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

Ziyan Chu
First Street Foundation

Yichen Christy Zhou*
Clemson University

April 2019

Abstract

The U.S. Federal Aviation Administration (FAA) has undergone a large scale multi-year modernization effort, called the Next Generation Air Transportation System (NextGen), and continues to invest in NextGen to improve airspace efficiency. To assess the efficacy of NextGen investments, we estimate how NextGen projects affect air travel time and delays by exploiting the high-frequency air flight on-time performance data from 2010 to 2017. Using a difference-in-differences design, we find that adopting one additional category of NextGen projects in both departure and arrival airports would improve air travel time by 2.4 minutes, with most time savings resulting from reductions in departure delays and 14 percent from reductions in taxi time. The effect of NextGen is much stronger for flights on the right-tail of the distribution of air travel delays due to unexpected shocks such as poor weather and prior delays. Preliminary calculations suggest that the NextGen projects have lead to passenger time saving of 221 dollars per flight and the airline fuel saving of 45 dollars per flight, amounting to 1.3 billion dollars of private benefits in 2017.

Keywords: Air transportation, air travel time and delay, infrastructure upgrade, cost and benefit analysis of air transpiration

JEL codes: L93, R4, Q4

*Corresponding author. John E. Walker Department of Economics, Clemson University, Sirrine Hall #223, Clemson, SC 29634, USA. Contact: yichen2@clemson.edu.

Zhou thanks the National Bureau of Economic Research for financial support.

1 Introduction

The US air transportation system serves a large volume of passengers and represents most commercial air traffic worldwide. In 2017 the number of flights departing from US airports accounted for 27 percent of all flights worldwide according to the World Bank data.[1] The commercial aviation industry accounted for 5.4 percent of US GDP in 2012 (FAA, 2014). Moreover, air traffic has been increasing in the past few decades, and the FAA forecasts that air traffic shall continue to grow at an average rate of 2.4 percent per year between 2016 and 2037 (FAA, 2018). Despite its importance, US air traffic in the 2000s has experienced more congestion, and its infrastructure has been underdeveloped when compared to other developed countries (FAA and Eurocontrol, 2010).

This paper studies how improving air transportation infrastructure affects commercial air travel performance. An important source of air travel inefficiency stems from the market failure of failing to account for external costs on competitors (e.g., delay, fuel consumption, cancellation, etc.), and therefore economists would usually prescribe a well-designed Pigouvian tax such as peak-time pricing or more sophisticated tax accounting for the market power and the network effect as the first-best solution (e.g., Daniel, 1995, 2001; Brueckner, 2002, 2005; Mayer and Sinai, 2003). However, when airport facilities and air traffic infrastructure create binding constraints for airlines' optimal scheduling problem to maximize their profits, and because public good is usually under-provided, there may be important private and social gains from adopting modern technologies in the air transportation system.

The U.S. government has planned to implement new airspace infrastructure under a multi-year effort on the Next Generation Air Transportation System (NextGen) since 2004, with a budget of 20 billion dollars (GAO, 2017).[2] The first wave of projects began construction in 2013. In 2014 the first wave projects were completed in multiple large airports such as the Hartsfield-Jackson Atlanta International Airport (ATL) and San Francisco International Airport (SFO). Despite the ex-ante projections produced by the agency (e.g., FAA, 2016b) and the magnitude of public expenditure on NextGen, there has been little retrospective analysis on the effectiveness of the NextGen program using detailed air flight data. The most relevant ex-post study finds that FAA expenditures have reduced the number of air traffic delays using cross-sectional aggregate airport-level data in 2000 (Morrison and Winston, 2008).

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

[2] We describe the background of the legislation that initiated NextGen in Section 2.

This paper assesses the effectiveness of NextGen by using high-frequency flight-level data from 2010 to 2017. The nature of the NextGen projects and the suggestive evidence of air travel performance over time motivates us to adopt a difference-in-differences approach. The Joint Planning and Development Office (JPDO) has planned to target 39 airports, most of which represent an important market share of US air traffic (JPDO, 2004). This planning allows us to distinguish between treated and untreated airports. Also, suggestive evidence shows that air travel time improves sharply in 2014 in treated airports, away from the trend, compared to untreated airports. Moreover, the gap between the two groups continues to widen gradually after 2014 for various measures of air travel performance (details in Section 3).

To estimate the effect of NextGen on air travel performance, we assemble various sources of air flight data, the most important of which is the On-Time Performance data from the US Department of Transportation (DOT) from 2010 to 2017. The On-Time Performance data allow us to observe detailed air travel performance (e.g., airborne time, taxi time, delay departure, etc.) for each air flight repeated every day over the sample period. The rich information of high-frequency flight data allows us to include an exhaustive set of fixed effects at the route, carrier, and aircraft level, an extensive set of time fixed effects, as well as additional interactions to remove airport-specific trends. Doing so allows us to account for unobservables such as the fact that treated airports tend to be larger airports, hub airports, and therefore more likely to be congested.

To quantify the extent to which a flight traveling from or to a given airport is treated, we collect the completion history of NextGen projects at each airport for each quarter. Treated airports have been scheduled to implement one or more of the four categories of NextGen projects (details in Section 2). For each treated airport we observe the quarter in which a particular FAA project is completed as well as the name and the category of the NextGen project completed. For each flight, the infrastructure upgrades at its origin airport (i.e., the departure airport) and destination airport (i.e., the arrival airport) may have affect air travel performance by a different magnitude and may affect different measures (e.g., an upgrade at the origin airport might not affect the taxi-in performance). We therefore construct two treatment variables for our baseline estimation – the number of categories of NextGen projects completed at its origin airport and the number of categories of NextGen projects completed at its destination airport.

Even with an extensive set of fixed effects and interactions of fixed effects with time trend, our estimates can still be vulnerable to omitted variables. Our identification rests upon the assumption that the timing of completion is as good as random and that an airline can not systematically manipulate air travel performance in response to upgrades at its origin and destination airports. Later in section 4 we explain how the uncertainty and variation of the duration of multiple stages prior to the implementation stage allow us to assume that the quarter in which a project is completed can be as good as random, and that the uncertainty and variation of duration it takes for air flight re-scheduling to take effect removes the concern of endogenous manipulation of air travel performance from the airline.

We find that adopting NextGen has lead to an improvement in air travel performance. In particular, increasing one category of projects completed at both ends of a route would reduce air travel time (the elapsed time plus the departure delay) by 2.4 minutes for the average flight. This change is a sizable improvement since there are roughly nine thousand operations per day and 92 percent of flights are travel from or to a NextGen airport.[3] We find that 87 percent of this reduction is due to reducing departure delays and 14 percent is due to more efficient taxi stages at both airports. Airborne time worsens slightly, but the estimates are not statistically significant.

We then further examine the heterogeneous effects of NextGen on different flights. This information can be valuable for informing better policy design since it is likely that most economic benefits of NextGen come from improving the flights that are subject to the greatest delays and shock in air travel time, i.e., the flights on the thin-tail. We find that the positive effect of NextGen is disproportionately greater for flights that are subject to severe weather and delay from the previous operation. Our findings confirm that NextGen has improved the flights that are most in need. Further, we examine how NextGen projects affect hub and non-hub airlines. We find that implementing projects affects both hub and non-hub airlines at hub airports but disproportionately more for non-hub airlines. This finding reveals that when a pre-existing congestion externality is present (Mayer and Sinai, 2003) and when infrastructure constraints have been plausibly binding, lifting the infrastructure constraints via implementing NextGen technologies can help close the wedge in air travel performance created by externalities.

To quantify the private benefits of adopting NextGen on airline and passengers, we focus on the effect of NextGen in the 2017 market. In 2017, treated airports have implemented 1.1 categories of NextGen projects and total 1.3 NextGen projects on average (description of the four categories of projects in Section 2). These treated airports cover a great volume of traffic – 92 percent of flights travel from or to a NextGen airport. We begin by producing the counterfactual changes in air travel time and delays in 2017 if treated airports had not

[3]DOT On-Time Performance only collect flight that lands in a US airport. We observe 15 thousands flights per day. In Section 3 we explain why our final sample represents about 60 percent of all flights in On-Time Performance.

been treated. Using assumptions in the FAA Cost and Benefit Guideline (FAA, 2016a), we further calculate the private cost savings on passengers and airlines (See Section 5.3 for the assumptions that we use). We find that adopting NextGen projects had lead to sizable private benefits to airlines and passengers in 2017. If the NextGen airports were not treated in 2017 (i.e., if the infrastructure were at the beginning of 2014), the 3.6 million flights in our sample in 2017 (representing 67 percent of all flights in 2017) would be associated with private welfare loss of 1.3 billion (2017 US dollar). Passenger time-saving accounts for the majority of the benefit at 809 million dollars. Fuel-saving also accounts for an important share of the benefit at 165 million dollars. Further in Section 5.3 we discuss how we plan to quantify the effect on social welfare such as noise, pollution emissions, and greenhouse gas emissions.

This paper contributes to the literature on air transportation efficiency and policy evaluation regarding the NextGen program by examining how adopting NextGen technologies affects air travel performance. Given that NextGen is a multi-year program that has been extended to 2030, our findings deliver a timely evaluation to the policymaker at the interim of the program by using high-frequency flight data rather than engineering simulations. Moreover, this paper also contributes to the literature on how improving infrastructure and increasing capacity affect congestion. Theories in congestion predict ambiguous results since it is possible that improving capacity (e.g., highway capacity) could decrease the marginal cost to join the traffic during peak hours and in the equilibrium the congestion would be the same in the absence of the capacity improvement. Our results imply that at least in the short to the medium run of the implementation of NextGen, air travel performance has experienced an improvement that results primarily from reductions in departure delays and time spent on the aircraft.

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 the implications in private welfare benefits in Section 5. We discuss alternative specifications and robustness in Section 6.

2 The NextGen Program

The U.S. government has spent and plans to continue to spend billions of dollars on improving air transportation efficiency. 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 Reauthorization Act (CARA).[4] According to the United States Federal Aviation Administration (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."[5] The FAA received 7.4 billion dollars from Congress to facilitate NextGen programs from 2004 to 2014 and has projected 20.6 billion USD in total spending to NextGen programs through 2030 (GAO, 2017).

Policymakers and the airline industry support the program because investing in NextGen can pay for itself by helping to address the increased demand for air transportation. It will do this by reducing the frequency and length of delays that are costly to airline and consumers, reducing cruising time by allowing planes to fly closer and to take straighter routes, reducing waiting time on runways, and reducing fuel consumption via the above channels (FAA, 2011).[6] In addition to private benefits, policymakers support NextGen because it has the potential to reduce greenhouse gas emissions (GHG) from air transportation (FAA, 2011).[7] In 2015, the transportation sector contributed 27 percent of the total CO2 emissions in the US, among which jet fuel consumption alone contributed 9 percent (EPA, 2016).

The potential private and social gains from improving the air transportation system have motivated JPDO to selected 39 airports to undergo NextGen upgrades (JPDO, 2011). In Figure 1, we plot all airports in our analysis and highlight the 39 airports that have undergone NextGen projects (hereafter *NextGen airports*). Most of these airports represent a significant portion of air traffic volume. Table 1 Panel 1 shows that 92 percent of flights travel from or to a NextGen airport. Appendix Table A.1 further shows that most of the treated airports are hub airports and each of them accounts for a sizable share in the total air traffic.

Despite the need to improve the air transportation infrastructure and the fact that initial planning started in 2004, implementation is a decade behind schedule (GAO, 2016; Morrison and Winston, 2007). The first wave of projects did not complete their pre-implementation stages until 2012 and 2013. Appendix Table A.1 shows that the first sets of projects were

[4] The CAR Act (P.L. 108-176) was signed by President G.W. Bush. Congress authorized and created the Joint Planning and Development Office (JPDO) to coordinate NextGen programs with the FAA.

[5] Source: https://www.faa.gov/nextgen/what_is_nextgen/.

[6] It also has other benefits such as providing safer flights and improving national security. The Vision 100 – The CAR Act was signed after the events of September 11, 2001.

[7] NextGen may have other public benefits such as national security benefits which are beyond the scope of this project.

only completed in 2014 at a small number of large airports such as Atlanta (ATL), San Fransico (SFO), and Houston (IAH).

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, the Atlanta International Airport (ATL) implemented wake recategorization (a type of MRO project) on July 2014, which allows aircraft to safely takeoff and land closer to each other, increasing capacity and flight efficiency. PBN can improve the flight path during the cruising time and increase the predictability of arrival time. SO projects primarily improve logistics on the ground at the gate, between gates, and between the gate and the runway (for takeoff or landing). As a result, SO 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 in turn benefit both surface operations and the ascending and descending stages.

To evaluate the effect the NextGen projects, we need to construct a variable that represents the treatment of airport upgrades. The FAA records and publishes the completion history of NextGen projects on their website and within published documents for each airport by quarter.[9]

Based on the completion history, we construct four dummies variables for the above four categories of NextGen projects to indicate if an airport has completed a particular category of project(s) in a given quarter. To measure the degree of treatment, we use a continuous variable that equals the cumulative total number of project categories completed at an airport, ranging from 0 to 4. In Figure 2 we plot the total categories of projects completed from 2010 to 2017. Treated airports started to complete NextGen projects in 2014, and the greatest variation is found in 2014 and 2015. Consistent with our measure, the ex-ante cost and benefit analysis of NextGen by FAA (2016c) reports almost zero benefits from 2010 to 2014, and the majority of the simulated benefits begin in 2015.[10] Appendix Table A.1 further shows the first year of NextGen completion for the 39 NextGen airports and 67

[8] Description for each category of project: https://www.faa.gov/nextgen/snapshots/priorities/.

[9] Completion history: https://www.faa.gov/nextgen/snapshots/priorities/completion_history/. We cross-checked the completion history with the priority records published by the FAA, e.g., the record of MRO https://www.faa.gov/nextgen/snapshots/priorities/?area=mro and projects at various metroplex area http://metroplexenvironmental.com/oapm.html.

[10] See Table 1 of FAA (2016c). We think their benefits from 2010 to 2014 mostly come from 2014 since some airports have completed first wave of projects in 2014.

percent of the airports began to have at least a project completed by 2014–2015. By the last quarter of 2017, 49 categories of projects were completed in the US.[11]

3 Data and Suggestive Evidence

Air travel time data We assembled the main data set from various sources, the most important of which is the On-Time Performance data from the U.S. Department of Transportation (DOT) for 2010 to 2017. US airports began reporting detailed air travel and delay information for all domestic and international non-stop segments starting in 1987 (14 CFR § 234.4 1987).[12] This data set allows us to observe several variables for each flight (defined as a non-stop segment of travel), including the actual and scheduled departure and arrival times, the actual total elapsed time, airborne time, taxi-out time, taxi-in time, and total taxi time, all of which provide important information in terms of measuring air travel performance, fuel consumption, and emissions.[13]

Table 1 Panel A shows that taxi time, airborne time, and departure delays account for an important share of total travel time, as measured by total elapsed time plus departure delays, and there is a good amount of variation we need to account for in our empirical model.[14] Figure 3 shows that from 2010 to 2017, the total elapsed time plus the departure delays have increased by about 15 minutes over the eight years in the sample. Departure delays have increased by only for 2 minutes, the airborne time increased for about 10 minutes, and the taxi-time has increased for 3 minutes.

In addition to air travel time, if a flight is delayed (e.g., the actual departure is 15 minutes later than the scheduled departure, this data set also reports the number of minutes of delay by cause of delay such as aircraft (previous operation), weather, etc., which we later use to estimate heterogeneous effect of NextGen. Lastly, this data set reports other flight

[11] Some airports have completed several projects within a main category. An alternative measure of treatment is the total number of projects completed. In our data, this variable ranges from 0 to 5. Appendix Figure A.1 shows the total number of projects completed, which illustrates a similar trend as Figure 2. By the last quarter of 2017, 89 projects were completed in the US. Later in Section 6 we show that our baseline is robust to this alternative measure of treatment.

[12] According to U.S. Federal Register, rule 14 CFR § 234.1-4 1987 requires airlines with at least 1 percent of domestic flights to report metrics of On-Time Performance. More details of the rule at 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/14/234.4.

[13] The total elapsed time is the number of minutes from gate-out to gate-in, the airborne time is the number of minutes from wheel-off (taking-off) to wheel-on (landing), the taxi-out time is the number of minutes from gate-out to wheel-off mostly on the runway, the taxi-in time is the number of minutes from wheel-on to gate-in mostly on the runway, and the taxi time is the summation of taxi-out and taxi-in time.

