NET Institute*

www.NETinst.org

Working Paper #13-07

September 2013

**The Effect of the Internet on Performance and Quality:
Evidence from the Airline Industry**

Itai Ater
Tel Aviv University

Eugene Orlov
Compass Lexecon

* The Networks, Electronic Commerce, and Telecommunications (“NET”) Institute, <http://www.NETinst.org>, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, “virtual networks” comprised of computers that share the same technical standard or operating system, and on network issues in general.

The Effect of the Internet on Performance and Quality: Evidence from the Airline Industry*

Itai Ater (Tel Aviv University) and Eugene Orlov (Compass Lexecon)

September 15, 2013

Abstract

How did the diffusion of the Internet affect performance and product quality in the airline industry? We argue that the shift to online distribution channels has changed the way airlines compete for customers - from an environment in which airlines compete for space at the top of travel agents' computer screens by scheduling the shortest flights, to an environment where price plays the dominant role in selling tickets. Using flight-level data between 1997 and 2007 and geographical growth patterns in Internet access, we find a positive relationship between Internet access and scheduled flight times. The magnitude of the effect is larger in competitive markets without low-cost carriers and for flights with shortest scheduled times. We also find that despite longer scheduled flight times, flight delays increased as passengers gained Internet access. More generally, these findings suggest that increased Internet access may adversely affect firms' performance and firms' incentives to provide high quality products.

JEL classification: D83; L15; L93. Keywords: Internet; Search; Air Travel; Quality

I. Introduction

How do improvements in information availability affect markets? Economists have long emphasized that price information is critical for the efficient functioning of markets. The growth of electronic commerce and Internet marketplaces has been considered as contributing towards market efficiency, because it has enabled consumers to compare prices across hundreds of vendors with much less effort than would be required in the physical world (e.g. Brown and Goolsbee 2002). Indeed, the traditional economic view suggests that the Internet makes markets more competitive, thereby leading firms to reduce prices. But, besides setting lower prices, are there other ways in which firms have responded to the changes in the competitive landscape (Ellison

* We are thankful for comments and suggestions made by Jim Dana, Liran Einav, Silke Forbes, David Genesove, Vardit Landsman, Sarit Markovich, Oren Rigbi and seminar and conference participants at Hebrew University, Mannheim University, Tel Aviv University, Ben-Gurion University, 2012 Paris-Tech Economics of ICT, 2012 FTC Microeconomic and Toulouse ICT Conferences. We gratefully acknowledge financial support from the NET Institute, www.NETinst.org and from the ISRAEL SCIENCE FOUNDATION (grant no. 363/11). Remaining errors are our own. Comments are welcome at ater@post.tau.ac.il and eorlov@compasslexecon.com.

and Ellison 2009)? Could more price information also lead firms to offer lower product quality? Could more information change the market equilibrium in ways that make performance less important? Our paper aims to address these questions by highlighting subtle ways in which the diffusion of the Internet has affected performance and product quality in the airline industry. To our knowledge, this is the first empirical study that has examined the relationship between performance and product quality and Internet access or search costs in general.

The US airline industry provides a natural setting for examining the effect of the Internet. The distribution of airline tickets has undergone dramatic changes over the past 15 years, as more air passengers have gained Internet access. In 1997 only 0.5 percent of U.S. domestic airline tickets were purchased online, whereas in 2007 between 50 percent and 60 percent of tickets were purchased online (Berry and Jia 2010). Previous research has also shown that Internet usage is strongly associated with lower airfares (Orlov 2011), and that over 35 percent of traditional brick-and-mortar travel agencies exited their markets between 1997 and 2003 as more air passengers began to use the Internet to purchase airline tickets (Goldmanis et al. 2010).

We use the scheduled elapsed flight time, i.e. the amount of time between the scheduled departure and the scheduled arrival, as our main measure for performance. Figure 1 plots a time series for the mean scheduled flight times between 1990 and 2007. The time series fluctuates in the early 1990s and then shows a strong upward trend starting in the second half of the 1990s. We argue that the shift in airfare distribution channels from traditional travel agencies to online distribution channels can explain about a quarter to a half of this upward trend. In the early 1990s, when brick-and-mortar travel agencies sold over 80 percent of airline tickets (Borenstein 1992), flights typically appeared on the travel agent's screen in ascending order according to their scheduled durations (GAO 1995). Given that travel agents booked over 80 percent of flights from the first screen and that the majority of the tickets were sold from the first line of search results, airlines had strong incentives to maintain short scheduled flying times (Guerin-Calvert 1992). This incentive changed as the Internet became the most common channel for purchasing airline tickets, and flight duration was no longer the main criterion to sort flights. In other words, the Internet has changed the way airlines compete for customers, creating a shift from an environment in which airlines compete to appear at the top of travel agents' computer screens by scheduling the shortest flights, to an environment where price plays the dominant role in selling tickets, and scheduled time is not as important. Consistent with this view, a senior executive at American Airlines said: "[before the Internet] you had to be on the first screen [of the travel agent]. All the booking came off the first screen. You'd bring your airplanes in as fast as you can and you'd let them go as fast as you can. [When most tickets are sold online] it's important still to have competitive elapsed time, because there's a point at which consumers will say

