation also relates to the literature that studies two-way fixed effects and the difference-in-differences estimator that either propose robust estimators that consider the variation in treatment timing and effect (see e.g., de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Borusyak et al., 2022). Our setting also considers varying treatment intensity over time within a unit, and heterogeneity arises from different types of technologies (discussed later in Section 5.1.)

[18] We later explore either shutting down the continuous treatment channel and estimating equation (3), or shutting down the channel of over-time improvement and estimating equation (4). Results suggest the baseline is robust, and it is important to model and include both channels.

The key parameters of interest are $\beta_1$ and $\beta_2$. Similar to our event study specification in equation (1), the main identifying fixed effect is $\phi_{ojb}$. It interacts the origin airport by carrier by time block (measured by the day of the week and the hour of the day). This fixed effect allows us to compare flights that take off from the same airport (e.g., ATL) operated by the same carrier (e.g., Delta Airlines) during a similar time (e.g., Wednesday 10 am, a peak hour targeted for a business traveller), yet in which one flight took off in 2013 (before ATL was treated) and another in 2016 (after ATL underwent the initial treatment). We include quarterly fixed effects $\eta_{yq}$, as in equation (1), to remove the time-series unobservables that are common to all flights. Also, our control variables $X_{ijot}$ are the same as in the event studies. In summary, our exhaustive sets of fixed effects and controls allow us to compare how air travel performance has been improved for a flight that travels from a NextGen airport over time, compared to its counterpart that travels from other airports. We cluster the standard errors at the level of the origin airport by the hour of the day to allow for correlation within a departing airport by a similar time in a day.

The key identifying assumption is that the *specific quarter* when a project is completed at an airport is exogenous. The airports selected to be treated are not random. Appendix Table A.1 documents that most NextGen airports are hub airports. Mayer and Sinai (2003) show that hub airports are more likely to experience worse air travel time and delays. Also, NextGen airports tend to be large airports. Therefore, prioritizing improving air travel performance in those airports may have been the policymaker's priority, making the selection of NextGen airports non-random. Similarly, the choice to make certain airports as earlier adopters is not random as well.

The plausible randomness in the treatment timing (down to a particular quarter) stems from the uncertainty in the duration of each stage of a NextGen project, which is common to many infrastructure programs. FAA frequently updates their scheduled, rescheduled, and actual milestones (i) on their website and (ii) in their "NextGen Implementation Plan Report" over time FAA (2014, 2015, 2016b, 2018b, 2021) when a planned milestone has a new proposed quarter of completion and when the actual milestone is completed earlier or later than the proposed quarter due to implementation difficulties, bureaucratic challenges, and other unanticipated causes.[19] For example, in April 2015, Los Angeles International Airport (LAX) planned to complete its first MRO project in 2016 Quarter 1, and in March 2016

$$y_{ijot} = \beta_1 s \cdot \mathbf{1}(s > 0) + \beta_2 s^2 \cdot \mathbf{1}(s > 0) + \phi_{ojb} + \eta_{yq} + \gamma X_{ijot} + u_{ijot} \tag{3}$$

$$y_{ijot} = \beta_1 n_{oq} + \phi_{ojb} + \eta_{yq} + \gamma X_{ijot} + u_{ijot} \tag{4}$$

[19]Examples of regulatory, funding, and bureaucratic challenges: https://www.rpc.senate.gov/policy-papers/nextgen-delayed-just-like-your-plane and https://www.govinfo.gov/content/

FAA changed the planned completion to 2016 Quarter 3 with a "new/revised" marking, and it was indeed completed in 2016 Quarter 3.[20] Therefore, while there are endogenous reasons that some airports may be treated earlier and airlines may form predictions of the broad timing (e.g., ATL would be treated early on), the specific quarter that the implementation completes is not deterministic. The exogeneity of the specific completion time comes from the randomness of each project stage which is common to many implementation programs. In the case of NextGen, before a project is marked as completed (i.e., entering the "post-implementation stage"), a project has to finish multiple stages: the "budget impact stage", which takes about 2 quarters; the "study stage" which takes 1 to 3 quarters; the "facility resource issue stage", which takes 1 to 3 quarters; the "evaluation stage", which takes 2 to 7 quarters; an optional "En Route Automation Modernization (ERAM) resource impact stage", which takes 4 to 5 quarters; and finally the "implementation stage" which takes 2 to 5 quarters (e.g., FAA, 2016b); and the timing of each stage can vary and is subject to uncertainty. Therefore, the actual completion of each stage does not always conform to the scheduled completion date and is exogenous.

Nevertheless, one may worry that airlines can respond to NextGen projects by strategically re-optimizing their flights at treated airports before and after the treatment. For example, after ATL completed its first NextGen project in 2014, the hub airline at ATL, Delta Airlines, might have clustered more flights during peak hours to serve more consumers, or they could have decreased their strategic scheduling paddings (see e.g., Forbes et al., 2019), both of which could bias estimates upwards. The plausible friction for airlines to make such changes in the short run mitigates such concern. Airlines typically schedule their flights 9-12 months ahead of time on a Computer Reservation System (CRS), and it takes time to reschedule flights (see Rule 14 CFR § 255 for regulation details and Borenstein (1999) and Forbes (2008) for discussions). While they can divert and cancel flights, carriers have to pay the CRS platform fees to make changes, and these changes may be subject to constraints. Also, both economic and engineering literature that study schedule recovery (i.e., strategically changing operations on the fly) suggests that re-optimization is limited in what it can achieve in

pkg/CHRG-113shrg95362/html/CHRG-113shrg95362.htm. Examples of implementation difficulties: https://www.uas.aero/issues-nextgen-implementation/.

[20] In Online Appendix https://zhouyc.files.wordpress.com/2023/03/nextgenfaa_20230301_onlineappendix_a_grandfatherweb.pdf, we provide historical FAA websites of the same page recorded at different times. The LAX MRO example can be find on pages 25 (webpage in 2015), 26 (webpage in 2016), and 9 (webpage in 2018). For another example of MRO, in April 2015, San Francisco International Airport (SFO) planned to complete its first MRO project in 2015 Q3 (page 25), and in March 2016, FAA changed the planned completion to 2016 Q2 (page 26), and the project was indeed completed in 2016 Q2. For an example of SO, in April 2016, Charlotte Douglas International Airport (CLT) planned to complete its first SO project in 2017 Q4 (page 27). Still, CLT has not implemented it at the end of our sample period.

the short run and is usually quite costly (Rupp et al., 2005; Evler et al., 2021). Lastly, it is possible that the treatment received at the treated airports may benefit departing operations indirectly via the network in which case our estimates can be slightly on the conservative side.[21]

# 5 Estimation Results

In this section, we first present our baseline results in Section 5.1. We then discuss the heterogeneous effects of implementing NextGen technology in Section 5.2. Finally, we discuss the welfare implications of our baseline estimates in Section 5.3.

## 5.1 The Impact of NextGen on Air Travel Performance

**NextGen programs in general.** For our baseline, we estimate equation (2) on the total air travel time (minutes of elapsed time plus the departure delay). Table 2 column 1 reports the results. We find strong evidence of a reduction in the total air travel time with some diminishing returns over time. In particular, the point estimate $\hat{\beta}_1$ is -1.41 and statistically significant; $\hat{\beta}_2$ is 0.12, smaller in magnitude and less precisely estimated. To interpret the coefficients, we compute the marginal effect of adopting an additional category project ($\Delta n_{oq} = 1$) in various years after the initial adoption - doing so would result in a reduction in the total air travel time by 1.3 minutes in the first year, 2.3 minutes in the second year, and 3.2 minutes in the third year.

To understand the source of time savings, we also examine the following outcomes: elapsed time, airborne time, total taxi time, taxi-out time, taxi-in time, and departure delay. We show the results in Table 2 columns 2 to 7. We find time savings across all these components of the base measure except for the taxi-in time. Most coefficients, especially the linear term coefficients, are precisely estimated. We do not obtain a precise estimate of the effect on airborne time. Still, its magnitude is relatively large and consistent with the magnitude of other measures (i.e., marginal effects from columns 3, 5, 6, and 7 add up to column 1). As for taxi-in time, not only are the coefficients imprecisely estimated, their magnitudes are close to zero. We interpret this result to indicate that the infrastructure upgrade at the origin airport did not indirectly benefit the taxi-in operations at the destination airports. The operations that directly relate to the origin airport contributed the most to the air travel time savings: the airborne time reduction contributed 54% to the total air travel time

[21] At the flight level, when (and if) it happens, part of the benefits may indirectly propagate to an immediate successor flight at the next node (the destination of the inbound flight), and the propagated benefit can even be smaller when the ground buffer is greater (e.g., Brueckner et al., 2022). At the airport level, a similar effect can happen to the next node, where a subset of flights may experience the propagated benefits, depending on the composition of inbound flights and the corresponding ground buffer of the outbound flights.

savings in the first year after the initial adoption and became more important over time, the taxi-out time reduction contributed 13% in the year one but became less important over time, and the reduction in departure delay consistently contributed 33%.

The above evidence suggests that NextGen technologies likely benefit (i) the navigation, approaching and climbing operations which shorten the airborne time, (ii) ground logistics and operations which reduce the taxi-out time, and (iii) air traffic tower efficiency which reduces congestion and leads to shorter taxi-out time and departure delay. (We will discuss specific technologies next.) Lastly, we also find time savings in the arrival delay, as reported in column 8.

**Specific NextGen projects.** The FAA has implemented four categories of NextGen projects - Multiple Runway Operations (MRO), Performance-based Navigation (PBN), Surface Operation and Data Sharing (SO), and Data Communication (DC). Our baseline equation (2) uses $n_{oq}$ as the main variable of interest where we construct $n_{oq} = \sum_c d^c_{oq}$, in which $c$ indicates each of the four categories of NextGen project $c = \{mro, pbn, so, dc\}$. Appendix Figure A.8 shows the temporal variation in implementation for each category of projects from 2014 to 2017.

To further assess how adopting individual categories of projects affects air travel performance outcomes, we repeat our baseline model by replacing $n_{oq}$ with four dummy variables $d^c_{oq}$. As the quadratic term estimates are relatively small in the baseline, we only estimate the linear post-trend for better interpretability in this exercise.

We report the results in Table 3. We find that while some technologies lead to consistent travel time savings across almost all measures, other technologies appear to result in tradeoffs across performance measures. For example, we find Multiple Runway Operations (MRO) led to all performance improvements at the origin airport (i.e., all measures except for taxi-in time and arrival delay). MRO is designed to improve efficiency by allowing simultaneous same or opposite direction operations on parallel runways, and it is the technology with the highest rate of implementation (27 airports adopted MRO). Our findings of time savings in taxi-out time and delay departure are consistent with the plausible efficiency gain from MRO. We also find sizable time savings from Surface Operation and Data Sharing (SO). SO was adopted by 12 airports. It includes technologies such as system-wide information management (SWIM), traffic flow management system (TFMS), and time-based flow management (TBFM) which improve both queueing efficiency, departure operations, and surface logistical efficiency. Our findings are again consistent with the nature of the technology.

In contrast, we find data communication (DC) led to some tradeoffs. 10 airports adopted DC. This technology is designed to provide a direct link between pilots, air traffic controllers, and ground automation. We consistently observe sizable time savings in departure delay and some small savings in taxi-out time (imprecise). Yet the additional data communication and safety clearance service provided by DC seem to lead to greater airborne time. Lastly, we also find some tradeoffs generated by adopting Performance-based Navigation (PBN). As this technology focuses more on improving navigation precision and accuracy, safety, and arrival rate, we see some evidence of shorter delays (imprecise). However, we also observe that adopting PBN leads to tradeoffs in greater airborne time and taxi-time. Nevertheless, estimates for this technology lack power as it was adopted by the fewest airports (6 in total).

FAA (2016c) suggests that FAA spent much more on certain technologies in PBN and DC than other technologies (such as technologies in SO). Combining FAA's expenditure reports and our results, it is likely that PBN and DC are the technologies that produce some time savings in some parts of air travel at the cost of other parts of air travel while inflicting greater costs; in contrast, SO and MRO can be more appealing.

We caution readers and policymakers against over-extrapolating the result in that (i) our estimates do not directly produce a cost-benefit analysis for each type of project, and (ii) PBN and DC may provide other benefits as well such as safety, or benefit time-savings indirectly in the long run. Also, the policy recommendation will depend on which part of air travel performance the policymaker is most interested in improving. For example, if the policymaker is mostly interested in reducing departure delay, DC will provide the most benefit, and MRO and SO the next; if the overall air travel time is more important for the policymaker, SO will provide the most benefits instead.

This exercise explains how the adoption of each technology contributes to the estimates that we obtained in Table 2. Many of these technologies lead to improvements in (i) the air traffic tower logistics that arrange and optimize the lineup of flights to leave the gate, resulting in shorter departure delays and taxi time, and (ii) the take-off and terminal climb operations, resulting in shorter airborne time with some tradeoffs in airborne time. Nevertheless, we caution readers against over-extrapolating implications about the effect of individual categories of technologies, as the four individual treatments are modestly correlated with each other (Appendix Table A.3). In summary, we find that NextGen infrastructure projects lead to sizable improvements in air travel performances.

## 5.2 Heterogeneous Effects of NextGen

It is of great policy interest to evaluate how the benefits are distributed and whether the benefits fall on the targeted recipients. In the case of air traffic, one group of targeted recipients is arguably the flights that experience the most severe unexpected delays (such as a snowstorm). These flights experience thin-tail (low-probability) but high-impact events. Benefiting these flights disproportionately more would lead to higher total benefits and plausibly lower variation in the extreme outcome. The other possible groups of targeted recipients are the flights that are (endogenously) pre-distorted the most by other market failures - e.g., non-hub flight congestion at a hub airport during peak time created by hub airlines (Mayer and Sinai, 2003). In this case, benefiting those flights disproportionately more would restore some of the market efficiency. In this section, we examine the distributional effects of adopting NextGen projects across flights that are delayed the most for exogenous reasons and pre-distorted the most for endogenous factors.

**Heterogeneity by weather conditions.** Many flights experience delays due to bad weather. The FAA reports that 69% of delays are due to severe weather that impairs visibility.[22] Despite the fact that modern weather predictions enable airlines to plan for the weather ahead of time to a certain degree, the variance of the severity and uncertainty in the precise timing of the weather shock can still affect air travel performance (Morrison and Winston, 2008).

First, we examine the effect of NextGen when visibility is impaired. We restrict the estimation to the subsample when visibility was under 10, 9, 7, and 4.5 miles. These numbers represent the worst 25th, 15th, 10th, and 5th percentile in terms of visibility among airports over time, which we select using NOAA's NECI dataset (we exclude the hours from 1 am to 5 am). Figure 5 Panel A shows the results. We plot $\hat{\beta}_1$ as the second-order coefficient $\hat{\beta}_2$ is in general relatively small. We find $\hat{\beta}_1$ for the base measure, the total air travel time, grows from -1.41 to -2.10 as we further restrict the subsample to the hours when the visibility is poor. Although we see some increasing time savings in taxi-out time, most of this change comes from a shorter departure delay. (Full results in Table A.6.)

Similarly, we study the case when the sky ceiling is low. Specifically, we examine the airport and hour when the sky ceiling was less than 6,000, 3,000, 1,750, 1,400, 900, and 750 feet, which represent the worst 40th, 30th, 25th, 20th, 15th, and 12.5th percentile among airport over time (they account for 24%, 16%, 10%, 8%, 5%, 4% of the flights that experienced the

[22] Source: https://www.faa.gov/nextgen/programs/weather/faq/.

worst sky ceiling). In Figure 5 Panel B, we observe a similar pattern. As we further restrict the subsample to the time when the sky ceiling is low, $\hat{\beta}_1$ grows from -1.41 to -2.35 with most of the changes coming from a shorter delay departure. (Full results in Table A.7.)

In summary, we find that the total air time savings are greater for flights that experience severe weather conditions. We also identify the components of the total air time saving that led to this outcome: adopting NextGen led to a slightly shorter taxi-out time likely by improving ground operation efficiency, and a much shorter departure delay likely by improving airport terminal and air traffic tower efficiency during bad weather conditions.

**Heterogeneity by prior delay.** In addition to weather conditions, unexpected prior delays can have a large impact on air travel performance. On average, an aircraft is used for five operations throughout each day (aka 5 flights), 26% of flights experience a prior delay; for flights with a previous operation, the average prior delay amounts to 29 minutes (Table A.1 Panel A). Moreover, 52% of delayed flights report "late aircraft" as the reason to the DOT (Appendix Table A.1 Panel B).

These flights are potential treatment targets because they have significant delays, and part of their delays have an unexpected component from previous operations. While the carrier can control the expected prior delay through strategic scheduling and buffering, the realized minutes of prior delay is unexpected. Moreover, the realized prior delay may propagate throughout the day within an aircraft over its numerous operations (see Brueckner et al., 2022), making the prior delay of a later operation more unpredictable. Moreover, pre-distortion such as congestion created by hub airlines at the hub in those previous operations may further exacerbate the unexpected part of the prior delay.

We repeat our estimation using different subsamples when the arrival delay of the previous operation of the current aircraft is greater than 1, 10, 20, 40, 75, and 120 minutes. These observations represent flights that experienced the worst 25th, 15th, 10th, 5th, 2.5th, and 1st percentile in terms of minutes of prior delay. Figure 5 Panel C shows the result. As in the previous panels A and B, here we only show $\hat{\beta}_1$. The point estimate $\hat{\beta}_1$ for overall air travel time grows sharply from -1.41 to -5.28 as we examine the flights with worse and worse prior delays. This change comes from a greater $\hat{\beta}_1$ in almost all components of the total air travel time: the $\hat{\beta}_1$ for taxi-out time almost triple and $\hat{\beta}_1$ for the departure delay grows 8-fold as we move to the worst 1st percentile. $\hat{\beta}_1$ for the airborne time also sees an increase (See the full results in Table A.8.)

