Incentives and Competition in the Airline Industry

November 1, 2018

Abstract

We examine how performance changes at airlines in response to a change in executive incentives. Airlines with executive bonuses contingent on on-time arrival do improve on-time performance. Competitors on the same routes also improve their on-time performance, even when their executive bonuses are not contingent on on-time performance. We find limited evidence of strategic gaming such as carriers increasing scheduled flight times. Carriers also do not decrease the frequency of flights or the number of passengers to make it easier to be on-time, but they do slightly decrease fares. Our results suggest that incentives heighten competition in on-time performance, consistent with competition in strategic complements.

Keywords: Incentives, Product Market Competition, Airlines.

JEL classification : G30, G34, G32

1 Introduction

Much of the managerial incentives literature focuses on grants of stock and stock options, and then examines performance measures such as stock or accounting returns. Relatively less explored is how managers respond to incentives to affect specific actions and performance measures. We address this question by looking at the airline industry, and at specifically delineated incentives that have measurable outcomes closely tied to actions. We focus on bonuses provided by airlines to executives for on-time arrival performance. On-time arrival is a key metric for customer satisfaction in the airline industry.[1] Over the past 20 years, virtually every airline has adopted on-time arrival as a measure used in the provision of executive bonuses. Importantly for our empirical study, the introduction of bonuses tied to on-time arrival has been staggered in time across airlines. This allows us to study the effect of introducing a bonus on the performance of the firm that adopts the bonus.[2]

Incentive schemes may also be gamed; that is, managers can take actions to make it easier to achieve the bonus but that do not actually result in improved performance. Bennett et al. (2017) show that managers do try to strategically achieve specific compensation targets. In the context of airlines, there are several ways to game an on-time arrival incentive. The most obvious way is to increase the scheduled amount of time for a flight, thus making it easier for flights to arrive on-time and the executives to achieve their bonus. The number of flights scheduled could also be reduced, and flights could be flown with fewer passengers, thus making it easier to arrive on-time. We are able to study the extent to which managers at airlines might game an on-time incentive.

More generally, executives receiving an on-time performance bonus might change the number of scheduled and performed flights in an attempt to turn around the planes faster and meet departure and arrival times. Executives could also introduce changes in the number of passengers and seats available as a way to expedite the boarding and deplaning of passengers, making it easier for the carrier to depart on-time. Thus, we examine changes in capacity, both the number of flights as well as passengers and seats. It is also possible that introducing an on-time incentive might signify a general toughening in competition by an airline. Such an airline might also increase capacity

[1] Dresner and Xu (1995) show a strong correlation between delays and customers filing complaints.

[2] Our evidence on the actual actions taken by managers to influence on-time arrival is still somewhat indirect. However, the outcomes we observe (whether the flights are on-time) are very closely tied to the incentive (bonuses for improved on-time performance), and thus have less noise between incentive, action, and outcome.

and lower prices perhaps anticipating that rivals will accommodate the tougher competitor. We therefore examine changes in prices as well as changes in quantities.

When a firm changes managerial incentives, this may induce or provoke a strategic response from industry competitors. Indeed, firms may alter incentives strategically for precisely this reason (see Aggarwal and Samwick (1999)). Since on-time performance is a key metric of customer satisfaction, it is also an area in which airlines can and do compete with each other. We are able to see how competitors respond to the introduction of an on-time arrival incentive by rivals. For example, competition in on-time performance may be competition in either strategic complements or strategic substitutes. If competition is in strategic complements, then instituting an on-time arrival incentive may cause both the initiating firm and its rivals to toughen competition, which would increase consumer surplus.

In addition, incentives may be more or less effective in different market environments. For example, incentives may be less effective in routes where an airline faces competitor airlines, as the competition already disciplines the airline to be more on-time. Similarly, incentives may be more effective in monopoly routes where there is no competitive pressure. We also study the responsiveness of performance to incentives by market environment.

We hand collect data on the specifics of an airline's CEO bonus from the carrier's proxy statements (Form DEF 14A). Data on the airline's on-time performance is from The Bureau of Transportation Statistics (BTS). We merge the two datasets to study how an airline's on-time performance has varied in response to changes in incentives based on on-time arrival. We are able to disaggregate on-time performance into departure times, scheduled flight times, actual flight times, and arrival times, thus more precisely seeing how the airline responds to the introduction of the incentive for executives.

We find that when airline executives are given bonuses based on on-time arrival performance, on-time arrival does indeed improve. Delays decrease by up to 2.4 minutes in our preferred specification, or 20 percent, of the average non-negative arrival delay (negative delays are flights arriving early) relative to a carrier without an incentive after saturating our specifications with fixed effects. The fraction of flights delayed fifteen minutes or more (an industry standard definition for a flight being late) decreases by eleven percent relative to the frequency of delayed flights for a carrier without an incentive. Because we see both arrival delays and departure delays, we show that most of the decrease in arrival delays is due to airlines decreasing departure delays–getting departing flights out of the gate closer to on-time. In a separate analysis, when we restrict our sample to problem routes (those routes where more than thirty percent of flights in a quarter are late by fifteen minutes or more), we find that arrival delays shrink by over two minutes, and the fraction of flights delayed by fifteen minutes or more falls by 12.5 percent.

We exploit the staggered introduction of incentives across carriers to more precisely estimate each airline's response to the introduction of an incentive. We use a difference-in-difference specification where we consider an airline before and after introducing an incentive (treated) versus other airlines without an incentive (control) on the same routes. We use three different approaches. First, we examine the quarter by quarter response to the treatment for the four quarters prior to the introduction of the incentive and the four quarters after using a dynamic staggered specification. Second, we pool the quarters before and after the treatment to present a staggered difference in difference specification. Third, we examine each airline separately as it introduces an incentive, thus allowing us to examine heterogeneity across airlines. Across each of these approaches, we consistently find that arrival delays fall after the introduction of an incentive.

For strategic gaming, such as increasing the scheduled time of the flights, we find aggregate evidence of such gaming that is generally small in magnitude. On average, carriers with an on-time arrival incentive increase their scheduled minus actual flight times by up to 30 seconds per flight. These aggregate estimates mask significant heterogeneity. Four carriers–Alaska Airlines, United Airlines, American Airlines, and Delta Airlines–substantially increase their scheduled minus actual time after introducing an incentive. We also find much larger strategic gaming effects as the on-time arrival incentive becomes a larger fraction of the total bonus, i.e., more high powered.[3]

In our difference-in-difference specifications, we are also able to examine the effect of the introduction of incentives by one carrier on its competitors that do not have an incentive. We find some evidence that competitors do respond to the introduction of incentives by reducing their delays on the routes on which they compete, even when the executives of the competitors did not receive

[3] Discussions with American Airlines executives revealed that carriers incur large unnecessary costs when padding flight times: pilots, flight attendants, gate agents and machinists have to be paid for the extra time, even though they are not really at work; and most importantly, the turnaround time for flights at the gate is severely affected: padding time on several or all flights out of a gate could cost, by the day's end, a flight slot at the gate. Gate slots are among the top airport expenses; losing a flight to padding could be very costly. Thus, there are trade-offs associated with padding flight times to achieve an on-time bonus.

bonuses based on their own on-time performance. This is consistent with competition in service quality (on-time performance) being competition in strategic complements. We caution that our results here are smaller in magnitude, and the effect is not present in all specifications.

When both executives of the own and rival airline receive bonuses based on on-time performance, we find that on-time performance on the routes on which they compete improves by a similar amount to when only the own airline has an incentive. Thus, there is little incremental effect from both competitors having an incentive. We also find that the market environment does not have differential effects on the responsiveness of arrival delays to on-time incentives–we find similar effects in both monopoly routes and competitive routes.

Capacity as defined by passengers and seats per flight increases for an airline with an on-time arrival incentive, while the effect on rival carriers without an incentive is generally negligible to negative. The key point here is that airlines with an on-time arrival incentive do not strategically decrease capacity or utilization to make it easier to achieve bonuses.[4] We also find that airlines with on-time arrival incentives lower fares by one percent. Consistent with the increase in capacity and passengers, on-time incentives seem to signify a toughening in competition on several dimensions, but these effects are small in magnitude.[5]

We contribute to the literature exploring the relation between managerial incentive-based compensation and non-accounting or stock-based outcomes.[6] The outcomes that we examine are a relatively noise-free transformation of actions managers would take to meet the incentive goal (Holmstrom and Milgrom (1987), Oyer (2004)). The paper closest to ours is Forbes et al. (2015). While their primary focus is on the response of airlines to the introduction of the US Department of Transportation's mandatory disclosure program of flights delayed 15 minutes or more, as part of their analysis, they also consider the introduction of employee bonus programs for reducing arrival delays below the 15 minute threshold. They find heterogeneous effects across carriers: Continental

[4] These results are consistent with Dana and Orlov (2014), who find that over our sample period, load factors for airlines have consistently increased due to Internet use by consumers.

[5] Schmidt (1997), Hermalin (1992), and Golan et al. (2015) provide theoretical examinations of how increasing product market competition affects a manager's incentive to exert effort. Our paper is also related to the literature that examines the relationship between product quality and product market competition. Mazzeo (2003) explores the relation between airline market concentration and flight delays, and shows that both the prevalence and duration of flight delays are significantly greater on routes where only one airline provides direct service. See also Rupp et al. (2006), Prince and Simon (2015), and Ciliberto and Schenone (2012a).

[6] Another example of non-accounting or stock returns-based compensation is the literature on corporate social responsibility contracting, which finds that companies increasingly put weight in executive bonus plans on metrics such as employee satisfaction, safety, or environmental outcomes (see Flammer et al. (2018)).

and Trans World Airways (TWA) show evidence of reducing delays around the threshold, while US Airways and United did not, and American showed very small effects.[7]

We focus on executive bonuses, rather than bonuses for the entire workforce, and find stronger effects, although we also find heterogeneity across carriers. Incentives for executives could be more effective than employee bonuses because executives can use a number of different levers to influence on-time arrivals beyond employee bonuses, from changing schedules and types of aircraft to concentrating resources where they matter most (e.g., problem routes). As we show below in our discussion of specific incentive plans, employee bonuses are small in magnitude while executive bonuses are typically large.

This paper is organized as follows. We describe our data and variable construction in Section 2. Section 3 presents our econometric specifications, details our identification strategy, and presents our results. Section 4 presents results on several aspects of competition, including the presence of Southwest Airlines, and the effect of incentives on quantities and prices. Finally, Section 5 concludes.

2 Data

We select the airline industry as we have measurable outcomes very specifically tied to actions that respond to the particular incentive given to a manager. Furthermore, the airline industry yields a vast amount data for studying competitive reactions among industry rivals. Airlines compete over many different routes, and over different time periods. At any given point in time, a carrier faces different competitors on different routes; and for any given route, a carrier is likely to face different competitors at different times as carriers enter and exit routes. For example, United Airline's competitors between Denver and San Francisco in the first quarter of 2001 are not the same set of competitors United faces in the route between Dallas-Forth Worth and Newark in the first quarter of 2001. Thus, a carrier's business strategy at a point in time and on one route might differ from its strategy on a different route at the same time; and the strategy on one route might differ across time depending on which carriers enter or exit that route (see for example, Berry (1990), Brueckner et al. (1992), and Gerardi and Shapiro (2009)). Therefore, even though we focus on a few airlines,

[7] TWA introduced an incentive program for employees in 1996, but that program was only for employees and did not include top executives.

the strategic decision each one makes on any one route during a specific time period is likely to differ from the strategic response chosen at the same time but on a different route where the set of product market competitors is different.

The airline data in our work is a compilation of different databases from the Office of Airline Information at the Bureau of Transportation Statistics (BTS). We focus only on domestic flights by publicly-traded US based carriers. The sample ranges from January 1^{st}, 1993 to December 31^{st}, 2010 (72 quarters). During this time window, most of the large US carriers instituted an incentive. Data on executive contracts and bonuses are gathered manually using firm proxy statements.

2.1 Data Sources and Definitions

2.1.1 Performance Based Incentives

We collect data on CEO bonuses from firm proxy statements (Form DEF 14A accessed from the SEC's website). We read all proxy statements to see if bonuses depend on on-time arrival or performance. Firms generally disclose key measures for firm bonuses. Firms are not required to disclose how much of the bonus depends upon each measure, and there is variability in how much they do so. For those airlines that do report the fraction of the bonus that depends on on-time performance, we collect these data as well. Table 1 shows when each airline in our sample initiated an on-time arrival incentive. These incentives are quite sticky, and once instituted are not rescinded for most carriers (the exceptions are Delta and JetBlue). It is important to recognize that the introduction of incentives has been staggered across time and carriers. This allows us to isolate the effect of incentives on individual carriers. The first carrier to introduce an on-time performance incentive was Continental in 1995; by the end of our sample period, 2010, the only surviving carriers without an incentive are Northwest (which was acquired by Delta at the end of our sample period) and JetBlue (which rescinded its incentive at the end of 2008). America West did not introduce an incentive but was absorbed by US Airways in 2005, which then introduced an incentive in 2008.

We provide two examples of the type of on-time performance incentives that have been introduced by carriers. The first is from American Airlines, which introduced its incentive in 2004.