no. But it isn't important to have the fastest elapsed time".¹ As flight duration times became less important in affecting revenues, airlines could schedule longer flights and reduce costly fuel consumption² and other related costs (Forbes, Lederman and Tombe 2012).

To substantiate our claim, we investigate the relationship between increases in Internet usage and scheduled elapsed flight times using several sources of flight level data over the years 1997–2007. Our analysis exploits four main sources of variation in the data: (1) differences in Internet penetration over time; (2) differences in Internet penetration across geographical locations; (3) differences in competition levels across flight segments (*flight segment* refers to a directional non-stop route between two airports); and (4) differences in flight times within flight segments. These sources of variation allow us to (a) to estimate the marginal effect of Internet penetration on flight duration, and (b) investigate whether the size of the effect varies across flight segments with different levels of competition. We expect that the effect of the Internet on scheduled flight duration will be greater in competitive markets because airlines were especially likely to schedule short flights in these markets when traditional travel agents were the dominant distribution channel. For instance, consider two Boeing 727 Delta flights departing on the same day from Atlanta International Airport. One flight flies to Boston and faces market competition while the other flight is to Albuquerque and is a monopoly. We expect that a similar increase in Internet access for passengers on these flights will have a greater impact on flight duration in the competitive market to Boston than the corresponding effect on flight duration for the flight to Albuquerque.

Our identification strategy also exploits the fact that connecting passengers or passengers returning to their home town likely bought their ticket in the metropolitan area where they began their trip, which is not necessarily the same area in which a given flight departs from. Thus, we can compare the effect of the Internet on flights departing from the same airport on the same day but also carry different fractions of passengers who purchased their tickets elsewhere, i.e., in areas with different levels of access to the Internet. For instance, assume that half of the passengers from Atlanta to Boston and from Atlanta to Albuquerque purchased their tickets in Atlanta and that the other half bought their tickets at their origin town (Boston and Albuquerque, respectively). Because the Internet grew faster in Boston than in Albuquerque, we expect that a larger fraction of passengers on the flight to Boston purchased their ticket online compared to the corresponding fraction of passengers who

¹ See at http://www.nesdb.go.th/specialWork/suvarnabhumi/ceo_talk/No%20Peaking.pdf. The Wall Street Journal also reports that scheduled time has increased dramatically over the years, sometimes by up to 39 percent for similar flights. The Middle Seat: Why a Six-Hour Flight Now Takes Seven. Scott McCartney (2010, February 4), *Wall Street Journal* (Eastern Edition), p. D.1. For instance, the scheduled duration time of United Airlines' before-noon flight from San Francisco (SFO) to New York City (JFK) increased from 323 minutes in July 1997 to 347 minutes in July 2007.

² Flying at higher altitudes is more fuel-efficient but requires flights to fly longer to reach the higher altitude. See, <http://online.wsj.com/article/SB10001424052748704901104575423261677748380.html>

bought tickets online on the flight to Albuquerque. Consequently, we expect that the effect of the Internet on the duration of the flight from Atlanta to Boston will be greater than the corresponding effect on flight duration of the flight from Atlanta to Albuquerque. By also incorporating various measures of congestion at the flight and the airport levels and a rich set of fixed effects, we believe that we are able to identify the impact of the Internet on scheduled elapsed flight times.

In our empirical analysis, we find that Internet usage is associated with a significant increase in scheduled elapsed flight times. This effect is particularly large in competitive markets and can explain about one-half of the total increase in scheduled elapsed flight times. To further support our argument and to rule out alternative explanations for the effect of the Internet on flight times, we also investigate two other settings in which the Internet likely affected airlines' incentives to use flight duration as a competitive tool. First, we focus on the flights with the shortest scheduled times in their segments. We expect that the effect of the Internet on elapsed scheduled time will be larger for flights that previously appeared at the top of traditional travel agents' screens. Second, we focus on markets in which low-cost-carriers offer service and exploit the fact that low-cost carriers did not primarily rely on the traditional computer reservation systems to sell airfares. Accordingly, in markets with low-cost carriers we expect scheduled elapsed flight time to be less important in the first place and the effect of the Internet on scheduled elapsed flight time to be weaker. In both settings, our regression results confirm our expectations.