These results suggest that adopting NextGen improves air travel performance for previously delayed flights disproportionately more than other flights. The benefit comes from many components of air travel performance especially taxi-out time and departure delay, suggesting disproportionate gains for those flights in terms of ground efficiency and terminal logistics.

**Heterogeneity by slot administration status.** We similarly analyze the heterogeneous effect by whether the airport runway capacity is restricted by plausible exogenous regulations from the FAA and the International Air Transport Association (IATA). We repeat the baseline estimations for 7 slot-administrated airports (JFK, LGA, DCA, ORD, LAX, EWR, and SFO) with the same control groups. We also repeat the exercise for non-slot-administrated airports.

We report our results in Table 4. To best interpret the results, we compute the marginal effect of $\Delta n = 1$ in various years after the initial adoption. While we find greater time savings at the slot administrated airports in the first two years after the initial adoption, time savings in non-slot-administrated airports almost caught up by year 3 (although point estimates are less precise). This evidence suggests that NextGen infrastructure benefits the capacity-constrained airports more so than other airports but only temporarily.

**Heterogeneity by hub airports and hub airlines.** Next, we examine how the infrastructure improvements affect flights by the hub status of an airline at a hub airport. In our sample, 67% of flights travel from a hub airport and 56% of NextGen airports are hub airports (see Table A.1). Hub airlines create a pre-existing market inefficiency at their hubs by clustering flights during peak hours which in turn congests other airlines, and the external costs can be quite sizable (Mayer and Sinai, 2003; Morrison and Winston, 2007). Mazzeo (2003) documents that less competitive routes are associated with worse air travel performance and delays.

In Table 5 panels A and B, we re-estimate our baseline model on subsamples depending on the hub status of the origin airport.[23] As in the previous tables, we report the marginal effect of $\Delta n = 1$ in various years after the initial adoption for better interpretability. We find that flights from a hub airport reaped more time savings in the initial two years than those that departed from a non-hub airport; flights from a non-hub airport experienced smaller gains initially but the gap between these two types of flights closed up and reversed over time. This result seems to suggest that NextGen initially restored some market efficiency to the congestion caused by hub airlines in hub airports in comparison to non-hub airports, but eventually that treatment effect evened out across all airports.

[23]We define hub airports using the Herfindahl-Hirschman Index, similar to Mayer and Sinai (2003) and rule out non-hub airports by cross-checking the hub status for each airport.

We further examine the outcome within hub airports by the carrier hub status in Table 5 Panels C and D. We find flights operated by non-hub carriers at a hub airport (e.g., Southwest at ATL) received greater initial time savings than the hub carriers, but the difference tapered off in year two, and eventually reversed by year 3. This result is consistent with Panel A and B as most externalized delays and congestion at hub airports are experienced by the non-hub carriers. Our findings of no long-term greater gain for the non-hub carriers is consistent with the theoretical implications as in Zhang and Zhang (2006), who find no benefit of airport capacity expansion when carriers have market power. In summary, we find NextGen led to some disproportionate initial gains for non-hub carriers at the hubs and recovered some pre-existing distortions temporarily; yet the gains became more equalized over time and distributed back to hub carriers disproportionately in the longer term.

**Heterogeneity by legacy and low-cost carrier.** To further study the effect of NextGen by market power, we examine the outcome by whether a flight is provided by a legacy carrier or a low-cost carrier (LCC). In our sample, 54% of flights are operated by a LCC (see Table A.1). Li et al. (2022) estimate that legacy airlines have a greater connecting quality than the LCC airlines, and Mayer and Sinai (2003) document the connectability correlates with externalized delays from peak-hour clustering.

In Table 6 we repeat the baseline by legacy and low-cost carrier status. In our sample period, the legacy airlines include American Airlines (AA), Delta Air Lines (DL), United Airlines (UA), US Airways (US), Alaska Airlines (AS), and Hawaiian Airlines (HA); the low-cost airlines are the remaining airlines in the sample such as Southwest. Similar to our results by hub status, we find flights by low-cost carriers experienced a greater time saving initially following the NextGen adoption, but legacy carriers started to see greater time savings in year 3 (despite being imprecisely estimated).

In summary, while we find adopting the modern infrastructure under the NextGen program has disproportionately benefitted flights that experience the worst air travel performance due to unexpected reasons (e.g., severe weather, prior delay, and slot-constrained airport), we also find that the NextGen program has benefitted flights that were distorted by pre-existing efficiencies the most for a year or two after the initial adoption. Eventually flights provided by carriers with a greater market power benefit more from the NextGen.

## 5.3 Implications for private and social benefits

In this section, we discuss the implications for private benefits based on time-saving estimates obtained in Section 5.1. We provide social benefits from carbon reduction as well.

**Benefits of NextGen in 2017.** We begin by calculating the benefits of adopting NextGen technologies on flights in 2017 alone. In particular, we compare the actual air travel performance in 2017 for flights that departed from a NextGen airport to their performance in the hypothetical scenario if none of the NextGen airports were treated.

Using the observable degree of treatment $n_{oq}$ in 2017 and our baseline estimates in Table 2, we produce the counterfactual outcome if no NextGen airports were ever treated. We present the actual and counterfactual air travel time in Table 7 Panel A. In Row A of Panel A, we report the actual 2017 air travel performance, and in Row B we report the counterfactual air travel performance. We compute the difference between Row A and Row B, i.e., the changes in air travel performance $\Delta y$, in Row C. Our calculation suggests that there are an average of 4.2-minute total air travel time savings in 2017 compared to the counterfactual scenario if NextGen technologies were never introduced.

Next, we compute the private benefits on airlines and passengers via the channel of time savings (aka our estimates of $\Delta y$ in Row C). To provide the most conservative results, our baseline calculation only includes fuel cost savings and passenger time savings. For the passenger time savings, we evaluate passenger time savings at $48.71 per hour per passenger using the parameter of an all-purpose traveler (business or casual) recommended by the FAA's Cost and Benefit Guideline (FAA, 2016a, Section 1).[24]

As for fuel cost savings, we use AEDT data on matched airports and aircraft models to estimate the per-minute benefit of fuel savings. We report the estimation result in Appendix Table A.12 Column 1. For example, the marginal fuel use savings from a one-minute reduction in taxi-out time is 18.8 kilograms. In our baseline counterfactual, we use the fuel-saving parameter from taxi-out operation (Row A.1) to compute the taxi-out fuel savings, we use the fuel-saving parameter from taxi-in operation (Row B.1) to compute the taxi-in fuel saving, and we use the average of take-off climb (Row A.4) and approach (Row B.2) to compute the airborne fuel savings to be the most conservative. We evaluate the fuel cost savings using the 2017 jet fuel price from the EIA.

We report our baseline private benefits in Table 7 Panel B.1. We find the reductions reported in Panel A result in a per-flight savings of passenger time of $308 (2017 USD) and fuel cost savings of $60, making a total savings of $368. The overall private benefits from the treated flights in our sample in 2017 amount to a total of $0.93 billion (2017 USD). Passenger time savings contribute most to the private savings, accounting for 84% of the overall private benefits; fuel cost savings account for 16% of the overall private benefits.

[24] We obtain the number of passengers per flight from the DOT 2017 Air Traffic Data report. https://www.bts.gov/newsroom/2017-traffic-data-us-airlines-and-foreign-airlines-us-flights.

Our baseline fuel cost savings are calculated based on certain sets of assumptions using AEDT estimates from various departing and arrival operations. In Appendix Table A.13 Panel A.1 Columns 1 to 3, we calculate private benefits under alternative fuel use assumptions using AEDT data. Our estimates range from \$0.89 to \$1 billion, close to our baseline of \$0.93 billion. Also, we use fuel use and fuel cost parameters suggested by the FAA Guideline FAA (2016a) in Columns 4 and 5. The private benefits range from \$0.93 to \$1.06 billion. These calculations suggest our baseline private benefits calculations are quite conservative and robust to alternative fuel use assumptions.

Moreover, we also present less conservative calculations with additional considerations and report them in Table 7 Panel B.2. If we were to consider flights that are not matched in Form-B43 and if their fuel use estimates are similar to Appendix Table A.12, the overall private gain would be \$1.26 billion. If we were to consider the delay multiplier used in the FAA's guidelines, the overall private gain would amount to \$1.39 billion. If we also consider potential crew time savings using parameters in the FAA's guidelines, the private gains would be \$1.11 billion. Finally, if multiple possibilities presented above were considered at the same time, the private gains could be even greater.[25] Our results are also conservative in that we do not study co-benefits that NextGen may lead to in stimulating other local industries and its effect on local labor markets (Lakew and Bilotkach, 2017).

Lastly, we calculate the plausible social benefits via reducing $CO_2$ emissions. We take estimates from AEDT data (see Appendix Table A.12 Column 2) and calculate the implied $CO_2$ emission reduction in the same way that we calculate the fuel use savings. We report the social benefits in Table 7 Panel B.3. Evaluating the social cost of carbon at \$40 per ton in 2017, the social benefits from carbon reduction would be \$15.3 per flight and \$40 million (2017 USD) in total. This social benefit is relatively small, equivalent to only 4.2% of the private gain. Also, as NextGen technologies may have increased air traffic, the aggregated carbon emission may have increased. Examining this rebound effect is beyond the scope of this study and will be important to address for future research. In summary, we find sizable private benefits from reducing air travel time and some weak evidence on social benefits.

**Benefits of Fully Implementing NextGen in 2017.** We proceed to examine the hypothetical scenario in 2017 if all NextGen airports had fully adopted all four categories of NextGen technologies and if they had adopted those technologies in 2014. This would be equivalent to evaluating the outcome of air travel time in year three after fully adopting all

[25] Also, our calculation excludes second-order private gains from shorter travel time, such as a lower probability to miss connection flights (e.g., Bratu and Barnhart, 2005), which will imply even greater private gains.

four technologies. We then compare that to the scenario in which the NextGen airports were never treated in 2017. This exercise is also more related to our empirical findings than the previous exercise: given our parametric specification, the most informative counterfactuals we can produce takes the form of "evaluating the private benefits of adopting the NextGen technologies in year X". This is also relevant to inform policymakers about a potential pay-back period from investing in those technologies.

We begin by simulating counterfactual air travel time in Table 8 Panel A. Rows A-B repeat Rows A-B in Table 7 Panel A. In Row C, we add the changes in air travel time to Row B assuming the NextGen airports were fully treated in 2014. In Row D we report the difference between Rows B and C, $\Delta y$. Our estimates suggest that after implementing all categories of NextGen technologies, in event year 3 alone we should expect a 12.6-minute reduction in total air travel time savings, compared to the scenario in which NextGen technologies were never implemented.

In Panel B.1 of Table 8, we report the associated private benefits. We find the air travel time savings $\Delta y$ reductions reported in Panel A Row D would lead to per-flight savings in passenger time of \$928 (2017 USD) and fuel cost savings of \$191, making a total of \$1,119 in savings. The overall private gains from the treated flights in our sample in 2017 would be \$2.8 billion (2017 USD). Passenger time savings contribute most to the private savings (83%), as in our previous exercises.

Similar to the previous exercise reported in Table 7, our baseline private gains are robust and conservative relative to alternative ways of calculating fuel cost savings (see Panel B of Appendix Table A.13). Private gains would be much greater under additional considerations (see Panel B.2 of the current Table 8). Lastly, we find some evidence of social benefits from carbon reduction (\$ 120 million), but the magnitude is much smaller than the private gains and subject to the same critique on the rebound effect as explained earlier.

As we will show later in Section 6, our results are robust to many other factors that may affect air travel performance, such as weather and the level of competition. Therefore, our results in Tables 7 and 8 can also be used for policymakers to infer the average benefits of adopting NextGen technologies conditional on those factors as well.

Our results in Tables 7 and 8 suggest that adopting NextGen has led to a sizable increase in private welfare in 2017 alone (\$0.93 billion) and it would lead to sizable private gains in event year 3 (\$2.8 billion) if all treated airports were fully treated with all categories of NextGen technologies in 2014. These results imply that the \$20 billion budget of NextGen would plausibly be justified and paid off by event year 7-8 based on its private benefits, under the assumption that the air travel time savings would evolve in the linear slope that we document. Similarly, if the aviation infrastructure program under IIJA continues to improve air transportation similar to NextGen for both current NextGen and a broader base of airports (such as BWI, DCA, and FLL), it could provide sizable private gains.

# 6 Robustness

In this section, we examine the robustness of our baseline model by investigating alternative treatment measures, alternative specifications of the model, and other robustness checks.

**Alternative specifications.** In our baseline model, we consider both the effect from (i) a continuously repeated treatment $n_{oq}$ and (ii) the possibility that the effect of adopting NextGen can intensify over the course of time given the number of $n_{oq}$. Here we assess the results if we only consider one channel.

In Table A.4 Panel A, we estimate equation (3) without only the quadratic post-trend variables $s \cdot \mathbf{1}(s > 0)$ and $s^2 \cdot \mathbf{1}(s > 0)$. This specification is most similar to Dobkin et al. (2018) in which there is no repetitive treatment. The results are qualitatively consistent with the baseline. The point estimates are much more sizable since we contribute the varying intensity of treatment $n_{oq}$ into the effect of the linear and quadratic post-trend variables.

In Table A.5 Panel A, we estimate equation (4) with $n_{oq}$ as the only variable of interest. Again, we find a consistent result compared to the baseline results. The point estimates are much larger which comes from averaging out the point estimates after years of the initial adoption. These exercises suggest it is important to include both varying treatment intensity using $n_{oq}$ and allowing the effect of NextGen to intensify over time in the baseline.

Lastly, we find a similar pattern if we estimate the equation on individual categories of projects (see Table A.5 Panel B). Similar to Section 5.1, we find MRO and SO lead to time savings across all measures while PBN and DC lead to some tradeoffs across performance measures.

**Earlier versus later adopters.** We examine the effects of NextGen for earlier adopters and later adopters. It is plausible that the effects for early adopters are very different from later adopters. While some large airports were treated later in the sample (e.g., Washington Dulles (IAD) in 2017), many large airports were treated earlier in the sample (e.g., Atlanta (ATL) in 2014). Although our identification does not rely on the exogeneity of the general treatment timing (e.g., ATL may be treated earlier for endogenous reasons) but on the timing of the specific quarter an airport completes its treatment (e.g., the specific quarter that ATL was treated earlier in the sample), it is still important to compare earlier versus later adopters. In Table 9 Panel A we examine earlier adopters by excluding airports that were first treated in 2016 and 2017 (e.g., exclude IAD and others). Similarly in Panel B, we examine later adopters by excluding airports that were first treated in 2014 and 2015 (e.g., exclude ATL and others). We find while effects for both groups are comparable to the baseline, the earlier adopters have slightly greater point estimates. Due to a lack of the number of quarters in the post-treatment period for later adopters, the effects in Panel B are less precise than Panel A. In summary, our baseline is robust to the consideration of earlier versus later adopters.

**Size of the treated airport.** As most of the control group airports are midsize and medium-large airports, we investigate this factor by dropping the last two quarters of 2017 and restricting the sample from 2010 to 2017 Quarter 2. Doing so effectively moves Austin Bergstrom International Airport (AUS) and Washington Dulles International Airport (IAD) from the treatment group to the control group and increases the representation of large airports in the control group. We report the result in Table 10 Panel A. The implied marginal effects from $\hat{\beta}_1$ and $\hat{\beta}_2$ are comparable to the baseline.

Furthermore, we repeat the baseline by interacting a dummy variable that indicates midsize airports. We define an airport to be midsize if it is a non-top-15 airport.[26] Table 10 Panel B shows the result. Most of the interaction terms are insignificant. Again, the implied marginal effects from $\hat{\beta}_1$ and $\hat{\beta}_2$ are comparable to the baseline. In summary, we find our baseline robust to the consideration of the size of the treated airports.

**Consolidated carriers.** Next, we consider a subsample that excludes major carriers that were consolidated during our time period. We exclude three carriers: Continental Airlines (CO) consolidated by United Airlines (UA) after 2011; Airtran Airways (FL) consolidated by Southwest (WN) after 2012; and US Airway (US) consolidated by American Airlines (AA) after 2014. We report the results in Table 11 Panel A and find the results are similar to the baseline. Similarly, we investigate the results if we restrict to airlines that go through major mergers. In particular, we include only six carriers: CO, UA, FL, WN, US, and AA. We report results in Table 11 Panel B, and they are also similar to the baseline.

**Alternative samples.** Additionally, we consider a less restricted sample that includes flights that do not match the DOT Form-B43. We report the results in Appendix Table A.9 and find that the results are quantitatively similar to the baseline.

[26] The choice of choosing non-top-15 is that five control airports were in the top 16-25 using the number of domestic flights in our datasets from 2010 to 2017. Note we only count the number of flights rather than the number of passengers, and we only use domestic flights. Also note that the current ranking is not the same as it was in 2010–2017.

Also, it may be helpful to include additional control groups although it may come with the tradeoffs of including smaller airports (e.g., see Appendix Table A.2). Nevertheless, our results are robust if we further include the departure airports in the top 50 and top 60 airports (see Appendix Table A.9 Panels A.2 and A.3) or if we extend the sample to 2018 (see Appendix Figure A.7).