American's proxy statement (DEF 14A) from 2004 states on page 32:

> The Company and its labor unions have agreed upon a structure for future annual bonus plans. The annual bonus plan is styled the **"Annual Incentive Plan" (the "AIP")**. Eligible participants under the AIP include all U.S. based employees of American (including the named executive officers). The **AIP has two primary components: (a) a financial component; and (b) a customer service component**. For the 2004 AIP, the financial component required a pretax earnings margin of 5% before any payments would accrue. The customer service component contemplated payments ranging from $25 to $100 per month, per employee, predicated upon **one of two customer service measures: (i) on-time performance; and (ii) customer satisfaction**. [bold added]

Note that for American, the on-time performance incentive extends throughout the workforce beyond the executive officers, although the executive officers are included as well. Note also that the bonuses for employees are small in magnitude per employee. Further, American does not state the relative weight placed on each objective.

The second example is from US Airways, which introduced its incentive in 2008. US Airways' proxy statement from 2008 states on page 28:

> The Committee further adjusted the weighting of the financial and operational goals with 50% tied to financial performance targets and **50% to operational performance targets**. The **operational performance targets were weighted 67% to relative on-time flight performance** and 33% to internal baggage handling improvements. [bold added]

US Airways provides more explicit information about the on-time performance incentive. In this case, 33.5% of an executive's annual bonus depends on on-time performance. In our empirical work, if an airline introduces an on-time performance incentive for its CEO, then we code that airline as having an incentive an all of its routes.

2.1.2 Airline Performance Data

On-time Arrival and Departure Outcomes

The Bureau of Transportation Statistics (BTS) collects on-time performance data reported by US certified air carriers that account for at least one percent of domestic scheduled passenger revenues, and reports it in the database named: "Airline On-Time Performance Data." Carriers report daily flight information for the routes they serve.[8] Specifically, it contains scheduled, and actual, departure and arrival times, and scheduled and actual flight times.[9] As we explain below, we collapse this daily flight data to quarterly data for our empirical analysis.

The BTS defines actual arrival time as, "the instance when the pilot sets the aircraft parking brake after arriving at the airport gate or passenger unloading area," and scheduled arrival time as "the scheduled time that an aircraft should cross a certain point (landing or metering fix)." For each flight, we calculate the minutes of arrival delays as the difference between actual and scheduled arrival times. We aggregate the data to average arrival delays, $MeanArrDelay_{jrt}$, by a carrier in a route in a quarter. We sum all of the arrival delays across all flights by a carrier in a route in a quarter and divide by the total number of flights by that carrier in that route in that quarter. For most of our analysis, we also adjust for early arrivals, which would have a delay of less than zero. Early arrivals are presumably welcome by passengers, but are not relevant for on-time performance incentives. As a result, we do not include them in the effective measure of delays. In order to avoid confounding our measure of flight delays with early arrivals, we replace negative arrival delays with 0 delayed minutes before we aggregate up to the route-carrier-quarter unit of observation. Departure delays, $MeanDepDelay_{jrt}$, are defined similarly.[10]

The BTS collects data on actual flight times (length of the flight), and also on scheduled flight times. The latter are reported to the Computer Reservation System (CRS) to allow consumers and travel agencies organize and compare trips. Scheduled and actual flight times are recorded for each

[8] A route is a segment in the passenger's itinerary–meaning it is nonstop point-to-point travel. For example, a passenger flying from Boston to Los Angeles via the itinerary BOS-SFO-LAX would have two routes, BOS-SFO and SFO-LAX. Data on on-time performance and capacity is at the route level.

[9] These data, as well as the capacity data, are available through the BTS maintained website: http://www.transtats.bts.gov/DataIndex.asp

[10] There are several issues we deal with using the raw data from the BTS. First, we eliminate observations for which arrival, departure, as well as actual and scheduled flight times is missing. Second, we drop observations for which arrival delays are more than 600 minutes (10 hours) and early arrivals are more than 180 minutes (3 hours). We also drop observations for which departure delays are more than 600 minutes, and for which early departures exceed 30 minutes.

carrier, for every flight in a given route on a daily basis. For every flight, we calculate the difference between the scheduled flight time and the actual flight time, and call this variable Pad. It denotes the extent to which the scheduled length of a flight exceeds the actual length of a flight in minutes. The greater is the padding, the easier it is for a flight to arrive on-time. We aggregate the data to get quarterly padding by route and by carrier, Pad_{jrt}, by summing the padding across all flights by a carrier on a route during a quarter and dividing by the number of flights by that carrier on that route in that quarter.[11]

Capacity, Utilization, and Price Data

Executives receiving an on-time performance bonus might strategically change the frequency of flights served, or the number of transported passengers, to aid in turning planes around in a more efficient way, thus improving arrival and departure times. To account for this possibility, we study changes in capacity around the time the carriers introduced the performance incentive.

The capacity data are from the T100 Domestic Segment of Form 41-Traffic data provided by the BTS and includes items such as passengers enplaned, seats available, load factors, and number of departures performed and scheduled. This database is not a sample of flights, but a record of *all monthly flight segments* between an origin and destination airport located within the US boundaries or its territories for US certified air carriers that account for at least one percent of domestic scheduled passenger revenue. These data are at the aircraft type level, i.e., a carrier, on a route, on a given day, flying a specific aircraft (a route can be served with different aircraft depending on the time of day the flight occurs). As with the on-time performance data, we collapse the data to a quarterly basis and merge it with the on-time performance measure.[12] To collapse the

[11] As before, several adjustments need to be done to the raw BTS Data. First, in calculating actual and elapsed times we account for any time zone differences between origin and destination airports. Specifically, we adjust the origin departure time to be the equivalent of the destination time to compute actual flight times as the difference between the actual arrival times and the time-zone-adjusted departure time. Second, we drop flights for which the actual elapsed time is less than 15 minutes, or more than 720 minutes (12 hours).

[12] We identify routes that are not served in a consistent manner (e.g., sporadic service between an origin and destination). We drop route-carrier-year-quarter observations for which the quarterly sum of scheduled or performed departures is less than 24 (routes served less than 8 times a month by a carrier), and drop observations for which the quarterly sum of passengers, or seats, is less than 100. We also drop route-carrier-year-quarter observations for which the quarterly sum of scheduled or performed departures is greater than 2,7000 (routes served more than 30 times a day by a carrier), as these observations are likely to be data entry errors. We drop observations for which the route includes either an origin or destination that is not ranked, by the total number of passengers transported, in the top 100 origin and destination airports. These are standard filters in the literature to avoid small, unrepresentative routes that are sporadically served.

data, we take the average of the seats available per scheduled flight across all flights on a route by a carrier in that quarter, and the average of the actual number of passengers enplaned per scheduled flight across all flights on a route by a carrier in that quarter. We also take the sum of departures scheduled and the sum of departures performed on a route by a carrier in that quarter. We define quarterly load factor as the ratio of revenue passenger per mile to the available seats per mile on a route by a carrier in that quarter.[13]

To examine the impact of incentives on ticket prices, we use itinerary fares obtained from the Airline Origin and Destination Survey (DB1B) database, maintained by the BTS. The DB1B is a quarterly, 10% random sample of domestic airline tickets. We apply the standard filters used in the existing literature on airline pricing to construct our price sample.[14] These data include information on the itinerary fare each flown passenger paid during a year-quarter. As in prior literature, we compute the average fare charged by a carrier on a route in a quarter as the carrier's total revenue on that route in that quarter divided by the total number of passengers flown by that carrier in that route in that quarter (itinerary fares are adjusted using the 2008 CPI).

2.2 Summary Statistics and Econometric Controls

2.2.1 Summary Statistics

Table 2, Panel A, reports summary statistics on scheduled minus actual flight time, as well as arrival and departure delays, where we do not yet adjust early arrivals/departures by coding them as 0 minute delays. We split our observations into four mutually exclusive categories:

1) $None$–no carrier on a specific route r in quarter t has an on-time incentive;

2) Own–a specific carrier j on a specific route r in quarter t has an on-time incentive, and no other carrier on route r in quarter t has an incentive;

3) $Rivals$–at least one competitor carrier to carrier j on route r in quarter t has an on-time incentive, but carrier j does not;

[13]The BTS defines a passenger as: "Any person on board a flight who is not a member of the flight or cabin crew," and available seats as "Installed seats in an aircraft (including seats in lounges) exclusive of any seats not offered for sale to the public by the carrier; provided that in no instance shall any seat sold be excluded from the count of available seats."

[14]For a detailed explanation of the price data construction, see Dennis et al. (2018). See also Borenstein (1989), Berry (1990), Brueckner et al. (1992), Borenstein and Rose (1995), Gerardi and Shapiro (2009), Ciliberto and Schenone (2012b), and Snider and Williams (2015).

4) *Joint*–both a specific carrier j and at least one competitor to j on route r in quarter t have on-time incentives.

Panel A shows that if there are no on-time incentives for any carrier on a route, then arrival delays are on average 6.12 minutes. If a carrier has an on-time incentive, then its arrival delays are 6.27 minutes. If a carrier's rival on a route has an incentive and it does not, then the carrier's delays are 6.85 minutes. Finally, if both carriers on a route have an incentive (*Joint*), then arrival delays are 7.80 minutes.

Departure delays are presented in the second row of Panel A, and show a similar pattern, although the departure delays are a little larger in magnitude. The difference between the CRS reported scheduled flight time and the actual elapsed flight time, Pad_{jrt}, is presented in the third row. If carriers wanted to pad the flight time as a response to on-time arrival incentives, then we should see a substantial increase in $CRS\_Scheduled_{jrt}$, and we do see some evidence of that in the second through fourth columns. As should be true, within each column, $MeanArrDelay_{jrt} = MeanDepDelay_{jrt} - (Pad_{jrt})$, up to rounding error. The total number of route-carrier-quarter observations in our sample is given in the last row of Panel A.

There has been a significant time trend in airline performance. Through time, airlines have scheduled more flights per day, and capacity utilization or load factor has increased (see Dana and Orlov (2014)). We remove the effects of the time trend by controlling for year-quarter fixed effects, and present adjusted summary statistics in Panel B. Panel B shows that if there are no on-time incentives for any carrier on a route, then arrival delays are on average 8.22 minutes. If a carrier has an on-time incentive, then its arrival delays are 6.83 minutes, a decrease in delays relative to carriers without an incentive. This is in contrast to Panel A, which did not adjust for time effects. Interestingly, if a carrier's rival on a route has an incentive, then the carrier's delays are 6.49 minutes, possibly suggesting that there is a competitive response to a rival carrier having an incentive.

Finally, if both carriers on a route have an incentive (*Joint*), then arrival delays are 8.57 minutes. This increase in delays for *Joint* suggests that there is something different about these routes. Routes on which two carriers compete and both have an incentive are typically routes that involve at least one major hub, such as Atlanta to Dallas, where Delta and American have hubs, respectively. Hubs are busier and more subject to delays. In our subsequent empirical specifications, we control for hubs by controlling for route-carrier fixed effects, $origin \cdot year$ fixed effects, and $destination \cdot year$ fixed effects. When we do, we no longer find that delays on $Joint$ routes are larger than $None$ routes–in fact, delays decrease.

Departure delays are presented in the second row of Panel B, and again the pattern for departure delays is similar to that of arrival delays (although now the departure delays for $Joint$ are less than the departure delays for $None$). The difference between the CRS reported scheduled flight time and the actual elapsed flight time, Pad_{jrt} is presented in the third row. Interestingly, in the fourth column for $Joint$, routes on which both an airline and at least one competitor have an incentive exhibit negative Pad_{jrt} times, indicating that these are routes that are chronically delayed, since actual elapsed times are on average greater than scheduled times. As noted before, these $Joint$ routes typically involve hub airports, with high arrival delays seemingly caused by chronically delayed flights en route.

The first four columns of Panel C of this table report arrival and departure delays where early arrivals and departures have been coded as no delays (negative minutes delayed are replaced by 0 delayed minutes). The summary statistics in Panels C and D are after controlling for year-quarter fixed effects, as we did in Panel B. The results in the left four columns are consistent with those in Panel B and, not surprisingly, larger in magnitude. Pad_{jrt} times are not reported in Panel C as they would be identical to those in Panel B since we do not re-code negative values for Pad_{jrt} times, which are flights that consistently take longer to actually fly than their scheduled times.

The industry standard definition of a flight being delayed is if the flight arrives at least 15 minutes late relative to its scheduled time. Reducing the frequency of flights that are delayed by this definition (15 minutes late or more) is the typical goal of an on-time incentive in an executive bonus. The right four columns of Panel C show the frequency with which flights are delayed 15 minutes or more. If no carrier on a route has an incentive, flights are delayed 15 minutes or more 23% of the time upon arrivals. If a carrier has an incentive, this drops to 20%, and also drops to 20% if a rival carrier has an incentive. If both a carrier and its rival have an incentive, the frequency of flights arriving 15 minutes late or more is 21%, which again is a hub airport effect. Departure delays show a similar pattern, although smaller in magnitude.

We define problem routes for each carrier as those routes on which a carrier's flights are delayed more than 15 minutes more than 30% of the time *prior* to the carrier's introduction of the incentive.

We then look at differences in arrival delays before and after the implementation of the incentive–Panel D reports these statistics. The left four columns show the average minutes delayed for problem routes, and the right four columns show the frequency with which flights are delayed more than 15 minutes in a problem route. Note that these average delays can be less than 15 minutes and the frequency of delays less than 30% after the introduction of the incentive (the columns for Own and $Rivals$), since the definition of a problem route is that at least 30% of the flights on that route were delayed 15 minutes or more in the quarter prior to the introduction of an incentive. Delays decrease with the introduction of an incentive by a carrier, and to a lesser extent in minutes, when a rival introduces an incentive. Nonetheless, delays in problem routes seem to be quite persistent, as delays are still elevated after the introduction of an incentive as compared to Panel C. Note that the number of observations in problem routes is much smaller than in the full sample.