We also use the same empirical strategy to examine the effect of the Internet on actual flight times and flight delays, a common measure for quality in the airline industry (e.g. Forbes 2008, Mazzeo 2003, Prince and Simon 2009).³ Our findings indicate a strong positive effect of the Internet on actual flight times. We also find that flight delays (the difference between actual and scheduled time) increased as more passengers likely purchased their tickets online. Taking into account the total effect of the Internet on flight times and delays, our estimates suggest that in competitive markets the increase in Internet access contributed about 5 to 8.2 additional minutes in arrival delays and actual flight times, respectively. We discuss possible interpretations for these findings in the Results Section.

Our paper contributes to the voluminous literature on the effect of information and communication technologies on productivity and performance (e.g. Bloom et al. 2012). Naturally, the existing literature predominantly focuses on the large benefits associated with the adoption of information and communication technologies. Our findings, however, highlight a potentially negative

³ The U.S. Department of Transportation publishes monthly statistics on airlines' on-time performance. A flight is "on-time" if it arrives no more than 15 minutes after its scheduled arrival time. In addition, Bowen and Headley (2011) devise a quality metric composed mostly of on-time performance statistics and factors that are exacerbated by delayed flights (e.g., customer complaints, lost baggage).

effect of the diffusion of online distribution channels on firms' incentives to invest in product quality and in improving their performance. Furthermore, although our findings come from the airline industry, we believe that exploring how firms respond to consumers' tendency to purchase more visible options is relevant in many other settings of online search. Indeed, recent theoretical (Arbatskaya 2007; Salant 2011) and empirical (Oestreicher-Singer and Sundararajan 2011) papers have shown that the prominence or visibility of a firm product affect firms' performance. Our findings also contribute to a growing literature on the effect of competition on performance and product quality (e.g. Hubbard 2003; Matsa 2011; Syverson 2011). Unlike most of the existing evidence, our results suggest that effect of Internet access on measures of performance and product quality are negative, particularly in more competitive environments.

The paper proceeds as follows. In Section II we discuss the data sources. In Sections III and IV we describe our estimation and identification strategy and present the results. In Section V we discuss alternative explanations for our findings. We discuss and conclude in Section VI.

II. Data

Our data come from several sources. Data on Internet usage over time come from the Computer Use and Ownership Supplement to the Consumer Population Survey (CPS), which asks respondents about Internet usage in 1997, 1998, 2000, 2001, 2003 and 2007. We use the CPS to measure Internet penetration for each of the 243 major metropolitan areas in our data. In particular, we use sample weights provided by the CPS to compute the percentage of respondents with Internet access. Table 1 provides descriptive statistics for this variable.

Our primary source of flight-level data is the on-time performance database from the U.S. Bureau of Transportation Statistics. This database reports scheduled and actual flight information for all flights on each airline with at least 1 percent of US domestic passenger revenues. In particular, the data include measures of scheduled and actual departure and arrival times as well as identifying information for each aircraft and flight. Table 2 reports the average scheduled flight times and delays for the 9 carriers that reported this information throughout the period we are investigating.

Another data source we use is the Origin and Destination Survey (DB1B) market database, which contains a 10 percent random sample of all domestic passenger tickets purchased in each quarter for each year of our sample. Importantly, the DB1B database includes the outbound and return portions of round-trip tickets, so we can identify each passenger's home airport (and therefore the associated metropolitan area in which he or she is likely to have purchased the ticket).⁴

⁴ Before 2000, Southwest Airlines reported all of its round-trip ticket sales as two one-way tickets, so we cannot identify the home airport for these Southwest passengers, and thus we exclude Southwest Airlines from our analysis prior to 2000.

The distinction is important because our hypothesis is that the share of passengers who buy their tickets online (or have the option of doing so) affects airlines' scheduled elapsed flight times. Therefore, we calculate the following weighted measure of Internet penetration among the passengers on airline i serving flight segment j in quarter t : (1) $IP_{ijt} = \sum_k \omega_{ijkt} I_k^m IP_{mt}$, where ω_{ijkt} is the share of all passengers on the flight who began their travel itinerary in airport k , IP_{mt} is the Internet penetration in metropolitan area m , and I_k^m is an indicator that is equal to 1 if airport k is located in metropolitan area m and 0 otherwise. For example, consider a flight from Atlanta to Boston. Assume that 45 percent of the passengers are flying round trip from Atlanta to Boston (so they are on the outbound portion of their round-trip itinerary), another 35 percent of the passengers are flying round trip from Boston to Atlanta (so they are on the return portion of their round-trip itinerary) and another 20 percent of the passengers are flying round trip from Tampa to Boston with a stop each way at Atlanta (so they are on the second segment of the outbound portion of their round-trip itinerary). The weighted average Internet penetration for passengers on this particular airline's flight from Atlanta to Boston is equal to $0.45 IP_{Atlanta} + 0.35 IP_{Boston} + 0.2 IP_{Tampa}$, where IP_i denotes the Internet penetration in the metropolitan area in which the relevant airport is located.⁵ Because different flights that originate from the same airport (e.g., Atlanta) carry passengers that began their journey in different locations we expect that the Internet variable will be different. We also use the DB1B to calculate the mean airfare which in some specifications we include as a control variable.⁶