[14] We remove some outliers where the minimum is a very high negative number, e.g., if departure delay is less than -45. This exercise removes less than 0.1 percent of overall flights.

characteristics such as departure and arrival airports, airline companies (not the local operating carrier), date of the flight, and importantly, the tail number, which later allows us to obtain aircraft-specific information from other data sets.

Suggestive evidence of air travel performance In Figure 4 we plot annual air travel performance by whether or not the flight is en route from a NextGen airport in Panel A, and whether or not the flight is en route to a NextGen airport in Panel B. For each graph, we regress the specific air travel variable on year fixed effects and plot the residual air travel measure. For example, the first graph shows the residual elapsed time plus the departure delay after the effects captured in the year fixed effects are removed. To make comparison easier, we removed the level in the initial year for both groups.

Figure 4 motivates us to use a difference-in-differences approach and graphically illustrates our identification strategy. For example, Figure 4 Panel A suggests that flights which travel from a NextGen airport have improved significantly across many dimensions, including elapsed time, airborne time, taxi-out time, and plausibly departure delay compared to flights travel from other airports. Moreover, most of the improvements show up in the data around 2014 and 2015, and for some variables the gap between the control and treatment group continues to widen, consistent with the timeline of NextGen project completion. For taxi-in time and arrival delays, we see the effect goes the other direction which can due to other confounding factors not picked up by year fixed effects in this suggestive exercise. Similarly, Panel B also shows that flights en route to treated airports have experienced reductions in air travel time in many dimensions, including elapsed time, airborne time, and taxi-in time, compared to flights travel to other airports.

Information for heterogeneous effects We collected additional data to examine the heterogeneous effect of NextGen projects when air travel performance is plausibly in the tail of the distribution. Table 1 Panel A suggests that the maximum delay departure, taxi time, and airborne time are orders of magnitude higher than their averages. Factors such as weather can increase unexpected delays and increase air travel time. Therefore, we collected hourly weather information at each airport (e.g., the height of sky ceiling and the distance of visibility) reported in the National Center for the Environmental Information (NCEI) database from the National Oceanic and Atmospheric Administration (NOAA). Also, Morrison and Winston (2008) suggest that shocks in local air traffic control resources can affect air travel performance. We therefore collected the number of overflights (the number of aircraft that fly through the airspace of an airport) and seconds (the number of takeoff and landing operations at another local airport that uses the same terminal air traffic control tower of a given airport) from the Tower Operation data set and the

Terminal Radar Approach Control Facilities (TRACON) Operations data set from the Operations Network (OPSNET) database of the FAA Operation and Performance (ASPM) database.

Other Information We collected monthly fuel consumption data for all carriers from the DOT Form-F41 Schedule-P52. Figure 5 shows total fuel use and fuel cost for each quarter from 2010 to 2017. Panel A of Figure 5 suggests that the total amount of fuel used in commercial aviation is a huge volume and has almost doubled between 2010 to 2017. Later we plan to use the actual fuel use data as a benchmark to calibrate our fuel use predictions, which allows us to match the predicted fuel use that we produce from an engineering model to the moments of actual fuel use. Panel B shows the fuel cost increased in the first three years of our sample as fuel use increased and declined later possibly because that jet fuel price declined from 3 dollars per gallon from 2012 to 2014 to 1.5 dollars per gallon in 2016. Also, we collected monthly national jet fuel prices from the U.S. Energy Information Administration (EIA) to calculate the benefits of fuel savings.

In order to predict fuel consumption savings from air travel time savings, we need to collect aircraft information. Specifically, we collected aircraft make, model, and trim name as well as other model information from the DOT Form-B43. Each aircraft has to register for a tail number, i.e., the identification number painted on an aircraft, frequently on the tail (similar to the serial number of a phone or vehicle identification number of a vehicle). We add the aircraft information to our main data set using the tail number of each aircraft. Table 1 shows that our sample includes 6,957 unique aircraft, 44 aircraft model (such as Boeing 737), and 262 aircraft trims (such as Boeing 737-200).

Lastly, to generate fuel consumption and emission predictions, we applied for the Base of Aircraft Data (BADA) license from the European Organization for the Safety of Air Navigation (EuroControl) which allows us obtained the Aviation Environmental Design Tool (AEDT) from the FAA and the Aviation Emission Model (AEM) from the EuroControl. Both AEDT and AEM simulators allow us to obtain simulated data of air travel time, fuel use, and emissions for an aircraft at an airport by different stages of a flight such as the taxiing stage and the cruising stage. These estimates allow us to generate a mapping from travel time savings to reductions in fuel use and emissions, which further allows us to infer the private welfare using the counterfactual travel time savings produced from our baseline.[15] Our current version of private benefits are evaluated using parameters

[15]More information about BADA: https://simulations.eurocontrol.int/solutions/bada-aircraft-performance-model/; More information about AEDT: https://aedt.faa.gov/; More information about AEM: https://www.eurocontrol.int/services/aem-advanced-emission-model.

of FAA's cost and benefit report (FAA, 2016a). We plan to further improve the fuel savings and add to our calculation emission reductions using estimates from AEDT and AEM (Details in Section 5.3).

Our final dataset includes data on 25 million domestic and international flights from 2010 to 2017 in the US.[16] As shown in Table 1, this final data set includes 19 airlines that account for at least 1 percent of domestic air traffic, covers 5,968 routes that depart from and to 275 airports, among 29 which have undergone FAA NextGen projects during our sample period.

4 Empirical Strategy

In this section, we describe our empirical strategy to examine the effect of NextGen on air travel performance. Later in Section 5.3, we describe how we use our baseline estimates to analyze private and social benefits gained from adopting NextGen technologies.

Motivated by Figure 4 as discussed in the previous section, one viable approach is to quantify the effect of NextGen on air travel performance by using a difference-in-differences (DID) strategy.[17] For a flight i of airline j flying from origin airport o to destination airport d on date t and time h, we estimate the following equation

$$\begin{aligned} travel_{ijodtm} &= \beta_o NextGen_{ot} + \beta_d NextGen_{dt} \\ &\quad + \alpha_{od} + \alpha_j + \alpha_o \times y + \alpha_d \times y + \boldsymbol{\phi}_{tm} + u_{ijodtm} \end{aligned} \tag{1}$$

[16]Our final data set represents 55 percent of all flights in the sample period, about 46 million. First, we lose 26 percent of the original data points when merging the On-Time Performance Data with aircraft model data from DOT Form-B43, which is necesarry to calculate fuel consumption and emissions using the AEDT model. When we add back flights without aircraft information matched to Form-B43 which covers 83 percent of all flights, we find similar results (see Appendix Table A.5). Second, as explained later in Section 4, we exclude a flight if the origin or the destination airport is within the implementing period of a NextGen project, similar to Burlig et al. (2018). In particular, we remove a flight if it travels to or from an airport that is within 2 quarters of the completion of a project. We also remove the year 2013 to tease out the effect of system-wide preparation in 2013. Doing so removes 20 percent of the data. Also, we lose about 1 percent when we match the fuel consumption data in DOT Form-F41 Schedule-P52, 0.3 percent by dropping Comiar Airline, 1 percent when matching to airport weather information from the NOAA NCEI database, and another less than 0.1 percent from other types of outliers. Finally, since our main regression includes an extensive set of fixed effects, we lose about 1 percent of data due to the singularity problem.

[17]An alternative approach is to conduct a DID-version of quantile regression. However, Appendix Figure A.3 suggests that it is unclear whether the adoption of NextGen changes the distribution (in particular the second or higher moments) of air travel performance in a systematic way. For example, it is unclear whether the adoption of NextGen has systematically improved the departure delay on the right-tail. However, it is still possible that NextGen has improved some special cases of flights further on the right tail. We analyze the heterogeneous effect in section 5.2.

where $travel_{ijodtm}$ is a measure of air travel time including the following eight variables - minutes of elapse time plus delay departure, elapse time, airborne time, taxi time, taxi-in time, taxi-out time, departure delay, or arrival delay. The key variables of interests are the total number of categories of NextGen projects completed at the origin airport o in a quarter of a year $NextGen_{ot}$ and the total number of categories of NextGen projects completed at the destination airport d at a quarter of a year $NextGen_{dt}$. We introduce two treatment variables because a flight can be treated at the departure airport and the destination airport. In robustness Section 6, we use a single treatment variable and find similar results.

Our goal is to identify the parameters β_o and β_d, the treatment effects of adopting NextGen projects. The richness and high-frequency nature of the data allow us to adopt a selection-over-unobservable approach to identify the treatment effects by using an extensive set of fixed effects. To remove potential confounding factors that are likely correlated with treatment and the air travel performance given a date and a time, we include the fixed effects of route α_{od} and airline α_j. For example, a busy route might be more likely to be treated and thus more likely to experience an improvement in air travel time. For another example, a hub airport might be more congested and more likely to be treated. Without controlling for these factors we might exaggerate the effect of NextGen on air travel performance. In robustness Section 6 we include route-by-airline fixed effects $\alpha_{od,j}$ and find similar results.[18]

To remove time-specific factors that are correlated with both NextGen projects and air travel performance, we add to the regression the vector $\boldsymbol{\phi}_{tm}$ which includes year-by-month fixed effects, day-of-month fixed effects, day-of-week fixed effects, and hour-of-day fixed effects. As shown in Figures 2 and 3, while air travel time has become worse, and the number of projects has increased over time. Doing so removes common factors varying over time such as macroeconomic conditions as well as seasonal factors such as holiday season. The day-of-week fixed effects remove factors such as the fact that most weekday flights are business travelers and weekend flights are casual travelers and airlines may schedule flights accordingly. The hour-of-day fixed effects pick up the effects on air travel performance during peak and non-peak hours that are common across all flights.

To relax the common trend assumption that is necessary for the identification of a difference-in-differences estimator, we further include the interactions of origin airport fixed effects with a linear trend and the interactions of destination airport fixed effects with a linear trend. Doing so allows us to capture the changes in air travel performance over time at a given airport. For example, consider a decrease in the demand for air travel in a metropolitan

[18]Table 1 shows that there are 15,232 routes by airline. Section 6 shows the magnitudes and significance of our estimates barely change.

area over time due to the improvements in the substitute options, such as improvements to the interstate and freeway network. The interaction terms may reduce congestion for flight travel from or to that city over time, and we would not like to attribute this kind of improvement to the adoption of NextGen. Therefore, the identifying assumption reduces to a parallel trend assumption of treated and untreated flights after removing a linear effect over time that is specific for each departure and arrival airport, which is a relatively weaker and more reasonable assumption. Further in the robust Section 6 we add to the regression the interactions of airline fixed effects with a linear trend and find similar results. On the last note, the interactions terms might capture some of the improvements caused by NextGen, which means that our estimates can be on the conservative side and represent a lower bound of the effect of NextGen.

The coefficients of interest β_o and β_d are therefore interpreted as the average treatment effects of increasing the categories of NextGen projects on air travel performance. They represent how air travel performance has been improved for a flight that travels from or to a NextGen airport over time, compared to its counterpart that travels from or to other airports.

In addition to the relaxed assumption of the parallel trend argued above, another important identifying assumption rests upon whether the interventions are as good as random. In the case of NextGen, the airports that are selected to be treated are not random. Appendix Table A.1 shows that most NextGen airports are hub airports, and Mayer and Sinai (2003) show that hub airports are more likely to experience worse air travel time and delays because hub airlines do not internalize the congestion externality they impose on non-hub airlines in a hub airport. Also, Appendix Table A.1 shows that most NextGen airports are large airports that account for an important share of total air traffic and therefore it is in the interest of the regulating agency to improve them first, then other airports second. Our route fixed effects α_{od} and interactions of airports with linear trends $\alpha_o \times y$ and $\alpha_d \times y$ remove unobserved factors that select the NextGen airports into the treatment group.

Nevertheless, our extensive fixed effects would not remove the unobserved factors that are correlated with the timing of treatment. For example, the Atlanta Hartsfield International Airport (ATL) completed its first NextGen project in 2014, and the New York John F. Kennedy International Airport (JFK) completed its first NextGen project in 2015. If there are unobserved factors that caused the FAA to start NextGen at ATL earlier than JFK, which indirectly caused projects in ATL to be completed earlier than at JFK, our estimates could be biased. Here we argue that the variation and the uncertainty of the duration of the pre-implementation stages work in our favor to allow the completion date to be as good as random. Before a NextGen project is marked as completed (i.e., entering the post-implementation stage), a project has to finish a "budget impact" stage which takes about 2 quarters, a "study stage" which takes 1 to 3 quarters, a "facility resource issue" stage which takes 1 to 3 quarters, an "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 an "implementation" stage which takes 2 to 5 quarters; and the actual completion of each stage does not always conform to the scheduled completion date (e.g., FAA, 2016b). The variation of each stage across airports as well as the uncertainty that makes the actual completion deviate from the scheduled completion of each stage of a project allows us to impose a rather weak identifying assumption – after accounting for unobservables that are captured by our fixed effects in Equation (1), the timing of implementation of a NextGen project at a given airport is exogenous.[19]

In addition to the timing of completion of NextGen projects, our estimate could still be biased if airlines respond to NextGen projects systematically by re-optimizing and re-scheduling their flights at treated airports, which would affect air travel performance. For example, responding to improvements in ATL after the completion of their first NextGen project in 2014, the hub airline at ATL, Delta Airlines, might respond by clustering more flights during peak hours to serve more consumers, which may bias our estimates towards zero. Similar to the duration of pre-implementation stages of NextGen projects, airline scheduling also has a fair amount of variation and uncertainty. It typically takes several months for an airline to reschedule their flights. The time it takes to reschedule flights vary across airports and airlines, and over time, and is subject to a fair amount of uncertainty (Forbes, 2008). Also, since airlines do not have perfect foresight to know precisely when a new NextGen project will be implemented at a given airport, the uncertainty regarding the timing of the infrastructure upgrades also mitigates the concern of biased estimates due to unobserved flight re-scheduling.

Lastly, to compare the performance before and after treatment, we should exclude flights that travel from or to an airport that is within the implementation stage, otherwise our estimate would be biased towards zero as argued in Burlig et al. (2018). In the case of NextGen projects, most of them take from 2 to 5 quarters to complete, and the national aviation system (NAS) started to evaluate and prepare to implement projects as early as 2013 (see e.g., page 54 of FAA, 2016b). Therefore, we remove the three quarters prior to

[19] In addition, recent econometric studies have presented new DID estimators in the case of Difference-of-Differences with variation in treatment timing (e.g., Callaway and Sant'Anna, 2018; de Chaisemartin and D'HaultfÆuille, 2018; Goodman-Bacon, 2018). Those estimators are most suitable if we want to allow the treatment effect to be heterogeneous over time and it is beyond the scope of this study.

the completion in our main analysis. Removing four quarters would be a more conservative approach. However, since many airports have underwent multiple projects from 2014 to 2017, doing so would mechanically remove q majority of the post-treatment data after 2014 given an airport, leaving us no variation to identify the treatment effect. Since airports started to complete their first projects in 2014, we remove observations in the year 2013 to at least remove the implementation stage of some airports for their first project.[20]

5 Estimation Results

In this section, we begin by presenting our baseline results of the effect of NextGen on various measures of air travel performance in Section 5.1. We then explore the heterogeneous effects of NextGen on flights with relatively poor air travel performance based on flight characteristics in Section 5.2. In particular, we examine whether NextGen improves air travel performance for flights that are subject to exogenous shocks such as severe weather. We then discuss the implications of our baseline on private and social welfare in Section 5.3. In the section 6, we examine the robustness of our baseline.

5.1 The Impact of NextGen on Air Travel Performance

For our baseline specifications, we estimate equation (1) with fixed effects of route, carrier, year-by-month, day-of-month, day-of-week, hour-of-day, as well as the interactions of origin airport (departing airport) fixed effects interacted with a linear year trend and the interactions of the destination airport (arrival airport) fixed effects interacted with a linear year trend using flight data from 2010 to 2017. We cluster our standard errors at the route level to allow unobservables to be arbitrarily correlated within a route. We estimate equation (1) on the following eight measures of air travel performance as the dependent variables: minutes off elapsed time plus the departure delay, elapsed time, airborne time, taxi time, taxi-out time, taxi-in time, departure delay, and arrival delay.