Executives receiving an on-time performance bonus might strategically reduce the number of scheduled and performed flights in an attempt to turn around planes faster and meet departure and arrival times. Executives could also introduce changes in the number of passengers and seats available as a way to expedite the boarding and deplaning of passengers, making it easier for the carrier to depart and arrive on-time. We use the capacity data reported in the T100 Domestic Segment database to examine these issues.

Table 3, Panel A, reports summary statistics for capacity, capacity utilization, and fares. Panel A, Part 1, includes all routes while Panel A, Part 2, focuses on problem routes. The carriers that introduce the incentive (Own) see an increase in the number of seats per scheduled flight available as well as the number of enplaned passengers per performed flight and the capacity utilization (load factor) relative to carriers without an incentive. We see similar effects for competitors of the carrier that introduced the incentive ($Rivals$), although there is a small decrease in the number of seats per scheduled flight. We also see similar effects when a carrier and at least one rival have an incentive ($Joint$).

Regarding departures scheduled and performed, we see a decrease when a carrier introduces an incentive, but this decrease is smaller than the increase in number of passengers per performed flight. In effect, the greater capacity utilization after the introduction of an incentive more than offsets the reduction in the number of flights performed. When a carrier without an incentive competes with one that has an incentive ($Rivals$) and when both a given carrier and at least one competitor on a specific route in that quarter have on-time incentives ($Joint$), both departures scheduled and performed are much larger in magnitude relative to carriers without an incentive. Panel A, Part 2, reports summary statistics for the sample restricted to problem routes. The data are noisier here due to the smaller sample size. One feature here is that carriers that introduce an incentive generally increase capacity and utilization on the problem routes.

Panel B reports ticket-price statistics. For each route-carrier and year quarter observation, we have a random sample of 10 percent of *all itinerary* fares. The average fare reported here under each categorical variable is the average $Price_{jrt}$ across carriers, routes and year quarter observation falling in each category. Fares in all categories are around 200 dollars per itinerary, and fares in problem routes are generally higher than in all routes.

2.2.2 Econometric Controls: Route Presence and Origin and Destination Airport Specific Shocks

To control for the effects that a particular carrier j flying a particular route r can have on its on-time performance we include route-carrier fixed effects u_{rj}. For example, a carrier might fly more modern planes that could impact its on-time performance on one route, and older planes on other routes. This is a route specific effect for a carrier and the route-carrier fixed effect absorbs this variation. As another example, route-carrier fixed effects will incorporate variation due to hub airports, specifically non-time-varying hub effects. Furthermore, if there are sample selection issues that arise when the unobservables that affects a carrier's choice of entering route r are the same unobservables that impact a carrier's on-time performance in that route, but not in others, then route-carrier fixed effects control for this carrier-route selection.[15]

We also need to control for whether there are competitors to carrier j serving that same route at the same time. As an example, consider routes on which carrier j competes with Southwest. The presence of Southwest will affect carrier j, regardless of whether carrier j has a performance-based incentive. We want to control for this effect for all competitors, not just Southwest, to carrier j. To control for competitor carrier k's presence in route r at time t, we define $kInRoute_{rt}$ as a categorical variable equal to one if competitor carrier k is active at time t in route r that carrier j

[15] See, for example, Nijman and Verbeek (1992) treatment of self-selection of firms into markets. The authors show that fixed effects estimators are more robust to non-response biases than random effects estimators.

flies. If competitor carrier k is not present in a route, then $kInRoute_{rt}$ equals zero. It is important to note that this set of controls is only identified when a carrier shifts into or out of a route–if a carrier k is always in a route or never in a route throughout our sample, then this effect will not be estimated for that carrier in that route. We refer to these fixed effects as competitor in route fixed effects. The route-carrier fixed effects u_{rj} are distinct from the competitor in route, $kInRoute_{rt}$, fixed effects. The route-carrier fixed effects pertain to a given carrier j, while the $kInRoute_{rt}$ fixed effects encompass all of $j's$ competitors on a route at time t. The route carrier fixed effects will absorb substantially more variation than the $kInRoute_{rt}$ fixed effects in our data.

There could be origin- and/or destination-specific shocks that can alter the on-time performance of flights involving such origins and destinations. For example, airports undergoing construction or maintenance could see changes in on-time performance or utilization measures associated with the shock to the airport, but unrelated to whether a carrier serving that airport has an incentive. Therefore, we include $origin \cdot year$, u_{oy}, and $destination \cdot year$, u_{dy}, fixed effects. Further, to the extent that either the origin or destination is a hub airport for any carrier, the $origin \cdot year$ and $destination \cdot year$ fixed effects will absorb any time-varying hub effects.

We also include year-quarter fixed effects, u_t, that control for performance changes that stem from seasonal and other exogenous shocks that can affect performance outcomes, and can therefore confound the effect of incentives on performance. Year-quarter fixed effects also control for any serially correlated industry-specific shock to demand that can impact performance measures. Note that we are saturating our specifications with thousands of fixed effects when we include the sets of $kInRoute_{rt}$, route-carrier, $origin \cdot year$, $destination \cdot year$, and year-quarter fixed effects. These fixed effects absorb a large amount of variation in our specifications.

Over our sample period, the airline industry saw three important mergers, US Airways-America West, Delta-Northwest, and United-Continental. Consolidation impacts competition and airline behavior (see Kin and Singal (1993), and Oliver (2003)). In addition, six of our ten airlines have operated under Chapter 11 bankruptcy protection during our sample period: Continental, America West, United, Delta, Northwest, and US Airways (twice). Prior work has shown that Chapter 11 protection alters airline performance and product market competition (see, for example, Borenstein and Rose (1995), Borenstein and Rose (2003), Ciliberto and Schenone (2012a)). The combination of route-carrier fixed effects and year-quarter fixed effects fully controls for mergers (specifically, the time between when a merger is completed and when the carriers start jointly reporting on-time statistics to the BTS) and bankruptcy-related effects.

3 Results

3.1 Main Specifications

Our initial specification examines the extent to which the introduction of an on-time incentive influences arrival delays, departure delays, and scheduled - actual times of flights (the amount which the airline pads its schedule). In order to demonstrate the dynamic effects of introducing an incentive, we estimate the following difference in difference specification:[16]

$$\begin{aligned} P_{jrt} \quad = \quad & \sum_{q=-4}^{q=4} \alpha_q \cdot Quarter_q \cdot Treated_{jq} + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ & + \sum_k \gamma_{krt} \cdot k\_InRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt}; q \neq 0 \end{aligned} \tag{1}$$

where P_{jrt} is one of three on-time measures for carrier j in route r during year-quarter t: non-negative arrival or departure delays and padding time (CRS scheduled time minus actual elapsed flight time). $Quarter_q$ is an indicator equal to one for the q^{th} quarter relative to the introduction of the on-time incentive. The time window is defined by four quarters *before* the the incentive was implemented and four quarters *after* the incentive was introduced (the quarter the incentive was implemented is therefore excluded). In this specification, we control for total CEO compensation. To control for the presence of competitors in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. We include route-carrier fixed effects u_{rj}, year-quarter fixed effects u_t, $origin \cdot year$ fixed effects u_{oy}, and $destination \cdot year$ fixed effects u_{dy}. Standard errors are clustered at the route-carrier level. The error term is ζ_{jrt}. The coefficient estimates of α_k represent the change in minutes performance due to the introduction of an incentive q quarters away by treated carriers relative to the change in minutes performance for carriers that did not introduce an on-time incentive (control carriers). The coefficients are dynamic difference-in-difference estimates, quarter-by-quarter, starting from four quarters prior to the introduction of an incentive by treated

[16] Similar to, for example, Stevenson and Wolfers (2006) and Lin et al. (2018).

carriers and going until four quarters after the introduction of the incentive. In this specification we are pooling all of the carriers that introduce an on-time incentive (treated) and all carriers that do not have an incentive at the same calendar time (control), and treat the quarter in which the incentive is initiated as the event quarter (quarter 0). Because we include year-quarter fixed effects, the traditional post treatment indicator in a difference in difference regression is eliminated. Because we include route carrier fixed effects, the traditional treated indicator is also eliminated. The results are presented in Table 4.

To illustrate the dynamic effect of introducing the incentive on on-time performance, Figure 1 plots the coefficients α_q. The dashed lines around the coefficient estimates represent the 95 percent confidence interval, adjusted for route-carrier clustering. The first column of Table 4 and the top panel of Figure 1 present results for arrival delays. Prior to the introduction of the incentive (quarters -4 to -1), arrival delays are essentially no different than arrival delays at quarter 0 for carriers introducing an incentive relative to those that do not. None of the coefficients are significant. This is a test for pre-trends prior to the introduction of the incentive. As can be seen in Table 4 and Figure 1, there is no pre-trend for carriers introducing an incentive relative to those that do not, alleviating some endogeneity concerns associated with the introduction of the incentive.

After the introduction of the incentive, arrival delays steadily decrease, reaching a reduction in arrival delays of 2.39 minutes in quarter 4 relative to quarter 0 for carriers that introduce an incentive versus those that do not. This reduction in arrival delays is consistent with the incentive working as intended, and is large in magnitude. Recall from Table 2, Panel C, that the mean non-negative arrival delay for a carrier without an incentive is 11.75 minutes. After four quarters, a 2.39 minute reduction is a 20.3 percent reduction starting from delays of 11.75 minutes.

The second column of Table 4 and the middle panel of Figure 1 present results for departure delays. The results are very similar to those for arrival delays, and reinforce the general point that arrival and departure delays are highly correlated. Flights that depart late will arrive late. The third column of Table 4 and the bottom panel of Figure 1 present results for padding (changes in scheduled - actual flight times). If an airline wanted to strategically game an incentive, increasing the amount of time scheduled per flight would be the most straightforward way to do so. By quarter 4, we do see an increase in padding of 0.50 minutes, or 30 seconds. While this is evidence of strategic gaming, it is not large in magnitude.

Table 4 and Figure 1 illustrate the effect of introducing an incentive using a dynamic, staggered difference-in-difference estimator. We can also pool the quarterly data from before and after the introduction of the incentive to present just a staggered difference-in-difference estimator. This estimator gives us in effect a time series average of the coefficients for quarters one through four after the introduction of the incentive. While obviously less rich than the dynamic results in the previous specification, this specification is a useful transition to later carrier-by-carrier results that demonstrate the heterogeneity in our data. We estimate a generalized staggered difference-in-difference specification of the type:

$$\begin{aligned} P_{jrt} &= \alpha_{DID} \cdot (Incentive_j \cdot Post_{jt}) + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ &\quad + \sum_k \gamma_{krt} \cdot k\_InRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt} \end{aligned} \quad (2)$$

P_{jrt} is one of our three on-time measures for carrier j, in route r, at time t. $Post_{jt}$, equals 1 if the observation year-quarter t is *after* carrier j implemented the incentive. $Incentive_j$ equals 1 if observation carrier j implemented an on-time arrival incentive within our sample period. The difference-in-difference variable is $DID_{jt} = Incentive_j \cdot Post_{jt}$. To control for the presence of competitors in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. We also control for the CEO's total compensation. We include year quarter fixed effects and route carrier fixed effects. Note that including year quarter fixed effects eliminates the Post dummy, and including route carrier fixed effects eliminates the Incentive dummy, as it is measured at the carrier level. This leaves the interaction term, which is the DID coefficient. The error term is ζ_{jrt}.

The results are in Table 5. Panel A includes $origin \cdot year$ and $destination \cdot year$ fixed effects. Panel B replaces $origin \cdot year$ and $destination \cdot year$ fixed effects with origin and destination time trends ($Origin_{rt} \cdot Timetrend$ and $Dest_{rt} \cdot Timetrend$), where $Timetrend$ takes a value between 1 and 72 corresponding to each quarter in our sample. Panel A corresponds to our previous results in Table 4. The results in both panels are very similar. Using $origin \cdot year$ and $destination \cdot year$ fixed effects is much more stringent than using origin and destination time trends, as the former uses dramatically more degrees of freedom. This binds later when we report carrier-by-carrier results. The similarity in these two panels suggests that our subsequent carrier-by-carrier results are unlikely to be driven by the use of origin and destination time trends. We report the difference-in-difference coefficient, α_{DID}. The left two columns in both panels present the estimates for arrival and departure delays. Treated carriers reduce their arrival and departure delays from before to after the introduction of the incentive relative to control carriers by over a minute. Depending upon specification, this is about ten percent of the magnitude of delays for carriers without an incentive.

The third and fourth columns show the percentage of flights delayed 15 minutes or more. The fraction of flights delayed 15 minutes or more is reduced by 2.6 percentage points for treated carriers relative to control carriers from before to after the introduction of the incentive. From Table 2, Panel C, carriers without an incentive have arriving flights delayed 15 minutes or more 23 percent of the time. A 2.6 percentage point reduction thus represents a decrease in arrival delays of 11.3 percent. Further, unconditionally arrival delays of 15 minutes or more decrease from 23 percent to 20 percent if a carrier has an incentive in Table 2, Panel C. Our staggered difference-in-difference estimator suggests that most (87 percent) of this reduction is due to the introduction of the incentive (2.6/3.0=0.87). The fifth column shows that padding increases by about 15 seconds.