We supplement these data with additional sources. First, we use the Official Airline Guide (OAG), which publishes a complete set of scheduled flights for *all* airlines between all U.S. airports. We use the OAG data to construct the measures of competition for each flight segment in addition to several measures of aircraft operations in each airport and at different times throughout the day. In Section IV we provide further details on these measures which enable us to control for potential changes in the number of flights and congestion patterns over time and across airports. In addition, we use these data to define an airline as a low-cost-carrier⁷ as well as to investigate the effect of Internet access on the shortest scheduled flight times. Figure 2 uses the OAG data and focuses on the flights with the shortest scheduled times to show the large increase in the minimum scheduled elapsed flight time between the years 1997 and 2007. Figure 3 shows how the ratio

⁵ In the paper we use the terms Internet access, Internet and Internet penetration (IP) interchangeably. See also Dana and Orlov (2013) for a similar use of this measure of Internet penetration.

⁶ The DB1B dataset reports airfares at the itinerary level. We allocate this fare to each of the itinerary's segments proportional to segments' distances. We then compute the average of airfares on each segment.

⁷ The list of low-cost carriers includes Southwest Airlines, AirTran, JetBlue, Frontier, ATA Airlines, Allegiant Air, Skybus Airlines, Spirit Airlines, Sun Country Airlines, ValuJet Airlines, National Airlines, and Independence Air.

between the shortest scheduled time and the average scheduled time fell as more passengers used online channels to purchase flight tickets. The figure illustrates that the change in scheduled time was larger for flights with the fastest flying time.

Second, we use data from the Bureau of Economic Analysis to construct demographic and economic measures at the metropolitan-area level. These measures control for expected and unexpected demand changes that may be spuriously correlated with Internet penetration. Finally, we use the T100 (Form 41) database from the Bureau of Transportation Statistics to derive a directional carrier–flight-segment load factor variable — i.e., how full that airline’s flights are — on a particular flight segment for each month in our sample. Table 3 lists descriptive statistics for the variables that we construct from these datasets and which are later used in the estimation.

After matching these datasets, we limit our sample to traffic during one Thursday in each quarter in the years CPS data are available. We do not use data for the years for which CPS data are not available. In what follows, we present estimation results for nine carriers for which data on on-time performance were available for all years between 1997 and 2007. We also limit our analysis to flight segments between the 100 airports with the highest numbers of passengers. This leaves us with 286,004 observations. The results are qualitatively similar if we do not constrain the data.

III. Identification and Empirical Specification

The explosion of e-commerce in recent years illustrates how the Internet is shaping the way consumers and businesses interact, affecting industry structure, and accelerating diffusion of information across firms and consumers. Nevertheless, identifying causal effects of the Internet on firms’ behavior has proved to be difficult. A common research design uses cross-sectional data with the key right-hand-side variable being a measure of Internet usage or Internet access. Such analyses suffer from potentially severe omitted-variable bias, because the adoption of the Internet is likely to be correlated with unobserved characteristics of Internet adopters. Using panel-data design in this study we can potentially address this concern because we control for both cross-sectional differences correlated with Internet access as well as common changes over time. Identifying the marginal effect of the Internet using a panel-data design may still be difficult if the observed changes in Internet penetration are small. We believe that the airline industry is well-suited to address this concern. Our study focuses on a period in which Internet access grew rapidly, from an average metropolitan-area penetration level of 19.4 percent in 1997 to an average penetration level of 76.3 percent in 2007. This change is also reflected in the share of airline tickets sold online, rising from 0.5 percent in 1997 to 50–60 percent in 2007 (SITA 2008). Importantly, substantial variation in changes in Internet penetration across U.S. metropolitan areas over time (see Table 1) provides additional identification power for our

results. Nevertheless, the panel data results may still be biased if areas that experienced faster Internet growth also experienced other changes that resulted in longer flight times. For example, an increase in local Internet usage could lead to lower prices, which may result in more flights operating at the local airport. This additional air traffic may translate into congestion and longer flight times. Thus, our Internet estimates may capture the impact of the Internet on the number of aircraft operations and on airport congestion.