Table 2 reports the estimated coefficients. This table shows the positive benefits of improving airport infrastructure under the NextGen program on almost all measures of air travel time performance and almost all our coefficients are precisely estimated.[21] Table 2 column 1 implies that implementing an additional category of NextGen projects at both the origin

[20] In robustness Section 6, we add the year 2013 back to the analysis and remove four quarters (instead of three) and find similar results.

[21] The only exception that we find is that adopting an additional NextGen project at the origin airport would reduce taxi-out time by 0.1 minutes (compared to the average of taxi-out time). One possibility is that since NextGen upgrades at origin and destination airports are positively correlated, and NextGen priorities focus on reducing the travel time via reducing the taxi time and improving the operational logistics to minimize departure delay, it is likely NextGen at both airports have a net zero effect on airborne time.

and destination airports would lead to a reduction in total travel time (elapsed time plus departure delay) by 2.4 minutes which is equivalent to a 2-percent improvement from the average (143 minutes). This effect is a sizable improvement considering that 92 percent of flights are en route from or to NextGen airports as shown in Panel A of Table 1 and there are on average nine thousand flights per day.

It is important to understand the composition of the above improvement. A fair share of this reduction is through the departure delay, as in Table 2 column 7. Adopting an additional category of NextGen projects at both ends of the route would reduce departure delay by 2.1 minutes, which is sizable compared to the average departure delay (9 minutes). The improvement in departure delay could in turn affect the arrival delay as well. Table 2 column 8 implies that the same marginal change would lead to a reduction in arrival delay by 2.6 minutes, the magnitude of which is higher than one half of the average arrival delay (4.7 minutes).

Compared to departure delay and arrival delay, the reduction in various stages of air travel during a flight is much smaller. In Table 2 columns 2 to 6, we examine how the NextGen treatment affects the elapsed, airborne, and taxi times. The same marginal change would lead to a reduction in elapsed time by 0.32 minutes (compared to an average of 134 minutes), a small and insignificant increase in airborne time (0.01 minutes and insignificant), a reduction in taxi time by 0.33 minutes (compared to an average of 23 minutes and s.d. of 11 minutes), a reduction in taxi-out time by 0.31 minutes (compared to an average of 16 minutes and s.d. of 9 minutes), and a small and insignificant reduction taxi-in time by 0.02 minutes (compared to an average of 7 minutes and s.d. of 5 minutes).

Our estimates suggest that most welfare gains of air travel time savings for both passengers and airlines come from reducing departure delay, which accounts for 88 of reduction in total elapsed time plus departure delay. This finding is consistent with the priorities of NextGen which focus on improvements on ground logistics and operation with airport terminal towers (see Section 2). Later in Section 5.3, we quantify the private gains of welfare under a various set of assumptions of the value of time. In contrast to departure delay, the improvement in taxi time accounts for a smaller share, about 14 percent of the reduction of the above marginal change. Nevertheless, this reduction would still have an important effect on fuel consumption and emissions which we discuss later in Section 5.3.[22]

Consistent with this possibility, our estimates suggest that we would fail to reject the hypothesis that the combined effect of NextGen at origin and destination airports have an effect different from zero.

[22] The numbers add up to be more than 100 percent since the above marginal change would lead to an increase in airborne time (although insignificant).

Also, our model captures a large amount of variation of the dependent variables except for the departure delay and arrival delay. Table 2 column 7 and 8 show the R-squared are only 0.04 for both specifications. The reason is that the departure delay is calculated using the difference between the actual departure and the scheduled departure. By construction, this procedure removes a lot of important unobserved factors that are specific to a route, airline, time block, etc. and a portion of the remaining variation is noise and other factors that can affect departure delay. Therefore, it is not surprising that our model only explains a small share of the variation in columns 7 and 8.[23]

Table 2 shows that almost all coefficients are precisely estimated. However, it is possible that the significance level in our estimates is driven by the sample size and the relatively sparse fixed effects, the combined effect of which would drive down the standard error. In particular, our sample includes 25,037,569 observations, and we only have 5,819 routes and 19 airlines (more information about the panel in Table 1 Panel B). We introduce many sets of fixed effects separately (e.g., fixed effects of the route and fixed effects of the airline) into our estimation equation due to the computation constraint.[24] Later in the robustness section 6, we show our baseline results are robust when we introduce fixed effects of route by airline by hour-of-day by day-of-week (resulting in 667,389 cells with an average of 36 observations per cell), and when we introduce fixed effects of route by airline by aircraft model (resulting in 38,681 cells with an average of 647 observations per cell).

5.2 Heterogeneous Effects of NextGen on Air Travel Performance

We proceed to explore how implementing NextGen projects affects flights with different characteristics. Results presented in this section reveal a considerable degree of heterogeneity. This information could be valuable to policymakers that no only care about overall welfare gains from infrastructure upgrades, but also are interested in identifying the types of flights whose air travel performance NextGen could feasibly improve.

Examining differential impacts not only reveals interesting heterogeneity and distributional effects, but demonstrates that it is likely that a sizable share of economic benefits from improving infrastructure NextGen comes from reducing the loss of low-probability events with high economic impacts. Air travel performance resembles the thin-tail high-impact distribution similar to the distribution of catastrophes due to climate change. Pindyck (2011)

[23]Also, as we examine fights with severe delays due to various reasons in Section 5.3, we find that the fit improves for departure delay and arrival delay, suggesting that the FAA upgrades do explain important variation for those flights even though the reason of the delay can be quite random.

[24]Even by applying the Frisch-Waugh-Lovell (FWL) algorithm, it still takes a day to produce eight columns for our baseline model in Table 2. The estimation time increases 2- to 3- fold when interacting more fixed effects together, and many folds when introducing more fixed effects interacting with a linear trend.

and Weitzman (2009) have argued the importance of both the impact and the probability for events that fall on the tail of distributions when investigating the expected outcome and the distributional impact of climate change.

In the case of air travel, the economic impact of flights that are subject to an extensive degree of delay and long wait time and are located on the very end of the right-tail distribution can be orders of magnitude higher. For example, the cost of delaying a flight from 15 to 30 minutes might be much lower than delaying a flight from 45 to 60 minutes since the latter might cause passengers to miss their connecting flights, make the crew members and the aircraft late for their next operation, make the aircraft sharing the same gate and runway to wait longer, and the butterfly effect might be further propagated over the flight network. Since the economic loss associated with air travel time and delay can increase at a non-linear (and maybe exponential) rate, the efficacy of the NextGen upgrade could crucially depend on how NextGen projects affect those flights on the right-tail distribution of air travel performance. Therefore, we next assess the heterogeneous effects of NextGen, in particular, the effect of NextGen for a few cases of thin-tail events.

Effect of NextGen during bad weather First, the FAA reports that 69 percent of delay is due to severe weather that impairs the visibility.[25] Although weather predictions can allow airlines to account for the weather factor ahead of time when scheduling their flights to the Computer Reservation System (CRS), the variance of the severity, as well as uncertainty in the precise timing of the weather shock can still affect air travel performance (Morrison and Winston, 2008). A.2 shows that on average 5 percent of flight delays are reported due to weather with a rather high standard error (50 percent), and that the reported time due to weather delay is in average 2 minutes with a large standard error too (28 minutes). It is therefore valuable to examine if implementing NextGen projects improves air travel performance under severe weather shocks.

In Table 5 we present our estimates for flights during bad weather using hourly weather information from the NCEI data set from NOAA (details in Section 3). In particular, we restrict our sample when visibility is impaired (when miles of the visibility are less than 5 miles) and the sky ceiling is low (when the sky ceiling is less than 1,500 feet). Alternatively, we could interact dummies that represent severe weather with our key variable of interest. Because our purpose is not to test if various flights have different parameters, but to quantify the effect of NextGen for some potential thin-tail events, and importantly, introducing interactions may increase our estimation time, we re-estimate our baseline using sub-samples.

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

Table 5 Panel A repeats the baseline for convenience. In Panels B.1 and B.2, we restrict our sample when the sky ceiling at the origin or the destination airport is lower than 1500 feet (compared to the average of 41,592 and standard deviation at 31,698 feet), either case of which represents about 9 percent of the full sample. Panel B.1 shows that the effects of NextGen at the origin airport are roughly 2- to 3- fold higher compared to our baseline when the weather is severe at the origin airport. The reduction in total air travel time shows improvements from all channels, including reduced departure delay, airborne time, and taxi-out time. Our estimates for NextGen at the destination airport are similar compared to the baseline. Our result in Panel B.2, however, is similar to the baseline, and some estimates also become insignificant, likely because we lose a large number of observations for us to identify the effect of NextGen. The results in B.2 are intuitive with the possibility that most severe weather affects the departing flights rather than arriving flights.

In Table 5 Panels B.3 and B.4, we examine the cases in which visibility is severely impaired. In particular, we restrict our sample to flights in which the visible distance is less than 5 miles (compared to the average of 9.3 miles and standard deviation of 2.0 miles), either case of which represents about 5 percent of the full sample. Panel B.3 suggests that NextGen at the origin airport has no effect on total travel time (compared to baseline) and would increase taxi time. We find our results are similar in Panel B.4 when the visibility is severe at the destination airport. We further re-estimate our baseline when the airline reported to DOT a flight is delayed due to poor weather in Appendix Table A.4 and we find similar results as in Panel B.3 and B.4. We do find that NextGen projects reduce departure delay and arrival delay, however, the effects are not statistically significant. Nevertheless, we do not rule out the possibility is that the extensive fixed effects that we include in our baseline do not leave us enough variation to identify the effect of NextGen.

In summary, we find mixed evidence of the effect of NextGen on air travel performance. We do find that the positive effect of NextGen on reducing air travel time and delay is stronger when a flight experiences low sky ceiling at both ends of the route. The new technologies of NextGen reduce air travel time and delay to a greater extent for flights when the sky ceiling is low at the departing airport compared to other flights. However, we do not find similar effects by examining using other measures of severe weather.

Effect of NextGen for late aircraft Next, we examine how NextGen projects affect flights that have been previously delayed. Prior delay can have a large impact on air travel performance both because hub airlines tend to cluster flights in a short window at their hub airports with minimum buffer time between operations to increase connect-ability (Mayer and Sinai, 2003). In our sample, 56 percent of flights have a buffer time of less than 60 minutes and 12 percent have a buffer time of less than 30 minutes.[26] Moreover, exogenous shocks on the previous operation can be passed through to the next operation and the short buffer time can exacerbate delays and air travel time.

Table 1 Panel B shows that there are 6,957 unique aircraft in our sample. Appendix Table A.2 shows that on average an aircraft has five operations per day. Using the DOT definition of delay (i.e., the actual arrival time is at least 15 minutes later than the scheduled arrival time), we show in Appendix Table A.2 that 36 percent of an aircraft's previous operation is delayed. Although the average previous delay is only 16 minutes, the standard deviation is quite high at 38 minutes. Moreover, DOT requires airlines to report the cause of delays. Appendix Table A.2 shows that 51 percent of flight delays are self-reported due to late aircraft, and the reported delayed time due to late aircraft is 23 minutes with a rather high standard error 42 minutes.

In Table 6 we show how implementing NextGen projects affects air travel performance by estimating equation (1) under the restricted sample when the aircraft of a flight is delayed. Similar to the previous exercise, we estimate equation (1) using sub-samples instead of introducing interaction terms. Panel A we repeat our baseline for convenience.

In Table 6 Panels B.1 through B.4, we report our estimates if the previous aircraft is delayed for more than 15 minutes, 30 minutes, 60 minutes, 90 minutes, and 120 minutes. These cases present 12, 7, 3, 2, and 1 percent of all flights in our sample, respectively. Panels B.1 through B.4 shows that the effect of NextGen on the total travel time (measured by elapsed time plus departure delay) increases 2- to 3-fold when the aircraft has been late. This increased impact proportionately comes from all stages, including the departure delay and taxi time. In addition, different from the baseline, we also find NextGen projects at the origin airport have a much larger and significant effect on airborne time. Similar to the previous exercise for severe weather, we could further restrict our sample to more extreme cases, but the extensive fixed effects do not leave us enough variation to identify the effect of NextGen.

These results suggest that NextGen projects have helped reduce previously delayed flights by both improving ground logistics and possibly the efficiency of the take-off. Our findings suggest that infrastructure improvements have improved air travel performance for the flights that are in the most need and with potentially higher economic costs of delays.

[26] Author's calculation using the computer reservation system (CRS) scheduled arrival of previous operation and scheduled departure of the current operation.

Effect of NextGen by airlines at hub airports Next, we examine how NextGen projects affect flights by the market power status of an airline at a hub airport. Mayer and Sinai (2003) have documented that hub airlines at hub airports fail to internalize the costs that arise for other airlines when clustering flights during peak hours, and Morrison and Winston (2007) quantify these external costs and find that the absolute magnitude of externality can be rather high. Since a large share of flights, 84 percent, travel from or to hub airports (see Appendix Table A.2) and many NextGen airports are also hub airports (see Appendix Table A.1), we are interested to know whether NextGen infrastructure improvements close the gap of air travel performance created by the congestion externality and whether NextGen affects hub and non-hub airlines differently at hub airports. The answers to these questions can be relevant for policymakers since NextGen may create a redistribution of benefits across airlines.

In Table 7 we address the above question by restricting our sample to hub and non-hub airlines at hub airports. In Panel A we repeat the baseline for convenience. In Panels B.1 and B.2 we report results for hub and non-hub airlines separately if the origin airport is a hub airport.[27] Comparing Panels B.1 and B.2, we find the NextGen upgrade at the origin airport affects both hub and non-hub airlines at a hub airport, and proportionately higher for non-hub airlines. Similarly, we re-estimate equation (1) for hub and non-hub airlines separately if the destination airport is a hub airport. Our results in Panel B.3 and B.4 echo the results in Panel B.1 and B.2. We find that the NextGen technologies affect both hub and non-hub airlines but disproportionately greater for non-hub airlines if the destination airport is a hub airport.

The benefits of NextGen on air travel time and delay are not equally distributed across hub and non-hub airlines at a hub airport. Considering the pre-exiting distortion that is created from failing to internalizing congestion externalities of hub airlines to non-hub airlines, NextGen projects not only create welfare gains but also close the gap in terms of air travel time and delay between the hub and non-hub airlines.

In addition to weather, prior delay, and hub status, we also find that implementing NextGen projects improves air travel time and delay during peak hours more than any other time of the day if a flight is en route from or to a busy airport.

[27] We define an airport is a hub airport 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 define a hub-airline at a hub airport similarly using market concentration and cross-checking the status of an airline at a given airport.

5.3 Implications on private and social welfare

In this section we discuss the implications of the reduction in air travel time documented in Section 5.3 on private and social welfare.

We begin by assessing how adopting NextGen projects has affected air travel performance in 2017. In particular, we examine how air travel time and delays would be if the treated airports had not been treated in 2017. This scenario is equivalent to the hypothetical case in which the treated airports had been using the pre-treatment technologies and infrastructure at the beginning of 2014. By the end of 2017, the treated airports have completed 49 of NextGen project categories and 89 of NextGen projects. Using estimates from our baseline in Table 2, we produce the counterfactual travel performance in Table 3. This table suggests the NextGen technologies has reduced air travel time and delays by 2.5 minutes per flight on average, with a sizable reduction from departure delay and taxi time.

Using the counterfactual travel time from Table 3 , we proceed to examine the private benefits of implementing NextGen technologies in the 2017 market. We infer the airline cost savings and passenger cost savings using assumptions from the Cost and Benefit Guideline from FAA (2016b). All our estimates are in 2017 US dollars. For airline cost savings, we evaluate the crew cost at 1,039.58 dollars per hour, fuel and oil cost at 2,443.23 dollars per hour, and maintenance at 793.37 dollars per hour using parameters recommended by FAA (2016b) Section 4 on variable costs in the operating costs. Because of the nature of the definition of fuel costs, we only apply them to airborne and taxi times. Similarly, we only apply maintenance cost to taxi time. Alternatively, for fuel consumption, we could use the fuel use parameter (gallon per hour) in FAA (2016b) Section 3 and infer the fuel cost savings using EIA jet fuel price in 2017 at about 1.5 dollars per gallon. This alternative approach leads to a smaller number.[28] For passengers benefits, we evaluate the time-saving at 48.71 dollar per hour per passenger using the parameter of all-purpose traveler recommended in FAA (2016b) Section 1.

We present the itemized private benefits for each flight excluding propagated benefits to other flights in Table 4 Panel A. This panel reveals that a majority of benefits of NextGen upgrades fall on consumers in the form of time savings. This calculation of passenger benefits could be relatively conservative since our model does not include the benefits of reducing cancellation.