Our results in Tables 4 and 5 and Figure 1 demonstrate that carriers introducing an on-time incentive show substantial reductions in arrival delays. Most of this reduction comes from reducing departure delays. Anecdotally, departure delays are reduced by better coordination between ground (baggage) personnel, gate attendants, maintenance crews, servicing crews, flight crews, and better scheduling of times between flights. Interestingly, while we see evidence of gaming by increasing scheduled times of flights, the magnitudes are not large on average.

3.2 Carrier-by-Carrier Differences-in-Differences

Our previous results use pooled and staggered difference in difference specifications to isolate the effect of an introduction of an incentive on treated carriers. As an alternative, we can use the richness of our data to examine heterogeneous carrier-by-carrier changes in on-time performance around the introduction of an incentive. Each carrier introduces an on-time arrival incentive at a different point in time (indeed, Delta introduces and then removes its incentive twice before instituting it permanently). We use this fact to estimate the reduction in arrival delays for a carrier around the introduction of the incentive relative to carriers that do not have an incentive at that time. The advantage of this approach is that it allows us to estimate the reduction in arrival delays by carrier at the time of the introduction of the incentive. The disadvantage of this approach is that each carrier's specification uses less of the data, thus limiting our ability to saturate the specifications with fixed effects.

For each of the carriers that institutes an incentive, we select a time window of four quarters before and four quarters after the carrier introduces an incentive. We then select as the control all those carriers that have not introduced an incentive at any time prior to the selected time window. For example, Delta introduced an incentive in the third quarter of 1997. The only carrier that had an incentive in place before Q3 1997 is Continental, so we include in the control group for Delta all carriers excluding Continental. For carriers that instituted an incentive in the later part of our full sample period, such as US Airways, the set of control carriers is limited to those carriers that have not yet introduced an incentive within our sample years of 1993-2010, which is essentially Southwest. When Southwest introduces its incentive in 2009, there are no control carriers for the before period (Q1-4 of 2008), and so we cannot estimate our difference-in-difference specification for Southwest. We also deal with carriers that introduce an incentive in overlapping time windows as follows: United, American, and Delta2 all introduce an incentive in Q1 2004. Therefore, none of them is used as a control carrier for the others in this specification. We run a difference-in-difference specification of the type:

$$\begin{aligned} P_{jrt} \;=\; & \alpha_j \cdot Carrier_j + \alpha_{After} \cdot After_{jt} + \alpha_{DID} \cdot (Carrier_j \cdot After_{jt}) + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ & + Origin_{rt} \cdot Timetrend + Dest_{rt} \cdot Timetrend + \zeta_{jrt} \end{aligned} \tag{3}$$

P_{jrt} is one of three on-time measures for carrier j, in route r, at time t: minutes of arrival and departure delays and scheduled minus actual flight times. $Carrier_j$ is a categorical variable equal to 1 if the observation corresponds to carrier j; $After_{jt}$ equals 1 if the observation corresponds to the time window *after* carrier j implements the incentive; $Carrier_j \cdot After_{jt}$ is the interaction term that captures that effect on arrival delays following the implementation of the on-time performance incentive, and is the main variable of interest. In this specification, we use origin and destination time trends ($Origin_{rt} \cdot Timetrend$ and $Dest_{rt} \cdot Timetrend$) in place of $origin \cdot year$ and $destination \cdot year$, as the latter uses too many degrees of freedom when estimating separate carrier-by-carrier regressions. The error term is ζ_{jrt}.

Note again that we run specification (3) separately, one for each carrier that receives an incentive, and we include only the carrier that receives the incentive and the carriers that have not yet received the incentive. As we advance in calendar time, the set of carriers without an incentive drops, and hence the set of control carriers falls.

Table 6 contains the summary results. In Panel A, the first row contains the difference-in-difference coefficient (α_{DID}) for arrival delays for each treated carrier's introduction of an incentive ($Carrier_j \cdot After_{jt}$). Each entry in this table is one difference-in-difference coefficient from a separate regression–we are not reporting each full regression, just the coefficient of interest. All carriers, except Delta2 (DL2) and JetBlue (B6), show significant reductions in arrival delays after the introduction of an incentive. For example, the difference-in-difference coefficient for Continental of -1.94 implies that Continental reduces its arrival delays by 1.94 minutes from the four quarters after the introduction of the incentive relative to the four quarters prior to the incentive and relative to rival carriers without an incentive (in this case, all of them, as Continental was the first to introduce an incentive).

The second row reports the difference-in-difference coefficient for departure delays and finds similar effects, although Alaska, United, and Delta3 (in addition to JetBlue) do not show significant reductions in departure delays, while Delta2 does. In the third row, we look at scheduled - actual times. When looking at the difference-in-difference coefficients, some carriers decrease their scheduled - actual times after the incentive and some increase them (strategic gaming). In particular, Alaska, United, American, and Delta3 stand out for strategic gaming by increasing their scheduled - actual times substantially. Interestingly, Delta2 decreases its scheduled - actual time, thus offsetting its reduction in departure delays and resulting in no change in arrival delays. This heterogeneity across carriers in strategic gaming shows that responses to the introduction of incentives can be quite nuanced.

Summing up, we can easily see how much of the change in arrival delays is due to reductions in departure delays versus changing the scheduled - actual times (strategic gaming). For example, Delta1 dramatically reduces its arrival delays. Most of this is due to decreasing departure delays–getting the departing flight out of the gate on-time. However, some of the reduction in arrival delays is also due to increasing the scheduled - actual time of the flights.

Panel B presents similar difference-in-difference coefficients to those in Panel A, but looks at the fraction of flights delayed 15 minutes or more. The results are similar to those in Panel A, although interestingly, Alaska and United show some improvement in departure delays around the 15 minute threshold, suggesting that not all of the improvement in arrival delays is due to strategic gaming of scheduled times.

We next examine how competition shapes the response of carriers to incentives. Our conjecture is that routes on which carriers face competition prior to the introduction of an incentive will show improved on-time performance from the incentive, but less than on routes that are monopoly routes. In effect, competition provides some inherent discipline, while incentives then interact with competition to further improve performance. Monopoly routes, by contrast, are only affected by the incentive, and therefore have more scope for improvement from the incentive.

Table 7 again shows the difference-in-difference coefficient for each specification. Panel A of Table 7 considers the effects of incentives on a carrier on those routes where that carrier faces competition from carriers without an incentive. We are excluding monopoly routes, and routes in which both carriers have an incentive. For arrival delays, we generally find similar effects to those in Table 6 with the exceptions that American and Delta3 do not significantly reduce arrival delays. Departure delays are also similar to what we find in Table 6 with the exceptions of American and Delta2. Interestingly, we find a little more evidence of padding when a treated carrier is on a competitive route, as US Airways also exhibits padding behavior.

Panel B considers what happens on routes in which the carrier has a monopoly. Again, we find similar effects to our overall results. While we initially conjectured that incentives would show larger effects on monopoly routes (since on these routes, there is no competition to discipline performance), we do not see this consistently across carriers. One possibility is that as carriers have relatively more monopoly routes, the disciplinary effect of incentives dominates the disciplinary effect of competition in all subsamples (see Golan, Parlour, and Rajan (2015). Thus, the effects of the incentive seem to be fairly uniform across different types of competition (monopoly versus competitive) that carriers face, while the carriers themselves exhibit substantial heterogeneity in their response to the incentive.

In Panel C, we ask how carriers that do not have an incentive respond to a carrier introducing an incentive. In response to the first carriers introducing an incentive, rival carriers reduce their arrival delays, as can be seen in how rivals respond to Continental, Delta, and Alaska introducing incentives. For example, when Delta first introduces an incentive, rivals on the same routes as Delta reduce their arrival delays by 1.41 minutes relative to the rival's routes where they do not compete with Delta. This effect fades for later introductions of incentives, as the number of competitors without incentives shrinks. This can be seen in the reduction of observations across the columns in the panel.

In Table 8, we repeat the prior difference-in-difference analysis, but focus on the frequency of flights delayed 15 minutes or more. The results in Table 8 are similar to those we reported in Table 7 for competition, monopoly routes, and the effects of introducing an incentive on competitor delays. While we observe some competitive responses from the introduction of on-time incentives, we do not observe substantial differences based on type of competition (monopoly versus oligopoly), while still observing substantial heterogeneity across carriers.

3.3 Panel Fixed Effects Specifications

We further examine the nature of responses to the introduction of incentives by using all of our panel data, and not just the four quarters before and four quarters after the introduction of incentives. We estimate a panel fixed effects regression of the type:

$$\begin{aligned} P_{jrt} &= \alpha_{Own} \cdot Own_{jrt} + \alpha_{Rival} \cdot Rival_{jrt} + \alpha_{Joint} \cdot Joint_{jrt} + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ &\quad + \sum_{k} \gamma_{krt} \cdot kInRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt} \end{aligned} \tag{4}$$

P_{jrt} is one of the performance measures for carrier j, in route r, at time t: arrival/departure delays in minutes or percent of flights delayed 15 minutes or more, and scheduled minus actual flight times. We will later also examine available seats, enplaned passengers, load factors, scheduled and actual performed flights, and fares or prices. Own, $Rival$, and $Joint$ refer to the three categorical variables for whether a carrier on a route has an incentive, whether a rival carrier has an incentive, or both a carrier and its rival have an incentive. To control for a competitor's presence on a route we include $kInRoute_{rt} = 1$ if carrier k serves route r at time t. We include route-carrier fixed effects u_{rj}, year-quarter fixed effects u_t, $origin \cdot year$ fixed effects u_{oy}, and $destination \cdot year$ fixed effects u_{dy}. Finally, ζ_{jrt} is an unobservable error term.

Table 9, Panel A, provides results from this specification. The left two columns examine the reduction in the minutes of arrival and departure delays. The omitted category is observations with no incentive at all ($None$), so coefficients should be interpreted as reductions or increases in delays in minutes relative to routes with no airlines having an incentive in that quarter. Airlines that institute an on-time arrival incentive significantly decrease their arrival delays by 0.43 minutes. While this is smaller than our difference-in-difference estimates, here we use all of our sample data, and not just the data in the four quarters before and four quarters after the introduction of the incentive. It is important to note that we are saturating this specification with thousands of fixed effects and controls and still find a significant incentive effect.

For competitors, if a rival airline in the same route initiates an on-time arrival incentive, then the competitor airline that did not introduce the incentive still significantly reduces its arrival delays by 0.34 minutes after controlling for route-carrier fixed effects, $origin \cdot year$ and $destination \cdot year$ fixed effects, total CEO compensation, bankruptcy, mergers, and other carriers flying the same route. If both airline competitors have an on-time arrival incentive ($Joint$), then the reduction in delays is 0.56 minutes, similar to the effect of Own. Since our four categories–$None, Own, Rivals$,and $Joint$–are mutually exclusive, there is a small incremental benefit from both an airline and its rival having an incentive relative to only an airline or its rival having an incentive. The competitive effect on rival airlines is interesting–in principle, if one airline introduces an incentive, its rival airlines still get the benefit of a reduction in delays even though they do not have an incentive. Recall that the summary statistics in Table 2, Panels A through C, showed an increase in delays for $Joint$ relative to $None$. This increase is reversed in the regression results by controlling for route-carrier fixed effects and $origin \cdot year$ and $destination \cdot year$ fixed effects.[17]

The second column of Panel A repeats the analysis for departure delays and finds similar results. This again demonstrates that arrival delays are mostly caused by departure delays–the reduction in departure delays feeds through to reductions in arrival delays. The third and fourth columns of Table 9, Panel A, consider the fraction of flights delayed more than 15 minutes. Instituting an on-time arrival incentive is associated with a reduction in the fraction of flights delayed more than 15 minutes by a significant 1 percentage point. For rivals, the reduction in the fraction of flights

[17]In the raw data, the large delays for $Joint$ are primarily due to United and American Airlines. These two airlines compete directly at their joint hub O'Hare Airport in Chicago, and both instituted their incentives at the same time. Our regression results control for these effects.

delayed more than 15 minutes is not significant. If both competitors on a route have on-time arrival incentives, then the frequency of arrival delays decreases by 2 percentage points for both airlines. We find similar results for departure delays in the fourth column.

Another possibility is that airlines might try to game the on-time arrival incentive. When faced with a new incentive, executives might strategically increase the scheduled amount of time for a flight. To examine this, we look at the difference between the CRS scheduled flight time and the actual elapsed time for flights, in the fifth column of Panel A, Table 9. An airline instituting an on-time arrival incentive decreases its scheduled flight time relative to the actual flight times by a statistically insignificant amount, after all of our controls. In other words, introducing an incentive does not appear to result in much strategic gaming in this specification.

For rival airlines, however, there is a statistically significant increase in the amount of time a flight is scheduled for minus the actual flight time of 0.20 minutes, or 12 seconds. Therefore, there is some evidence of strategic gaming by rivals in response to an airline introducing an incentive. If both have an on-time arrival incentive, then both will incrementally increase their scheduled flight times by 0.34 minutes. That both airlines increase their scheduled times on a route in response to both having an incentive could suggest some degree of collusive behavior, but this effect is small in magnitude.

In Panel B of Table 9, we restrict attention to problem routes. Here we find effects that are larger in magnitude. Having an on-time arrival incentive is associated with a reduction in arrival delays of 2.38 minutes. The effect for rivals is much smaller, and is statistically insignificant. When both have an incentive, there is a reduction in delays of 1.46 minutes, somewhat smaller in magnitude than when only a given carrier has an incentive. All of these effects are similar for departure delays.

In the third and fourth columns, we find that having an incentive is associated with a decrease in the frequency of delayed flights of four percentage points for arrivals and three percentage points for departures. There are no significant effects on rivals. We find smaller effects when both carriers have incentives relative to when a given carrier has an incentive. The effects for *Own* here are larger in magnitude, and suggest that the incentive is particularly effective in reducing delays below the 15 minute threshold in problem routes.