There are three main reasons why we think this concern is mitigated in our study. The first main reason is that our data include direct measures of aircraft and airport operations. In the regressions below we include a wide range of flight and airport congestion variables that are supposed to control for changes in the level of operations from each airport (X^C in the regression below). In particular, we include the number of departures and arrivals at the flight's destination airport within an hour of the flight's arrival; the number of departures and arrivals at the flight's origin airport and, separately, at the flight's destination airport on that day; the number of departures and arrivals at the flight's origin airport within an hour of the flight's departure; and the Herfindahl-Hirschman index (HHI) at the flight's origin airport and at the flight's destination airport. The airport HHI measures should reflect potential situations in which airlines operating in concentrated airports gain advantages in handling flight operations and avoiding delays. Second, our empirical strategy implies that we are comparing the effect of the Internet on flights departing from the same airport but carry different fractions of passengers who bought their tickets online. Thus, while aircraft operations at the origin airport are likely to affect flights that depart from that airport similarly, our analysis focuses on the effect of the Internet, which was, at least partially, determined elsewhere. The third reason why we do not think that congestion or other unobserved factors which are correlated with Internet penetration are driving the results is our findings regarding flights with the fastest scheduled times and the effect of low-cost-carriers on flight duration. It is difficult to see why congestion should affect the fastest flights more than other flights on the same segment. Likewise, it is unclear how airport congestion is related to our finding that the effect of the Internet on markets in which low-cost carriers operate is smaller than its effect in other markets.

To control for other socio-economic changes at the metropolitan-area level we also include the following demographic variables (X^D in the regression below): metropolitan-area population, level of unemployment and average per capita income. These variables are included in order to control for both short- and long-run variations in demand growth across airline segments and thus are constructed from the metropolitan-area data in the same way that the Internet penetration variable is constructed. Importantly, we also include a set of fixed effects that are supposed to capture any unobservable changes along the carrier, regional and time dimensions. Below we further explain how the inclusion

of these fixed effects can help us identify the effect of the Internet on scheduled and actual flight times.

A. Main Specification for Scheduled Flight Time

The basic framework for the analysis is a fixed-effects regression of the form:

$$(2) Y_{fijt} = \beta_1 IP_{ijt} + \beta_2 IP_{ijt} \cdot HHI_{jt} + \beta_3 HHI_{jt} + \beta_D X_{ijt}^D + \beta_C X_{ijt}^C + \alpha_j + \theta_{act} + \delta_{it} + \mu_{ki} + \varepsilon_{fijt}$$

where an observation is a flight f of airline i on directional segment j departing from airport k on date t . Y_{fijt} is an outcome measure of interest measured in minutes. The primary independent variables are the traffic-weighted measure of metropolitan-area Internet penetration (IP_{ijt}) and the interaction between this measure and the competition level, measured by the HHI for the corresponding flight segment ($IP_{ijt} \times HHI_{jt}$). We also add the segment's HHI, measured based on the share of each carrier flights on a segment, as a separate regressor.

To control for unobservables we include the following set of fixed effects: directional segment, aircraft type, carrier/date and airport/carrier fixed effects. The directional segment fixed effects, α_j , control for time-invariant airport-pair characteristics, such as distance, the flying direction and landing and departing conditions for each segment. The aircraft-type fixed effects, θ_{act} , account for unobservable technical differences among different aircrafts that may allow aircrafts to fly at different speeds and recover lost time differently during flight. By including the carrier/date fixed effects, δ_{it} , we account for seasonal, technological and other trends that may affect airlines' tendency to change their schedules. We allow this effect to differ across airlines. Hence, for example, if American Airlines implemented a nationwide change in its schedules whereas United did not implement such a change, then these carrier/date fixed effects should control for that. In the regressions in which we use measures of actual flight times as the dependent variable, these carrier/date fixed effects may also capture unobservable events that have different effects on different airlines on a given day, such as extreme weather conditions that generate ripple effects on a carrier entire network performance. Including the carrier/date fixed effects implies that our cross-sectional variation comes from comparing two flights by the same carrier on the same day. In general, these two flights can depart from the same or different airports (e.g., one Delta flight from Salt-Lake City and another Delta flight from Chicago). Still, such a comparison across airports may not be completely accurate. There are many unobserved differences across airports that may affect airlines' incentives to adjust their scheduling decisions. If these factors, such as the number of gates a carrier leases or the scale of operations in a

particular airport, are correlated with the growth of Internet usage in the area that the airport serves, then we may actually obtain biased Internet estimates. To control for such bias, in some specifications we also add airport/carrier fixed effects, μ_{ki} , which should capture unobserved differences across carriers that operate in the same airports. Hence, in these latter specifications we actually test whether differences in the fraction of passengers who likely purchased their flights online affected scheduled elapsed flight times on the same carrier's flights on the same day, on the same type of aircraft, and originating from the same airport (e.g., comparing a Boeing 727 Delta flight from Atlanta to Boston with a Boeing 727 Delta flight on the same day from Atlanta to Albuquerque).