Using the estimates from Panel A and the recommended delay multiplier recommended by FAA (2016b) Section 10, we compute the private welfare gain per flight in Panel B column

[28]The alternative approach would suggest 54 percent of the reported fuel saving.

1. In this column, the crew cost change is based on arrival delay (Panel A column 8), fuel and oil cost change is based on the total saving from the off-ground stages (Panel A column 3, 5, and 6), and the maintenance cost is based on the taxi stages (Panel A column 5 and 6).[29] Lastly, we adjust the cost savings by how much a flight would lead to a delay of other flights using the delay multiplier in Panel B column 1, and compute the inferred cost savings for all flights in our sample in 2017 in Panel B column 2.

Table 4 Panel B suggests that adopting NextGen from 2014 to 2017 would lead to an increase in private welfare in 2017 along by 1.2 billion dollars, with a majority of benefits from passenger time savings by 62 percent, and a fair proportion from fuel saving by 13 percent, and reduction in other variable costs for airlines. Our estimates are rather conservative because our sample only covers about 87 percent of observation of 5 million observations from the On-Time performance in 2017 (see footnote 16 for details), and that the On-Time performance only includes domestic flights. Also, compared to the ex-ante benefit predictions from (FAA, 2016a), our estimate is smaller but still in the same order of magnitude. Lastly, compared to the overall approved budget of NextGen, 20 billion dollars over two decades, our estimates imply that overall present value of benefits from 2014 to 2030 has a good potential to justify the budget since we only account for the benefit in 2017 during the interim of the NextGen program.

To refine the fuel consumption implications and to add to our accounting exercise implications on social benefits, we have matched our aircraft model to the AEDT simulator. AEDT allows us to generate predictions of multiple stages of air travel time duration, fuel consumption, and emissions (such as SO_2, NO_x, and CO_2), and noise for each aircraft model at an airport at a given time.[30] Appendix Figure A.4 shows that we can observe for each stage of the operation, the duration it takes as well as the projected emission and fuel consumption. The taxi stages in the DOT data correspond to the taxi (climb taxi and descend taxi) and the ground stages (climb ground and descend ground) in FAA AEDT data, and the airborne stage in the DOT data correspond to the off-ground stages in FAA AEDT database.[31] The next step is to produce a set of parameters that allow us to infer fuel savings and emission reduction (and maybe noise reduction as well) from taxi and airborne time-savings. Doing so allow us to refine fuel savings in Table 4 and produce additional implications for social benefits.

[29] Alternatively we could exclude maintenance cost and will find a similar magnitude of private benefits.
[30] We are able to match about 90 percent of airport-model pair to the simulator.
[31] Based on the Ruby Gem auto_click: https://rubygems.org/gems/auto_click

6 Robustness

In this section, we examine whether our model is robust to alternative measures of the treatment, alternative specifications of the model, and other potential sources of bias.

Alternative measures of treatment First, we examine whether our results are robust to an alternative approach to measuring the treatment. In our baseline, we define the treatment variables using the number of categories of completed projects (0 to 4) at each end of the route. Since some airports have implemented multiple projects within a category (MRO, PBN, SO, or DC), we could alternatively quantify the treatment of all four categories of NextGen projects separately. Appendix Figure A.2 shows that there is some variation in each category of projects from 2014 to 2017. In Table 2 Panel B we report our results. We find most travel time savings from PBN, SO, and DC. In contrast, we find some mixed evidence for the MRO treatment. Some estimates are imprecise and some are even positive, which is likely due to how four treatment types are correlated with each other (see Appendix Table A.3). Although the number in Panel B appeal to be different from Panel A, our counterfactual in Section 5.3 would be in a similar magnitude if we were to use our estimates from Panel B (see Appendix Table A.6).

Also, we could define treatment variables using the total number of projects completed at each end of a route. Appendix Figure A.1 shows the number of NextGen projects completed from 2010 to 2017. By the end of 2017, a total of 89 projects were completed, compared to 49 categories of projected completed in Figure 2. We report our estimation results in Table 8. We repeat the baseline in Panel A, and our results using the alternative measure in Panel B. Our results are qualitatively similar. The magnitude of most estimates is smaller because the magnitude of the alternative treatment variables is roughly twice that of our baseline treatment variables. Panel C.1 and C.2 suggests our baseline is robust if we only evaluate the treatment at one end of the route. Lastly, in Panel D, we repeat our exercise in Table 2 Panel B and using the number of projects completed for each category of NextGen. Our results are qualitatively and quantitatively similar.

Richer fixed effects Next, we examine whether our results are robust to richer fixed effects. In our baseline, we introduce fixed effects from airlines and routes separately. One would argue that we should use variation within a well-defined type of flight defined by route, airline, the day of the week, and the hour of the day. For example, we should treat flights differently by the day of week, since flights on airlines are more likely to target passengers on a business trip on weekdays and passengers on a personal trip on weekends. Also, as we argued in Section 5.1, we consider that our precision may be driven by the sample size and the relatively small set of fixed effects. By interacting four sets of fixed effects, we assign all flights into 667,389 types and in practice have 36 observations per type of flight. We represent our results in Table 9 Panel B.1. Our estimates are quantitatively similar when compared to the baseline, as repeated in Panel A. Also, the precision of our estimates is similar, which supports the argument that our main results are not driven by the relatively large sample size.

The efficiency of air travel may also depend on the aircraft model. It is likely that air traffic control towers may prioritize a flight with a large capacity, which means the economic costs of increasing air travel time and delay is greater for that flight. It is also likely that a relatively more modern aircraft model is more likely to be equipped with better technologies and works better with the terminal tower, therefore has a shorter air travel time and delay. We compensate for this concern by interacting route by airline and further with the 262 aircraft model trims that we observe in the data. We report our results in Table 9 Panel B.2. Again, our results are quantitatively similar to the baseline in Panel A. This exercise allows us to define 105,253 types of flights which leaves us on average of 226 observations per type of flight. It is again reassuring that our precision barely changes from the baseline.

Further relax the parallel trend assumptions Last, we further relax the assumption of the parallel trend by introducing the interaction of carrier fixed effects with a linear year trend. Doing so further allows us to assume flights from different airlines would evolve at different rates over the years in our sample. Table 9 Panel C shows that our baseline is robust. In addition, our results are similar if we interact linear trend with dummies that indicate if at each end of a route, if the flight is in carried by a hub airline in a hub airport, a non-hub airline in a hub airport, or any airline in a non-hub airport. This exercise controls for the possibility if hub airlines respond to infrastructure upgrades systematically different from other airlines over time. In summary, we find that our results are robust to alternative specifications of the model and other potential sources of bias.

References

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.

Burlig, F., C. Knittel, D. Rapson, M. Reguant, and C. Wolfram (2018). Machine learning from schools about energy efficiency. Technical report.

Callaway, B. and P. H. C. Sant'Anna (2018, August). Difference-in-Differences with Multiple Time Periods and an Application on the Minimum Wage and Employment. Working paper.

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'HaultfÆuille (2018, August). Two-way fixed effects estimators with heterogeneous treatment effects. Working paper.

EPA (2016). 2014 National emissions inventory (NEI), Technical support document (TSD). Technical report, Office of Transportation and Air Quality, U.S. Environmental Protection Agency.

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 (2016a). Economic eavlues for faa investment and regulaotry 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). 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 (2018). FAA aerospace forecasst fiscal year 2018–2038. 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.

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. (2018, September). Difference-in-Differences with Variation in Treatment Timing. Working Paper 25018, National Bureau of Economic Research.

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.

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.

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.

Pindyck, R. S. (2011, 07). Fat Tails, Thin Tails, and Climate Change Policy. *Review of Environmental Economics and Policy 5*(2), 258–274.

Weitzman, M. L. (2009, 02). On modeling and interpreting the economics of catastrophic climate change. *Review of Environmental Economics and Policy 91*(1), 1–19.

Figures

Figure 1: **The Map of NextGen Airports**

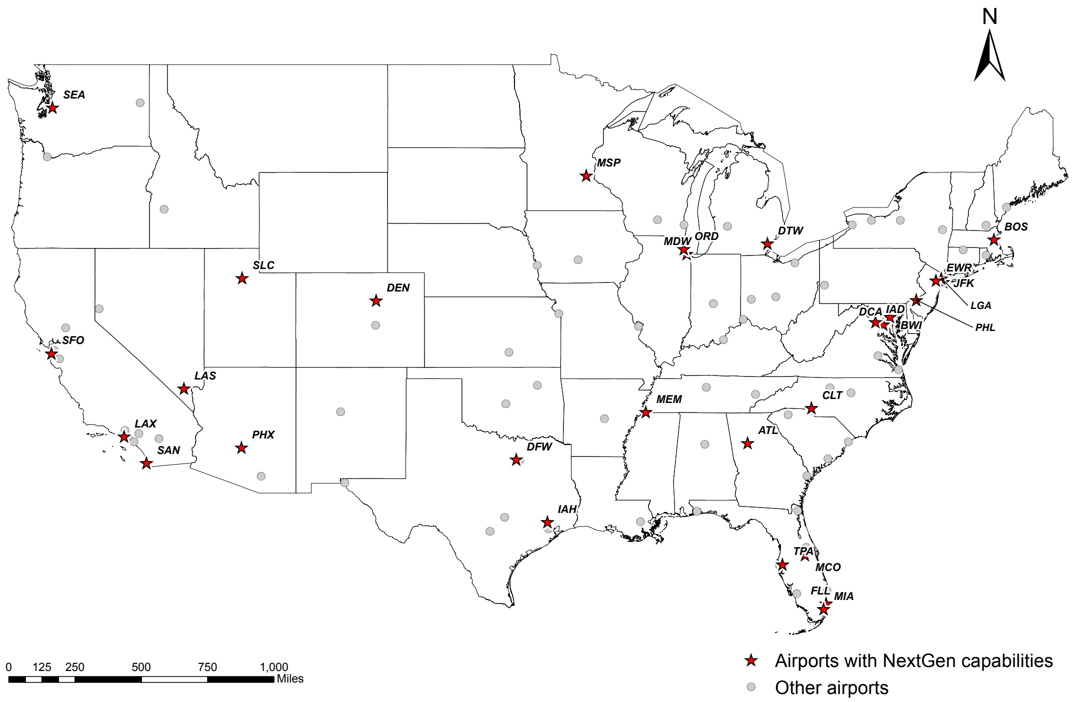
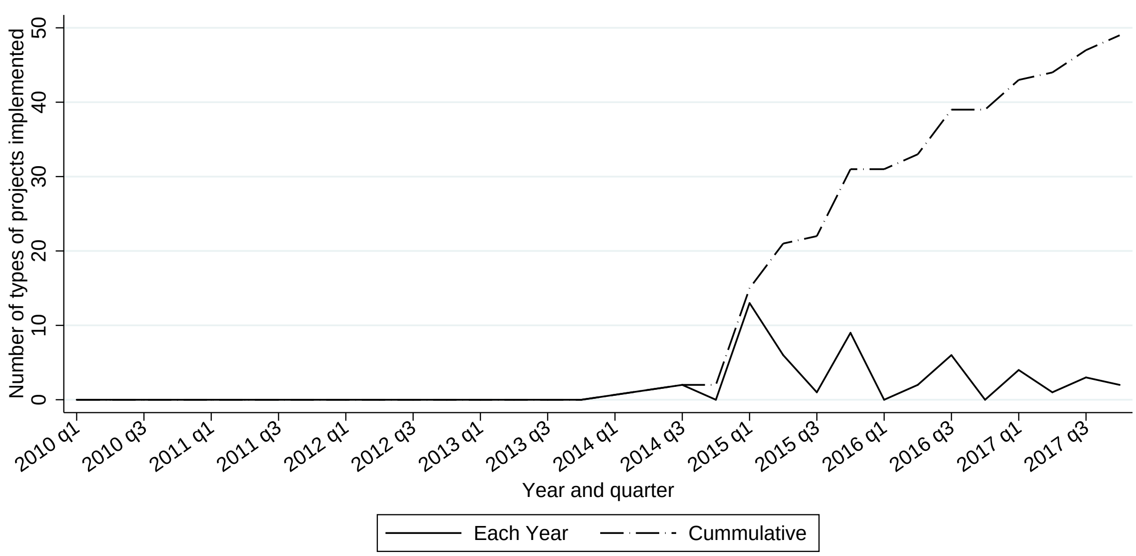


Figure 2: **Categories of NextGen Projects Completed, 2010–2017**



Notes: We count the number of categories of NextGen projects completed for each airport. For example, in 2014, ATL completed two projects in MRO and none of other types of projects. This means, among four types of projects (MRO, PBN, SO, and DC), ATL has completed one type of projects. In Appendix Figure A.1, we plot the number of total projects completed.