Finally, for the difference between scheduled and actual flight times, there is an increase associated with having an incentive of 0.82 minutes, so there is some evidence of strategic gaming on problem routes. However, the reduction in departure delays of 1.99 minutes is 2.4 times larger, suggesting that the incentive is also working as intended on these routes by inducing reductions in departure delays that then account for the reductions in arrival delays. For rivals and when both carriers have an incentive, the difference between scheduled and actual flight times is insignificant.

3.4 The Effects of the Bonus Magnitude

In Table 10, Panel A, we make use of variation in the size of the incentive. For a subset of our sample, we have data on the fraction of the overall bonus that depends on the on-time incentive. This ranges from as little as 0.1% up to 33.5% of the overall bonus, with a mean of 10.9% and a standard deviation of 11.8%. In this analysis, we use a continuous measure of the own bonus, and so omit the indicator variables for $Own, Rivals$, and $Joint$.[18] The effects here are substantial. In the first column, the coefficient of -13.18 implies that an increase in the fraction of the bonus tied to on-time performance of one standard deviation, 11.8 percentage points, reduces arrival delays by 1.56 minutes. This effect is smaller for departure delays (in the second column), and we find a similar pattern for the percentage of flights delayed 15 minutes or more (in the third and fourth columns). The fifth column shows why: as the size of the incentive increases, carriers increase the scheduled amount of time for flights relative to the actual time required to complete the flights. A one standard deviation increase in the fraction of the bonus tied to on-time performance increases the padding time by $17.07 \cdot 0.118 = 2.01$ minutes. This is evidence of substantial strategic gaming, as executives have more direct control over the scheduled length of flights, and they increase the scheduled time as more of their bonuses depend on on-time arrival.

We further examine variation in the size of the bonus tied to on-time performance by partitioning the sample into large bonuses and small bonuses. In several cases (American and Alaska), the CEO's bonus tied to on-time arrival was less than one percent of the total bonus. In these cases, the airline instituted an on-time bonus incentive for all employees in which the CEO also participated, but the magnitude of the incentive for all employees was small (typically several hundred dollars per employee). As the CEO received no additional bonus for on-time performance, the standard bonus for all employees amounted to a small fraction (less than one percent) of the CEO's total bonus,

[18] Because we do not have the continuous measure for all carriers (some carriers do not report the fraction of the bonus that depends on on-time arrival), we do not include continuous versions of $Rivals$ and $Joint$.

which depended more heavily on other factors such as operating income or revenue growth. These cases allow us to examine the impact of small bonuses, as well as whether bonuses for all employees matter versus a large bonus for the CEO. Table 10, Panel B, contains the results. The first column shows that arrival delays are strongly decreasing in the bonus percentage for bonus percentages greater than one percent, consistent with Panel A, Column 1. The second column of Table 10, Panel B, examines bonus percentages less than one percent. As these magnitudes are small, rather than using the bonus percentage as the independent variable, we simply use the categorical variable $Own = 1$ if the CEO has an on-time arrival incentive but it equals less than one percent of the CEO's total bonus. The effect of the incentive on arrival delays is insignificant. These results are confirmed for departure delays and padding in the subsequent columns. These results imply that the magnitude of the incentive matters, and importantly, that in this case, top management incentives matter more than incentives for all employees.

3.5 Placebo Test

To assess the robustness of our primary results, we perform a placebo test. For each of the carriers, we randomly assign an on-time arrival incentive in any of the 72 quarters that comprise our sample period. The carrier maintains the placebo on-time incentive for all quarters after the randomly assigned quarter of incentive initiation. We then pool the carriers with their placebo random treatments and estimate the effect of the on-time incentive (Own) using the specification from Table 9, but we omit the categorical variables Rival and Joint. We run 1000 Monte Carlo simulations and report the coefficients and standard errors for this placebo test in Table 11, Panel A. For none of the columns–arrival delays or departure delays in minutes or the fraction of flights delayed 15 minutes or more or padding–do we find that the coefficient on the on-time incentive (Own) is statistically different from 0. Panel B reports the analogous results from our actual data for ease of comparison (the specification is Equation 4 excluding the categorical variables Rival and Joint).

3.6 Mean Reversion

A potential concern with our results is a specific form of endogeneity. In particular, carriers may introduce an incentive when on-time performance is especially poor. If performance mean-reverts, then we may incorrectly ascribe the performance improvement to the incentive rather than simple mean reversion.

To account for the mean reversion explanation, we use an "Ashenfelter Dip Strategy", where we estimate specifications similar to those in Table 9, but drop observations from two quarters prior to and two quarters after the introduction of an incentive.[19] The idea is that we compare on-time performance post the introduction of an incentive to at least three quarters prior to the introduction of an incentive, thus removing the effect of the particularly poor quarters that may have induced the introduction of the incentive. The results are in Table 12. In all three panels, the results are similar to those in Table 9, although a little smaller in magnitude. Thus, we conclude that mean reversion is unlikely to be driving our results.

4 Competition

4.1 Presence of Southwest Airlines

Another concern with our results is that they might be driven by a competitive response to an entering airline, specifically Southwest Airlines. Goolsbee and Syverson (2008) show that incumbent carriers respond to the threat of entry by Southwest Airlines by cutting fares preemptively. Prince and Simon (2015) examine whether incumbents adjust their quality provision in response to entry and entry threats in the airline industry: specifically, they assess how incumbent airlines adjust their on-time performance in response to entry and entry threats by a particularly active (potential) entrant, Southwest Airlines. They find that incumbents' on-time performance actually worsens in response to entry and entry threats by Southwest. Since Southwest is known to have the best on-time performance of major carriers, it is also possible that other carriers might improve on-time performance in response to the threat of entry by Southwest. To address this possibility, in Table 13, Panel A, we exclude Southwest and all routes on which Southwest operates. The results are essentially unchanged relative to Table 9, although the magnitude of the coefficient when a carrier has an incentive is larger. Thus, our results are not driven by the presence of Southwest Airlines.

To further examine the impact of Southwest Airlines, in Table 13, Panel B, we introduce an

[19] See Ashenfelter (1978). In his study of the effect of training programs on earnings Ashenfelter noted that trainees suffered an unexpected drop in earnings in the year prior to entering the training program. To deal with this, Ashenfelter dropped the year prior to entering a training program. Addressing stylized facts in this way became known as the "Ashenfelter dip" and has been applied in different studies since.

additional independent variable for the year that Southwest entered a route.[20] In the first column, in the year that Southwest entered a route, arrival delays drop by 0.64 minutes independent of all other effects. Having an incentive also reduces arrival delays by 0.40 minutes, a similar magnitude to what we find in Table 9. We find similar effects for departure delays and the percentage of flights delayed fifteen minutes or more. In sum, while Southwest's presence does reduce delays, Southwest Airlines does not explain our on-time incentive results. Instead, the effect of an on-time incentive seems to be similar in magnitude to the competitive pressure other carriers experience when Southwest enters a route.

4.2 Capacity and Prices

In addition to the strategic gaming of scheduled times discussed before, it is also possible that carriers could reduce capacity in response to on-time arrival incentives. The idea here is that having more flights and passengers increases the likelihood of delays, either because it takes longer to board more passengers, longer to fuel flights that are more fully loaded, or because with more flights scheduled, any delays will compound throughout the airline system. Therefore, in order to make it easier to achieve an on-time bonus, carriers might reduce capacity. Table 14 considers the response of capacity and utilization to on-time arrival incentives.

The left three columns consider seats per scheduled flight, passengers per performed flight, and quarterly load factors (a measure of the number of passengers per number of seats per flight-mile). Instituting an on-time arrival incentive is associated with an increase in seats per scheduled flight and passengers per performed flight of one percent on each route. For the rival airline on the same route, passengers per performed flight shrink by two percent with a negligible change in seats per scheduled flight. If both have an incentive, the effects are similar to a rival carrier having an incentive. The load factor does not change for the incentive-initiating, while it decreases for the rival airline and if both have an incentive.[21]

The fourth and fifth columns consider departures scheduled and departures performed. An

[20] In this specification, we do not impose the restriction that a carrier operate on a route before and after the introduction of an incentive, as this would eliminate our ability to examine entry into routes by Southwest.

[21] These results are sensitive to the inclusion of $origin \cdot year$ and $destination \cdot year$ fixed effects, which absorb much of the variation. Excluding these fixed effects, we find that load factors increase for all three incentive categories relative to no incentive, consistent with Dana and Orlov (2014), who find that over our sample period, load factors for airlines have consistently increased.

incentive-initiating airline insignificantly decreases its departures scheduled and performed. The rival airline on the same routes decreases its departures scheduled and performed by three percent. If both have an incentive, then departures scheduled and performed decrease by ten percent. These results suggest that the increases in flights scheduled and performed in Table 3, Panel A, for *Rivals* and *Joint* are explained by our control variables, in particular $origin \cdot year$ and $destination \cdot year$ fixed effects, and are not attributable to the introduction of on-time incentives. The primary conclusion from Table 14 is that airlines do not reduce capacity in response to the introduction of an incentive. If anything, airlines with incentives increase capacity and passengers, and yet still show improvement in on-time arrivals.

To complete the picture of how on-time incentives affect competition, we also consider the impact on prices. Our previous results suggest that competition between airlines in on-time performance is competition in strategic complements, and that introducing an on-time incentive toughens competition. We also see that airlines with an incentive increase capacity on their routes.

Column 6 of Table 14 shows the effect on fares by route when a carrier has an incentive. In this specification, we additionally control for the presence of Southwest and other low cost carriers. Here we find that carriers with an incentive have fares that are lower by one percent. Rivals with an incentive have no impact on fares, and there is no effect if both a carrier and its rivals have an incentive. While we do find that the incentive is associated with a carrier having lower fares, consistent with the incentive being part of toughening competition by a carrier, the effect is small in magnitude. A carrier introducing an incentive is associated with a reduction in delays, a small increase in quantities, and a small reduction in prices.

5 Conclusion

We find that airlines reduce the length and frequency of arrival delays in response to the introduction of on-time arrival bonuses for their executives. Because these bonuses were staggered in their introduction across airlines over our 18 year sample period, we are able to identify the effect of the bonus on arrival delays. Because on-time arrival is an important quality metric for passengers and airlines, we can directly observe the impact of an incentive on the outcome variable it is meant to influence.

We find evidence that part of the improvement in on-time performance comes from strategic gaming by airlines, especially when incentives are more high-powered; that is, when the on-time arrival incentive is a larger fraction of the total bonus. We are able to isolate the strategic gaming effect most precisely in our difference-in-difference analysis. While most airlines reduce departure delays in response to the introduction of incentives (and therefore reduce arrival delays), we do find some instances of airlines simply adjusting the scheduled length of their flights. Nonetheless, for most airlines, most of the improvement comes from the reduction in departure delays (getting airplanes out of the gate on-time).

We find some evidence that competitors increase their on-time arrival performance in response to a rival introducing an incentive, even if the competitor does not have an on-time arrival incentive in place. This suggests that competition in quality is competition in strategic complements. Overall, our evidence suggests that incentives work as intended when there is clear observability and measurability of performance goals.

References

Aggarwal, Rajesh K., and Andrew A. Samwick, 1999, Executive compensation, relative performance evaluation, and strategic competition: Theory and evidence, *Journal of Finance* 54, 1999–2043.

Ashenfelter, Orley, 1978, Estimating the effect of training programs on earnings, *The Review of Economics and Statistics* 60, 47–57.

Bennett, Benjamin, Carr J. Bettis, Radhakrishnan Gopalan, and Todd Milbourn, 2017, Compensation goals and firm performance, *Journal of Financial Economics* 124, 307–330.

Berry, Steven, 1990, Airport presence as product differentiation, *American Economic Review Papers and Proceedings* 80, 394–399.

Borenstein, Severin, 1989, Hubs and high fares: Dominance and market power in the us airline industry, *RAND Journal of Economics* 20, 344–365.

Borenstein, Severin, and Nancy Rose, 1995, Bankruptcy and pricing behavior in the U.S. airline markets, *American Economic Review Papers and Proceedings* 85, 397–402.

Borenstein, Severin, and Nancy Rose, 2003, The impact of bankruptcy on airline service levels, *American Economic Review Papers and Proceedings* 93, 415–419.

Brueckner, Jan K., Nichola J. Dyer, and Pablo T. Spiller, 1992, Fare determination in airline hub-and-spoke networks, *RAND Journal of Economic* 23, 309–333.

Ciliberto, Federico, and Carola Schenone, 2012a, Are the bankrupt skies the friendliest?, *Journal of Corporate Finance* 18, 1217–1231.

Ciliberto, Federico, and Carola Schenone, 2012b, Bankruptcy and product-market competition: Evidence from the airline industry, *The International Journal of Industrial Organization* 20, 564–577.

Dana, James D., and Eugene Orlov, 2014, Internet penetration and capacity utilization in the us airline industry, *American Economic Journal Microeconomics* 106–137.

Dennis, Patrick J., Kristopher Gerardi, and Carola Schenone, 2018, Common ownership does not have anti-competitive effects in the airline industry, University of Virginia Working Paper.

Dresner, Martin, and Kefeng Xu, 1995, Customer service, customer satisfaction, and corporate performance in the service sector, *Journal of Business Logistics* 16, 23–40.