Two key features of the airline industry allow us to strengthen this identification strategy even further. The first is that airlines face different levels of competitive pressure across different segments originating from the same airport. This implies that the identification of the effect of the Internet on an airline's flights from the same airport and the same day comes at least partially from comparing flights by the same airline that depart from the same airport on the same day but in different competitive environments. As mentioned above, we expect that the effect of the Internet on flight duration will be greater in markets in which airlines were more likely to use duration times as means to sell tickets.

The second feature that can help identify the differential effect of the Internet on an airline's flights (same airport, same day, same aircraft) is that these flights likely carry different fractions of passengers who purchased their tickets online. This occurs because of the hub and spoke network structure of the airline industry. The network structure implies that many of the passengers who, for example, depart on a Delta flight from Atlanta actually began their journeys in different airports in different metropolitan areas. The share of connecting passengers is substantial. Ater (2012) reports that more than 50 percent of the passengers at the 17 largest U.S. hub airports are using these airports to connect to other flights. As we discussed above, our weighted Internet variable reflects this feature of the airline industry and our assumption that passengers who purchase their tickets online do so in the areas where they begin their trips. Consequently, our comparison across an airline's flights from the same airport on the same day does not necessarily imply that we are comparing flights that carry passengers living in the same metropolitan area and who have similar levels of access to the Internet.

The above discussion also applies to the actual flight time measure as well as the measure of flight delays (see Section III.A below).

B. Presence of Low-Cost Carriers and Business vs. Non-Business Flights

Our main argument is that the Internet changed the way airlines compete, making the flight duration characteristic less important than it were when brick-and mortar travel agents booked most flight tickets. Another way to examine this argument is to investigate the relationship between Internet access and scheduled elapsed flight time under alternative competitive market conditions. To do so, we exploit two main characteristics of the airline industry: first, we investigate how scheduled elapsed flight times differ between segments in which low-cost carriers operate and those in which such carriers do not operate. Our conjecture is that legacy airlines are more likely to compete on scheduled elapsed time in segments in which low-cost carriers *do not* offer service. This is both because the presence of low-cost carriers already alters the nature of competition towards price competition (e.g. Goolsbee and Syverson 2008) and also because low-cost-carriers, such as Southwest Airlines, the largest and most successful low-cost carrier, historically have not relied on the traditional computer reservation systems to sell tickets. Specifically, we test whether scheduled elapsed times are longer for flights in segments where a low cost-carrier operates, and whether the magnitude of the effect of the Internet on scheduled elapsed flight times is smaller in these segments. Thus, we estimate again the equation above and add a dummy variable for whether a low-cost carrier operates on a given segment and its interaction with Internet penetration.

Second, in a separate specification we distinguish between flights that are more versus less likely to carry business passengers. Our hypothesis is that airlines are particularly interested in attaining short scheduled flight times when they can increase their sales to business passengers. We consider two alternative definitions for business flights. First, we define business flights as flights that depart between 6 a.m. and 9 a.m., and, second, we consider the flights that connect the 30 largest U.S. cities as business flights (Gerardi and Shapiro 2009). We enhance the base specification presented above by adding a dummy variable for morning flights, a dummy variable for segments from or to airports in the 30 largest cities, and their interactions with the variable for Internet penetration. Adding these two dummy variables, along with the set of fixed effects introduced above, implies that we are now exploiting two additional sources of variation: the distinction between ‘morning versus non-morning’ flights operated by the same carrier on the same day and the same segment, as well as the distinction between flights heading to ‘large versus non-large’ cities. We conjecture that before the Internet became highly accessible, morning flights heading to larger cities were likely to be shorter than non-morning flights heading to non-large cities, and that this difference declined once more passengers started using the Internet to purchase airline tickets.