**Additional controls and richer fixed effects.** Lastly, we consider additional potential omitted variables. Our results are robust if we control for weather conditions and weather-induced delays (see Appendix Table A.10), or if we control for the level of competition that a carrier experiences at a route level (see Appendix Table A.11) which may affect the quality of service (Mazzeo, 2003; Greenfield, 2014).[27] Our results are robust if we include additional destination fixed effects and aircraft model fixed effects (available upon request).

# 7 Conclusion

In this paper, we conducted a retrospective evaluation of a recent large-scale infrastructure investment called the Next Generation Air Transportation System (NextGen) which began to be implemented in 2014. The NextGen program has a similar scope and budget as the air transportation program proposed in the recent Infrastructure Investments and Jobs Act passed by Congress and signed into law by the Biden Administration. We focus on various performance measures of air travel time and delays.

Our results suggest that implementing NextGen has led to shorter air travel times and shorter delays. These improvements imply that implementing all four categories of NextGen programs will pay back the $20 billion investment in 7-8 years by generating private benefits from reduced air travel time for passengers and reduced fuel costs. The reduction in airborne time and taxi time would also lead to small-scale social benefits such as lower carbon emissions, subject to caveats such as the rebound effect depending on the elasticity. Our calculations on private gains are conservative in that we do not investigate the benefits from fewer cancellations (which requires more assumptions about costs for airlines and consumers) and other general equilibrium effects such as local labor market benefits and other forms of spillover effects to other industries. Our results imply that infrastructure programs of a similar nature and design can plausibly justify the budget and provide sizable benefits for both passengers and carriers.

[27] Table A.10 implies that more severe weather (lower visibility and sky ceiling) and a weather-delayed flight are associated with greater delay and air travel time. Table A.11 columns 4-8 imply a more concentrated route is associated with less time spent on the runway and shorter delays. Columns 1-3 are not inconsistent with columns 4-8. A more concentrated route is likely a longer route and therefore greater airborne time and overall air travel time.

Our study demonstrates one particularly valuable feature of NextGen - flights that experience exogenous shocks, which are likely the targeted recipients of NextGen, tend to receive more gains from NextGen than other flights (e.g., severe weather and greater prior delay). However, we do not find the same results for flights that are pre-distorted by market power. While we find flights by non-hub carriers and low-cost carriers received greater time savings than hub-carriers and legacy carriers in the first two years after the initial adoption, the gap shrinking effect disappeared by year 3 or was fully reversed. Our results caution us about potential unintended consequences of policies and technological upgrades on existing market distortions, which while intended to restore the efficiency of the public provision of modern infrastructure, may instead exacerbate existing market failures. It is important for future studies to look at the interactions of various forms of inefficiencies and how mitigating one would affect another.

# References


Borenstein, S. (1999). The dominant-firm advantage in multiproduct industries: Evidence from the U. S. Airlines. *The Quarterly Journal of Economics 106*(4), 1237–1266.

Borusyak, K., X. Jaravel, and J. Spiess (2022). Revisiting event study designs: Robust and efficient estimation.

Bratu, S. and C. Barnhart (2005). An analysis of passenger delays using flight operations and passenger booking data. *Air Traffic Control Quarterly 13*(1), 1–27.

Brueckner, J. (2002). Airport congestion when carriers have market power. *American Economic Review 92*(5), 1357–1375.

Brueckner, J. (2005). Internalization of airport congestion: A network analysis. *International Journal of Industrial Organization 23*(7-8), 599–614.

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

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. Technical Report 2. Themed Issue: Treatment Effect 1.

Daniel, J. (1995). Congestion pricing and capacity of large hub airports: A bottleneck model with stochastic queues. *Econometrica 63*(2), 327–70.

Daniel, J. I. (2001). Distributional consequences of airport congestion pricing. *Journal of Urban Economics 50*(2), 230 – 258.

de Chaisemartin, C. and X. D'Haultfœuille (2020, September). Two-way fixed effects estimators with heterogeneous treatment effects. Technical Report 9.

Dobkin, C., A. Finkelstein, R. Kluender, and M. J. Notowidigdo (2018, February). The economic consequences of hospital admissions. *American Economic Review 108*(2), 308–52.

Evler, J., E. Asadi, H. Preis, and H. Fricke (2021). Airline ground operations: Schedule recovery optimization approach with constrained resources. *Transportation Research Part C: Emerging Technologies 128*, 103129.

FAA (2011). Concept of operations for the next generation air transportation system. Version 3.2. Technical report, Joint Planning and Development Office, Next Generation Air Transportation System (NextGen).

FAA (2014). The economic impact of civil aviation on the US economy. Technical report, Air Traffic Organization, Federal Aviation Administration.

FAA (2015). NextGen implementation plan 2015. Technical report, Office of NextGen, Federal Aviation Administration.

FAA (2016a). Economic values for FAA investment and regulatory decisions, a guide. Technical report, FAA Office of Aviation Policy and Plans, Federal Aviation Administration.

FAA (2016b). NextGen implementation plan 2016. Technical report, Office of NextGen, Federal Aviation Administration.

FAA (2016c). Office of inspector general audit report: Total cost, schedules, and benefits of FAA's NextGen tranformational programs remain uncertain. Technical report, FAA Office of Aviation Policy and Plans, Federal Aviation Administration.

FAA (2016d). Update to the business case for the next generation air transportation system based on the future of the NAS report. Technical report, Office of NextGen, Federal Aviation Administration.

FAA (2018a). FAA aerospace forecassts fiscal year 2018–2038. Technical report, Federal Aviation Administration.

FAA (2018b). NextGen implementation plan 2018-2019. Technical report, Federal Aviation Administration.

FAA (2021). NextGen implementation plan 2019-2021. Technical report, Federal Aviation Administration.

FAA and Eurocontrol (2010). 2010 US Europe comparison of ATM-related operational performance. Technical report, Air Traffic Organization System Operations Service of Federal Aviation Administration and the European Organization of the Safety of Air Navigation (EUROCONTROL) Performance Review Commission.

Forbes, S. J. (2008). The effect of air traffic delays on airline prices. *International Journal of Industrial Organization 26*(5), 1218 – 1232.

Forbes, S. J., M. Lederman, and Z. Yuan (2019). Do airlines pad their schedules? *Review of Industrial Organization 54*, 61–82.

Freyaldenhoven, S., C. Hansen, and J. M. Shapiro (2019, September). Pre-event trends in the panel event-study design. *American Economic Review 109*(9), 3307–38.

GAO (2016). Next generation air transportation system: Information on expenditures, schedule, and cost estimates, fiscal years 2004–2030. Technical Report GAO-17-241R.

GAO (2017). Air traffic control modernization: Progress and challenges in implementing NextGen. Technical Report GAO-17-450, United States Government Accountability Office, Report to Congressional Requests.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Technical Report 2. Themed Issue: Treatment Effect 1.

Greenfield, D. (2014). Competition and service quality: New evidence from the airline industry. *Economics of Transportation 3*(1), 80–89. Special Issue on Airlines and Airports.

JPDO (2004). Next generation air transportation system integrated plan. 2004. Technical report.

JPDO (2011). The concept of operations for the next generation air transportation system (NextGen). version 3.2. Technical report.

Lakew, P. A. and V. Bilotkach (2017, March). Airport delays and metropolitan employment. *Journal of Regional Science 58*(2), 424–450.

Lam, C. and Y. C. Zhou (2021). Network propagation and air traffic congestion policies. Technical report, Working Paper.

Lee, S., Z. Liu, and H. Yu (2021). Heterogeneous price effects and increased price dispersion from quantity-based congestion management. *Economic Inquiry 59*(3), 1378–1402.

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2022). Repositioning and market power after airline mergers. *The RAND Journal of Economics 53*(1), 166–199.

Mayer, C. and T. Sinai (2003, September). Network effects, congestion externalities, and air traffic delays: Or why not all delays are evil. *American Economic Review 93*(4), 1194–1215.

Mazzeo, J. M. (2003). Competition and service quality in the U.S. Airline industry. *Review of Industrial Organization 22*, 275–296.

Morrison, S. A. and C. Winston (2007, December). Another look at airport congestion pricing. *American Economic Review 97*(5), 1970–1977.

Morrison, S. A. and C. Winston (2008). The effect of FAA expenditures on air travel delays. *Journal of Urban Economics 63*(2), 669 – 678.

Roth, J. (2022). Pre-test with caution: Event-study estimates after testing for parallel trends. Technical report, American Economic Review: Insight (forthcoming).

Rupp, N. G., G. M. Holmes, and J. DeSimone (2005). Airline schedule recovery after airport closures: Empirical evidence since September 11. *Southern Economic Journal 71*, 800–820.

Schimidheiny, K. and S. Siegloch (2020). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. Technical report, Working Paper.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199. Themed Issue: Treatment Effect 1.

Yan, J. and C. Winston (2014). Can private airport competition improve runway pricing? The case of San Francisco bay area airports. *Journal of Public Economics 115*, 146–157.

Zhang, A. and Y. Zhang (2003). Airport charges and capacity expansion: Effects of concessions and privatization. *Journal of Urban Economics 53*, 54–75.

Zhang, A. and Y. Zhang (2006). Airport capacity and congestion when carriers have market power. *Journal of Urban Economics 30*, 229–247.

# Figures

Figure 1: **NextGen and Control Airports**

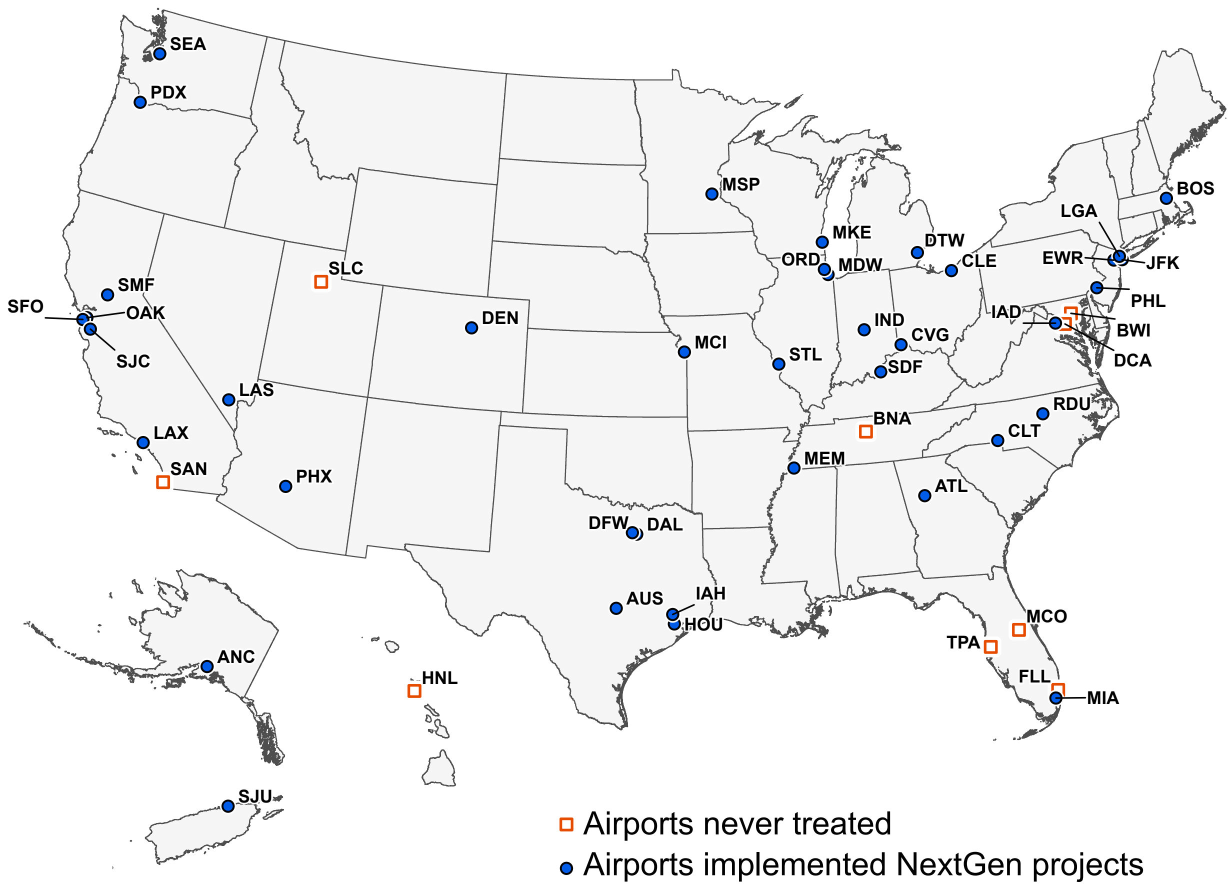
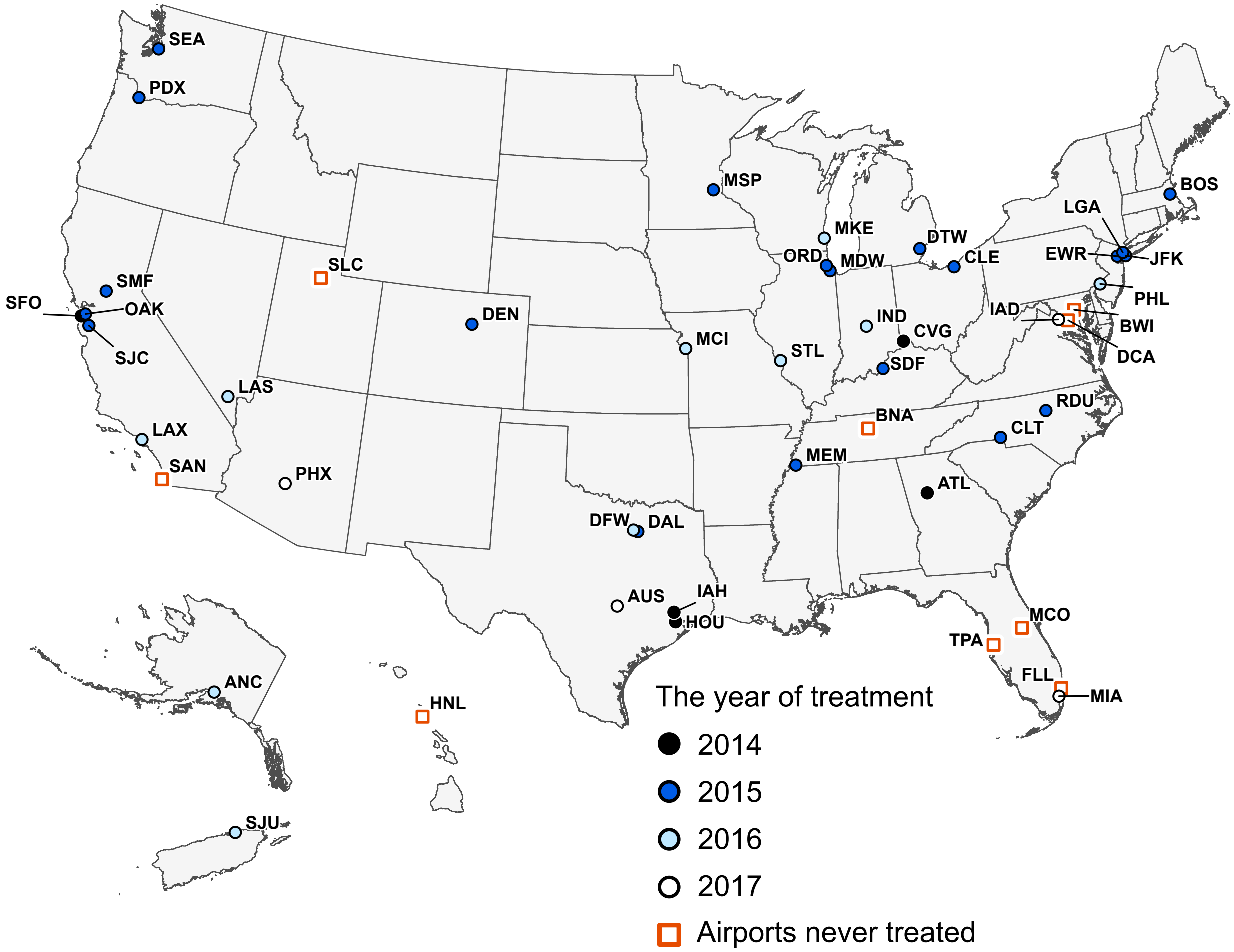


Figure 2: **Airports by the Initial Treatment Year**




*Notes*: The initial treatment year is the year when the first NextGen project was completed.