Flammer, Caroline, Bryan Hong, and Dylan Minor, 2018, Corporate governance and the rise of integrating corporate social responsibility criteria in executive compensation: Effectiveness and implications for firm outcomes, *https://ssrn.com/abstract=2831694* .

Forbes, Silke J., M. Lederman, and T. Tombe, 2015, Quality disclosure programs and internal organizational practices: Evidence from airline flight delays, *American Economic Journal: Microeconomics* .

Gerardi, Kristopher S., and Adam Hale Shapiro, 2009, Does competition reduce price dispersion? New evidence from the airline industry, *Journal of Political Economy* 117, 1–37.

Golan, Limor, Christine A. Parlour, and Uday Rajan, 2015, Competition, managerial slack, and corporate governance, *The Review of Corporate Finance Studies* 4, 43–68.

Goolsbee, Austan, and Chad Syverson, 2008, How do incumbents respond to the threat of entry? Evidence from the major airlines, *Quarterly Journal of Economics* 123, 1611–1633.

Hermalin, Benjamin, 1992, The effects of competition on executive behavior, *RAND Journal of Economics* 23, 350–65.

Holmstrom, Bengt, and Paul Milgrom, 1987, Aggregation and linearity in the provision of intertemporal incentives, *Econometrica* 55, 303–328.

Kin, Han E., and Vijay Singal, 1993, Mergers and market power: Evidence from the airline industry, *American Economic Review* 83, 549–469.

Lin, Chen, Liu Sibo, and Gustavo Manso, 2018, Shareholder litigation and corporate innovation, University of California at Berkeley Working Paper.

Mazzeo, Michael, 2003, Competition and service quality in the us airline industry, *Review of Industrial Organization* 22, 275–296.

Nijman, Theo, and Marno Verbeek, 1992, Nonresponse in panel data: The impact on estimates of a life cycle consumption function, *Journal of Applied Econometrics* 7, 243–257.

Oliver, Richard, 2003, Flight frequency and mergers in airline markets, *International Journal of Industrial Organization* 21, 907–922.

Oyer, Paul, 2004, Why do firms use incentives that have no incentive effects?, *The Journal of Finance* 59, 1619–1649.

Prince, Jeffrey, and Daniel Simon, 2015, Do incumbents improve service quality in response to entry?, *Management Science* 61, 372390.

Rupp, Nicholas, D. Owens, and L. Plumly, 2006, Does competition influence airline ontime performance?, *Advances in Airline Economics: Competition Policy and Antitrust, ed. D. Lee* .

Schmidt, Klaus M., 1997, Managerial incentives and product market competition, *The Review of Economic Studies* 64, 191–213.

Snider, Connan, and Jonathan Williams, 2015, Barriers to entry in the airline industry: A multidimensional regression-discontinuity analysis of air-21, *Review of Economics and Statistics* 97, 2897–2929.

Stevenson, Betsey, and Justin Wolfers, 2006, Bargaining in the shadow of the law: Divorce laws and family distress, *The Quarterly Journal of Economics* 4, 267–288.

Table 1: **Timing of Incentives**

Year Quarters of Incentives by Carrier

Air Carrier	**Introduce Incentive**	**Withdrew Incentive**
Continental	1995- Q1	No
Delta	1997- Q3	2002-Q4
Alaska	2002- Q1	No
United	2004- Q1	No
American	2004- Q1	No
Delta	2004- Q1	2005-Q4
JetBlue	2007- Q1	2008-Q4
Delta	2007- Q2	No
US Airways	2008- Q1	No
Southwest	2009- Q1	No

For our sample years, 1993-2010, we report the year quarters during which the carrier had an on-time incentive with starting and ending quarters. If a carrier does not withdraw its incentive, the ending quarter is 2010-Q4.

Table 2: **Arrival and Departure Delays, Scheduled Minus Actual Flight Times**

Panel A: Raw Summary Statistics
Early Arrivals (departures) are coded as negative minutes

	None	Own	Rivals	Joint
Arrival Delays	6.12	6.27	6.85	7.80
	(0.02)	(0.04)	(0.07)	(0.07)
Departure Delays	7.94	7.98	9.38	10.40
	(0.02)	(0.03)	(0.05)	(0.06)
Scheduled-Actual	1.81	1.71	2.53	2.60
("Pad" times)	(0.01)	(0.02)	(0.04)	(0.04)
Observations	124,995	61,017	16,562	20,418

Panel B: Year-Quarter Adjusted Summary Statistics
Early Arrivals (departures) are coded as negative minutes

	None	Own	Rivals	Joint
Arrival Delays	8.22	6.83	6.49	8.57
	(0.13)	(0.35)	(0.62)	(0.83)
Departure Delays	8.64	7.45	8.26	6.82
	(0.10)	(0.27)	(0.43)	(0.66)
Scheduled-Actual	0.42	0.62	1.77	-1.76
("Pad" times)	(0.08)	(0.20)	(0.37)	(0.52)
Observations	124,995	61,017	16,562	20,418

Panel C: Minutes Delays ≥ 0 and Delays $\geq 15'$

	Minutes Delayed≥ 0				**Pct. Delays 15' or More**			
	None	Own	Rivals	Joint	None	Own	Rivals	Joint
Arrival Delays	11.75	10.90	10.59	11.82	0.23	0.20	0.20	0.21
	(0.11)	(0.29)	(0.45)	(0.77)	(0.00)	(0.01)	(0.01)	(0.01)
Departure Delays	9.18	8.59	8.77	8.20	0.17	0.15	0.15	0.14
	(0.09)	(0.26)	(0.41)	(0.57)	(0.00)	(0.00)	(0.01)	(0.01)
Observations	124,995	61,017	16,562	20,418	124,995	61,017	16,562	20,418

Panel D: At least 30% of flights are delayed 15 minutes of more

	Minutes Delayed≥ 0				**Pct. Delays 15' or More**			
	None	Own	Rivals	Joint	None	Own	Rivals	Joint
Arrival Delays	17.68	13.33	15.43	18.55	0.32	0.25	0.32	0.31
	(0.58)	(0.76)	(1.05)	(1.09)	(0.01)	(0.02)	(0.02)	(0.02)
Departure Delays	14.81	10.77	12.60	13.10	0.26	0.18	0.24	0.20
	(0.54)	(0.64)	(0.94)	(0.88)	(0.01)	(0.01)	(0.02)	(0.01)
Scheduled-Actual	1.34	2.21	-0.37	-2.86				
("Pad" times)	(0.47)	(0.66)	(0.69)	(0.81)				
Observations	6,947	6,631	1,782	2,236	6,947	6,631	1,782	2,236

Statistics for arrival and departure delays (minutes), and difference between scheduled and actual flight times. *None*= 1 for observations in which no carrier has an incentive. *Own*= 1 for observations in which the carrier has an incentive, and *Rivals*= 1 for observations in which a carrier faces a rival that has an incentive. *Joint*= 1 for observations in which a carrier and at least one rival in a route have an incentive. *For Own, Rival, and Joint, a carrier must serve a route before and after it implements the incentive.* Panel A reports summary statistics for early arrivals (departures) coded as a negative number. Panels B, C, and D control for year-quarter fixed effects. Panels C and D recode early arrivals (departures) as 0 minutes delayed. "*Pad*" times are not reported in Panel C as they are identical to those in Panel B. Panel D restricts the sample to "problem routes," those where the carrier that introduces an incentive reports at least 30% of flights arrive at least 15 minutes delayed *prior* to the introduction of the incentive. Flight time is from the On-Time Performance database maintained by the BTS.

Table 3: **Summary Statistics for Capacity Measures and Itinerary Fares**

Panel A: Capacity

	None	Own	Rivals	Joint
Part 1: In Route Both Before and After Incentive				
Seats per Scheduled	137.03	141.81	136.26	142.09
	(0.11)	(0.11)	(0.19)	(0.16)
Pass. per Performed	92.08	103.90	96.16	112.58
	(0.08)	(0.13)	(0.18)	(0.19)
Load Factor	0.67	0.72	0.70	0.78
	(0.00)	(0.00)	(0.00)	(0.00)
Deps. Scheduled	364.65	343.35	416.45	411.40
	(0.90)	(1.22)	(2.80)	(2.28)
Deps. Performed	359.84	337.94	407.96	403.66
	(0.83)	(1.20)	(2.67)	(2.22)
Observations	117,935	48,320	15,747	17,117
Part 2: In Route Before and After Incentive, *and* problem routes				
Seats per Scheduled	155.47	152.13	143.47	147.48
	(0.54)	(0.51)	(0.65)	(0.54)
Pass. per Performed	110.41	112.50	106.31	119.77***
	(0.44)	(0.53)	(0.67)	(0.59)
Load Factor	0.70	0.72	0.73	0.80***
	(0.00)	(0.00)	(0.00)	(0.00)
Deps. Scheduled	287.91	345.77	291.19	418.79***
	(3.03)	(4.36)	(5.75)	(7.17)
Deps. Performed	284.96	339.68	285.75	410.68***
	(2.97)	(4.24)	(5.61)	(6.92)
Observations	5,848	4,503	1,549	1,681

Panel B: Itinerary Fare

	None	Own	Rivals	Joint
Part 1: In Route Both Before and After Incentive				
Average Itinerary Fare	216.24	213.42	192.09	179.30
	(0.29)	(0.37)	(0.75)	(0.44)
Observations	121,099	47,977	16,188	17,131
Part 2: In Route Before and After Incentive, *and* problem routes				
Average Itinerary Fare	246.61	250.62	206.42	188.40
	(1.30)	(1.38)	(2.09)	(1.20)
Observations	5,843	4,479	1,547	1,681

The categorical variables *None*, *Own*, *Rivals*, and *Joint* are defined as in Table 2. Panel A summarize seats per scheduled flights, enplaned passengers per performed flights, load factor, and departures scheduled and performed. All carriers included in the sample serve the route at least 8 times a month (minimum number of scheduled and performed departures is 24 for a year-quarter-route-carrier), and serves a route both before and after the carrier introduced an incentive. Part 1 includes all such routes, and Part 2 restricts the sample to "problem routes," (those where the carrier that introduces an incentive reports at least 30% of flights arrive at least 15 minutes delayed *prior* to the introduction of the incentive). Panel B reports ticket-price statistics. Part 1 includes all routes, while part 2 restricts the sample to problem routes. For each route-carrier and year quarter observation, we have a random sample of 10% of *all itineraries* fares. The average fare reported here under each categorical variable is the average $Price_{jrq}$ across carriers, routes and year quarter observation falling in each category. Capacity data is from the T100-Domestic Segment of Form 41 merged with the Ontime Performance data. Price data is from the DB1B databases. All databases are from the BTS.

Table 4: **Delays and Padding: Quarters Surrounding the Introduction of Incentive**

	(1) Mins Arr Delays	(2) Mins Dep Delays	(3) Pad: CRS-Actual
Before:			
Quarter -4 * Treated	0.03	-0.11	-0.20
	(0.23)	(0.20)	(0.16)
Quarter -3 * Treated	-0.25	-0.18	0.11
	(0.20)	(0.17)	(0.13)
Quarter -2 * Treated	0.02	0.08	-0.04
	(0.18)	(0.17)	(0.14)
Quarter -1 * Treated	-0.15	0.01	0.09
	(0.16)	(0.15)	(0.13)
After:			
Quarter +1 * Treated	-0.16	-0.36*	-0.06
	(0.23)	(0.22)	(0.13)
Quarter +2 * Treated	-0.83***	-0.85***	0.03
	(0.17)	(0.15)	(0.14)
Quarter +3 * Treated	-1.43***	-1.56***	-0.23*
	(0.17)	(0.15)	(0.14)
Quarter +4 * Treated	-2.39***	-2.24***	0.50***
	(0.24)	(0.21)	(0.18)
Observations	95,428	95,428	95,428
R-squared	0.52	0.52	0.50
Origin and Dest Year FE	X	X	X
Competitor in Route FE	X	X	X
Route-Carrier FE	X	X	X
Year-Quarter FE	X	X	X
Cluster Route-Carrier	X	X	X

We estimate a dynamic staggered difference in difference regression of the type:

$$P_{jrt} = \sum_{q=-4}^{q=4} \alpha_q \cdot Quarter_q \cdot Treated_{jq} + \delta_{jy} \cdot Log(Total\_Comp)_{jy} + \sum_k \gamma_{krt} \cdot k\_InRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt}; q \neq 0$$

Where P_{jrt} is one of three on-time measures for carrier j in route r during year-quarter t: non-negative arrival or departure delays and padding time (CRS time minus actual elapsed flight time). $Quarter_q$ is an indicator equal to one for the q^{th} quarter relative to the introduction of the on-time incentive. Treated$_{jq}$ is an indicator variable equal to 1 if carrier j is treated (the incentive is implemented) at quarter 0. The time window is four quarters *before* the incentive was implemented to four quarters *after* it was introduced (the quarter the incentive was implemented is therefore excluded). To control for the presence of a competitor in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. u_{rj} are route-carrier fixed effects, u_t are year-quarter fixed effects, u_{oy} are $origin \cdot year$ fixed effects, and u_{dy} are $destination \cdot year$ fixed effects. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors are clustered at the route-carrier level. The coefficient estimates of α_k represent the change in minutes in performance due to the introduction of an incentive k quarters away. To illustrate the dynamic effect of introducing the incentive on on-time performance, Figure 1 plots the coefficients α_k. The dashed lines around the coefficient estimates represent the 95% confidence interval, adjusted for route-carrier clustering. Exogeneity of the on-time performance incentive is confirmed by $\alpha_q = 0, \forall q < 0$.