C. Including Price as a Control Variable

In additional specifications we also include as a separate control variable the logarithm of each airline's average airfare on a given flight segment in the corresponding quarter. We expect prices to fall as more passengers gain Internet access, and this fall may lead airlines to schedule longer flight times. By including the airfare variable we aim to isolate the two channels through which the Internet may have affected scheduled flight times: first, by affecting the way airlines compete with each other, making flight duration a less important means to sell tickets, and second by creating downward pressure on prices and, indirectly affecting airlines' incentives to arrive as fast as possible to their destination. If our argument that the Internet changed the mode of competition in the airline industry holds, then we expect that the coefficient on Internet variables will be significant and large also when we include the average price as a control variable in the regression analysis.⁸ Adding the airfare raises endogeneity issues. Passengers may care about both price and scheduled elapsed flight times, and an airline may jointly determine the price and elapsed flight time. To address this concern, we adopt an instrumental variable approach exploiting the variation in prices offered by the same airline on markets with similar characteristics. Specifically, we use the airline's average segment fare on all other segments of similar length (we divide segments into five quintiles according to length). This instrument is similar in spirit to the instruments adopted by Nevo (2001) and other studies that use data from different geographic markets.⁹

D. Using Actual Flight Times and Flight Delays as Dependent Variables

Following the increase in scheduled flight durations, airlines should consider adjusting their corresponding actual flying times. Notably, while passengers with online access can observe the flight scheduled duration at the time of purchase, the actual flight duration is observable only upon arrival at the destination airport. The difference in the transparency of these two flight attributes might drive airlines to increase actual flight times more than the corresponding increase in scheduled duration. This claim is particularly relevant as long as the difference in the two measures does not affect an airline's on-time performance statistics.¹⁰ Below, we directly examine this conjecture. First, we examine whether actual flying times indeed increased as more passengers gained Internet access.

⁸ This approach requires that the relationship between the average price and flight duration has certain properties which are not necessarily met. To further investigate the robustness of our approach, in separate analyses presented in the online appendix we verified that including non-linear price terms does not affect our results.

⁹ In the regressions in which we study the relationship between the presence of a low-cost-carrier in scheduled elapsed flight times we modified our definition of instruments: when calculating the airline's average segment fare on all other segments of similar length, we distinguish between segments where low-cost carriers are present and segments where low-cost carriers are not present.

¹⁰ The U.S. Department of Transportation publishes monthly statistics on airlines' on-time performance. A flight is "on-time" if it arrives no more than 15 minutes after its scheduled arrival time.

Second, we investigate whether this increase is greater or smaller than the increase in scheduled flight times. In other words, we investigate the relationship between Internet access and flight delays, i.e., the difference between actual flight times and scheduled flight times. In these regressions, we also add the following aircraft–flight–level variables: 1) an aircraft’s arrival delay (in minutes) on its previous flight; 2) the scheduled buffer time (in minutes) between a given flight’s scheduled departure and the scheduled time of its corresponding aircraft’s previous arrival; 3) a variable that represents the number of flights a given flight’s aircraft has performed on a given day prior to the flight under consideration; 4) an airline’s average load factor. In the instrumental variables regressions in which we use the load factor variable, we control for its potential endogeneity by using the average load factor in segments with similar lengths as instruments. Notably, omitting any or all of these additional variables does not qualitatively affect our results.

E. The Shortest Scheduled Elapsed Flight Times

To buttress the interpretation that the growth of online distribution channels has changed the competition mode in the airline industry, we also analyze the relationship between Internet access and the *shortest* scheduled elapsed flight time on a given segment. The effect of the Internet on scheduled flight times is likely to be greater for the flights with the shortest scheduled times, as these short times were presumably set in an effort to be displayed at the top of the travel agents’ screens in the pre-Internet era. To study the effect on the shortest scheduled elapsed time in a regression framework we estimate the following fixed-effects regression:

$$(3) Y_{jt} = \beta_I IP_{jt} + \beta_{I-MS} IP_{jt} \cdot HHI_{jt} + \beta_{MS} HHI_{jt} + \beta_C X_{jt}^C + \beta_D X_{jt}^D + \alpha_j + \delta_t + \varepsilon_{jt}$$

In this specification an observation is a directional segment-date pair, and Y_{jt} is an outcome measure of interest. IP_{jt} is the average of the Internet penetration variables across the flights that operate on segment j on a particular date t . Similarly, the demographic variables, X_{jt}^D , are obtained by taking the average of the demographic variables across the carriers that operate on the segment. The congestion measures X_{jt}^C correspond to the number of flights that depart from and arrive at the origin and destination airports. The fixed effects include segment fixed effects, α_j , and date fixed effects, δ_t . The dependent variable is the shortest scheduled elapsed flight time. To alleviate concerns regarding the effect of entry or exit into segments with few flights or into segments with longer or shorter flight distance, we restrict the analysis to segments in which at least five flights operated on each day included in our sample.