Figure 3: **Air Travel Performances, 2010–2017**

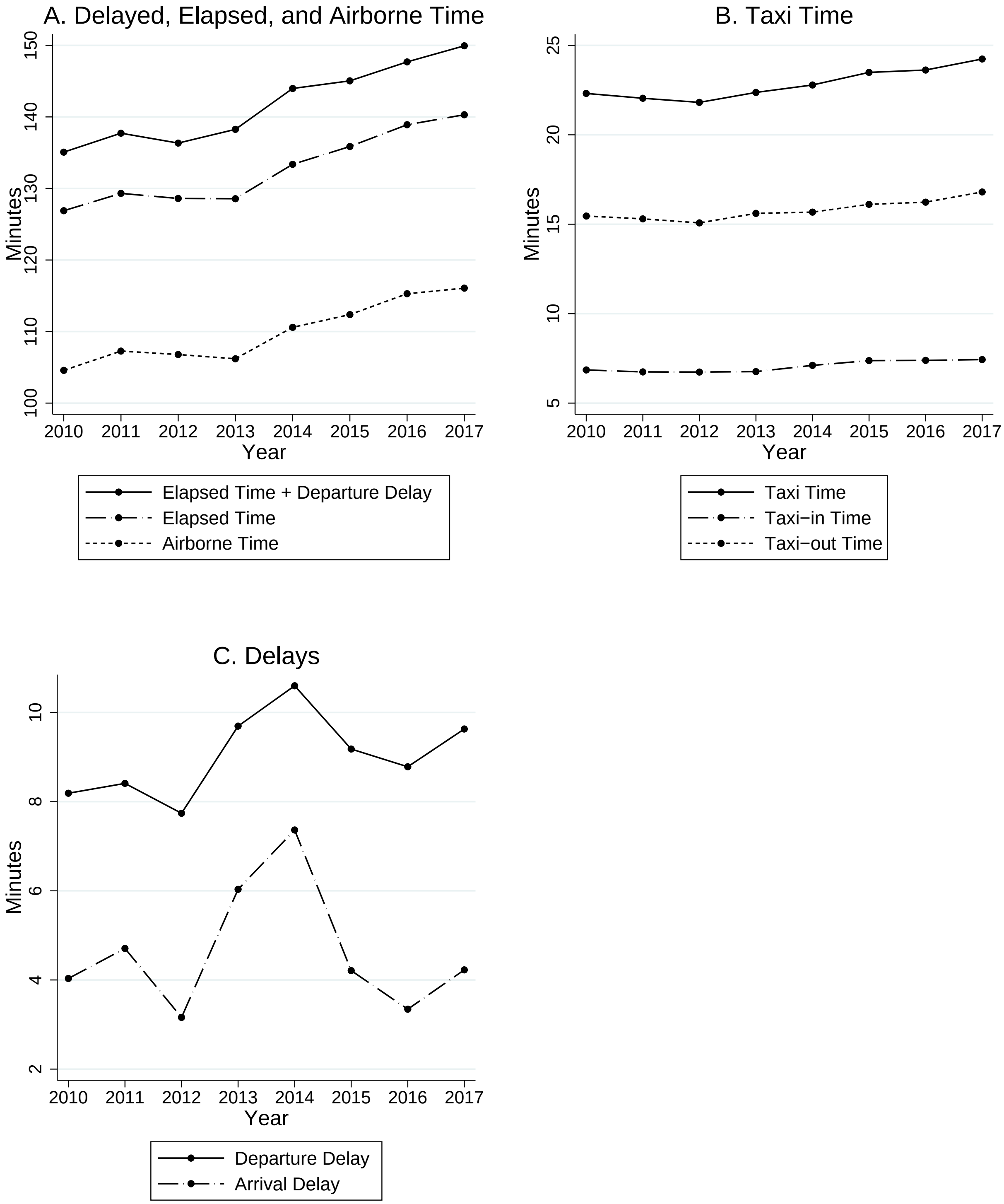
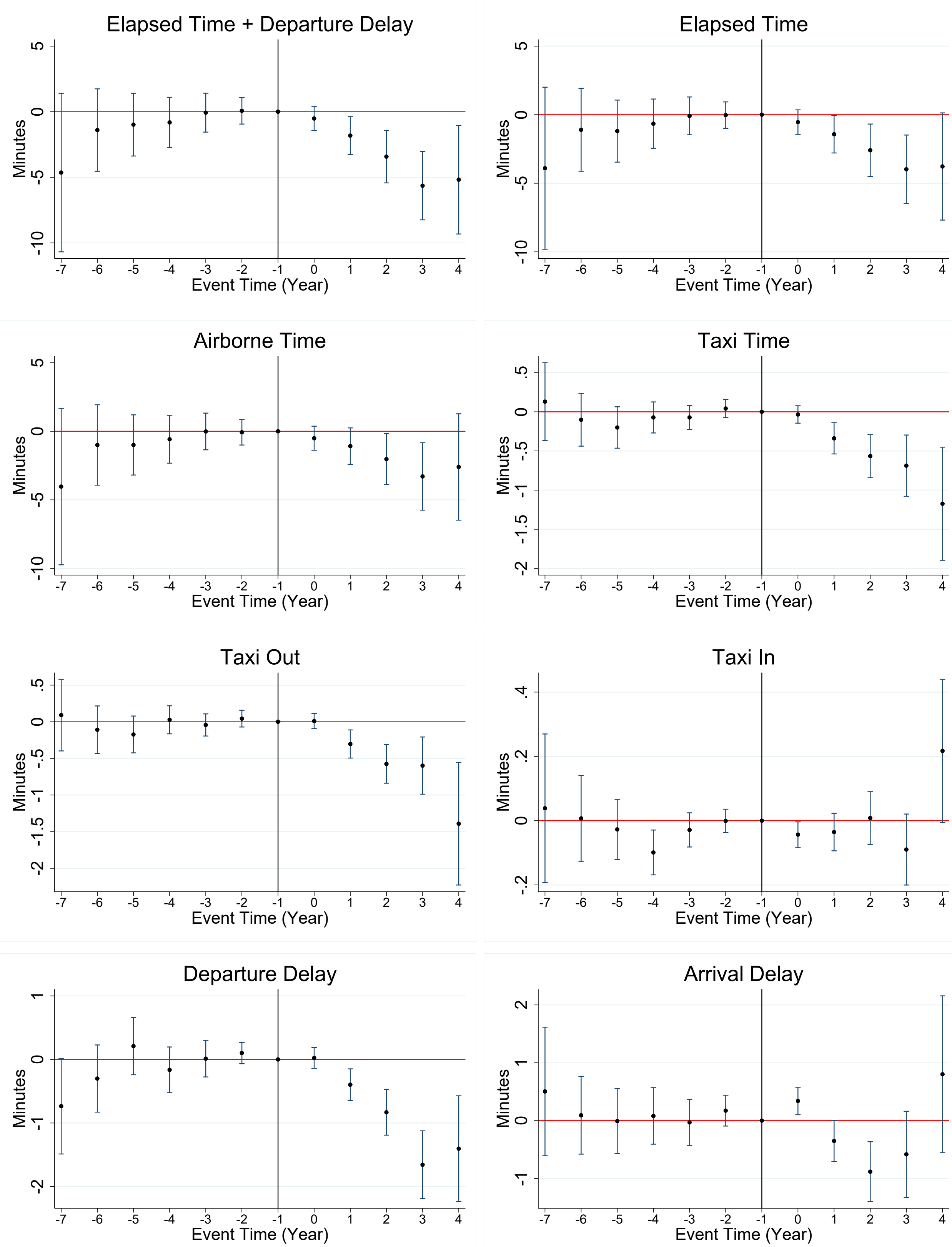


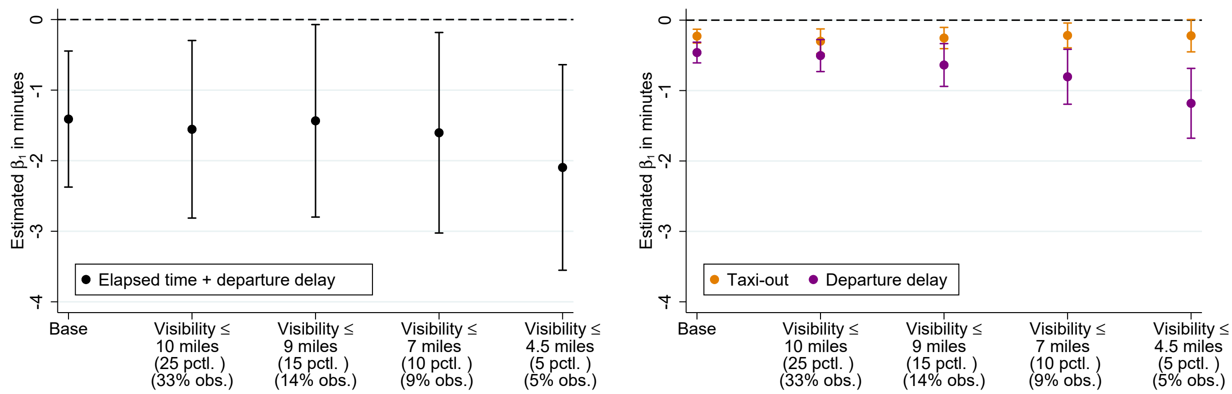
Figure 4: **Event Study for the Initial Treatment at the Origin Airport**



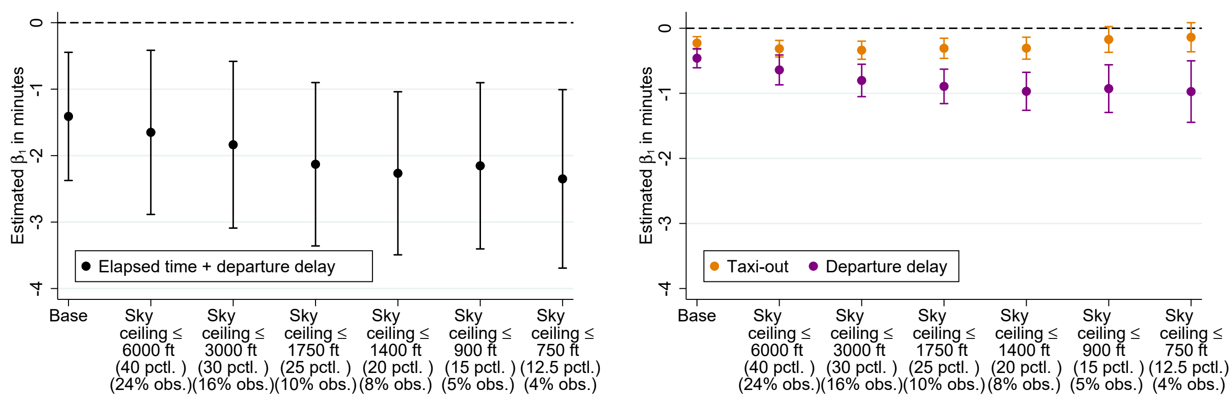

*Notes*: Results are from estimating equation (1) for each eight air travel performance variables.

Figure 5: **Conditional Estimate**

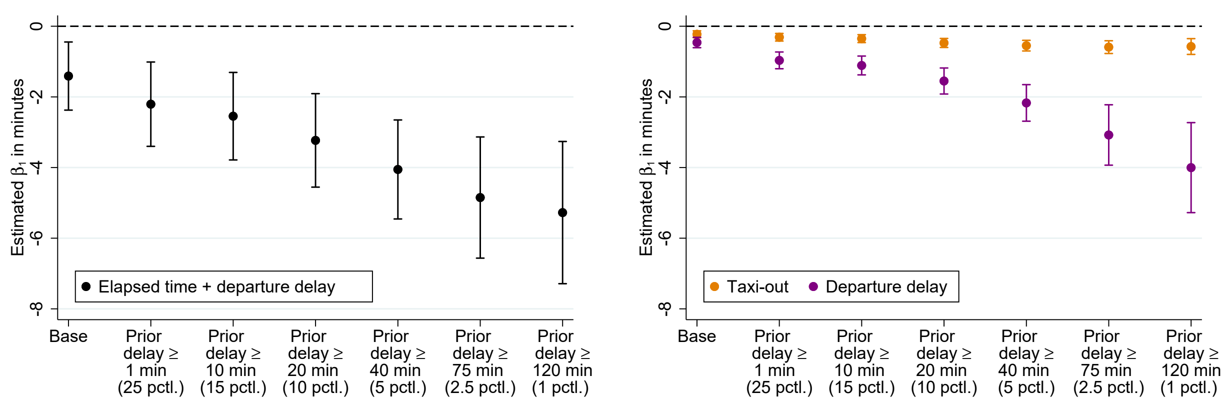
Panel A. $\hat{\beta}_1$ by weather condition (miles of visibility)




Panel B. $\hat{\beta}_1$ by weather condition (feet of sky ceiling)




Panel C. $\hat{\beta}_1$ by minutes of prior delay




*Notes*: Results are from estimating equation (1) in subsamples of severe weather conditions and prior delay. We only plot the first-order effect $\hat{\beta}_1$ and its 95% confidence interval as the second-order effect is small. Full results are in Appendix Table A.6-A.8. In panels A and B, we provide both the associated percentile of visibility and sky ceiling at the airport by hour level using NOAA data (excluding the range from 1 am to 5 am), as well as the percentile at the flight level in the main data. In Panel C, we report the associated percentile of prior delay at the flight level.

# Tables

Table 1: **Summary statistics of air travel time 2010–2017**

| Variable | Mean | SD. | Min. | Max. |
|---|---|---|---|---|
| Panel A. Air Travel Performance and Treatment Information | | | | |
| Actual elapsed time + departure delay (minutes) | 162.6 | 87.0 | 11 | 2,220 |
| Actual elapsed time (minutes) | 153.5 | 78.8 | 23 | 784 |
| Actual airborne time (minutes) | 129.4 | 76.9 | 8 | 723 |
| Actual taxi time (minutes) | 24.1 | 10.8 | 2 | 481 |
| Actual taxi-out time (minutes) | 17.0 | 9.3 | 0 | 278 |
| Actual taxi-in time (minutes) | 7.2 | 5.1 | 0 | 414 |
| Departure delay (minutes) | 9.1 | 35.2 | -45 | 1,863 |
| Arrival delay (minutes) | 4.1 | 37.9 | -108 | 1,845 |
| 1 = Travel from a NextGen (eventually treated) airport | 0.84 | 0.37 | 0 | 1 |
| Panel B. Panel Information | | | | |
| Number of origin airports | | | | 48 |
| Number of NextGen (eventually treated) airports | | | | 39 |
| Number of destination airports | | | | 75 |
| Number of carriers | | | | 18 |
| Number of origin airport by carrier by time-block dummies | | | | 61,159 |
| Number of observations | | | | 22,355,119 |

Table 2: **The effect of NextGen on air travel time**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | base measure | components of the base measure in (1) | | | | | | additional measure |
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.410*** | -0.950** | -0.726 | -0.224*** | -0.227*** | 0.004 | -0.460*** | -0.908*** |
| | (0.492) | (0.471) | (0.462) | (0.053) | (0.050) | (0.018) | (0.075) | (0.109) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.120 | 0.084 | 0.029 | 0.054*** | 0.062*** | -0.008 | 0.037** | 0.195*** |
| | (0.129) | (0.123) | (0.122) | (0.019) | (0.019) | (0.005) | (0.018) | (0.030) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.28 | 0.25 |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption (in minutes):* | | | | | | | | |
| *Percentage relative to the base measure (in %):* | | | | | | | | |
| Year 1 | -1.29 | -0.866 | -0.697 | -0.170 | -0.165 | -0.004 | -0.423 | -0.713 |
| | | | 54.0% | | 12.8% | 0.3% | 32.8% | |
| Year 2 | -2.34 | -1.564 | -1.336 | -0.232 | -0.206 | -0.024 | -0.772 | -1.036 |
| | | | 57.0% | | 8.8% | 0.1% | 33.0% | |
| Year 3 | -3.15 | -2.094 | -1.917 | -0.186 | -0.123 | -0.060 | -1.047 | -0.969 |
| | | | 60.1% | | 3.9% | 0.2% | 33.2% | |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table reports the estimated results of equation (2). All regressions include fixed effects of origin airport, carrier, and time block (measured by the day of a week and the hour of a day). All regressions also include origin airport fixed effects interacted with a linear year trend, year-by-quarter fixed effects, and a set of control variables. We measure treatment as $n_{oq}$, the total categories of NextGen (MRO, PBN, SO, and DC) implemented at the origin airport.

Table 3: **Effect of Specific Category of Projects**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| $d_{oq}$ (MRO) $\cdot$ | -0.961** | -0.605 | -0.479 | -0.126** | -0.132** | 0.006 | -0.356*** | 0.001 |
| $s \cdot \mathbf{1}(s > 0)$ | (0.396) | (0.378) | (0.369) | (0.064) | (0.066) | (0.018) | (0.070) | (0.112) |
| $d_{oq}$ (PBN) $\cdot$ | 0.227 | 0.350 | 0.190 | 0.161*** | 0.226*** | -0.065*** | -0.124 | -0.190 |
| $s \cdot \mathbf{1}(s > 0)$ | (0.482) | (0.464) | (0.457) | (0.051) | (0.050) | (0.021) | (0.095) | (0.117) |
| $d_{oq}$ (SO) $\cdot$ | -2.450*** | -2.054*** | -1.932*** | -0.122** | -0.101* | -0.021 | -0.396*** | -0.518*** |
| $s \cdot \mathbf{1}(s > 0)$ | (0.516) | (0.498) | (0.494) | (0.054) | (0.053) | (0.023) | (0.096) | (0.129) |
| $d_{oq}$ (DC) $\cdot$ | 0.220 | 0.829 | 0.887 | -0.058 | -0.035 | -0.023 | -0.609*** | -1.002*** |
| $s \cdot \mathbf{1}(s > 0)$ | (0.631) | (0.626) | (0.605) | (0.066) | (0.061) | (0.021) | (0.112) | (0.124) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.28 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 with specific categories of NextGen project $d_{oq}$ implemented at the origin airport, without estimating the quadratic effect.