Figure 3: **Total Air Travel Time, 2010–2017**

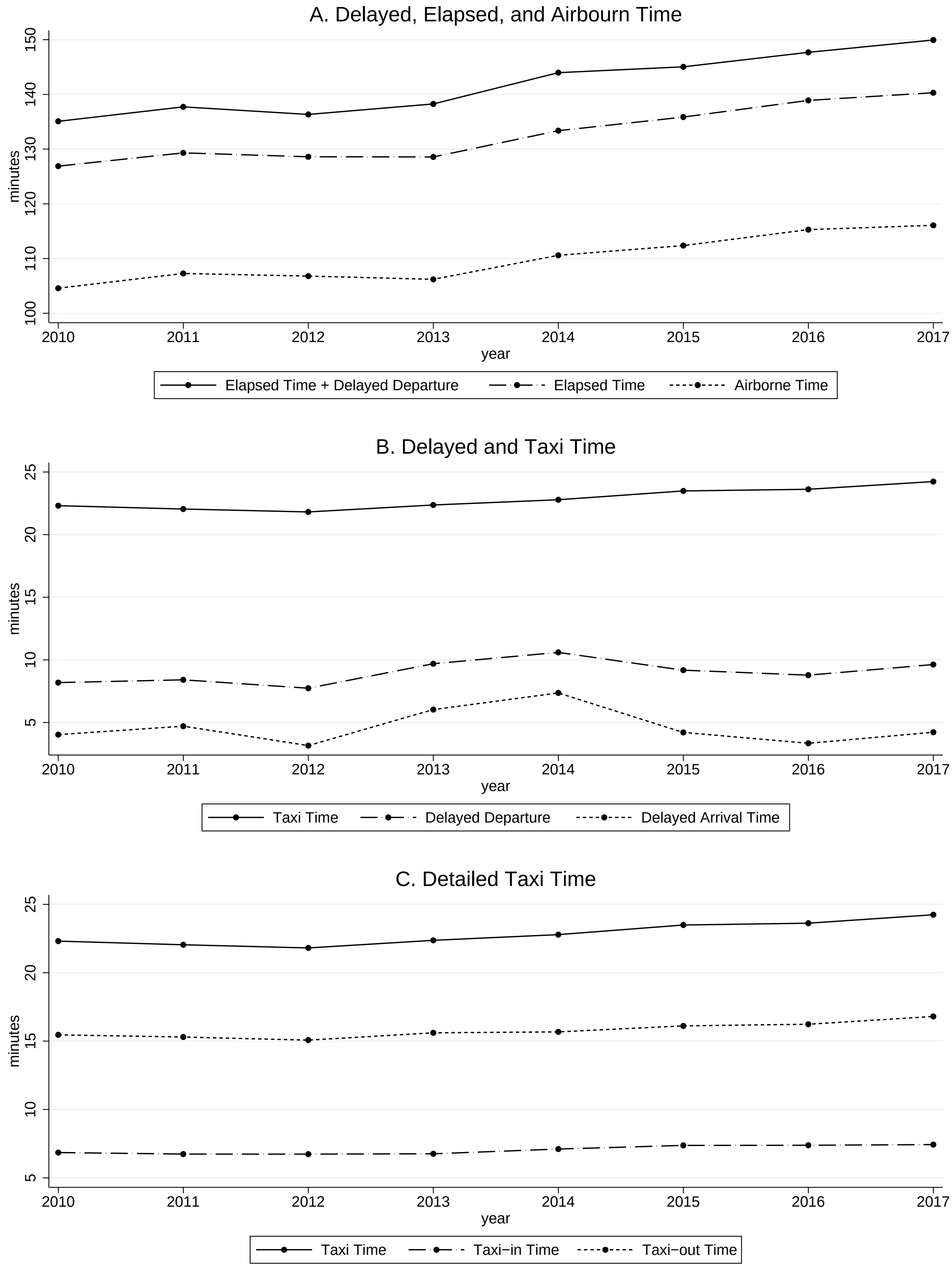
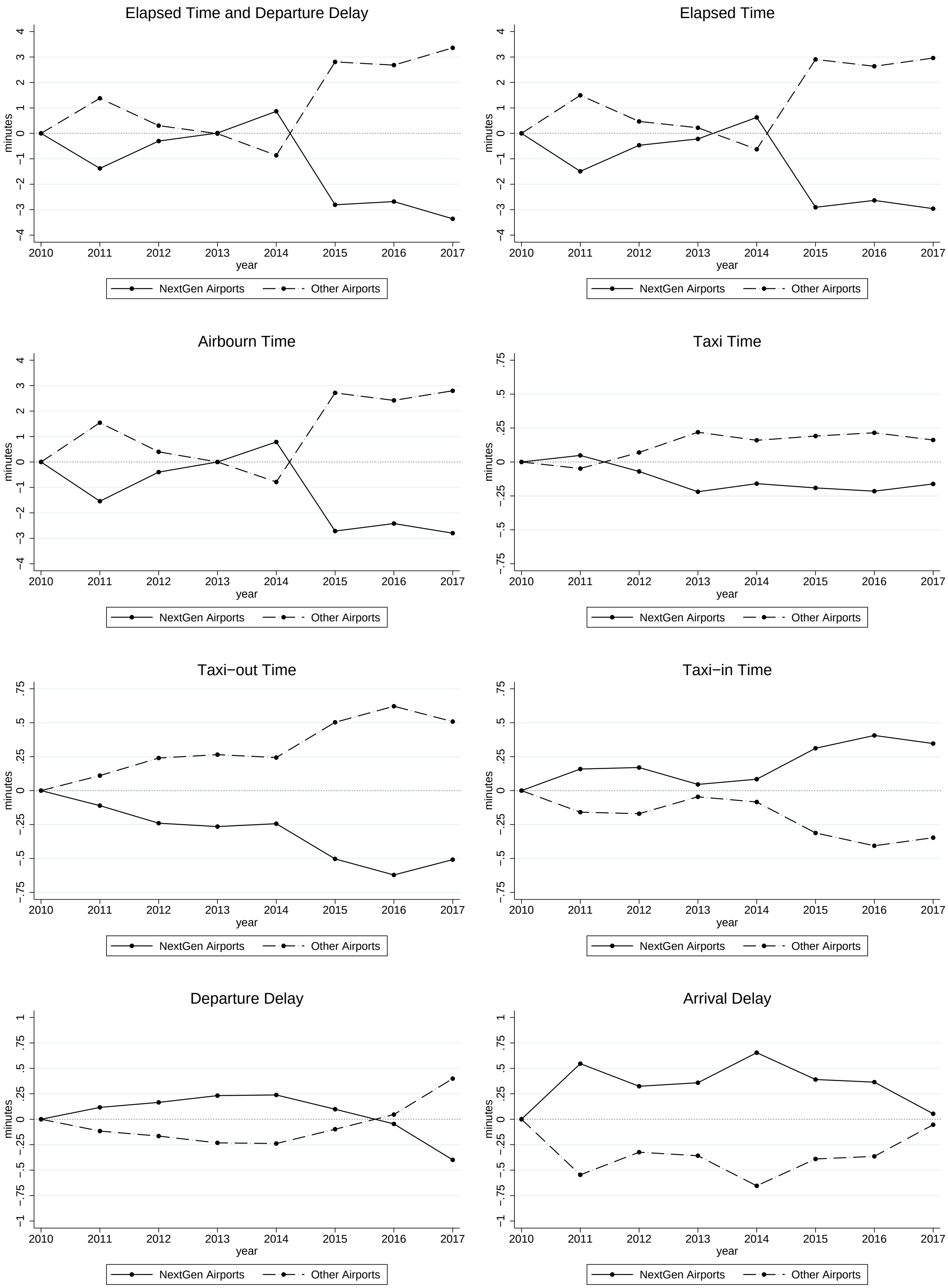


Figure 4: **Detrended Air Travel Time by Airport**

Panel A. Air Travel Time by Whether the Origin Airport is a NextGen Airport



Detrended Air Travel Time by Airport (Continued)

Panel B. Air Travel Time by Whether the Destination Airport is a NextGen Airport

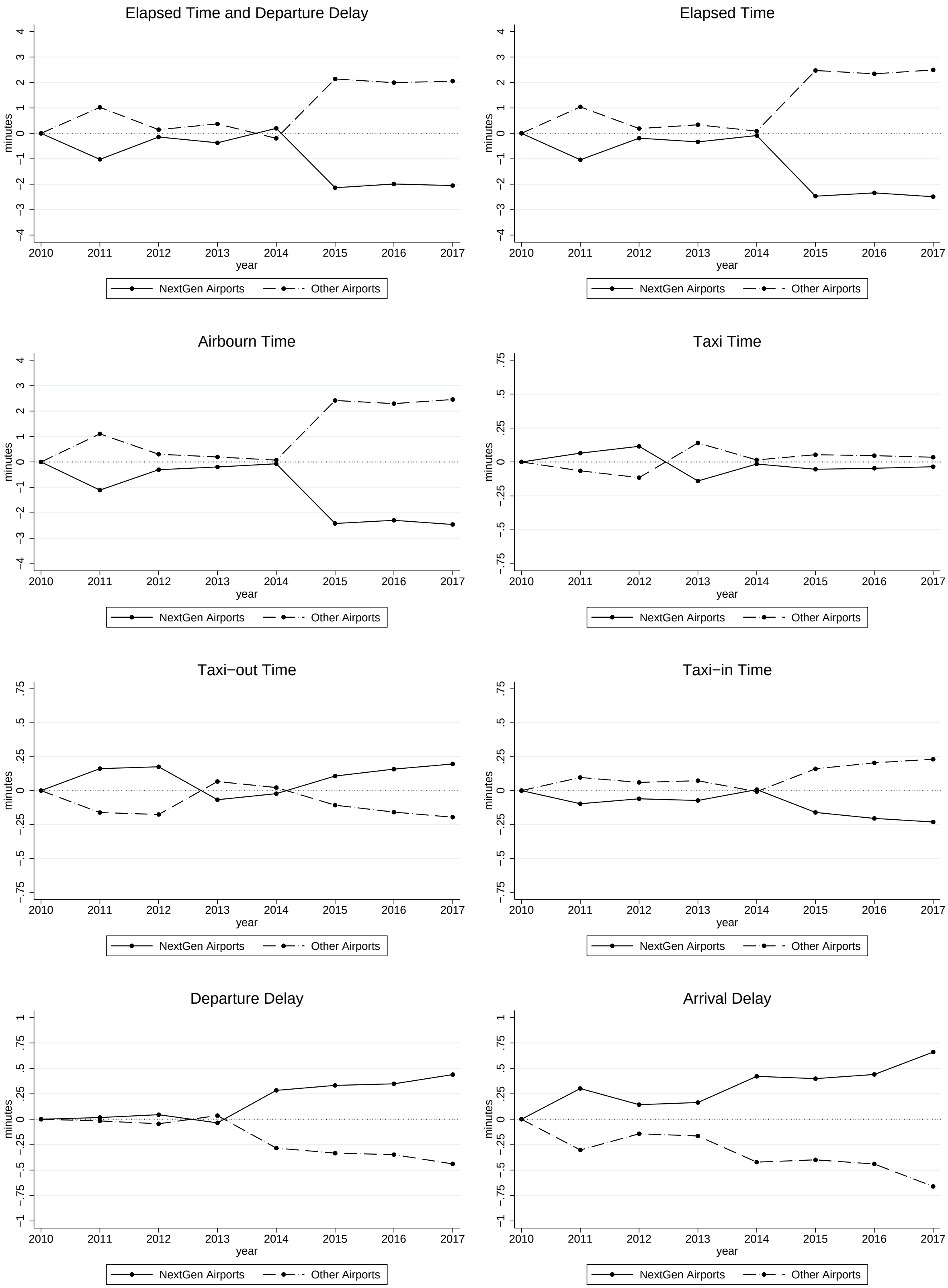
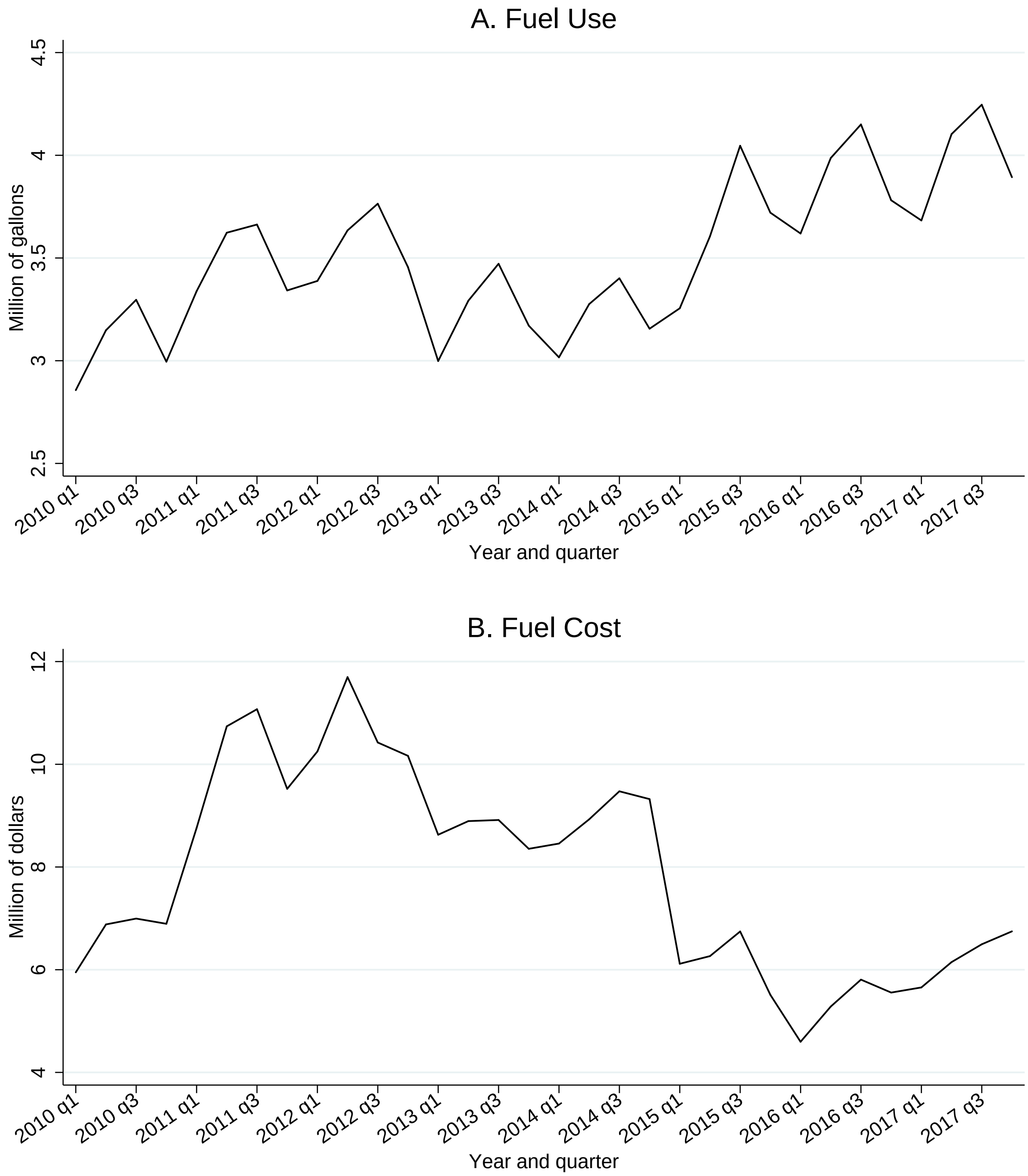


Figure 5: **Total Fuel Use and Fuel Cost, 2010–2017**



Tables

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

Variable	Mean	SD.	Min.	Max.
Panel A. Air Travel Performance and FAA Treatment				
Actual elapsed route time + departure delay (minutes)	142.7	82.5	-20	2,594
Actual elapsed route time (minutes)	133.7	73.2	20	784
Actual airborne time (minutes)	110.9	71.1	6	723
Actual taxi time (minutes)	22.8	10.5	2	481
Actual taxi-out time (minutes)	15.7	9.0	0	278
Actual taxi-in time (minutes)	7.1	5.2	2	414
Departure delay (minutes)	9.0	36.2	-45	2,402
Arrival delay (minutes)	4.7	38.6	-115	2,444
1 = Travel from or to an NextGen airport	0.92	0.27	0	1
Categories of NextGen projects completed	0.34	0.70	0	3
Panel B. Panel Information				
Number of airports			275	
Number of airports with FAA projects			39	
Number of airlines			19	
Number of routes			5,819	
Number of routes by airline			14,194	
Number of routes by airline by hour-of-day by day-of-week			667,389	
Number of routes by airline by aircraft model			38,681	
Number of aircraft models			44	
Number of aircraft model trims			262	
Number of aircraft			6,957	
Number of observations			25,037,569	

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

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 (MRO, PBN, SO, and DC) Implemented								
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
Panel B. Effect of Specific Category of Projects								
NextGen Origin MRO = 1	0.047	0.256**	0.184*	0.071	0.128**	-0.057	-0.208	0.064
	(0.208)	(0.122)	(0.098)	(0.076)	(0.065)	(0.038)	(0.153)	(0.206)
NextGen Origin PBN = 1	-2.119***	-1.409***	-0.784***	-0.625***	-0.683***	0.058	-0.710**	-1.843***
	(0.438)	(0.209)	(0.188)	(0.130)	(0.123)	(0.074)	(0.343)	(0.452)
NextGen Origin SO = 1	-3.460***	-0.781***	-0.453***	-0.328***	-0.371***	0.042	-2.679***	-4.645***
	(0.402)	(0.201)	(0.164)	(0.108)	(0.092)	(0.053)	(0.331)	(0.443)
NextGen Origin DC = 1	-1.555***	0.043	0.370***	-0.327***	-0.266***	-0.061	-1.598***	-2.063***
	(0.313)	(0.143)	(0.114)	(0.085)	(0.067)	(0.051)	(0.248)	(0.301)
NextGen Dest. MRO = 1	0.337	0.470***	0.112	0.358***	-0.047	0.405***	-0.133	-0.091
	(0.236)	(0.132)	(0.106)	(0.076)	(0.057)	(0.044)	(0.167)	(0.231)
NextGen Dest. PBN = 1	-1.336**	-0.235	0.129	-0.364**	-0.062	-0.302***	-1.101***	-1.035**
	(0.570)	(0.273)	(0.199)	(0.149)	(0.116)	(0.081)	(0.390)	(0.508)
NextGen Dest. SO = 1	-3.758***	-0.905***	-0.114	-0.791***	-0.394***	-0.397***	-2.853***	-2.516***
	(0.450)	(0.221)	(0.174)	(0.121)	(0.096)	(0.065)	(0.331)	(0.457)
NextGen Dest. DC = 1	-1.933***	-0.775***	-0.063	-0.711***	-0.313***	-0.399***	-1.158***	-1.563***
	(0.282)	(0.146)	(0.109)	(0.097)	(0.072)	(0.058)	(0.219)	(0.276)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. All regressions include fixed effects of carrier, route (a pair of origin and destination airports), year by month, day of a month, day of a week, and hour of a day. All regressions also include origin airport fixed effects interacted with a linear year trend, and destination fixed effects interacted with a linear year trend. All regressions exclude year 2013 as well as a flight if it travels to or from an airport within three quarters before the completion of a NextGen project. In Panel A, we measure treatment as the total categories of NextGen (MRO, PBN, SO, and DC) implemented at the origin and the destination airports. In Panel B, we include each category of NextGen completed at the origin and the destination airports. We present results from richer models and alternative assumptions in robustness Table 8 and 9.