Table 5: **Difference in Difference: Staggered Specification**

Panel A: Origin-Year, and Dest-Year, Fixed Effects

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	Arrival	Departure	Arrival	Departure	Pad Time
DID	-1.057***	-1.011***	-0.026***	-0.019***	0.268**
	(0.157)	(0.139)	(0.002)	(0.002)	(0.114)
Observations	98,996	98,996	98,996	98,996	98,996
R-squared	0.534	0.532	0.560	0.574	0.506
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Origin and Destination Time Trends

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	Arrival	Departure	Arrival	Departure	Pad Time
DID	-1.135***	-1.097***	-0.026***	-0.021***	0.251***
	(0.106)	(0.098)	(0.002)	(0.001)	(0.079)
Observations	99,192	99,192	99,192	99,192	99,192
R-squared	0.444	0.439	0.490	0.508	0.427
Origin and Dest Timetrends	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

We estimate a generalized (staggered) difference-in-difference specification of the type:

$$P_{jrt} = \alpha_{DID} \cdot (Incentive_j \cdot Post_{jt}) + \delta_{jy} \cdot Log(Total\_Comp)_{jy}$$

$$+ \sum_k \gamma_{krt} \cdot k\_InRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt}$$

P_{jrt} is an on-time measures for carrier j, in route r, at time t. $Post_{jt}$ equals 1 if the observation year-quarter t is *after* carrier j implemented the incentive. $Incentive_j$ equals 1 if observation carrier j implements an on-time arrival incentive within the sample window. The difference-in-difference variable is $DID_{jt} = Incentive_j \cdot Post_{jt}$. To control for the presence of a competitor in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. Both panels include route-carrier fixed effects, u_{rj}, and year-quarter fixed effects, u_t. Panel A includes $origin \cdot year$ fixed effects, u_{oy}, and $destination \cdot year$ fixed effects, u_{dy}. Panel B replaces u_{oy} and u_{dy} with origin and destination time trends ($Origin_{rt} \cdot Timetrend$ and $Dest_{rt} \cdot Timetrend$), where $Timetrend$ takes a value between 1 and 72 corresponding to each quarter in our sample. The reported coefficient is α_{DID}. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors, clustered at the route-carrier level, are in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 6: **Difference in Difference by Treated Carrier**

Base Specification

Panel A: Minutes of Delays or Padding Time

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrival	-1.94***	-3.25***	-2.43***	-0.66***	-1.29***	0.43	0.01	-2.27***	-3.93***
	(0.32)	(0.18)	(0.52)	(0.25)	(0.23)	(0.30)	(0.72)	(0.57)	(0.50)
R-squared	0.83	0.86	0.85	0.85	0.85	0.77	0.87	0.69	0.77
Departure	-2.43***	-2.78***	-0.50	-0.26	-0.72***	-0.44*	0.61	0.12	-3.66***
	(0.28)	(0.15)	(0.40)	(0.23)	(0.20)	(0.24)	(0.70)	(0.55)	(0.51)
R-squared	0.81	0.84	0.84	0.82	0.82	0.77	0.87	0.65	0.75
Pad	-0.61***	0.53***	2.76***	1.66***	2.19***	-1.38***	0.67	4.10***	0.22
	(0.23)	(0.17)	(0.33)	(0.28)	(0.22)	(0.26)	(1.18)	(0.44)	(0.26)
R-squared	0.15	0.17	0.47	0.50	0.49	0.41	0.56	0.48	0.63
Observations	23,797	22,095	16,877	10,759	12,385	13,389	9,762	11,037	8,424
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

Panel B: Delays of 15 minutes or more

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrivals	-0.04***	-0.07***	-0.07***	-0.02***	-0.04***	0.01	-0.02	-0.04***	-0.09***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
R-squared	0.83	0.87	0.87	0.86	0.86	0.86	0.88	0.85	0.88
Departures	-0.05***	-0.05***	-0.03***	-0.01***	-0.02***	-0.02***	0.01	0.00	-0.08***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
R-squared	0.79	0.84	0.85	0.83	0.83	0.83	0.87	0.84	0.88
Observations	23,797	22,095	16,877	10,759	12,385	13,389	9,762	11,037	8,424
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

For each treated carrier we run a difference-in-difference regression, using a sample restricted to the four quarters before and four quarters after the treated carrier introduces the incentive. We also restrict the sample to control carriers that do *not* have an incentive during the selected time window. A carrier must serve the route before and after the treated carrier introduces the incentive (a carrier not serving a route both before and after the incentive is dropped, but the route remains in the sample as long as some carrier serves it both before and after the incentive). Each regression specification is of the type:

$$\begin{aligned} P_{jrt} &= \alpha_j \cdot Carrier_j + \alpha_{After} \cdot After_{jt} + \alpha_{DID} \cdot (Carrier_j \cdot After_{jt}) + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ &\quad + Origin_{rt} \cdot Timetrend + \ Dest_{rt} \cdot Timetrend + \ \zeta_{jrt} \end{aligned}$$

P_{jrt} is one of three performance measures: Minutes of arrival and departure delays and padding time. $Carrier_j$ equals 1 if the observation corresponds to carrier j; $After_{jt}$ equals 1 if the observation corresponds to the time window *after* carrier j implements the incentive; $Carrier_j \cdot After_{jt}$ is the interaction term for a carrier following its implementation of an on-time performance incentive. The error term is ζ_{jrt}. The reported coefficient is the difference in difference α_{DID}. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors in parentheses are clustered by route carrier. Statistical significance at 1, 5, and 10 % level indicated by ***, **, and * respectively.

Table 7: **Diff-in-Diff: Selected Market Specifications**

Delays and Pad Times in Minutes

Panel A: Treated carrier faces competitors

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrival Delays	-1.28*	-3.70***	-2.12***	-1.09*	0.32	0.20	-0.46	-0.85	-4.97***
	(0.70)	(0.35)	(0.77)	(0.64)	(0.56)	(0.76)	(2.83)	(1.76)	(0.60)
R-squared	0.82	0.86	0.92	0.90	0.89	0.88	0.88	0.65	0.85
Departure Delays	-1.70***	-2.74***	-0.22	-0.13	0.75	-0.84	-0.05	2.45	-4.47***
	(0.57)	(0.29)	(0.64)	(0.63)	(0.47)	(0.73)	(3.19)	(1.64)	(0.56)
R-squared	0.79	0.84	0.92	0.86	0.86	0.86	0.83	0.62	0.86
Pad	-0.31	1.21***	2.48***	2.83***	1.83***	-1.22	1.78	6.13***	1.13**
	(0.49)	(0.33)	(0.65)	(0.97)	(0.65)	(0.80)	(3.63)	(1.19)	(0.57)
R-squared	0.14	0.19	0.65	0.42	0.52	0.40	0.42	0.59	0.61
Observations	1,632	3,957	560	526	942	925	56	595	1,646
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

Panel B: Each carrier is a monopolist in the route it serves

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrival Delays	-1.37***	-2.68***	-2.03**	-0.38	-1.32***	0.88**	-0.32	-1.90***	-4.41***
	(0.45)	(0.26)	(1.01)	(0.32)	(0.27)	(0.40)	(0.83)	(0.66)	(0.39)
R-squared	0.84	0.87	0.85	0.84	0.85	0.75	0.87	0.69	0.86
Departure Delays	-2.00***	-2.55***	-0.56	-0.08	-0.95***	0.30	0.61	-0.66	-4.10***
	(0.42)	(0.23)	(0.74)	(0.29)	(0.24)	(0.33)	(0.80)	(0.63)	(0.36)
R-squared	0.82	0.86	0.82	0.81	0.82	0.76	0.86	0.65	0.85
Pad	-0.99***	0.06	1.68***	1.41***	1.79***	-0.55*	1.21	2.04***	0.30
	(0.31)	(0.21)	(0.65)	(0.30)	(0.25)	(0.29)	(1.36)	(0.51)	(0.35)
R-squared	0.22	0.21	0.52	0.55	0.53	0.43	0.60	0.49	0.67
Observations	11,752	12,893	10,278	8,344	9,380	10,254	7,447	9,945	6,398
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

Panel C: Effect on Competitors

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrival Delays	-0.74	-1.41***	-0.62*	0.45	-0.22	0.88*	0.26	-0.66	0.84**
	(0.50)	(0.25)	(0.37)	(0.43)	(0.36)	(0.46)	(1.61)	(0.44)	(0.39)
R-squared	0.84	0.86	0.85	0.84	0.84	0.84	0.87	0.87	0.87
Departure Delays	-0.57	-1.21***	-0.45	0.01	0.05	0.16	-0.17	-0.82*	0.44
	(0.42)	(0.22)	(0.40)	(0.47)	(0.32)	(0.42)	(1.22)	(0.43)	(0.35)
R-squared	0.81	0.84	0.83	0.81	0.81	0.81	0.87	0.86	0.87
Pad	-0.00	0.06	0.48	-1.42**	0.19	-1.20**	-0.31	-0.06	-1.35***
	(0.32)	(0.23)	(0.49)	(0.61)	(0.42)	(0.47)	(1.12)	(0.40)	(0.48)
R-squared	0.16	0.20	0.47	0.52	0.53	0.52	0.58	0.66	0.72
Observations	21,630	17,247	16,513	9,647	10,008	9,982	9,464	6,334	6,370
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

We estimate the difference in difference specification defined in Table 6 for different sub-samples. The reported coefficient is α_{DID}. Panel A: Sample is restricted to routes served by the treated carrier before and after its incentive, and at least one other carrier serves that same route both before and after the treated carrier introduces the incentive (the treated carrier competes with the same rival(s) before and after implementing the incentive). Panel B: Sample restricted to routes served by only one carrier before and after a carrier introduces the incentive (routes where treated and control carriers are monopolists). In Panel C, the *treated sample* consists of routes served by a carrier (B) with no incentive on which it competes with another carrier (A) with an incentive, and the *control sample* consists of routes served by carrier (B) on which it does not compete with a carrier (A) with an incentive, before and after carrier (A) introduces an incentive. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors in parentheses are clustered by route carrier. Statistical significance at the 1, 5, and 10 % level indicated by ***, **, and * respectively.

Table 8: **Difference-in-Difference: Selected Market Specifications**

Percent of Flights Delayed at Least 15 Minutes.

Panel A: Treated Carrier Faces a Competitor

Treated Carrier	CO	DL1	AS	UA	AA	DL2	B6	DL3	US
Arrival Delays	-0.04***	-0.07***	-0.06***	-0.04***	-0.01	-0.00	0.04	-0.00	-0.10***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)	(0.02)	(0.01)
R-squared	0.84	0.86	0.92	0.91	0.89	0.88	0.80	0.87	0.89
Departure Delays	-0.04***	-0.04***	-0.02*	-0.03**	0.01	-0.03**	-0.01	0.08***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.06)	(0.02)	(0.01)
R-squared	0.79	0.83	0.91	0.86	0.87	0.85	0.84	0.87	0.88
Observations	1,632	3,957	560	526	942	925	56	595	1,646
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

Panel B: Each carrier is a monopolist in the routes it serves

Arrival Delays	-0.02***	-0.06***	-0.05**	-0.02***	-0.04***	0.01*	-0.03**	-0.03***	-0.09***
	(0.01)	(0.01)	(0.02)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
R-squared	0.84	0.88	0.87	0.86	0.87	0.87	0.88	0.85	0.89
Departure Delays	-0.04***	-0.05***	-0.02	-0.02***	-0.03***	-0.00	0.01	-0.02**	-0.08***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
R-squared	0.80	0.86	0.85	0.84	0.84	0.84	0.87	0.84	0.88
Observations	11,752	12,893	10,278	8,344	9,380	10,254	7,447	9,945	6,398
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

Panel C: Effect on Competitors

Arrival Delays	-0.01	-0.03***	-0.02**	0.02***	-0.01	0.02**	0.02	-0.01	0.02***
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)
R-squared	0.83	0.87	0.87	0.86	0.86	0.86	0.88	0.89	0.90
Departure Delays	-0.01	-0.02***	-0.01*	0.01	0.00	0.01	-0.00	-0.01	0.01
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
R-squared	0.79	0.85	0.85	0.83	0.83	0.83	0.87	0.88	0.90
Observations	21,630	17,247	16,513	9,647	10,008	9,982	9,464	6,334	6,370
Origin and Dest Timetrends	X	X	X	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X	X	X	X

We estimate a difference in difference model of the type defined in Table 6 for the different sub-samples defined in Table 7. The dependent variable is the fraction of flights delayed by 15 minutes or more. The reported coefficient is α_{DID}. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route-carrier are in parentheses. Statistical significance at the 1, 5, and 10 % level indicated by ***, **, and * respectively.