Table 4: **Effect of NextGen by Slot-administration Status**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| | | | | A. Depart from a slot-administrated airport | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.398*** | -1.611* | -1.226 | -0.385*** | -0.448*** | 0.062** | -0.787*** | -1.162*** |
| | (0.868) | (0.846) | (0.833) | (0.088) | (0.084) | (0.026) | (0.125) | (0.187) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.419* | 0.288 | 0.157 | 0.131*** | 0.153*** | -0.022*** | 0.131*** | 0.265*** |
| | (0.214) | (0.209) | (0.206) | (0.027) | (0.025) | (0.007) | (0.026) | (0.044) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.979 | -1.323 | -1.069 | -0.254 | -0.295 | 0.04 | -0.656 | -0.897 |
| Year 2 | -3.12 | -2.07 | -1.824 | -0.246 | -0.284 | 0.036 | -1.05 | -1.264 |
| Year 3 | -3.423 | -2.241 | -2.265 | 0.024 | 0.033 | -0.012 | -1.182 | -1.101 |
| | | | | B. Depart from a non-slot-administrated airport | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -0.461 | -0.164 | -0.190 | 0.026 | 0.063 | -0.037 | -0.296*** | -0.812*** |
| | (0.571) | (0.538) | (0.525) | (0.054) | (0.046) | (0.024) | (0.094) | (0.139) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | -0.165 | -0.152 | -0.097 | -0.056*** | -0.058*** | 0.003 | -0.012 | 0.167*** |
| | (0.172) | (0.160) | (0.156) | (0.019) | (0.018) | (0.008) | (0.030) | (0.047) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -0.626 | -0.316 | -0.287 | -0.03 | 0.005 | -0.034 | -0.308 | -0.645 |
| Year 2 | -1.582 | -0.936 | -0.768 | -0.172 | -0.106 | -0.062 | -0.64 | -0.956 |
| Year 3 | -2.868 | -1.86 | -1.443 | -0.426 | -0.333 | -0.084 | -0.996 | -0.933 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 and restricts the treated airports to either slot-administrated airports or non-slot-administrated airports, with the same control group.

Table 5: **Effect of NextGen by Hub Status**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | | | | A. Depart from a hub airport | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.310** | -0.740 | -0.513 | -0.227*** | -0.241*** | 0.014 | -0.569*** | -0.998*** |
| | (0.578) | (0.554) | (0.541) | (0.064) | (0.061) | (0.021) | (0.089) | (0.132) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.125 | 0.065 | 0.012 | 0.053** | 0.066*** | -0.013** | 0.060*** | 0.218*** |
| | (0.152) | (0.146) | (0.143) | (0.022) | (0.022) | (0.006) | (0.021) | (0.035) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.185 | -0.675 | -0.501 | -0.174 | -0.175 | 0.001 | -0.509 | -0.78 |
| Year 2 | -2.12 | -1.22 | -0.978 | -0.242 | -0.218 | -0.024 | -0.898 | -1.124 |
| Year 3 | -2.805 | -1.635 | -1.431 | -0.204 | -0.129 | -0.075 | -1.167 | -1.032 |
| | | | | B. Depart from a non-hub airport | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -0.789 | -0.445 | -0.382 | -0.063 | -0.109** | 0.046 | -0.344** | -0.546*** |
| | (0.874) | (0.815) | (0.807) | (0.061) | (0.054) | (0.029) | (0.142) | (0.170) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | -0.014 | -0.063 | -0.072 | 0.009 | 0.025 | -0.017** | 0.049 | 0.111** |
| | (0.252) | (0.233) | (0.232) | (0.022) | (0.019) | (0.008) | (0.041) | (0.053) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -0.803 | -0.508 | -0.454 | -0.054 | -0.084 | 0.029 | -0.295 | -0.435 |
| Year 2 | -1.634 | -1.142 | -1.052 | -0.09 | -0.118 | 0.024 | -0.492 | -0.648 |
| Year 3 | -2.493 | -1.902 | -1.794 | -0.108 | -0.102 | -0.015 | -0.591 | -0.639 |
| | | | | C. Depart from a hub airport and operated by a hub carrier | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.159 | -0.741 | -0.417 | -0.324*** | -0.341*** | 0.018 | -0.419*** | -1.110*** |
| | (0.893) | (0.865) | (0.848) | (0.085) | (0.076) | (0.031) | (0.126) | (0.176) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.071 | 0.076 | 0.013 | 0.063** | 0.084*** | -0.021** | -0.005 | 0.246*** |
| | (0.230) | (0.224) | (0.220) | (0.027) | (0.025) | (0.009) | (0.032) | (0.047) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.088 | -0.665 | -0.404 | -0.261 | -0.257 | -0.003 | -0.424 | -0.864 |
| Year 2 | -2.034 | -1.178 | -0.782 | -0.396 | -0.346 | -0.048 | -0.858 | -1.236 |
| Year 3 | -2.838 | -1.539 | -1.134 | -0.405 | -0.267 | -0.135 | -1.302 | -1.116 |
| | | | | D. Depart from a hub airport and operated by a non-hub carrier | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.504*** | -0.709 | -0.568 | -0.141* | -0.123* | -0.017 | -0.795*** | -0.969*** |
| | (0.578) | (0.564) | (0.559) | (0.072) | (0.063) | (0.027) | (0.089) | (0.128) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.203 | 0.050 | -0.010 | 0.060*** | 0.053*** | 0.007 | 0.153*** | 0.210*** |
| | (0.163) | (0.157) | (0.157) | (0.022) | (0.020) | (0.007) | (0.023) | (0.036) |
| *Marginal effect of $\Delta n_{oq} = 1$ in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.301 | -0.659 | -0.578 | -0.081 | -0.07 | -0.01 | -0.642 | -0.759 |
| Year 2 | -2.196 | -1.218 | -1.176 | -0.042 | -0.034 | -0.006 | -0.978 | -1.098 |
| Year 3 | -2.685 | -1.677 | -1.794 | 0.117 | 0.108 | 0.012 | -1.008 | -1.017 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 and restricts the treated group to flights that depart from a hub, a non-hub airport, flights operated by a hub airline or a non-hub airline from a hub airport, while keeping the same control group.

Table 6: **Effect of NextGen between Legacy and Low-cost Carriers**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| | | | | A. Flights provided by a legacy carrier | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.076 | -0.725 | -0.492 | -0.233*** | -0.297*** | 0.064** | -0.352*** | -0.845*** |
| | (0.798) | (0.773) | (0.762) | (0.075) | (0.069) | (0.030) | (0.110) | (0.159) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | -0.018 | -0.002 | -0.031 | 0.028 | 0.064** | -0.035*** | -0.015 | 0.176*** |
| | (0.209) | (0.202) | (0.201) | (0.025) | (0.025) | (0.008) | (0.029) | (0.045) |
| *Marginal effect of* $\Delta n_{oq} = 1$ *in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.094 | -0.727 | -0.523 | -0.205 | -0.233 | 0.029 | -0.367 | -0.669 |
| Year 2 | -2.224 | -1.458 | -1.108 | -0.354 | -0.338 | -0.012 | -0.764 | -0.986 |
| Year 3 | -3.39 | -2.193 | -1.755 | -0.447 | -0.315 | -0.123 | -1.191 | -0.951 |
| | | | | B. Flights provided by a low-cost carrier (LCC) | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.630*** | -0.998* | -0.825 | -0.173*** | -0.133*** | -0.040** | -0.632*** | -0.936*** |
| | (0.546) | (0.519) | (0.507) | (0.049) | (0.044) | (0.019) | (0.093) | (0.110) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.228 | 0.117 | 0.054 | 0.063*** | 0.050*** | 0.013** | 0.111*** | 0.205*** |
| | (0.145) | (0.137) | (0.134) | (0.017) | (0.016) | (0.005) | (0.024) | (0.030) |
| *Marginal effect of* $\Delta n_{oq} = 1$ *in variable years after the initial adoption:* | | | | | | | | |
| Year 1 | -1.402 | -0.881 | -0.771 | -0.11 | -0.083 | -0.027 | -0.521 | -0.731 |
| Year 2 | -2.348 | -1.528 | -1.434 | -0.094 | -0.066 | -0.028 | -0.82 | -1.052 |
| Year 3 | -2.838 | -1.941 | -1.989 | 0.048 | 0.051 | -0.003 | -0.897 | -0.963 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 and restricts the carriers to either legacy or low-cost carriers.

Table 7: **Benefits of NextGen in 2017 from Air Travel Time Savings**

Panel A. Air Travel Time Savings

| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|---|
| | | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| A. | Actual (minutes) | 164.41 | 153.99 | 129.64 | 24.34 | 17.11 | 7.23 | 10.42 | 4.31 |
| B. | Counterfacutal if never treated (minutes) | 168.58 | 156.77 | 132.14 | 24.63 | 17.33 | 7.31 | 11.81 | 5.76 |
| C. | $\Delta y$ from B to A (minutes) | -4.18 | -2.79 | -2.50 | -0.29 | -0.22 | -0.07 | -1.39 | -1.44 |
| D. | $\Delta y$% from B to A (percent) | -2.5% | -1.8% | -1.9% | -1.2% | -1.3% | -1.0% | -13.3% | -33.5% |
| Number of observation | | | | | | | | | 2,523,571 |

Panel B. Private and Social Benefits

| | Per-flight savings (2017 USD) | Total savings (Billions) | Percentage |
|---|---|---|---|
| B.1 Baseline Private Gains | | | |
| Fuel cost savings | \$60.0 | \$B 0.15 | 16.2% |
| Passenger time savings | \$307.6 | \$B 0.78 | 83.7% |
| Total private benefits | \$367.6 | \$B 0.93 | |
| B.2 Private Gains with Additional Considerations | | | |
| a. Total + flights not matched to Form-B43 | | \$B 1.26 | |
| b. Total + delay multipliers by FAA | | \$B 1.39 | |
| c. Total + crew time savings | | \$B 1.11 | |
| B.3 Social Benefits | | | |
| Social benefits of carbon reductions | \$15.3 | \$B 0.04 | |

*Notes*: Panel A compares air travel time and delays for flights that depart from a treated/NextGen airport in 2017 between two scenarios: (i) the actual air travel time in 2017 in Row A, and (ii) the counterfactual scenario if all NextGen airports had been treated in 2014 in Row B based on Table 2. Panel B reports the private and social benefits based on Panel A Row C. For fuel cost savings, we use per-minute fuel use estimates using AEDT data (see Appendix Table A.12) and the jet fuel price from the EIA. For passenger time savings, we evaluate cost per passenger (personal or business combined) at \$48.71 per hour following FAA guidelines (FAA, 2016a) and the number of passengers per flight from DOT's report in `https://www.bts.gov/newsroom/2017-traffic-data-us-airlines-and-foreign-airlines-us-flights`. For carbon benefits, we use \$40 per-ton social cost of carbon and the $CO_2$ estimates from AEDT data (see Appendix Table A.12). For plausible additional private games, we use a delay multiplier and crew cost estimates from the FAA guidelines.

Table 8: **Benefits of Fully Implementing NextGen Projects in 2017**

Panel A. Air Travel Time Savings

| | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| A. Actual (minutes) | 164.41 | 153.99 | 129.64 | 24.34 | 17.11 | 7.23 | 10.42 | 4.31 |
| B. Counterfacutal if never treated (minutes) | 168.58 | 156.77 | 132.14 | 24.63 | 17.33 | 7.31 | 11.81 | 5.76 |
| C. Counterfactual if fully treated in 2014 (minutes) | 155.99 | 148.38 | 124.48 | 23.91 | 16.85 | 7.06 | 7.60 | 1.89 |
| D. $\Delta y$ from B to C (minutes) | -12.59 | -8.39 | -7.66 | -0.73 | -0.48 | -0.24 | -4.20 | -3.87 |
| E. $\Delta y\%$ from B to C (percent) | -7.5% | -5.4% | -5.8% | -2.9% | -2.8% | -3.3% | -35.6% | -67.2% |
| Number of observation | | | | | | | | 2,523,571 |

Panel B. Private and Social Benefits

| | Per-flight savings (2017 USD) | Total savings (Billions) | Percentage |
|---|---|---|---|
| B.1 Baseline Private Gains | | | |
| Fuel cost savings | \$191.1 | \$B 0.48 | 17.1% |
| Passenger time savings | \$927.6 | \$B 2.34 | 82.9% |
| Total private benefits | \$1,118.7 | \$B 2.82 | |
| B.2 Private Gains with Additional Considerations | | | |
| a. Total + flights not matched to Form-B43 | | \$B 3.84 | |
| b. Total + delay multipliers by FAA | | \$B 4.23 | |
| c. Total + crew time savings | | \$B 3.37 | |
| B.3 Social Benefits | | | |
| Social benefits of carbon reductions | \$46.5 | \$B 0.12 | |

*Notes*: Panel A compares air travel time and delays for flights that depart from a NextGen airport in 2017 between two scenarios: (i) the scenario if all NextGen airports were fully treated with all technologies in 2014 in Row C, and (ii) the senario if the NextGen airports were never treated in Row B. Panel B reports the private and social benefits based on Panel A Row D. We use the same fuel use and cost estimates as in Table 7 Panel B.

Table 9: **Robustness: Early Adopters vs. Later Adopters**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | A. Effect for Earlier Adopters (Airports First Treated in 2014 or 2015) | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.598$^{***}$ | -0.993$^{**}$ | -0.834$^{*}$ | -0.159$^{***}$ | -0.156$^{***}$ | -0.003 | -0.605$^{***}$ | -1.146$^{***}$ |
| | (0.505) | (0.481) | (0.469) | (0.058) | (0.054) | (0.020) | (0.092) | (0.126) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.077 | 0.022 | -0.012 | 0.034 | 0.044$^{**}$ | -0.010$^{*}$ | 0.055$^{**}$ | 0.223$^{***}$ |
| | (0.145) | (0.139) | (0.137) | (0.021) | (0.021) | (0.006) | (0.021) | (0.033) |
| Num. of obs. | 16,923,714 | 16,923,714 | 16,923,714 | 16,923,714 | 16,923,714 | 16,923,714 | 16,923,714 | 16,923,714 |
| R-squared | 0.41 | 0.45 | 0.44 | 0.21 | 0.22 | 0.12 | 0.29 | 0.25 |
| | B. Effect for Later Adopters (Airports First Treated in 2016 or 2017) | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.159 | -0.773 | -0.569 | -0.204 | -0.081 | -0.123$^{*}$ | -0.387 | -0.592 |
| | (1.491) | (1.461) | (1.448) | (0.155) | (0.144) | (0.072) | (0.326) | (0.394) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | -0.143 | -0.172 | -0.166 | -0.007 | -0.109 | 0.102$^{**}$ | 0.029 | -0.092 |
| | (0.758) | (0.757) | (0.756) | (0.088) | (0.076) | (0.040) | (0.174) | (0.212) |
| Num. of obs. | 9,065,363 | 9,065,363 | 9,065,363 | 9,065,363 | 9,065,363 | 9,065,363 | 9,065,363 | 9,065,363 |
| R-squared | 0.46 | 0.50 | 0.50 | 0.19 | 0.19 | 0.13 | 0.29 | 0.26 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. $^{*}$, $^{**}$, and $^{***}$ indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats Table 2 and excludes later adopters that were initially treated in 2016 or 2017. Panel B repeats baseline and excludes earlier adopters that were first treated in 2014 or 2015.

Table 10: **Robustness: Examine the Size of Treatment Group Airports**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | A. Exclude 2017 Q3-4 (Move from Treatment to Control Group: AUS and IAD) | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.746*** | -1.186** | -0.904* | -0.283*** | -0.282*** | -0.001 | -0.560*** | -0.913*** |
| | (0.572) | (0.537) | (0.525) | (0.061) | (0.059) | (0.019) | (0.091) | (0.123) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.168 | 0.103 | 0.034 | 0.069*** | 0.071*** | -0.003 | 0.065** | 0.193*** |
| | (0.172) | (0.160) | (0.157) | (0.023) | (0.022) | (0.006) | (0.031) | (0.039) |
| Num. of obs. | 20,762,875 | 20,762,875 | 20,762,875 | 20,762,875 | 20,762,875 | 20,762,875 | 20,762,875 | 20,762,875 |
| R-squared | 0.42 | 0.46 | 0.45 | 0.21 | 0.22 | 0.13 | 0.29 | 0.25 |
| | B. Non-top 15 versus Top 15 Airports | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.591*** | -1.039* | -0.809 | -0.230*** | -0.220*** | -0.010 | -0.552*** | -1.057*** |
| | (0.587) | (0.568) | (0.555) | (0.065) | (0.061) | (0.021) | (0.077) | (0.122) |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ · non-top 15 | 0.518 | 0.336 | 0.261 | 0.076 | 0.043 | 0.033 | 0.181 | 0.487*** |
| | (0.939) | (0.900) | (0.887) | (0.118) | (0.112) | (0.029) | (0.131) | (0.176) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.160 | 0.109 | 0.049 | 0.060*** | 0.066*** | -0.005 | 0.051*** | 0.233*** |
| | (0.153) | (0.148) | (0.145) | (0.022) | (0.021) | (0.006) | (0.019) | (0.034) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ · non-top 15 | -0.079 | -0.097 | -0.053 | -0.044 | -0.042 | -0.002 | 0.018 | -0.110* |
| | (0.272) | (0.257) | (0.255) | (0.040) | (0.039) | (0.009) | (0.040) | (0.057) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.28 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats Table 2 and excludes 2017 quaters 3 and 4. Doing so moves the following airports from the NextGen airports to the control group: Austin (AUS) and Dullas Washington D.C. (IAD). Panel B repeats the baseline and interacts the main variables with a dummy variable that represents airports outside top 15.