Table 3: **The effect of NextGen from 2014 to 2017 on air travel time**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unit (minutes)	elapsed time + departure delay	elapsed time	airborne time	taxi time	taxi-out time	taxi-in time	departure delay	arrival delay
Actual	149.04	139.30	115.78	23.51	16.32	7.18	9.6	4.13
Change	2.54	0.33	-0.02	0.35	0.33	0.02	2.22	2.72
Change (percent)	1.71%	0.24%	-0.01%	1.47%	2.00%	0.26%	23.17%	65.94%
Counterfactual	151.42	139.62	115.77	23.85 3	16.65	7.20	11.80	6.85
Number of obs.							4,736,642	

Notes: The first row reports the actual travel time in 2017. This table reports the counterfactual travel time and delay if the treated airports in 2017 had been untreated, i.e., as if the airports were in the beginning of 2014. We use baseline in Table 2. In Appendix Table A.6 we use estimates in Table 2 Panel B and estimates in Table 8.

Table 4: **The benefit of reducing delay of NextGen in 2017**

Panel A. Itemized Benefit Per Flight

	(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.1 Variable Airline Cost Per Flight (2017 USD)							
Crew cost	44.15	5.70	-0.29	5.99	5.66	0.33	38.45	47.15
Fuel and oil	–	–	-0.29	14.08	13.31	0.77	–	–
Maintenance	–	–	–	4.57	4.32	0.25	–	–
	Panel A.2 Passenger Cost Per Flight (2017 USD)							
Time saving	128.10	16.53	-0.842	17.37	16.42	0.95	111.50	136.80
Num. of obs.							4,736,642	

Panel B. Benefit of NextGen via Reducing Delay and Air Travel Time in 2017

	Cost saving per flight (2017 USD)	Cost saving in 2017 (million 2017 USD)	Cost saving in 2017 (percentage)
Crew cost	70.72	258.58	21.4 %
Fuel and oil	41.20	150.64	12.5 %
Maintenance	13.71	50.13	4.15 %
Passenger	205.14	750.04	62.02 %
Total private benefits	330.77	1209.40	

Notes: This table reports the cost of the increase in the counterfactual travel time and delay if the treated airports in 2017 had been untreated using estimates from Table 3. All dollar in the above table is assessed using 2017 dollar. In Panel A.1, using FAA (2016a) Section 4, we evaluate crew costs at 1039.58 dollar per hour per flight, fuel and oil at 2443.23 dollar per hour per flight, and maintenance at 793.37 dollar per hour per flight. In Panel A.2, based on FAA (2016a) Section 1, we evaluate cost per passenger (personal or business combined) at 48.71 dollar per hour and extract number of passengers from DB1B.
For Panel B, we use arrival delay of Panel A.1 to compute the crew cost saving of airline, the summation of airborne, taxi-in, and taxi-out to compute the fuel saving for the airline, and the summation of taxi-out and taxi-in to compute the changes in maintenance cost, we use arrival delay to compute the time saving for passengers, and then we adjust them using the delay multiplier from FAA (2016a) Section 10 ranges from 1.4 to 1.9. Moreover, the second column uses the 3.6 million observations in 2017 in our regression. As we explained in footnote 14, our data represents roughly 58 percent of flight. In 2017, this 3.6 million observations represent 67 percent of flights.

Table 5: **Conditional Effect of NextGen on Air Travel Time by Weather Condition**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B.1 Sky Ceiling < 1,500 Feet (Origin)				
NextGen Origin	-1.779***	-0.363***	-0.141*	-0.223***	-0.182***	-0.041	-1.416***	-2.010***
	(0.274)	(0.115)	(0.085)	(0.070)	(0.063)	(0.031)	(0.235)	(0.269)
NextGen Dest.	-0.898***	-0.094	-0.045	-0.049	-0.113*	0.065*	-0.804***	-0.679**
	(0.285)	(0.128)	(0.098)	(0.077)	(0.066)	(0.038)	(0.239)	(0.276)
Number of obs.	2,255,847	2,255,847	2,255,847	2,255,847	2,255,847	2,255,847	2,255,847	2,255,847
R-squared	0.75	0.97	0.99	0.21	0.20	0.22	0.06	0.07
				B.2 Sky Ceiling < 1,500 Feet (Destination)				
NextGen Origin	-0.758**	0.037	0.090	-0.053	-0.050	-0.003	-0.795***	-1.053***
	(0.305)	(0.122)	(0.088)	(0.076)	(0.067)	(0.037)	(0.250)	(0.296)
NextGen Dest.	-1.206***	0.155	0.163*	-0.008	-0.073	0.065*	-1.361***	-0.987***
	(0.291)	(0.120)	(0.090)	(0.071)	(0.058)	(0.038)	(0.246)	(0.279)
Number of obs.	2,121,711	2,121,711	2,121,711	2,121,711	2,121,711	2,121,711	2,121,711	2,121,711
R-squared	0.73	0.96	0.98	0.21	0.18	0.23	0.08	0.09
				B.3 Visibility < 5 mile (Origin)				
NextGen Origin	-0.015	0.200	-0.119	0.319***	0.355***	-0.036	-0.215	-0.332
	(0.353)	(0.127)	(0.089)	(0.081)	(0.075)	(0.030)	(0.305)	(0.343)
NextGen Dest.	-1.022***	-0.020	0.082	-0.102	-0.042	-0.060	-1.002***	-0.826**
	(0.347)	(0.138)	(0.094)	(0.088)	(0.078)	(0.039)	(0.293)	(0.335)
Number of obs.	1,408,706	1,408,706	1,408,706	1,408,706	1,408,706	1,408,706	1,408,706	1,408,706
R-squared	0.70	0.96	0.98	0.20	0.19	0.21	0.07	0.07
				B.4 Visibility < 5 mile (Destination)				
NextGen Origin	-1.165***	-0.094	-0.037	-0.057	-0.047	-0.009	-1.070***	-1.573***
	(0.352)	(0.126)	(0.089)	(0.079)	(0.068)	(0.039)	(0.298)	(0.354)
NextGen Dest.	0.296	0.678***	0.390***	0.287***	0.088	0.200***	-0.382	0.196
	(0.361)	(0.134)	(0.093)	(0.086)	(0.066)	(0.049)	(0.302)	(0.365)
Number of obs.	1,355,643	1,355,643	1,355,643	1,355,643	1,355,643	1,355,643	1,355,643	1,355,643
R-squared	0.67	0.96	0.98	0.20	0.18	0.21	0.08	0.09

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. In Panels B.1 and B.2, we restrict to flights if the sky ceiling at the origin or the destination airport is less than 1,500 feet. In Panels B.3 and B.4, we restrict to flights if the visibility at the origin or the destination airport is less than 5 miles.

Table 6: **Conditional Effect of NextGen on Air Travel Time 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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B.1 Prior Delay > 15 minutes				
NextGen Origin	-2.563***	-0.452***	-0.173***	-0.279***	-0.274***	-0.005	-2.111***	-2.726***
	(0.298)	(0.093)	(0.067)	(0.061)	(0.052)	(0.029)	(0.279)	(0.293)
NextGen Dest.	-1.789***	-0.289***	-0.076	-0.213***	-0.175***	-0.037	-1.500***	-1.640***
	(0.286)	(0.100)	(0.076)	(0.067)	(0.053)	(0.036)	(0.269)	(0.280)
Number of obs.	3,045,782	3,045,782	3,045,782	3,045,782	3,045,782	3,045,782	3,045,782	3,045,782
R-squared	0.62	0.96	0.98	0.21	0.21	0.21	0.07	0.07
				B.2 Prior Delay > 30 minutes				
NextGen Origin	-3.012***	-0.630***	-0.236***	-0.394***	-0.388***	-0.006	-2.382***	-3.084***
	(0.371)	(0.105)	(0.075)	(0.071)	(0.060)	(0.033)	(0.352)	(0.362)
NextGen Dest.	-2.013***	-0.379***	-0.081	-0.299***	-0.221***	-0.077*	-1.633***	-1.783***
	(0.363)	(0.114)	(0.085)	(0.077)	(0.061)	(0.042)	(0.342)	(0.355)
Number of obs.	1,770,528	1,770,528	1,770,528	1,770,528	1,770,528	1,770,528	1,770,528	1,770,528
R-squared	0.58	0.96	0.98	0.20	0.21	0.20	0.08	0.08
				B.3 Prior Delay > 60 minutes				
NextGen Origin	-3.340***	-0.757***	-0.278***	-0.479***	-0.464***	-0.015	-2.583***	-3.322***
	(0.436)	(0.117)	(0.081)	(0.078)	(0.066)	(0.037)	(0.414)	(0.427)
NextGen Dest.	-2.241***	-0.359***	-0.056	-0.302***	-0.206***	-0.096**	-1.882***	-1.964***
	(0.434)	(0.126)	(0.091)	(0.087)	(0.067)	(0.047)	(0.412)	(0.428)
Number of obs.	1,188,996	1,188,996	1,188,996	1,188,996	1,188,996	1,188,996	1,188,996	1,188,996
R-squared	0.55	0.95	0.98	0.19	0.20	0.20	0.10	0.10
				B.4 Prior Delay > 90 minutes				
NextGen Origin	-4.456***	-0.717***	-0.306***	-0.411***	-0.414***	0.003	-3.739***	-4.342***
	(0.644)	(0.142)	(0.096)	(0.099)	(0.085)	(0.046)	(0.631)	(0.640)
NextGen Dest.	-2.323***	-0.263	-0.038	-0.226*	-0.186**	-0.040	-2.060***	-2.038***
	(0.690)	(0.161)	(0.110)	(0.116)	(0.088)	(0.065)	(0.673)	(0.683)
Number of obs.	480,384	480,384	480,384	480,384	480,384	480,384	480,384	480,384
R-squared	0.52	0.95	0.98	0.19	0.20	0.20	0.16	0.16
				B.5 Prior Delay > 120 minutes				
NextGen Origin	-4.505***	-0.523***	-0.198*	-0.325***	-0.319***	-0.007	-3.981***	-4.471***
	(0.871)	(0.161)	(0.109)	(0.111)	(0.099)	(0.053)	(0.856)	(0.870)
NextGen Dest.	-2.781***	-0.181	-0.027	-0.154	-0.107	-0.047	-2.600***	-2.508***
	(0.942)	(0.183)	(0.127)	(0.130)	(0.102)	(0.074)	(0.929)	(0.936)
Number of obs.	285,103	285,103	285,103	285,103	285,103	285,103	285,103	285,103
R-squared	0.51	0.95	0.98	0.19	0.20	0.20	0.21	0.20

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. Panel B.1–B.5 estimate the baseline when the previous operation of an aircraft is delayed (actual arrival compared to schedule arrival) by more than 15, 30, 60, 90, and 120 minutes.

Table 7: **Conditional Effect of NextGen on Air Travel Time by Hub Airlines**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B.1 Hub Airport and Hub Airline (Origin)				
NextGen Origin	-1.332***	-0.368**	-0.065	-0.304***	-0.199***	-0.104**	-0.964***	-2.094***
	(0.315)	(0.147)	(0.110)	(0.083)	(0.073)	(0.047)	(0.243)	(0.311)
NextGen Dest.	-0.812**	0.030	0.028	0.001	-0.026	0.028	-0.842***	-0.641*
	(0.339)	(0.190)	(0.150)	(0.111)	(0.083)	(0.069)	(0.245)	(0.328)
Number of obs.	6,539,157	6,539,157	6,539,157	6,539,157	6,539,157	6,539,157	6,539,157	6,539,157
R-squared	0.87	0.97	0.99	0.21	0.17	0.23	0.04	0.04
				B.2 Hub Airport and Non-hub Airline (Origin)				
NextGen Origin	-2.292***	-0.049	0.271***	-0.320***	-0.348***	0.029	-2.243***	-2.651***
	(0.220)	(0.109)	(0.084)	(0.072)	(0.058)	(0.039)	(0.176)	(0.223)
NextGen Dest.	-0.823**	-0.342**	-0.077	-0.265***	-0.362***	0.098*	-0.481*	-0.704**
	(0.335)	(0.161)	(0.125)	(0.101)	(0.076)	(0.057)	(0.264)	(0.344)
Number of obs.	6,998,051	6,998,051	6,998,051	6,998,051	6,998,051	6,998,051	6,998,051	6,998,051
R-squared	0.78	0.97	0.99	0.18	0.16	0.24	0.04	0.04
				B.3 Hub Airport and Hub Airline (Destination)				
NextGen Origin	-0.405	-0.210	-0.111	-0.099	-0.117	0.018	-0.195	-0.807**
	(0.343)	(0.177)	(0.137)	(0.103)	(0.086)	(0.061)	(0.274)	(0.337)
NextGen Dest.	-2.493***	-0.626***	-0.438***	-0.188*	-0.236***	0.048	-1.867***	-2.347***
	(0.277)	(0.139)	(0.103)	(0.100)	(0.070)	(0.063)	(0.209)	(0.274)
Number of obs.	6,522,845	6,522,845	6,522,845	6,522,845	6,522,845	6,522,845	6,522,845	6,522,845
R-squared	0.82	0.97	0.99	0.20	0.18	0.15	0.03	0.03
				B.4 Hub Airport and Non-hub Airline (Destination)				
NextGen Origin	-1.023***	-0.248	-0.070	-0.178*	-0.205***	0.027	-0.774***	-1.312***
	(0.362)	(0.160)	(0.118)	(0.099)	(0.079)	(0.054)	(0.287)	(0.369)
NextGen Dest.	-2.362***	-0.262**	-0.092	-0.170**	-0.158***	-0.012	-2.100***	-1.851***
	(0.228)	(0.113)	(0.086)	(0.073)	(0.056)	(0.042)	(0.181)	(0.231)
Number of obs.	6,986,915	6,986,915	6,986,915	6,986,915	6,986,915	6,986,915	6,986,915	6,986,915
R-squared	0.78	0.97	0.99	0.19	0.20	0.16	0.04	0.04

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. If a flight travels from a hub airport, we repeat the baseline if the airline of the flight is a hub airline at that airport in Panel B.1, and if the airline is not a hub airline at that airport in Panel B.2. If a flight travel to a hub airport, we repeat the baseline if the airline of the flight is a hub airline at that airport in Panel B.3, and if the airline is not a hub airline at that airport in Panel B.4.

Table 8: **Alternative Measure for the Treatment Variables**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B. Number of Projects Completed				
NextGen Origin	-0.946***	-0.171***	-0.127***	-0.044	-0.024	-0.020	-0.775***	-1.077***
	(0.118)	(0.060)	(0.048)	(0.033)	(0.027)	(0.019)	(0.098)	(0.122)
NextGen Dest.	-0.421***	0.145**	0.182***	-0.037	-0.112***	0.076***	-0.566***	-0.539***
	(0.111)	(0.063)	(0.045)	(0.038)	(0.028)	(0.021)	(0.084)	(0.121)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				C.1 Only Include NextGen at Origin				
NextGen Origin	-1.159***	-0.184**	-0.035	-0.150***	-0.125***	-0.025	-0.975***	-1.505***
	(0.155)	(0.077)	(0.058)	(0.044)	(0.037)	(0.024)	(0.126)	(0.158)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				C.2 Only Include NextGen at Destination				
NextGen Dest.	-1.049***	-0.101	0.049	-0.150***	-0.158***	0.008	-0.948***	-0.852***
	(0.150)	(0.078)	(0.061)	(0.048)	(0.037)	(0.028)	(0.116)	(0.154)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				D. Effect of the Number of Projects for Each NextGen Category				
NextGen Origin	-0.046	0.274***	0.119	0.155***	0.178***	-0.022	-0.320**	-0.115
Num. of MRO	(0.155)	(0.084)	(0.073)	(0.051)	(0.045)	(0.027)	(0.126)	(0.169)
NextGen Origin	-1.920***	-1.348***	-0.771***	-0.577***	-0.643***	0.066	-0.572*	-1.633***
Num. of PBN	(0.442)	(0.213)	(0.187)	(0.131)	(0.121)	(0.073)	(0.344)	(0.453)
NextGen Origin	-2.851***	-1.200***	-0.924***	-0.276***	-0.289***	0.013	-1.651***	-3.114***
Num. of SO	(0.370)	(0.212)	(0.169)	(0.101)	(0.078)	(0.061)	(0.291)	(0.379)
NextGen Origin	-1.628***	-0.046	0.318***	-0.363***	-0.311***	-0.052	-1.583***	-2.136***
Num. of DC	(0.309)	(0.143)	(0.116)	(0.086)	(0.069)	(0.049)	(0.249)	(0.300)
NextGen Dest.	0.678***	0.621***	0.333***	0.288***	-0.124***	0.412***	0.056	-0.324*
Num. of MRO	(0.186)	(0.100)	(0.074)	(0.059)	(0.044)	(0.031)	(0.132)	(0.187)
NextGen Dest.	-1.115**	-0.154	0.142	-0.296**	-0.032	-0.264***	-0.961**	-0.824
Num. of PBN	(0.567)	(0.267)	(0.196)	(0.148)	(0.117)	(0.077)	(0.390)	(0.521)
NextGen Dest.	-2.451***	-0.598***	-0.103	-0.495***	0.000	-0.495***	-1.853***	-0.464
Num. of SO	(0.382)	(0.187)	(0.157)	(0.102)	(0.079)	(0.054)	(0.301)	(0.396)
NextGen Dest.	-2.133***	-0.911***	-0.130	-0.781***	-0.284***	-0.497***	-1.222***	-1.511***
Num. of DC	(0.283)	(0.147)	(0.112)	(0.097)	(0.072)	(0.057)	(0.220)	(0.277)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. In Panel A we repeat the baseline. In Panel B, we measure the treatment using total number of projects completed at each end of a route as showed in Figure A.1. Panel C.1 and C.2 include only NextGen treatment variable at the origin or the destination airport. Panel D re-estimate Panel B of Table 2 using the number of project completed for each category of NextGen at the origin and the destination airports.