Table 9: **Effect of Incentives on Delays and Flight Times**

Panel Fixed Effects Regressions

Panel A: Each treated carrier serves the route before and after the incentive is introduced

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.43***	-0.37***	-0.01***	-0.01***	-0.02
	(0.13)	(0.12)	(0.00)	(0.00)	(0.08)
Rivals	-0.34***	-0.26**	-0.00	-0.00	0.20**
	(0.12)	(0.12)	(0.00)	(0.00)	(0.09)
Joint	-0.56***	-0.42**	-0.02***	-0.01***	0.34***
	(0.19)	(0.19)	(0.00)	(0.00)	(0.13)
Observations	210,324	210,324	210,324	210,324	210,324
Nu. Route-carrier	5,123	5,123	5,123	5,123	5,123
R^2-Adjusted	0.48	0.47	0.50	0.52	0.44
F	3.177	4.001	8.278	6.306	2.696
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Further restrict to problem routes

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	Arrival	Departure	Arrival	Departure	Pad
Own	-2.38***	-1.99***	-0.04***	-0.03***	0.82**
	(0.65)	(0.61)	(0.01)	(0.01)	(0.38)
Rivals	-0.55	-0.47	0.01	0.01	-0.04
	(0.73)	(0.68)	(0.01)	(0.01)	(0.44)
Joint	-1.46*	-1.33*	-0.02*	-0.01	0.37
	(0.84)	(0.79)	(0.01)	(0.01)	(0.56)
Observations	16,485	16,485	16,485	16,485	16,485
Nu. Route-carrier	561	561	561	561	561
R^2 -Adjusted	0.48	0.48	0.49	0.56	0.39
F	2.049	2.075	4.308	2.557	1.785
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

We estimate a panel fixed effect regression of the type:

$$P_{jrt} = \alpha_{Own} \cdot Own_{jrt} + \alpha_{Rival} \cdot Rival_{jrt} + \alpha_{Joint} \cdot Joint_{jrt} + \delta_{jy} \cdot Log(Total\_Comp)_{jy} + \sum_k \gamma_{krt} \cdot kInRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt}$$

P_{jrt} is one of five performance measures for carrier j, in route r, at time t: arrival and departure delays in minutes, arrival and departure delays as the fraction of flights delayed 15 minutes or more, and padding times. We include three categorical variables by route: *Own* for carriers that have an incentive, *Rivals*= 1 for carriers that face a rival that has an incentive, and *Joint*= 1 for carriers that have an incentive and at least one rival has an incentive. The omitted category is *None*. To control for the presence of a competitor in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. u_{rj} are route-carrier fixed effects, u_t are year-quarter fixed effects, u_{oy} are $origin \cdot year$ fixed effects, and u_{dy} are $destination \cdot year$ fixed effects. Panel B restricts the sample to "problem routes". *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route carrier are in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 10: **Variation in the Size of the Bonus Tied to On-time Performance**

Panel A: Variation in Percent of Bonus tied to On-Time Performance

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	Arrival	Departure	Arrival	Departure	Pad Time
Bonus Pct	-13.18***	-7.30***	-0.19***	-0.09***	17.07***
	(1.55)	(1.44)	(0.02)	(0.02)	(1.24)
Observations	37,138	37,138	37,138	37,138	37,138
Nu. Route-carrier	2,512	2,512	2,512	2,512	2,512
R^2 -Adjusted	0.490	0.466	0.544	0.567	0.402
F	8.208	5.270	8.105	5.456	20.43
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Distribution of Percent of Bonus tied to On-Time Performance

	Minutes Arrival Delays≥ 0		Minutes Departure Delays≥ 0		Pad:Scheduled-Actual Time	
	Bonus Pct > 1%	Bonus Pct ≤ 1%	Bonus Pct > 1%	Bonus Pct ≤ 1%	Bonus Pct > 1%	Bonus Pct ≤ 1%
Bonus Pct	-18.50***		-9.99***		24.21***	
	(2.17)		(2.03)		(1.84)	
Own		0.19		0.26		0.012
		(0.15)		(0.16)		(0.11)
Observations	20,811	145,637	20,811	145,637	20,811	145,637
R^2 Adjusted	0.455	0.522	0.435	0.508	0.406	0.4562
F	8.272	2.454	4.157	2.282	18.77	2.59
In Mkt Before & After Incentive	X	X	X	X	X	X
Origin and Dest Year FE	X	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X

We estimate panel fixed effects regressions of the type in Table 9 but now we exploit the fact that for a subset of our sample, we have data on the fraction of the overall bonus that depends on the on-time incentive, and thus replace the indicator variables for Own, $Rivals$,and $Joint$ with a continuous measure of the own bonus. Panel A uses the whole distribution of the percent of the overall bonus that depends on on-time performance, which ranges from as little as 0.001 to up to 0.335 of the overall bonus. Panel B examines the effect of bonus percentages above and below one percent (columns 1 and 2 respectively). *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route carrier are in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 11: **Placebo Test**

Panel A: Placebo Test

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	(1)	(2)	(3)	(4)	(5)
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.025	-0.015	-0.000	-0.000	0.024
	(0.025)	(0.018)	(0.000)	(0.000)	(0.019)
Nu. iterations	1,000	1,000	1,000	1,000	1,000
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Own Effects

	Minutes Delayed≥ 0		**Pct. Flights Delayed 15'**		**Sched-Actual**
	(1)	(2)	(3)	(4)	(5)
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.28**	-0.25**	-0.01***	-0.01***	-0.07
	(0.11)	(0.11)	(0.00)	(0.00)	(0.07)
Observations	210,324	210,324	210,324	210,324	210,324
Nu- Route-Carrier	5,123	5,123	5,123	5,123	5,123
R^2-Adjusted	0.457	0.440	0.477	0.499	0.414
F	3.083	4.268	9.099	6.954	1.869
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

We run a placebo test randomizing the quarter during which each carrier in our sample introduces an incentive and compare the results with those obtained using the actual data. In Panel A: For each of the carriers in our sample, we randomly assign an on-time arrival incentive in any of the 72 quarters that comprise our sample period (Northwest and America West included). The carrier maintains the placebo on-time incentive for all quarters after the randomly assigned quarter of incentive initiation. We perform a Monte Carlo simulation with 1000 repetitions for the specification from Table 9, omitting Rival and Joint. Panel B reports the analogous results from our actual data for ease of comparison. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route carrier are in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 12: **Ashenfelter's Dip Strategy to Address Mean Reversion**

Panel A: All routes where a carrier is in before and after it introduced the incentive

	Minutes Delayed ≥ 0		Pct. Flights Delayed 15'		Sched-Actual
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.31**	-0.26**	-0.01***	-0.01***	-0.06
	(0.12)	(0.12)	(0.00)	(0.00)	(0.09)
Rivals	-0.26**	-0.17	-0.00	-0.00	0.20**
	(0.12)	(0.12)	(0.00)	(0.00)	(0.10)
Joint	-0.49***	-0.34*	-0.02***	-0.01***	0.33**
	(0.18)	(0.18)	(0.00)	(0.00)	(0.13)
Observations	194,351	194,351	194,351	194,351	194,351
Nu. Route-carrier	5,058	5,058	5,058	5,058	5,058
R^2 -Adjusted	0.479	0.462	0.480	0.502	0.432
F	3.295	4.025	7.668	5.889	3.357
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Further Restrict sample to problem routes

	Minutes Delayed ≥ 0		Pct. Flights Delayed 15'		Sched-Actual
	Arrival	Departure	Arrival	Departure	Pad
Own	-2.12***	-2.01***	-0.05***	-0.05***	0.61
	(0.75)	(0.75)	(0.01)	(0.01)	(0.49)
Rivals	0.28	0.11	0.01	0.01	-0.37
	(0.75)	(0.70)	(0.01)	(0.01)	(0.47)
Joint	-1.69*	-1.67*	-0.04***	-0.04***	0.47
	(0.89)	(0.89)	(0.01)	(0.01)	(0.63)
Observations	14,451	14,451	14,451	14,451	14,451
Nu. Route-carrier	534	534	534	534	534
R^2 -Adjusted	0.501	0.502	0.489	0.548	0.395
F	1.728	2.198	3.769	3.040	1.364
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel C: Bonus Percent

	Minutes Delayed ≥ 0		Pct. Flights Delayed 15'		Sched-Actual
	Arrival	Departure	Arrival	Departure	Pad
Bonus Pct	-11.06***	-6.86***	-0.15***	-0.08***	12.30***
	(1.69)	(1.56)	(0.02)	(0.02)	(1.29)
Observations	34,048	34,048	34,048	34,048	34,048
Nu. Route-carrier	2,453	2,453	2,453	2,453	2,453
R^2 -Adjusted	0.480	0.454	0.540	0.560	0.407
F	4.492	3.205	4.961	3.614	9.336
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

We estimate fixed effects regressions of the type in Table 9 but now apply the Ashenfelter dip strategy: For each carrier that introduces an incentive we exclude the quarter when it implemented the incentive, the *two* preceding quarters and the quarter immediately after. All variables are as defined before. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route-carrier in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 13: **Effect of Southwest Airlines**

Panel A: Southwest Does Not Serve Any Route

	Minutes Delayed ≥ 0		Pct. Flights Delayed 15'		Sched-Actual
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.66***	-0.62***	-0.02***	-0.01***	-0.02
	(0.16)	(0.15)	(0.00)	(0.00)	(0.10)
Rivals	-0.36**	-0.20	-0.00	0.00	0.25**
	(0.16)	(0.16)	(0.00)	(0.00)	(0.13)
Joint	-0.48*	-0.36	-0.02***	-0.01***	0.38**
	(0.24)	(0.23)	(0.00)	(0.00)	(0.17)
Observations	153,670	153,670	153,670	153,670	153,670
Nu. Route-carrier	3,868	3,868	3,868	3,868	3,868
R^2 -Adjusted	0.456	0.455	0.494	0.532	0.305
F	4.556	5.655	9.370	9.042	3.706
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

Panel B: Routes where Southwest Entered During Sample

	Minutes Delayed ≥ 0		Pct. Flights Delayed 15'		Sched-Actual
	Arrival	Departure	Arrival	Departure	Pad
Own	-0.40***	-0.31**	-0.01***	-0.01***	0.02
	(0.13)	(0.13)	(0.00)	(0.00)	(0.08)
Rivals	-0.36***	-0.25*	-0.00	-0.00	0.24**
	(0.13)	(0.13)	(0.00)	(0.00)	(0.10)
Joint	-0.50**	-0.31	-0.01***	-0.01***	0.43***
	(0.20)	(0.19)	(0.00)	(0.00)	(0.13)
SW Entry	-0.64***	-0.47***	-0.01***	-0.01***	1.41***
	(0.13)	(0.13)	(0.00)	(0.00)	(0.15)
Observations	194,417	194,417	194,417	194,417	194,417
Nu. Route-carrier	4,807	4,807	4,807	4,807	4,807
R^2 -Adjusted	0.455	0.440	0.478	0.500	0.412
F	5.099	5.087	9.248	6.374	8.701
Origin and Dest Year FE	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X

We estimate fixed effects regressions of the type in Table 9. Panel A restricts the sample to routes not served by Southwest. Panel B instead restricts the sample to routes Southwest entered during the sample period, showing the effect of Southwest's entry on a route. Variables are defined as above. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route-carrier in parentheses. Statistical significance at 1, 5, and 10 % level: ***, **, and * respectively.

Table 14: **Effect of Incentives on Capacity and Prices**

	Origin-Year and Destination-Year Fixed Effects					
	Panel A: All routes where a carrier is active before and after introducing the incentive					
	Passengers and Seats			**Flights**		**Fares**
	Seats per Scheduled	Passenger per Performed	Load Factor	Scheduled Flights	Performed Flights	Ticket Fare
	(1)	(2)	(3)	(4)	(5)	(6)
Own	0.01***	0.01***	-0.00	-0.02	-0.02	-0.01**
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Rivals	-0.00	-0.02***	-0.02***	-0.03***	-0.03**	0.01
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Joint	-0.00	-0.01*	-0.01*	-0.10***	-0.10***	0.01
	(0.00)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	186,701	186,701	186,701	186,701	186,701	118,577
Nu. Route-carrier	4,473	4,473	4,473	4,473	4,473	4,046
R^2 -Adjusted	0.709	0.824	0.749	0.830	0.827	0.889
Origin and Dest Year FE	X	X	X	X	X	X
Competitor in Route FE	X	X	X	X	X	X
Route-Carrier FE	X	X	X	X	X	X
Year-Quarter FE	X	X	X	X	X	X
Cluster Route-Carrier	X	X	X	X	X	X

We estimate a panel fixed effect regression of the type:

$$\begin{aligned} P_{jrt} \quad = \quad & \alpha_{Own} \cdot Own_{jrt} + \alpha_{Rivals} \cdot Rival_{jrt} + \alpha_{Joint} \cdot Joint_{jrt} + \delta_{jy} \cdot Log(Total\_Comp)_{jy} \\ & + \sum_k \gamma_{krt} \cdot kInRoute_{rt} + u_{rj} + u_t + u_{oy} + u_{dy} + \zeta_{jrt} \end{aligned}$$

In Columns (1)-(5) the dependent variable, P_{jrt}, is one of five different measures of capacity: Log available seats per scheduled flight, log enplaned passengers per performed flight, load factor, log of scheduled, and performed, flights. The unit of observation is a route-carrier-year-quarter. In column (6) the dependent variable, P_{jrt}, is the log of the average ticket prices for carrier *j*, in route *r*, at time *t*. In column (6), we include standard controls for whether low cost carriers (LCC) serve the route. See Dennis, Gerardi and Schenone (2018) for a detailed description of ticket fare data, and relevant control variables. To control for the presence of a competitor in a route we include $kInRoute_{rt}$, which equals 1 if carrier k serves route r at time t. u_{rj} are route-carrier fixed effects, u_t are year-quarter fixed effects, u_{oy} are $origin \cdot year$ fixed effects, and u_{dy} are $destination \cdot year$ fixed effects. *Log(Total Compensation)* is included in all specifications but not reported. Standard errors clustered by route-carrier are in parentheses. Statistical significance at the 1, 5, and 10 % level is indicated by ***, **, and * respectively.

Figure 1: Dynamic Effect of Introducing Ontime Incentive