Table 11: **Robustness: Consolidated Carriers**

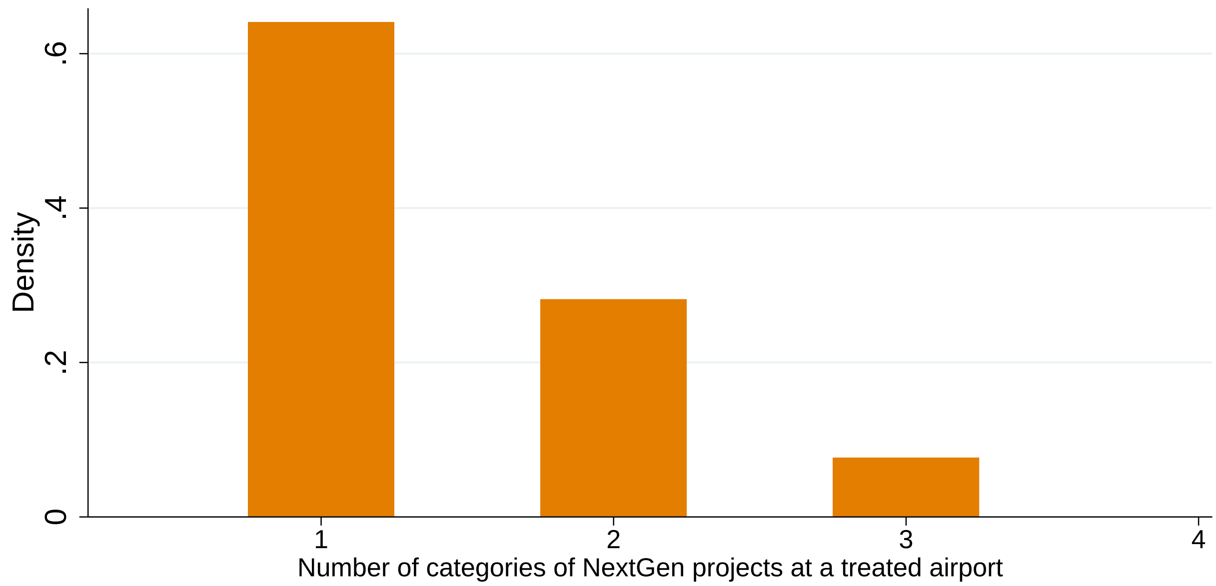
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| A. Subsample without 3 Major Consolidated Carriers (CO, FL, and US) | | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.395*** | -0.940** | -0.719 | -0.221*** | -0.226*** | 0.005 | -0.455*** | -0.901*** |
| | (0.491) | (0.471) | (0.461) | (0.053) | (0.050) | (0.018) | (0.074) | (0.109) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.116 | 0.081 | 0.027 | 0.054*** | 0.062*** | -0.008 | 0.035* | 0.193*** |
| | (0.129) | (0.123) | (0.122) | (0.019) | (0.019) | (0.005) | (0.018) | (0.030) |
| Num. of obs. | 19,875,638 | 19,875,638 | 19,875,638 | 19,875,638 | 19,875,638 | 19,875,638 | 19,875,638 | 19,875,638 |
| R-squared | 0.42 | 0.46 | 0.45 | 0.22 | 0.23 | 0.12 | 0.28 | 0.25 |
| B. Subsample with only 6 Major Merged Carriers (CO and UA; FL and WN; US and AA) | | | | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.274* | -1.281* | -0.969 | -0.312*** | -0.314*** | 0.003 | 0.007 | -0.233* |
| | (0.701) | (0.674) | (0.671) | (0.059) | (0.057) | (0.019) | (0.111) | (0.134) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.030 | 0.104 | 0.012 | 0.092*** | 0.103*** | -0.012** | -0.074*** | 0.035 |
| | (0.181) | (0.174) | (0.174) | (0.020) | (0.019) | (0.005) | (0.028) | (0.035) |
| Num. of obs. | 10,580,537 | 10,580,537 | 10,580,537 | 10,580,537 | 10,580,537 | 10,580,537 | 10,580,537 | 10,580,537 |
| R-squared | 0.36 | 0.38 | 0.37 | 0.23 | 0.23 | 0.12 | 0.30 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats Table 2 and excludes major carriers that were consolidated during the sample period: Continental Airlines (CO), Airtran Airways (FL), and US Airway (US). Panel B only includes major merged carriers: Continental Airline (CO) and United Airline (UA); Airtran (FL) and Southwest (WN); and US Airway (US) and American Airlines (AA).

# A Appendix Figures and Tables

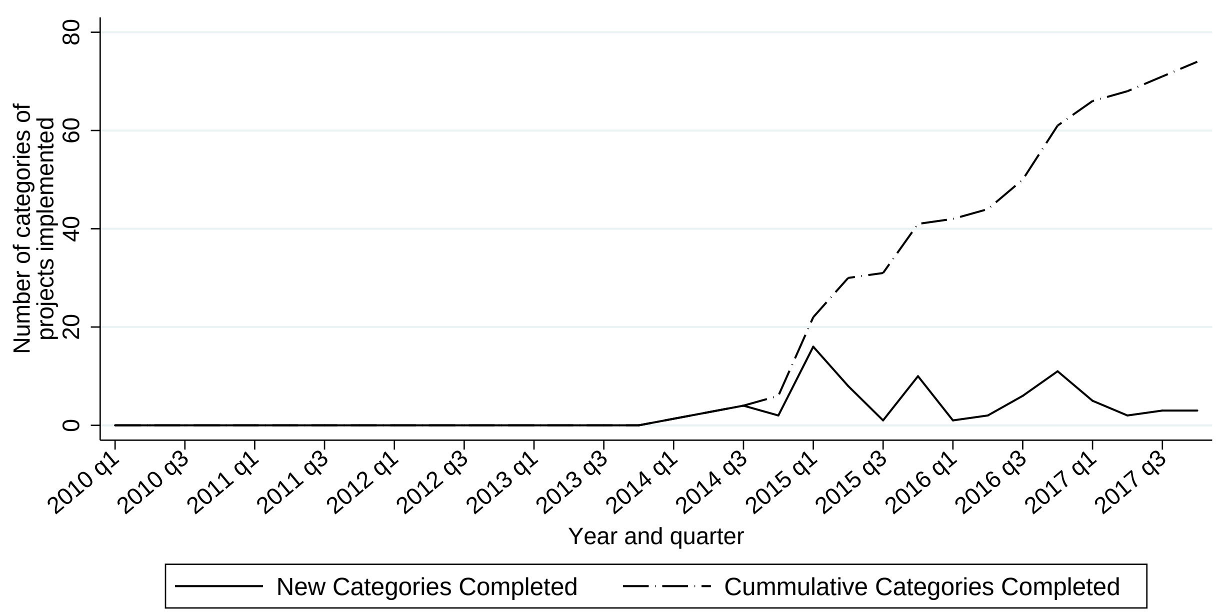
Figure A.1: **Destination Airports in the Sample**

Figure A.2: **Numbers of treatment at an airport, 2010–2017**



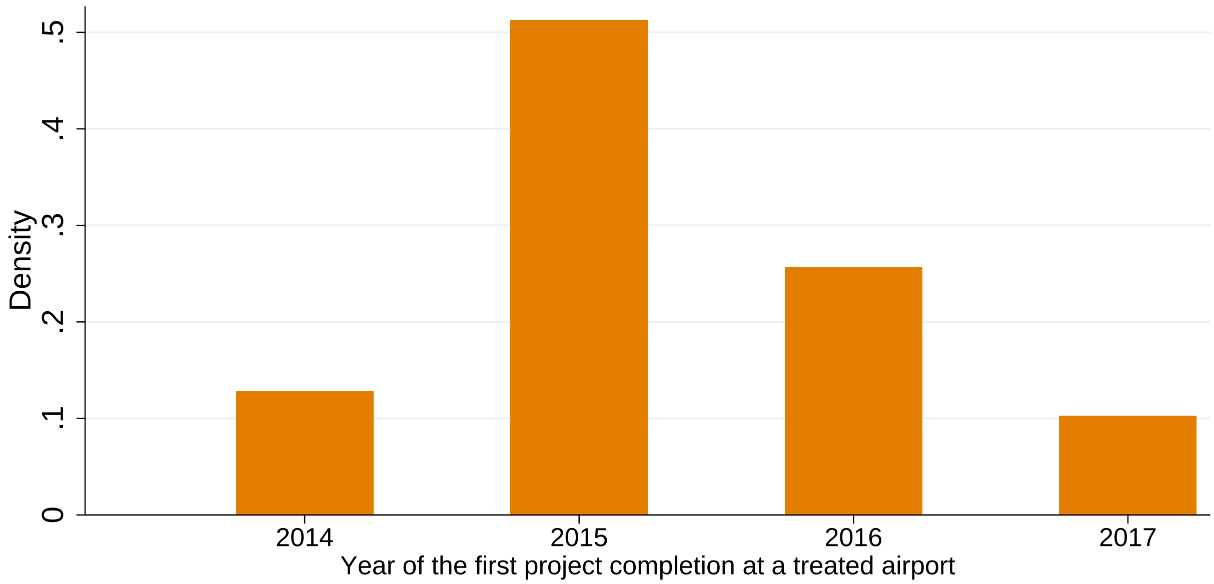

*Notes*: We tabulate and plot the number of categories of project (i.e., MRO, PBN, SO, and DC) completed at an airport from 2010 to 2017. 14 out of 39 treated airports (36%) completed more than one category of upgrade; no airport completed all four categories of treatment. The average number of treatments an airport receives is 1.4.

Figure A.3: **Categories of NextGen Projects Completed, 2010–2017**




*Notes*: We count the number of categories of NextGen projects completed at each airport in each quarter, and aggregate them at the quarterly level.

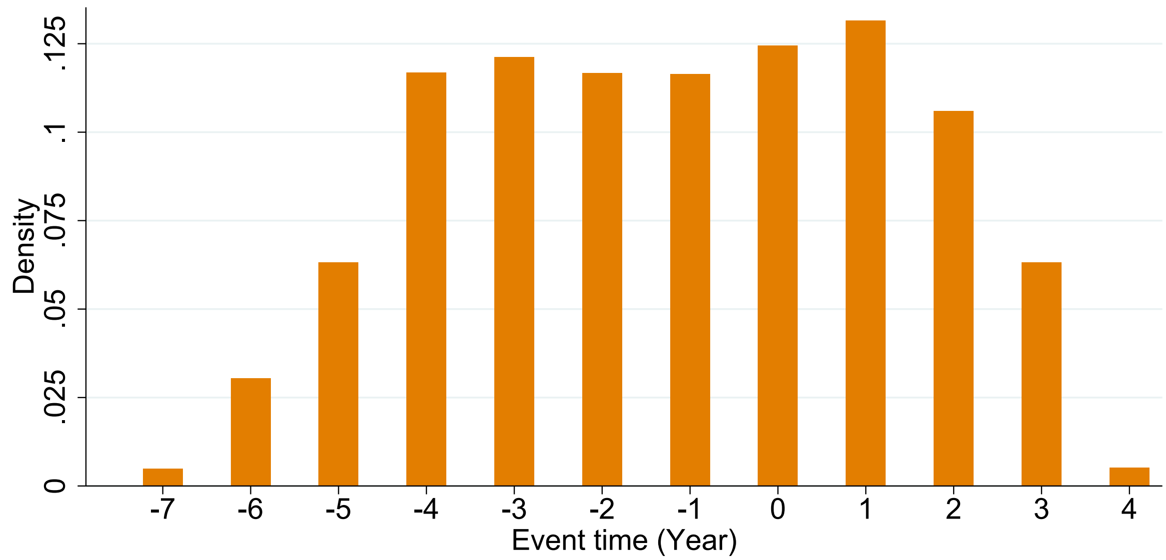
Figure A.4: **Year of the first treatment at the origin airport**



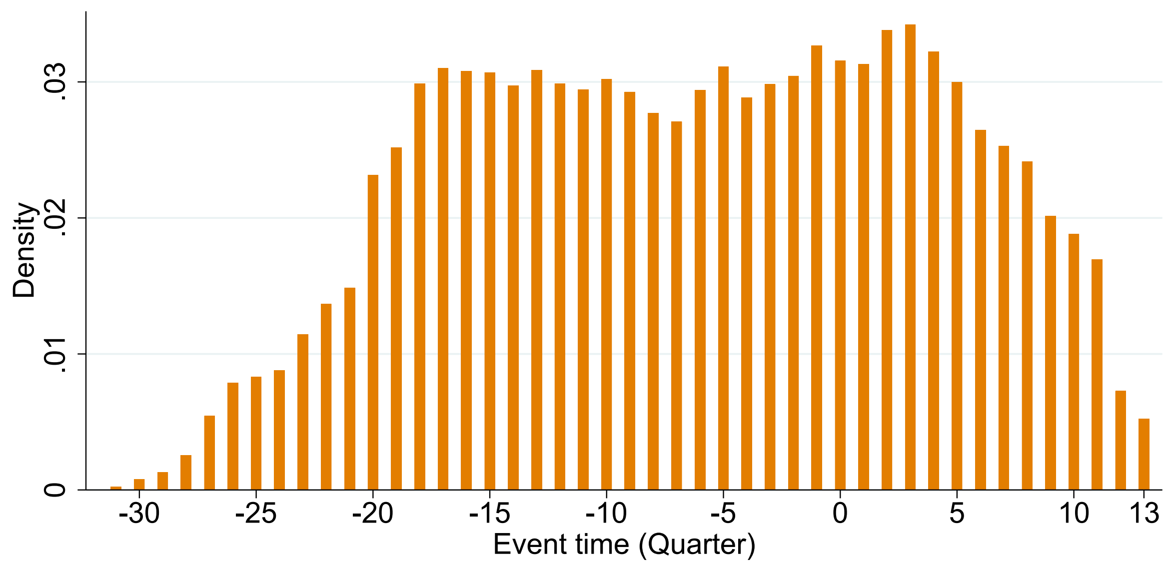

*Notes*: We tabulate the number of origin airports that have their first treatment in each year in the sample.

Figure A.5: **Distribution of Event Time for Flights Departed from Treated Airports**