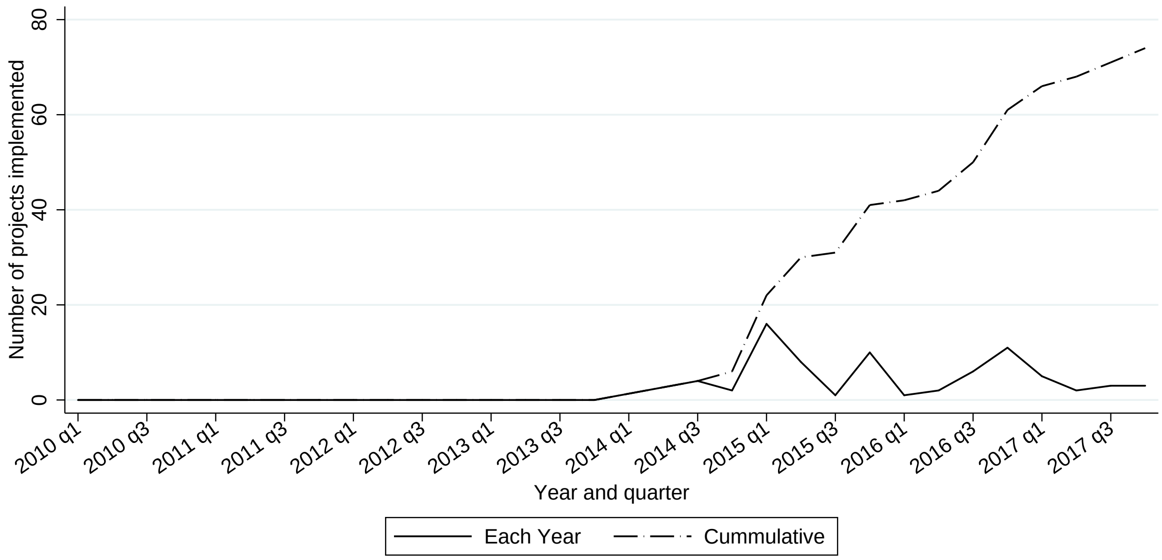
Table 9: **Richer Fixed Effects and Alternative Identifying Assumptions**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
B.1 Richer Fixed Effects: Route by Carrier by Day-of-week by Hour-of-day								
NextGen Origin	-1.223***	-0.232***	-0.055	-0.177***	-0.138***	-0.039	-0.992***	-1.585***
	(0.163)	(0.080)	(0.061)	(0.046)	(0.039)	(0.024)	(0.132)	(0.164)
NextGen Dest.	-1.183***	-0.108	0.036	-0.145***	-0.145***	0.001	-1.075***	-1.012***
	(0.156)	(0.082)	(0.065)	(0.051)	(0.038)	(0.031)	(0.119)	(0.159)
Number of obs.	24,997,532	24,997,532	24,997,532	24,997,532	24,997,532	24,997,532	24,997,532	24,997,532
R-squared	0.81	0.97	0.99	0.30	0.28	0.30	0.08	0.08
B.2 Richer Fixed Effects: Route by Carrier by Aircraft Model Trim								
NextGen Origin	-1.258***	-0.135*	0.011	-0.147***	-0.141***	-0.006	-1.123***	-1.547***
	(0.160)	(0.080)	(0.059)	(0.045)	(0.037)	(0.024)	(0.132)	(0.162)
NextGen Dest.	-1.237***	-0.103	0.040	-0.143***	-0.174***	0.031	-1.134***	-1.035***
	(0.156)	(0.082)	(0.063)	(0.050)	(0.038)	(0.029)	(0.122)	(0.157)
Number of obs.	25,034,463	25,034,463	25,034,463	25,034,463	25,034,463	25,034,463	25,034,463	25,034,463
r2	0.80	0.97	0.99	0.24	0.23	0.24	0.04	0.04
C.1 Further Relax Parallel Trend Assumption: Add Airline Fixed Effects by Linear Trend								
NextGen Origin	-1.138***	-0.166**	-0.019	-0.147***	-0.144***	-0.003	-0.972***	-1.498***
	(0.154)	(0.076)	(0.057)	(0.043)	(0.037)	(0.024)	(0.122)	(0.154)
NextGen Dest.	-1.037***	-0.089	0.058	-0.147***	-0.174***	0.027	-0.947***	-0.900***
	(0.148)	(0.077)	(0.059)	(0.048)	(0.038)	(0.028)	(0.113)	(0.149)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
C.2 Further Relax Parallel Trend Assumption: Add Airline Hub Status by Linear Trend								
NextGen Origin	-1.287***	-0.186**	-0.028	-0.158***	-0.133***	-0.025	-1.101***	-1.625***
	(0.153)	(0.078)	(0.058)	(0.044)	(0.038)	(0.024)	(0.123)	(0.155)
NextGen Dest.	-1.200***	-0.117	0.046	-0.162***	-0.174***	0.012	-1.084***	-1.034***
	(0.149)	(0.078)	(0.061)	(0.047)	(0.037)	(0.028)	(0.114)	(0.150)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. In Panel B.1, we use fixed effects of route by carrier by day-of-week by hour-of-day (667,389 of them) instead of introducing four sets of fixed effects separately. In Panel B.2, we use fixed effect of route by carrier instead of introducing two sets separately, and we further interact these fixed effects with the aircraft model trim fixed effects (resulting 38,681 cells). In Panel C.1, we add to the baseline the interactions of airline fixed effect and a linear year trend. In Panel C.2, we add the baseline the interactions of hub status fixed effects (hub airport hub airline, hub airport non-hub airline, and non-hub airport) at both each of the route and a linear trend.

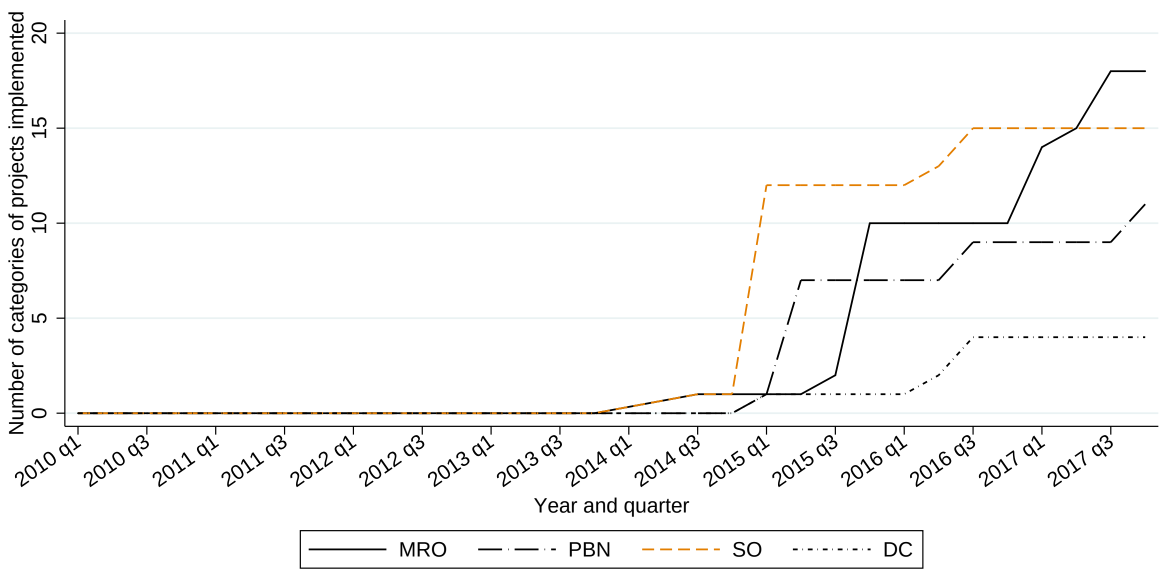
A Appendix Figures and Tables

Figure A.1: **Number of NextGen Projects Completed, 2010–2017**



Notes: We count the number of total NextGen projects completed for each airport. For example, in 2014, ATL has completed two projects in MRO and none of other types of projects. This means, ATL completed two projects in 2014.

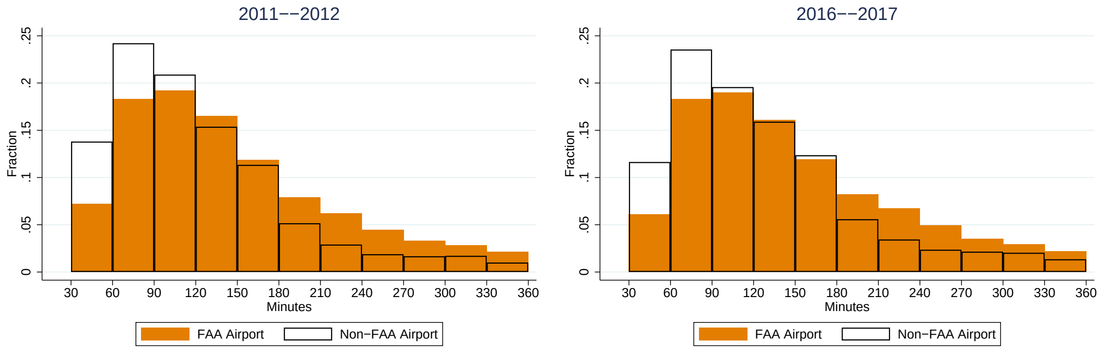
Figure A.2: **Number of Each Category of NextGen Projects Completed, 2010–2017**



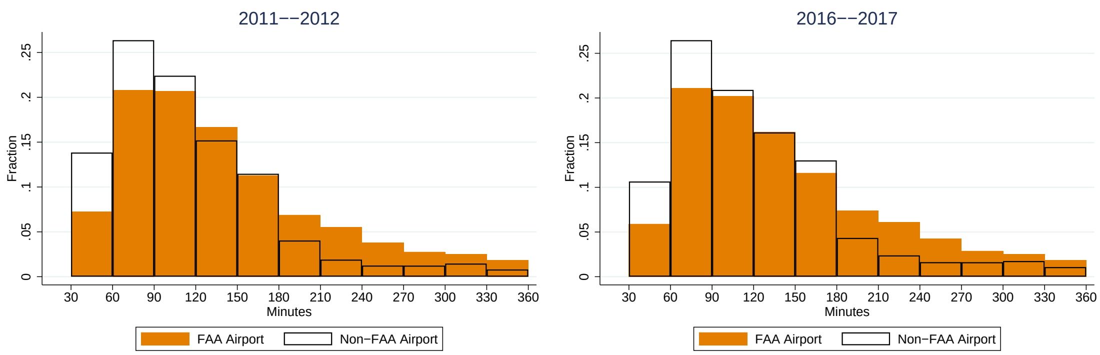
Notes: We count for each airport the number of total NextGen projects completed within each category of NextGen upgrade – MRO, PBN, SO, DC.

Figure A.3: **Distribution of Air Travel Time by Treatment**

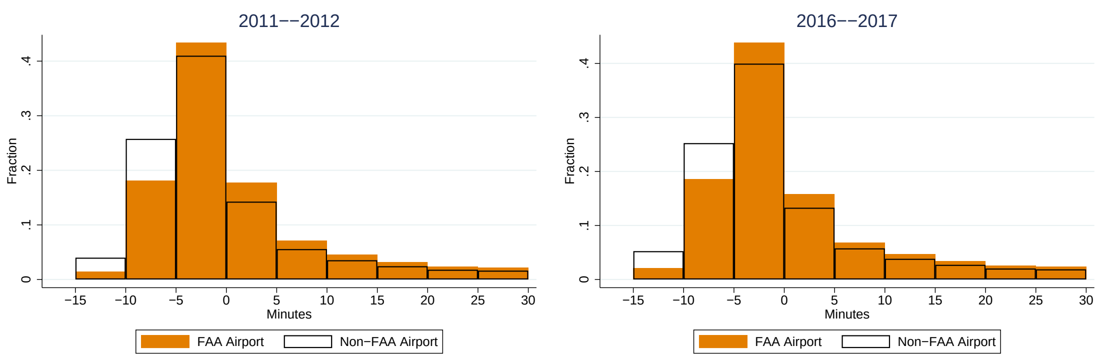
Panel A. Elapse Time and Departure Delay



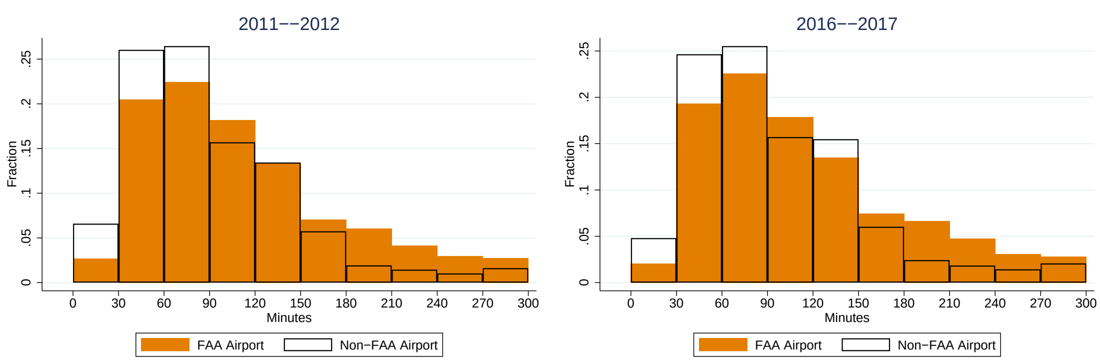
Panel B. Elapse Time



Panel C. Departure Delay Time

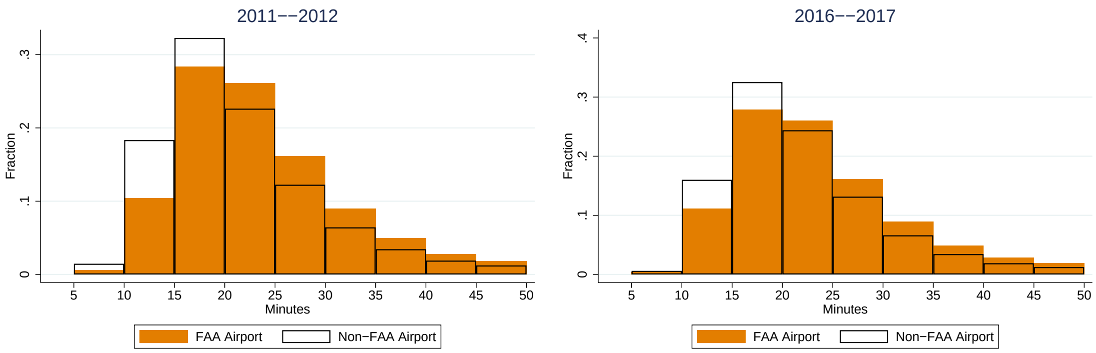


Panel D. Airborne Time

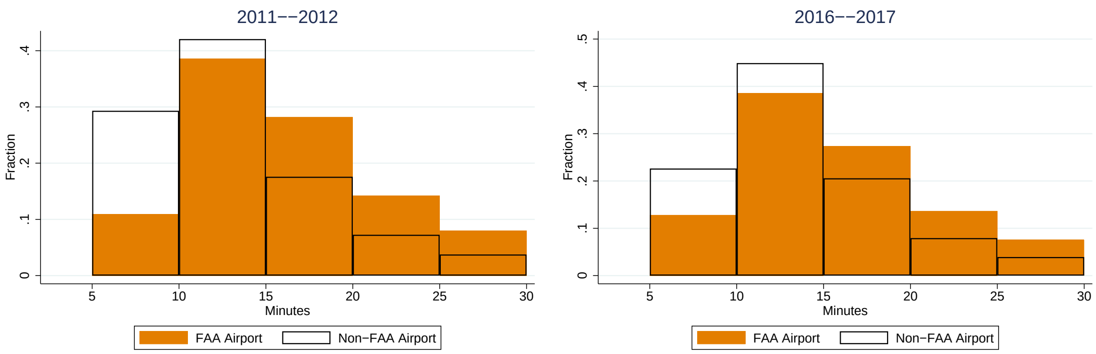


Distribution of Air Travel Time by Treatment (Continued)