Panel A. Yearly Distribution

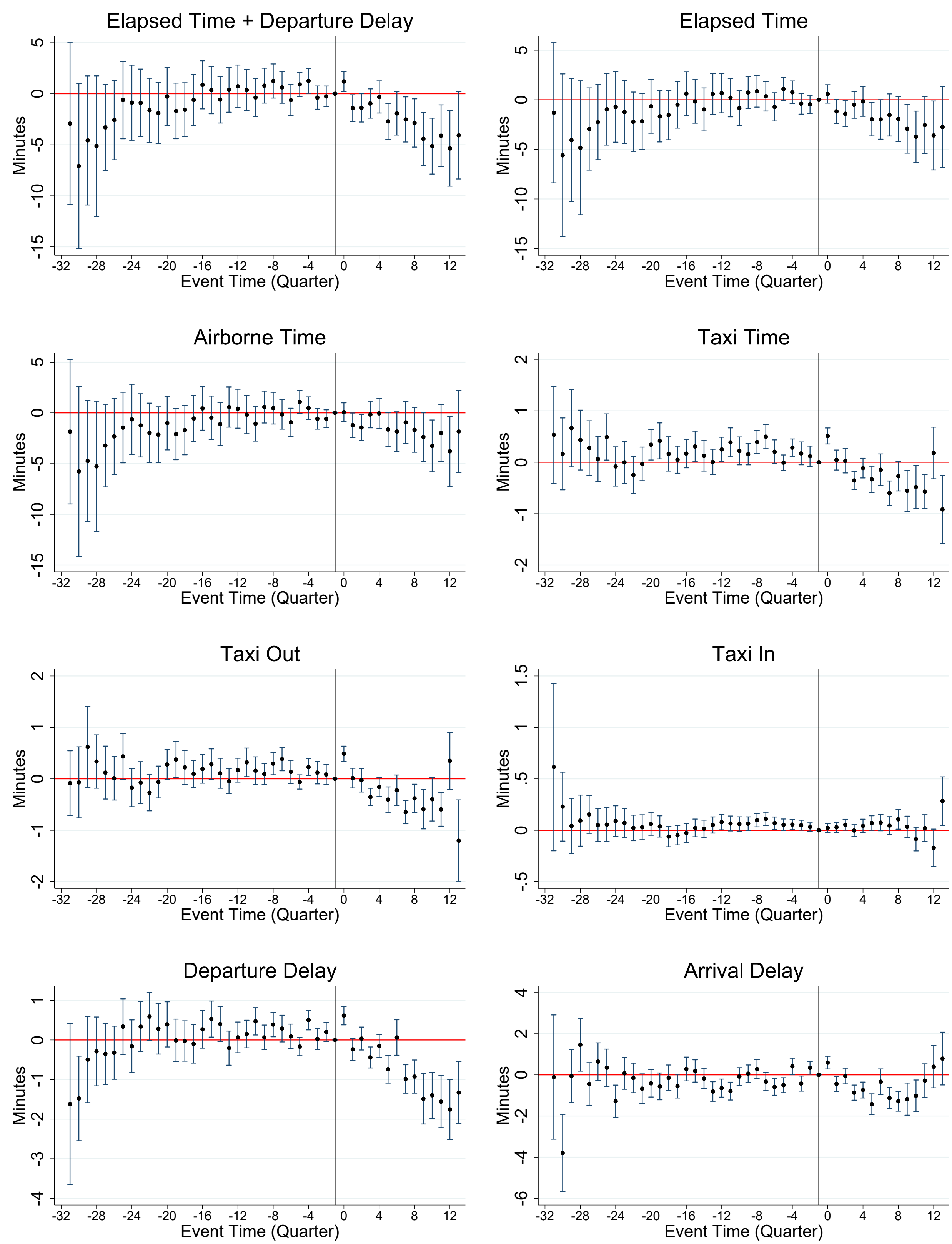



Panel B. Quarterly Distribution



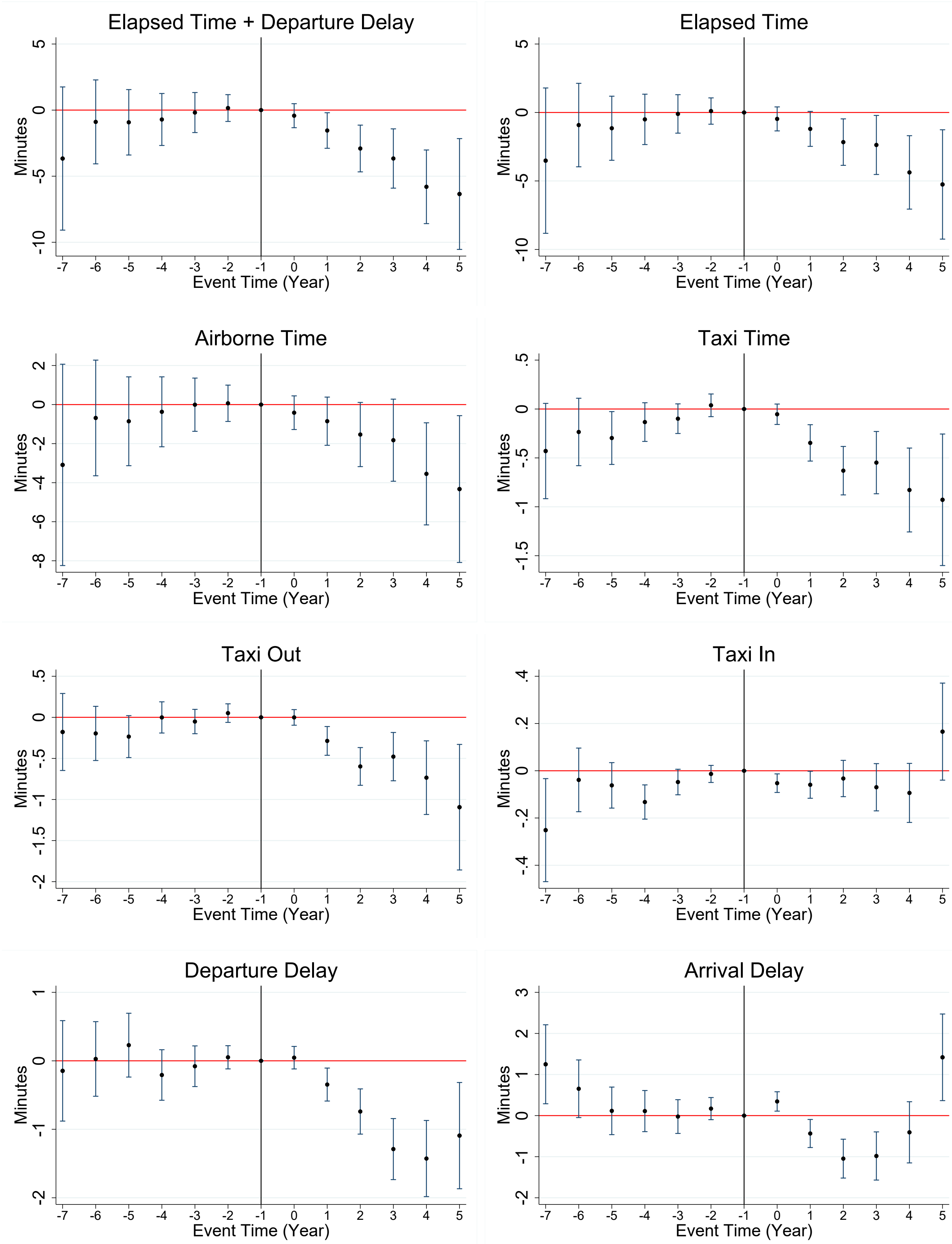

*Notes*: We count the number of flights at each event time and draw the distribution.

Figure A.6: **Event Study at the Quarterly Level**



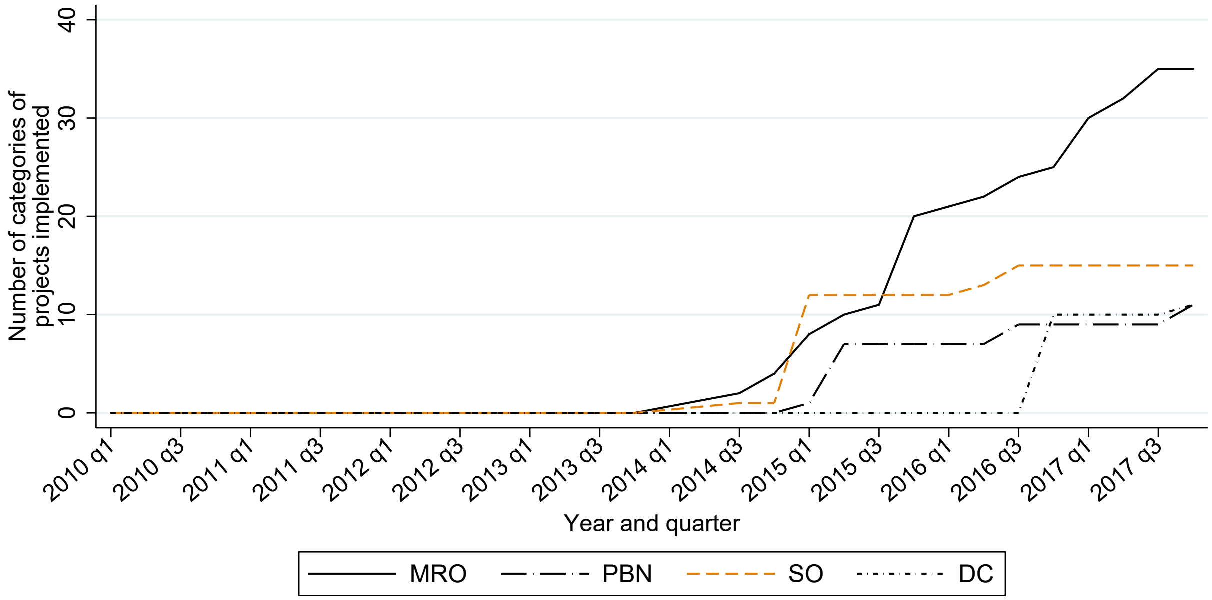

*Notes*: Results are from estimating equation (1) at the quarterly event-time level.

Figure A.7: **Robustness: Time Period 2010–2018**




*Notes*: Results are from estimating equation (1) with an extended sample period.

Figure A.8: **Number of Each Category of NextGen Projects Completed, 2010–2017**




*Notes*: We count the number of NextGen projects for each category (i.e., MRO, PBN, SO, and DC) completed at each airport each quarter, and aggregate them at the quarterly level.

Table A.1: **Summary Statistics of Subsample Variables 2010–2017**

| Variable | Mean | SD. | Min. | Max. |
|---|---|---|---|---|
| Panel A. Other Flight-level Information | | | | |
| Origin sky ceiling (1,000 feet) | 40.7 | 31.5 | 0 | 72.2 |
| Origin visibility (miles) | 9.3 | 1.9 | 0 | 91.0 |
| 1 = Prior flight is delayed | 0.26 | 0.44 | 0 | 1 |
| Prior delay (minutes) | 7.5 | 26.8 | 0 | 1,983 |
| Prior delay (minutes) if prior flight delays >= 1 min | 28.7 | 46.2 | 0 | 1,983 |
| Daily operations per aircraft | 4.7 | 1.8 | 1 | 17 |
| Flights traveled from a hub airport | 0.67 | 0.47 | 0 | 1 |
| Flights provided by a legacy carrier | 0.54 | 0.50 | 0 | 1 |
| Panel B. Self-reported Reasons of Delay for Delayed Flights | | | | |
| 1 = carrier delay | 0.52 | 0.50 | 0 | 1 |
| 1 = weather delay | 0.05 | 0.22 | 0 | 1 |
| 1 = NAS delay | 0.57 | 0.50 | 0 | 1 |
| 1 = security delay | 0.004 | 0.06 | 0 | 1 |
| 1 = late aircraft delay | 0.49 | 0.50 | 0 | 1 |
| Panel C. Additional Panel Information | | | | |
| Number of aircraft models | | | | 16 |
| Number of aircraft model trims | | | | 44 |
| Number of aircraft | | | | 5,333 |

*Notes*: In Panel B, we restrict the sample to delayed flights coded by DOT, which include flights with an arrival delay greater or equal to 15 minutes.

Table A.2: **Characteristics of NextGen and Control Airports**

| | NextGen airports (base treatment) | Other airports in Top 40 (base control) | Other airports in Top 50 | Other airports in Top 60 |
|---|---|---|---|---|
| Number of runways | 3.8 | 3.3 | 3.0 | 2.7 |
| Length of runways (k. ft.) | 9.3 | 8.7 | 8.7 | 8.5 |
| Visibility (mile) | 9.3 | 9.5 | 9.4 | 9.3 |
| Sky ceiling (k. ft.) | 39.7 | 42.4 | 42.9 | 43.1 |
| HHI | 0.57 | 0.60 | 0.62 | 0.63 |
| Number of rivals | 3.9 | 3.5 | 3.5 | 3.4 |

*Notes*: Runway data from Airnav.com. HHI and number of rivals are computed at route level and then aggregated to the airport.

Table A.3: **Correlation between the Adoption of Individual NextGen Technologies**

| | MRO | PBN | SO | DC |
|---|---|---|---|---|
| MRO | 1 | | | |
| PBN | 0.11 | 1 | | |
| SO | 0.42 | 0.26 | 1 | |
| DC | 0.28 | -0.04 | 0.12 | 1 |

Table A.4: **Alternative Measure for the Treatment Variables**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| $s \cdot \mathbf{1}(s > 0)$ | -1.796*** | -1.215* | -0.887 | -0.328*** | -0.333*** | 0.005 | -0.581*** | -1.185*** |
| | (0.692) | (0.655) | (0.638) | (0.089) | (0.083) | (0.029) | (0.128) | (0.165) |
| $s^2 \cdot \mathbf{1}(s > 0)$ | 0.015 | 0.009 | -0.019 | 0.027 | 0.028 | -0.001 | 0.006 | 0.283*** |
| | (0.193) | (0.180) | (0.179) | (0.029) | (0.030) | (0.009) | (0.044) | (0.058) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.38 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.07 | 0.06 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 without interacting quadratic post-trend terms with $n_{oq}$ and estimates equation (3).

Table A.5: **Alternative Specification with only Level Changes**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | Panel A. Effect of Total Number of Categories of Projects Implemented | | | | | | | |
| $n_{oq}$ | -2.239*** | -1.499*** | -1.300*** | -0.199*** | -0.165*** | -0.033** | -0.741*** | -0.920*** |
| | (0.506) | (0.485) | (0.475) | (0.048) | (0.047) | (0.017) | (0.090) | (0.111) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.28 | 0.25 |
| | Panel B. Effect of Specific Category of Projects | | | | | | | |
| $d_{oq}$ (MRO) | -1.933*** | -1.219* | -1.070* | -0.149 | -0.113 | -0.036 | -0.713*** | -0.566*** |
| | (0.692) | (0.653) | (0.624) | (0.106) | (0.104) | (0.027) | (0.114) | (0.162) |
| $d_{oq}$ (PBN) | -0.394 | 0.015 | -0.171 | 0.187* | 0.311*** | -0.124*** | -0.409** | -0.436* |
| | (1.130) | (1.094) | (1.078) | (0.100) | (0.098) | (0.045) | (0.203) | (0.242) |
| $d_{oq}$ (SO) | -4.603*** | -3.759*** | -3.321*** | -0.438*** | -0.443*** | 0.005 | -0.844*** | -1.333*** |
| | (1.088) | (1.051) | (1.038) | (0.101) | (0.099) | (0.039) | (0.158) | (0.222) |
| $d_{oq}$ (DC) | 0.520 | 1.554 | 1.830 | -0.276** | -0.277** | 0.001 | -1.033*** | -1.854*** |
| | (1.308) | (1.257) | (1.216) | (0.126) | (0.115) | (0.043) | (0.258) | (0.277) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.21 | 0.22 | 0.12 | 0.28 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats Table 2 without interacting $n_{oq}$ with quadratic post-trend terms and estimates equation (4). Panel B repeats Panel A and estimates the effect of the individual category of technologies.

Table A.6: **Conditional Effect of NextGen during Severe Weather Measured by Visibility**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| | | | | A.1 Origin Airport Visibility < 10 miles | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.555** | -1.052* | -0.727 | -0.325*** | -0.299*** | -0.025 | -0.503*** | -0.902*** |
| | (0.642) | (0.608) | (0.608) | (0.092) | (0.089) | (0.025) | (0.116) | (0.170) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.170 | 0.118 | 0.029 | 0.089*** | 0.090*** | -0.001 | 0.053* | 0.199*** |
| | (0.162) | (0.153) | (0.157) | (0.034) | (0.032) | (0.007) | (0.029) | (0.051) |
| | | | | A.2 Origin Airport Visibility < 9 miles | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.435** | -0.797 | -0.532 | -0.265*** | -0.254*** | -0.011 | -0.637*** | -1.014*** |
| | (0.696) | (0.647) | (0.624) | (0.076) | (0.077) | (0.021) | (0.155) | (0.191) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.114 | 0.066 | 0.025 | 0.041 | 0.046* | -0.005 | 0.047 | 0.175*** |
| | (0.181) | (0.167) | (0.160) | (0.025) | (0.025) | (0.006) | (0.043) | (0.057) |
| | | | | A.3 Origin Airport Visibility < 7 miles | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.604** | -0.801 | -0.564 | -0.237*** | -0.217** | -0.020 | -0.804*** | -1.227*** |
| | (0.725) | (0.663) | (0.638) | (0.090) | (0.090) | (0.023) | (0.199) | (0.239) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.116 | 0.057 | 0.036 | 0.021 | 0.025 | -0.004 | 0.059 | 0.198*** |
| | (0.192) | (0.173) | (0.166) | (0.031) | (0.031) | (0.007) | (0.057) | (0.071) |
| | | | | A.4 Origin Airport Visibility < 4.5 miles | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.096*** | -0.915 | -0.625 | -0.289** | -0.221* | -0.068*** | -1.181*** | -1.859*** |
| | (0.743) | (0.667) | (0.644) | (0.115) | (0.117) | (0.026) | (0.253) | (0.290) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.260 | 0.096 | 0.078 | 0.018 | 0.005 | 0.013 | 0.164** | 0.378*** |
| | (0.203) | (0.181) | (0.174) | (0.042) | (0.043) | (0.008) | (0.076) | (0.089) |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 Panel A in subsamples of severe weather conditions at the origin airport.

Table A.7: **Conditional Effect of NextGen during Severe Weather Measured by Sky Ceiling**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| | | | | A.1 Origin Airport Sky Ceiling < 6,000 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.650*** | -1.010* | -0.709 | -0.301*** | -0.314*** | 0.013 | -0.640*** | -1.106*** |
| | (0.630) | (0.590) | (0.578) | (0.066) | (0.065) | (0.020) | (0.117) | (0.163) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.130 | 0.057 | -0.008 | 0.066*** | 0.075*** | -0.009 | 0.073** | 0.226*** |
| | (0.170) | (0.158) | (0.155) | (0.024) | (0.024) | (0.006) | (0.033) | (0.047) |
| | | | | A.2 Origin Airport Sky Ceiling < 3,000 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.836*** | -1.034* | -0.703 | -0.331*** | -0.337*** | 0.006 | -0.802*** | -1.279*** |
| | (0.640) | (0.599) | (0.588) | (0.070) | (0.071) | (0.022) | (0.127) | (0.174) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.188 | 0.059 | -0.020 | 0.079*** | 0.087*** | -0.008 | 0.129*** | 0.300*** |
| | (0.172) | (0.161) | (0.159) | (0.026) | (0.027) | (0.006) | (0.035) | (0.050) |
| | | | | A.3 Origin Airport Sky Ceiling < 1,750 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.130*** | -1.237** | -0.919 | -0.318*** | -0.308*** | -0.010 | -0.893*** | -1.368*** |
| | (0.627) | (0.588) | (0.578) | (0.077) | (0.079) | (0.025) | (0.135) | (0.188) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.289* | 0.135 | 0.060 | 0.075*** | 0.079*** | -0.004 | 0.154*** | 0.346*** |
| | (0.175) | (0.162) | (0.158) | (0.028) | (0.029) | (0.008) | (0.038) | (0.055) |
| | | | | A.4 Origin Airport Sky Ceiling < 1,400 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.266*** | -1.297** | -0.965* | -0.331*** | -0.306*** | -0.025 | -0.969*** | -1.453*** |
| | (0.626) | (0.588) | (0.576) | (0.086) | (0.086) | (0.026) | (0.149) | (0.203) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.347** | 0.164 | 0.084 | 0.080*** | 0.080** | 0.000 | 0.183*** | 0.380*** |
| | (0.175) | (0.162) | (0.158) | (0.030) | (0.031) | (0.009) | (0.043) | (0.060) |
| | | | | A.5 Origin Airport Sky Ceiling < 900 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.153*** | -1.225** | -1.011* | -0.214** | -0.172* | -0.042 | -0.928*** | -1.463*** |
| | (0.638) | (0.591) | (0.575) | (0.099) | (0.101) | (0.028) | (0.187) | (0.226) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.363* | 0.196 | 0.152 | 0.044 | 0.037 | 0.007 | 0.166*** | 0.388*** |
| | (0.187) | (0.173) | (0.168) | (0.036) | (0.037) | (0.009) | (0.057) | (0.070) |
| | | | | A.6 Origin Airport Sky Ceiling < 750 Feet | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.350*** | -1.376** | -1.188** | -0.188* | -0.138 | -0.050* | -0.973*** | -1.617*** |
| | (0.685) | (0.619) | (0.597) | (0.113) | (0.114) | (0.028) | (0.241) | (0.284) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.411** | 0.240 | 0.219 | 0.021 | 0.011 | 0.010 | 0.171** | 0.415*** |
| | (0.196) | (0.180) | (0.173) | (0.039) | (0.040) | (0.009) | (0.074) | (0.087) |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 Panel A in subsamples of severe weather conditions at the origin airport.