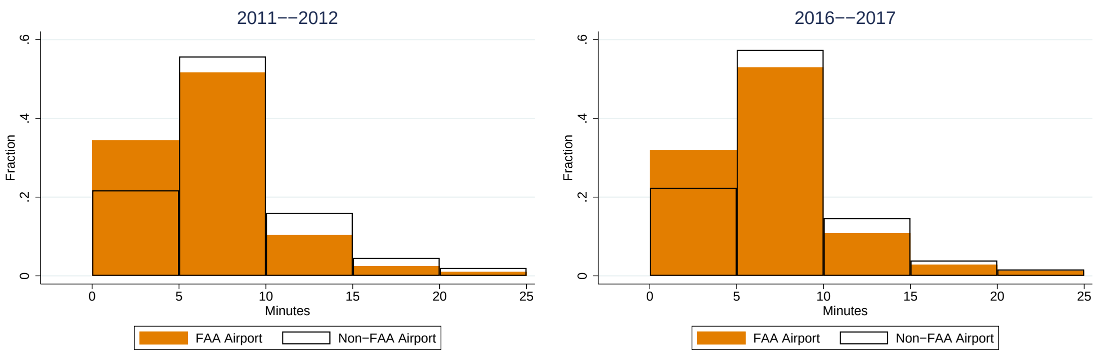
Panel E. Taxi Time



Panel F. Taxi-out Time



Panel G. Taxi-in Time



Panel H. Arrival Delay

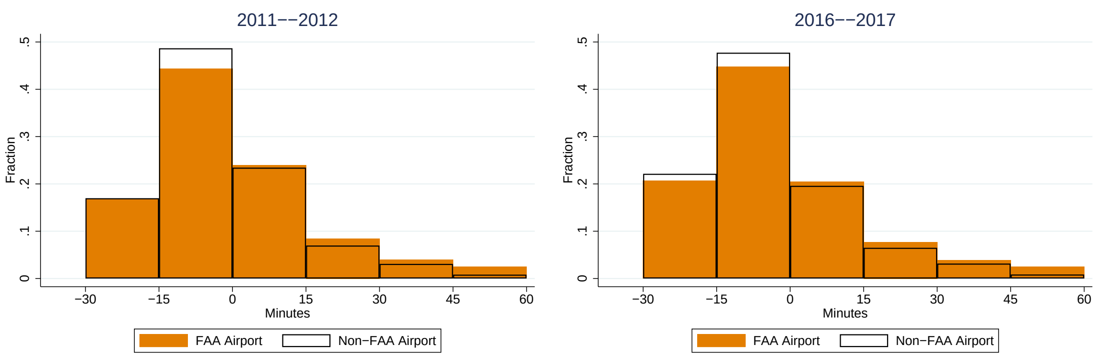


Figure A.4: **A Sample of AEDT Data**

Emissions | Speciated Organic Gases

Event ID	Equipment Type	Departure Airport	Arrival Airport	Mode	Fuel (g)	Distance (km)	Duration	CO (g)	HC (g)	TOG (g)	VOC (g)	NMHC (g)	NOx (g)
100000	H500D	ATL		Above 10000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Above 10000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Climb Below 1000	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100001	H500D	ATL		Climb Below 1000	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100000	H500D	ATL		Climb Below 10000	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100001	H500D	ATL		Climb Below 10000	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100000	H500D	ATL		Climb Below Mixing Height	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100001	H500D	ATL		Climb Below Mixing Height	5182.80	30.52	00:10:24.22	5714.51	89.32	87.51	74.56	77.93	17.3
100000	H500D	ATL		Climb Ground	504.00	0.00	00:01:00.00	548.77	8.63	8.45	7.20	7.53	1.7
100001	H500D	ATL		Climb Ground	504.00	0.00	00:01:00.00	548.77	8.63	8.45	7.20	7.53	1.7
100000	H500D	ATL		Climb Taxi	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Climb Taxi	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Descend Below 1000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Descend Below 1000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Descend Below 10000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Descend Below 10000	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Descend Below Mixing Height	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Descend Below Mixing Height	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Descend Ground	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Descend Ground	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100000	H500D	ATL		Descend Taxi	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0
100001	H500D	ATL		Descend Taxi	0.00	0.00	00:00:00.00	0.00	0.00	0.00	0.00	0.00	0.0

24 of 24 item(s) shown. 4 item(s) selected.

Notes: We collect additional data of air travel time on emission and fuel consumption for each airport, each aircraft model and engine number each year using AEDT simulator of FAA. FAA's ground time and taxi time correspond to DOT's taxi time, and FAA's climb below mixing height to descend below mixing height correspond to DOT'x airborne time. We explain additional details in the end of Section 5.3.

Table A.1: **List of NextGen Airports**

Airport Code	City	State	Hub airport	First year of NextGen	Percentage of flights departing from the airport
ANC	Anchorage	AK	1	2016	0.29
ATL	Atlanta	GA	1	2014	6.53
AUS	Austin	TX	0	2017	0.77
BOS	Boston	MA	0	2015	1.89
CLE	Cleveland	OH	0	2015	0.72
CLT	Charlotte	NC	1	2015	2.09
CVG	Cincinnati	OH	1	2014	0.42
DAL	Dallas	TX	1	2015	0.92
DEN	Denver	CO	1	2015	3.86
DFW	Dallas/Fort Worth	TX	1	2016	4.13
DTW	Detroit	MI	1	2015	2.00
EWR	Newark	NJ	1	2015	1.90
HOU	Houston	TX	0	2014	0.93
IAD	Washington	DC	1	2017	0.98
IAH	Houston	TX	1	2014	2.77
IND	Indianapolis	IN	0	2016	0.50
JFK	New York	NY	1	2015	1.64
LAS	Las Vegas	NV	1	2016	2.47
LAX	Los Angeles	CA	1	2016	3.64
LGA	New York	NY	0	2015	1.64
MCI	Kansas City	MO	0	2016	0.79
MDW	Chicago	IL	0	2015	1.48
MEM	Memphis	TN	0	2015	0.52
MIA	Miami	FL	1	2017	1.27
MKE	Milwaukee	WI	1	2016	0.61
MSP	Minneapolis/St. Paul	GA	1	2015	2.02
OAK	Oakland	CA	0	2015	0.79
ORD	Chicago	IL	1	2015	4.84
PDX	Portland	OR	0	2015	0.93
PHL	Philadelphia	PA	1	2016	1.32
PHX	Phoenix	AZ	1	2017	2.91
RDU	Raleigh	NC	0	2015	0.68
SDF	Louisville	KY	0	2015	0.23
SEA	Seattle	WA	0	2015	1.93
SFO	San Francisco	CA	1	2014	2.72
SJC	San Jose	CA	0	2015	0.72
SJU	San Juan	CA	0	2016	0.42
SMF	Sacramento	CA	1	2015	0.73
STL	St. Louis	MO	0	2016	0.94

Table A.2: **Summary statistics of other conditions associated with air travel performance 2010–2017**

Variable	Mean	SD.	Min.	Max.
Panel A. Conditions associated with air travel delay				
Number of operations per aircraft	5.0	2.1	1	17
1 = Prior flight is delayed	0.36	0.48	0	1
Prior delay (minutes)	15.5	37.5	0	2,028
Flights traveled from or to a hub airport	0.84	0.37	0	1
Origin visibility (km)	14.9	3.1	0	160,000
Origin sky ceiling (km)	12.7	9.7	0	22,000
Destination visibility (km)	15.0	3.1	0	160,000
Destination sky ceiling (km)	12.6	9.6	0	22,000
Number of observations			25,037,569	
Panel B. Reported reasons and time of delay if delay arrival == 1				
1 = carrier delay	0.50	0.50	0	1
1 = weather delay	0.05	0.21	0	1
1 = NAS delay	0.56	0.50	0	1
1 = security delay	0.003	0.06	0	1
1 = late aircraft delay	0.51	0.50	0	1
Reported carrier delay (minutes)	17.8	47.6	0	2,402
Reported weather delay (minutes)	2.4	18.8	0	1,934
Reported NAS delay (minutes)	13.8	28.4	0	1,605
Reported security delay (minutes)	0.08	2.3	0	827
Reported late aircraft delay (minutes)	23.4	42.2	0	1,756
Number of observations if delay arrival == 1			4,403,588	

Note: Panel A summarize statistics for the full sample. In Panel B, we restrict the sample to observations if a flight is delayed using DOT definition, i.e., if the actual arrival time is 15 minutes later than the scheduled arrival time.

Table A.3: **Correlation between Categories of NextGen Projects**

Panel A. Whether an airport have adopted a category of NextGen v.s. other categories

	MRO	PBN	SO	DC
MRO	1			
PBN	-0.197	1	1	
SO	-0.244	-0.181	1	
DC	-0.117	-0.09	0.480	1

Panel B. Num. of projects an airport adopted for a NextGen category v.s. other categories

	MRO	PBN	SO	DC
MRO	1			
PBN	-0.203	1	1	
SO	-0.237	-0.181	1	
DC	-0.113	-0.09	0.480	1

Table A.4: **Conditional Effect of NextGen by Reported Delay Reason**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B.1 DOT Delay Reason = Carrier Delay				
NextGen Origin	-1.357***	-0.156	0.024	-0.181**	-0.247***	0.066	-1.201***	-1.586***
	(0.317)	(0.108)	(0.082)	(0.072)	(0.061)	(0.041)	(0.318)	(0.294)
NextGen Dest.	-1.317***	-0.340***	-0.014	-0.326***	-0.228***	-0.098**	-0.977**	-1.014***
	(0.383)	(0.124)	(0.097)	(0.086)	(0.069)	(0.049)	(0.394)	(0.350)
Number of obs.	2,214,727	2,214,727	2,214,727	2,214,727	2,214,727	2,214,727	2,214,727	2,214,727
R-squared	0.57	0.96	0.98	0.21	0.21	0.20	0.05	0.05
				B.2 DOT Delay Reason = Weather Delay				
NextGen Origin	1.705*	1.384***	0.716***	0.668***	0.673***	-0.005	0.321	1.074
	(1.032)	(0.290)	(0.178)	(0.241)	(0.235)	(0.080)	(1.021)	(1.019)
NextGen Dest.	-2.424**	-0.483	0.074	-0.557*	-0.666**	0.109	-1.941*	-1.913*
	(1.086)	(0.337)	(0.196)	(0.296)	(0.269)	(0.099)	(1.103)	(1.074)
Number of obs.	202,238	202,238	202,238	202,238	202,238	202,238	202,238	202,238
R-squared	0.51	0.92	0.98	0.19	0.19	0.18	0.10	0.09
				B.3 DOT Delay Reason = National Airspace System (NAS) Delay				
NextGen Origin	-0.876***	0.414***	0.243**	0.171	0.051	0.120	-1.290***	-1.043***
	(0.332)	(0.135)	(0.100)	(0.111)	(0.104)	(0.075)	(0.322)	(0.293)
NextGen Dest.	-1.279***	0.026	0.389***	-0.363***	-0.301***	-0.062	-1.306***	-0.890***
	(0.369)	(0.144)	(0.109)	(0.120)	(0.103)	(0.082)	(0.370)	(0.327)
Number of obs.	2,424,774	2,424,774	2,424,774	2,424,774	2,424,774	2,424,774	2,424,774	2,424,774
R-squared	0.63	0.95	0.98	0.20	0.21	0.17	0.06	0.05
				B.4 DOT Delay Reason = Security				
NextGen Origin	-1.001	-1.101*	0.100	0.120	-0.021	-2.720	-3.745	
-3.721								
	(2.547)	(0.948)	(0.610)	(0.681)	(0.589)	(0.331)	(2.727)	(2.541)
NextGen Dest.	1.564	0.224	0.921	-0.696	-0.464	-0.233	1.339	1.575
	(2.486)	(0.961)	(0.699)	(0.683)	(0.590)	(0.375)	(2.571)	(2.468)
Number of obs.	14,753	14,753	14,753	14,753	14,753	14,753	14,753	14,753
R-squared	0.84	0.97	0.99	0.37	0.36	0.34	0.30	0.27
				B.5 DOT Delay Reason = Late Aircraft				
NextGen Origin	-0.646**	-0.447***	-0.132*	-0.315***	-0.319***	0.004	-0.199	-0.736***
	(0.303)	(0.105)	(0.072)	(0.073)	(0.063)	(0.035)	(0.299)	(0.278)
NextGen Dest.	-1.137***	-0.161	0.047	-0.208***	-0.142**	-0.066	-0.976***	-0.847***
	(0.329)	(0.111)	(0.084)	(0.075)	(0.062)	(0.042)	(0.325)	(0.292)
Number of obs.	2,260,676	2,260,676	2,260,676	2,260,676	2,260,676	2,260,676	2,260,676	2,260,676
R-squared	0.60	0.95	0.98	0.22	0.22	0.20	0.06	0.05

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. Panel B.1–B.5 estimate the baseline when the aircraft is delayed due to specific self-reported reasons listed in Appendix Table A.2.

Table A.5: **Effect of NextGen: Alternative Sample**

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. Baseline				
NextGen Origin	-1.267***	-0.195**	-0.030	-0.165***	-0.141***	-0.024	-1.072***	-1.598***
	(0.154)	(0.078)	(0.058)	(0.044)	(0.037)	(0.024)	(0.124)	(0.156)
NextGen Dest.	-1.166***	-0.119	0.046	-0.165***	-0.171***	0.006	-1.047***	-1.000***
	(0.149)	(0.078)	(0.061)	(0.048)	(0.038)	(0.028)	(0.114)	(0.151)
Number of obs.	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569	25,037,569
R-squared	0.80	0.97	0.99	0.22	0.21	0.22	0.04	0.04
				B. Alternative Sample with Aircraft Not Matched in B43				
NextGen Origin	-0.987***	-0.060	0.048	-0.108***	-0.089***	-0.019	-0.927***	-1.327***
	(0.133)	(0.067)	(0.050)	(0.039)	(0.033)	(0.021)	(0.105)	(0.135)
NextGen Dest.	-1.055***	-0.164**	-0.013	-0.151***	-0.134***	-0.017	-0.891***	-0.859***
	(0.128)	(0.068)	(0.054)	(0.042)	(0.033)	(0.026)	(0.098)	(0.131)
Number of obs.	33,447,478	33,447,478	33,447,478	33,447,478	33,447,478	33,447,478	33,447,478	33,447,478
R-squared	0.79	0.97	0.99	0.24	0.22	0.22	0.04	0.04

Notes: Robust standard errors clustered at route level in parenthesis. *, **, and *** indicate statistical significance at 10, 5 and 1 percent levels respectively. Panel A repeats the baseline results in Table 2. Panel B include aircraft with tail number not matched in the DOT Form-B43.

Table A.6: **The effect of NextGen from 2014 to 2017 on air travel time using alternative estimates**

	(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
Actual (minutes)	149.04	139.30	115.78	23.51	16.32	7.18	9.6	4.13
				Panel A. Baseline (Repeat Table 3)				
Change (minutes)	2.54	0.33	-0.02	035	0.33	0.02	2.22	2.72
Change (percent)	1.71%	0.24%	-0.01%	1.47%	2.00%	0.26%	23.17%	65.94%
				Panel B. Use Estimates from Table 2 Panel B				
Change (minutes)	2.71	0.43	-0.04	0.39	0.36	0.03	2.30	2.84
Change (percent)	1.82%	0.31%	-0.04%	1.64%	2.18%	0.41%	23.89%	68.90 %
Number of obs.							4,736,642	

Notes: The first row reports the actual travel time in 2017. This table reports the counterfactual travel time and delay if the treated airports in 2017 had been untreated, i.e., as if the airports were in the beginning of 2014. We use baseline in Table 2. In Appendix Table we use estimates in Table 2 Panel B and estimates in Table 8.