Table A.8: **Conditional Effect of NextGen by Prior Delay**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | | | | **A.1 Prior Delay > 1 minute** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.206*** | -1.240** | -0.933* | -0.307*** | -0.311*** | 0.005 | -0.966*** | -1.424*** |
| | (0.609) | (0.572) | (0.564) | (0.057) | (0.055) | (0.020) | (0.121) | (0.149) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.358** | 0.217 | 0.129 | 0.088*** | 0.094*** | -0.005 | 0.141*** | 0.303*** |
| | (0.158) | (0.147) | (0.145) | (0.020) | (0.020) | (0.006) | (0.029) | (0.038) |
| | | | | **A.2 Prior Delay > 10 minutes** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -2.852*** | -1.571*** | -1.176** | -0.395*** | -0.404*** | 0.009 | -1.281*** | -1.789*** |
| | (0.653) | (0.606) | (0.598) | (0.064) | (0.059) | (0.022) | (0.154) | (0.177) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.498*** | 0.293* | 0.175 | 0.118*** | 0.123*** | -0.006 | 0.205*** | 0.379*** |
| | (0.172) | (0.157) | (0.154) | (0.021) | (0.021) | (0.006) | (0.038) | (0.044) |
| | | | | **A.3 Prior Delay > 20 minutes** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -3.231*** | -1.681*** | -1.209** | -0.472*** | -0.473*** | 0.001 | -1.550*** | -2.121*** |
| | (0.675) | (0.617) | (0.610) | (0.070) | (0.065) | (0.023) | (0.188) | (0.210) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.597*** | 0.331** | 0.189 | 0.142*** | 0.146*** | -0.004 | 0.266*** | 0.454*** |
| | (0.177) | (0.160) | (0.157) | (0.023) | (0.023) | (0.006) | (0.047) | (0.052) |
| | | | | **A.4 Prior Delay > 40 minutes** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -4.054*** | -1.884*** | -1.319** | -0.565*** | -0.549*** | -0.015 | -2.170*** | -2.765*** |
| | (0.715) | (0.632) | (0.628) | (0.081) | (0.077) | (0.025) | (0.264) | (0.285) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.803*** | 0.393** | 0.221 | 0.172*** | 0.173*** | -0.001 | 0.410*** | 0.604*** |
| | (0.188) | (0.169) | (0.167) | (0.026) | (0.025) | (0.007) | (0.068) | (0.072) |
| | | | | **A.5 Prior Delay > 75 minutes** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -4.849*** | -1.771** | -1.175* | -0.596*** | -0.591*** | -0.005 | -3.078*** | -3.681*** |
| | (0.875) | (0.707) | (0.698) | (0.098) | (0.092) | (0.033) | (0.436) | (0.449) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.945*** | 0.328* | 0.147 | 0.181*** | 0.188*** | -0.007 | 0.617*** | 0.815*** |
| | (0.234) | (0.198) | (0.194) | (0.030) | (0.029) | (0.009) | (0.119) | (0.119) |
| | | | | **A.6 Prior Delay > 120 minutes** | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -5.275*** | -1.273* | -0.702 | -0.572*** | -0.574*** | 0.003 | -4.002*** | -4.687*** |
| | (1.027) | (0.761) | (0.749) | (0.118) | (0.114) | (0.041) | (0.650) | (0.651) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 1.011*** | 0.196 | 0.021 | 0.175*** | 0.188*** | -0.014 | 0.816*** | 1.039*** |
| | (0.278) | (0.222) | (0.217) | (0.035) | (0.034) | (0.011) | (0.174) | (0.171) |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table repeats Table 2 Panel A in subsamples such that the previous operation of an aircraft is delayed (actual arrival compared to schedule arrival) by more than 1, 10, 20, 40, 75, and 120 minutes.

Table A.9: **Effect of NextGen: Alternative Sample**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var.: air travel time (minutes) | elapsed time + departure delay | elapsed time | airborne time | taxi time | taxi-out time | taxi-in time | departure delay | arrival delay |
| | | | A.1 Include samples not matched with DOT Form-B43 | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.399*** | -1.004** | -0.754* | -0.249*** | -0.276*** | 0.026 | -0.395*** | -0.785*** |
| | (0.449) | (0.428) | (0.416) | (0.051) | (0.050) | (0.017) | (0.074) | (0.106) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.112 | 0.082 | 0.023 | 0.059*** | 0.070*** | -0.011** | 0.030 | 0.172*** |
| | (0.120) | (0.114) | (0.112) | (0.018) | (0.018) | (0.005) | (0.019) | (0.029) |
| Num. of obs. | 30,374,516 | 30,374,516 | 30,374,516 | 30,374,516 | 30,374,516 | 30,374,516 | 30,374,516 | 30,374,516 |
| R-squared | 0.40 | 0.44 | 0.43 | 0.24 | 0.24 | 0.14 | 0.27 | 0.24 |
| | | | A.2 Include top 50 departure airports | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.354*** | -0.905* | -0.692 | -0.213*** | -0.222*** | 0.009 | -0.449*** | -0.893*** |
| | (0.482) | (0.462) | (0.453) | (0.052) | (0.050) | (0.018) | (0.072) | (0.106) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.124 | 0.087 | 0.033 | 0.054*** | 0.062*** | -0.008 | 0.037** | 0.194*** |
| | (0.127) | (0.122) | (0.120) | (0.019) | (0.018) | (0.005) | (0.018) | (0.029) |
| Num. of obs. | 23,418,571 | 23,418,571 | 23,418,571 | 23,418,571 | 23,418,571 | 23,418,571 | 23,418,571 | 23,418,571 |
| R-squared | 0.42 | 0.46 | 0.45 | 0.22 | 0.23 | 0.13 | 0.29 | 0.25 |
| | | | A.3 Include top 60 departure airports | | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.288*** | -0.874* | -0.670 | -0.204*** | -0.225*** | 0.020 | -0.414*** | -0.842*** |
| | (0.474) | (0.455) | (0.446) | (0.052) | (0.049) | (0.018) | (0.069) | (0.102) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.109 | 0.082 | 0.029 | 0.053*** | 0.064*** | -0.011** | 0.027 | 0.180*** |
| | (0.126) | (0.121) | (0.119) | (0.018) | (0.018) | (0.005) | (0.018) | (0.029) |
| Num. of obs. | 24,407,121 | 24,407,121 | 24,407,121 | 24,407,121 | 24,407,121 | 24,407,121 | 24,407,121 | 24,407,121 |
| R-squared | 0.43 | 0.47 | 0.46 | 0.22 | 0.23 | 0.13 | 0.29 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A.1 repeats Table 2 and includes a greater sample such that aircraft with tail number is not necessarily matched in the DOT Form-B43, Panel A.2 includes flights that departed from top 50 airports, and Panel A.3 includes flights departed from top 50 airports.

Table A.10: **Robustness: Control for Weather**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | | | | A.1 Control for visibility and sky ceiling | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.427*** | -0.957** | -0.728 | -0.229*** | -0.233*** | 0.004 | -0.470*** | -0.923*** |
| | (0.492) | (0.471) | (0.462) | (0.053) | (0.051) | (0.018) | (0.075) | (0.109) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.126 | 0.086 | 0.030 | 0.056*** | 0.064*** | -0.008 | 0.040** | 0.200*** |
| | (0.129) | (0.123) | (0.122) | (0.019) | (0.019) | (0.005) | (0.018) | (0.029) |
| Visibility (mile) | -0.923*** | -0.410*** | -0.085*** | -0.325*** | -0.325*** | 0.000 | -0.513*** | -0.840*** |
| | (0.038) | (0.020) | (0.012) | (0.016) | (0.016) | (0.001) | (0.024) | (0.033) |
| Sky ceiling (1k ft) | -0.039*** | -0.016*** | -0.004*** | -0.012*** | -0.012*** | -0.000*** | -0.023*** | -0.035*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) |
| Num. of obs. | 22,355,055 | 22,355,055 | 22,355,055 | 22,355,055 | 22,355,055 | 22,355,055 | 22,355,055 | 22,355,055 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.22 | 0.23 | 0.12 | 0.29 | 0.25 |
| | | | | A.2 Control for DOT delay code being weather delay | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s > 0)$ | -1.380*** | -0.942** | -0.723 | -0.219*** | -0.223*** | 0.004 | -0.438*** | -0.880*** |
| | (0.491) | (0.471) | (0.462) | (0.053) | (0.050) | (0.018) | (0.072) | (0.106) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s > 0)$ | 0.117 | 0.083 | 0.029 | 0.054*** | 0.062*** | -0.008 | 0.034* | 0.192*** |
| | (0.128) | (0.123) | (0.122) | (0.019) | (0.018) | (0.005) | (0.018) | (0.029) |
| 1 = if delay code is weather | 69.694*** | 17.752*** | 6.470*** | 11.282*** | 10.645*** | 0.637*** | 51.943*** | 66.160*** |
| | (0.658) | (0.479) | (0.377) | (0.260) | (0.261) | (0.028) | (0.511) | (0.632) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.42 | 0.45 | 0.45 | 0.22 | 0.23 | 0.12 | 0.30 | 0.28 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A.1 repeats Table 2 and includes and controls for visibility (mile) and sky ceiling (1,000 feet) using data from NOAA, and Panel A.2 include a dummy variable that equals 1 if the DOT delay code is weather.

Table A.11: **Robustness: Control for Level of Competition**

| Dep. var.: air travel time (minutes) | (1) elapsed time + departure delay | (2) elapsed time | (3) airborne time | (4) taxi time | (5) taxi-out time | (6) taxi-in time | (7) departure delay | (8) arrival delay |
|---|---|---|---|---|---|---|---|---|
| | | | | A.1 Control for route-level HHI | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.443*** | -0.983** | -0.761* | -0.222*** | -0.227*** | 0.005 | -0.460*** | -0.907*** |
| | (0.479) | (0.459) | (0.449) | (0.053) | (0.051) | (0.017) | (0.075) | (0.109) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.136 | 0.099 | 0.046 | 0.054*** | 0.062*** | -0.009* | 0.037** | 0.195*** |
| | (0.126) | (0.121) | (0.119) | (0.019) | (0.019) | (0.005) | (0.018) | (0.030) |
| HHI | 46.468*** | 47.059*** | 49.440*** | -2.381*** | -0.486*** | -1.895*** | -0.591*** | -2.014*** |
| | (2.731) | (2.679) | (2.665) | (0.099) | (0.063) | (0.068) | (0.128) | (0.139) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.43 | 0.47 | 0.47 | 0.21 | 0.22 | 0.13 | 0.28 | 0.25 |
| | | | | A.2 Control for number of rival carriers within a route | | | | |
| $n_{oq} \cdot s \cdot \mathbf{1}(s>0)$ | -1.516*** | -1.055** | -0.834** | -0.221*** | -0.227*** | 0.006 | -0.461*** | -0.906*** |
| | (0.448) | (0.428) | (0.417) | (0.053) | (0.051) | (0.018) | (0.074) | (0.109) |
| $n_{oq} \cdot s^2 \cdot \mathbf{1}(s>0)$ | 0.148 | 0.111 | 0.057 | 0.054*** | 0.062*** | -0.009* | 0.037** | 0.195*** |
| | (0.118) | (0.113) | (0.111) | (0.019) | (0.019) | (0.005) | (0.018) | (0.030) |
| Num. of rivals | -12.249*** | -12.181*** | -12.487*** | 0.305*** | 0.044*** | 0.261*** | -0.068*** | 0.243*** |
| | (0.376) | (0.371) | (0.369) | (0.017) | (0.010) | (0.013) | (0.018) | (0.023) |
| Num. of obs. | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 | 22,355,119 |
| R-squared | 0.46 | 0.50 | 0.50 | 0.21 | 0.22 | 0.13 | 0.28 | 0.25 |

*Notes*: Robust standard errors clustered at origin airport by hour-of-day level in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A.1 repeats Table 2 and includes and controls for route-level HHI constructed from share of flights provided by a carrer in our sample period, and Panel A.2 controls for number of rival carriers within a route.

Table A.12: **AEDT Estimates of Fuel Use and Emissions**

| Dependent variable: | (1) Fuel consumption (kg.) | (2) $CO_2$ emissions (kg.) |
|---|---|---|
| *A. Departure Operations:* | | |
| A.1 Taxi-out (minutes) | 18.815*** | 57.756*** |
| | (0.028) | (0.104) |
| | $N$ = 1,454,871 | $N$ = 1,454,871 |
| A.2 Take-off ground roll (minutes) | 104.640*** | 330.140*** |
| | (0.232) | (0.732) |
| | $N$ = 10,396,573 | $N$ = 10,396,573 |
| A.3 Take-off airbourne (minutes) | 122.995*** | 388.050*** |
| | (0.186) | (0.586) |
| | $N$ = 12,214,894 | $N$ = 12,214,894 |
| A.4 Take-off climb (minutes) | 57.040*** | 179.961*** |
| | (0.145) | (0.457) |
| | $N$ = 23,187,256 | $N$ = 23,187,256 |
| *B. Arrival Operations:* | | |
| B.1 Taxi-in (minutes) | 19.155*** | 58.860*** |
| | (0.018) | (0.067) |
| | $N$ = 1,476,440 | $N$ = 1,476,440 |
| B.2 Approach (minutes) | 35.633*** | 112.421*** |
| | (0.025) | (0.078) |
| | $N$ = 22,143,593 | $N$ = 22,143,593 |

*Notes*: Robust standard errors in parentheses. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. This table reports the marginal effect of the duration of various operations on fuel use and $CO_2$ emissions using AEDT data for simulated operations that depart from or fly to a NextGen airport.

Table A.13: **Benefits under Alternative Assumptions**

Panel A. Benefits in Table 7 under Alternative Assumptions

| Fuel costs using | (1) AEDT Data (base) | (2) AEDT Data | (3) AEDT Data | (4) FAA Guidelines | (5) FAA Guidelines |
|---|---|---|---|---|---|
| A.1 Alternative Private Gains (Billions 2017 USD) | | | | | |
| Fuel and oil savings | \$B 0.15 | \$B 0.23 | \$B 0.12 | \$B 0.29 | \$B 0.15 |
| Passenger time savings | \$B 0.78 | \$B 0.78 | \$B 0.78 | \$B 0.78 | \$B 0.78 |
| Total private benefits | \$B 0.93 | \$B 1.00 | \$B 0.89 | \$B 1.06 | \$B 0.93 |
| A.2 Alternative Social Benefits (Billions 2017 USD) | | | | | |
| Benefits of carbon reductions | \$B 0.04 | \$B 0.06 | \$B 0.03 | N.A. | N.A. |

Panel B. Benefits in Table 8 under Alternative Assumptions

| Fuel costs using | (1) AEDT Data (base) | (2) AEDT Data | (3) AEDT Data | (4) FAA Guideline | (5) FAA Guideline |
|---|---|---|---|---|---|
| B.1 Alternative Private Gains (Billions 2017 USD) | | | | | |
| Fuel and oil savings | \$B 0.48 | \$B 0.74 | \$B 0.38 | \$B 0.86 | \$B 0.48 |
| Passenger time savings | \$B 2.34 | \$B 2.34 | \$B 2.34 | \$B 2.34 | \$B 2.34 |
| Total private benefits | \$B 2.82 | \$B 3.08 | \$B 2.72 | \$B 3.20 | \$B 2.83 |
| B.2 Alternative Social Benefits (Billions 2017 USD) | | | | | |
| Benefits of carbon reductions | \$B 0.12 | \$B 0.18 | \$B 0.09 | N.A. | N.A. |

*Notes*: Panel A repeats Table 7 under alternative assumptions; Panel B repeats Table 8 under alternative assumptions. For columns 1-3, taxi-out fuel savings use AEDT estimate for taxi-out operations (Appendix Table A.12 Panel A.1), and taxi-in fuel savings use AEDT estimates for taxi-in operations (Appendix Table A.12 Panel B.1). In column 1 the baseline, airborne fuel savings use the average of take-off climb and approach operations from the AEDT (Appendix Table A.12 Panel A.4 and B.2). In column 2, airborne fuel savings use the average of take-off airborne, take-off climb, and approach operations from the AEDT (Appendix Table A.12 Panel A.3, A.4 and B.2). In column 3, airborne fuel savings use only approach operations from the AEDT (Appendix Table A.12 Panel B.2). In column 4, we use FAA Guidelines (FAA, 2016a) fuel cost estimates at \$2,443 per hour. In column 5, we use FAA Guideline fuel use estimates at 874 gallons per hour. In columns 1-3 and 5, we use the jet fuel price in 2017 from the EIA. The FAA Guidelines do not have emission estimates so we do not have estimates of social benefits from $CO_2$ reduction in columns 4 and 5.