IV. Results

A. Scheduled Flight Time Measures

We report the results using scheduled flight time as the dependent variable in Table 4. The coefficient of the Internet variable in our base specification presented in column 1 is 4.2 and is highly significant, suggesting that airlines scheduled longer flights as Internet access among passengers increased. The HHI variable is positive and significant, consistent with our claim that airlines compete using scheduled elapsed flight times. The coefficient on the interaction term between Internet penetration level and HHI is negative and significant, suggesting that most of the effect of Internet penetration on scheduled elapsed time occurs in competitive markets. This finding lends support to our claim that airlines relied less on scheduled elapsed flight times to differentiate themselves as more passengers used the Internet to purchase airline tickets. Thus, the effect of the Internet on scheduled duration of flights from San Francisco International Airport to JFK international Airport in New York is considerably greater than its impact on flights that operate in a less competitive environment market such as the San Francisco – St. Louis market. Adding the fare variable in column 2 is associated with a small decline in the magnitude of the Internet coefficient estimates, although both variables are statistically significant. We obtain similar results when we control for the potential endogeneity of the price variable in column 3. As expected, in all specifications the coefficients that correspond to the number of flights departing from and arriving at the origin and destination airports within the same hour of the respective flight are positive and statistically significant. For example, our results suggest that the addition of 30 flights that depart within the same hour is associated with nearly 2 additional minutes in scheduled time. In columns 4–6 we present the regression results of similar specifications except that we also include the airport/carrier fixed effects. As seen in Table 4, the magnitude of the Internet coefficient decreases by about 20 percent, but it is still positive and significant. The estimate for the coefficient of the HHI slightly increases, and the magnitude of the coefficient of the Internet and HHI interaction term hardly changes. In fact, the coefficients for the Internet variable and for the Internet and HHI interaction term suggest that the effect of the Internet on flight duration is marginally negative for monopolistic markets.

B. More on Scheduled Flight Time Measures

Low-Cost Carrier Presence.—Table 5 reports estimation results for the analysis that examines the relationship between Internet access and the scheduled duration of legacy carriers' flights when a low-cost carrier either offers or does not offer service on the same segment. The coefficient of the low-cost carrier dummy variable indicates that flights by legacy carriers are more than 2 minutes

longer in segments in which low-cost carriers operate than in segments in which low-cost carriers do not operate. According to our interpretation, in these markets legacy carriers focus on the price dimension, and hence their scheduled elapsed flight times are relatively longer than in other flight segments. As passengers gained Internet access and price became more important in all markets, the distinction between the two types of markets disappeared.

“Business” and “Non-business” Flights.—Table 6 reports the estimation results that exploit our proxies for business and non-business flights. The results are consistent with our hypothesis concerning the effect of the Internet on business vs. non-business flights. In particular, the regression results suggest that morning flights used to be, *ceteris paribus*, shorter than “non-morning” flights (the difference is approximately 1 minute), and that as more passengers gained Internet access this difference disappeared. These two effects are both statistically significant and different from each other.¹¹

Changes in the Shortest Scheduled Flying Time. — Table 7 presents the regression results using the shortest scheduled elapsed flight time on a segment on a particular day as the dependent variable.¹² In column 1 we include only the Internet variable and segment fixed effects. The estimate on the Internet coefficient, 7.57, is large and highly significant. In column 2 we add the HHI variable and its interaction with the Internet variable. The coefficient on the Internet variable increases and, as before, we find that the effect of the Internet is considerably larger for more competitive segments. The magnitude of the Internet coefficient declines when we add date fixed effects and additional controls in columns 3 and 4, but it is still large and significant. The estimates of the HHI coefficient are positive and large, and their magnitudes are 2.5 times larger than those of the HHI coefficients shown in Table 4, in which we use the scheduled elapsed flight time as the dependent variable. The magnitudes of the coefficients in Table 7 suggest that when few passengers had Internet access, the shortest scheduled flight time dropped by more than 3 minutes once the market structure changed from a monopoly to an equal-size duopoly. As Internet access became more available to passengers, however, a similar change in market structure resulted in a considerably smaller effect – about 0.5 minutes – on the shortest scheduled flying time.

We also performed similar regression analyses for other measures of flight times on a given segment: 25th percentile, median and 75th percentile flight times (see the online appendix for the

¹¹ These and other results hold if instead of Thursdays in each quarter in each year in our sample we consider Mondays or Tuesdays. Furthermore, these results also hold if we add airport/carrier fixed effects to the regression analysis. Although the effect we find on the difference between business and non-business flights is not large, we believe that given the expansive set of fixed effects in our regressions, any effect is indicative of the presence of the relationship.

¹² To verify that these results are insensitive to the exclusion of flights by small carriers, we use the OAG data in this regression analysis. The regression results remain qualitatively the same if we use the On-Time Performance data instead of the OAG data.

regression results). Consistent with our interpretation, compared with the flights with the shortest flying time, the effect of the Internet on scheduled flying time was somewhat smaller for the 25th percentile and median flight times and was not significant for the 75th percentile and the maximum flight times.

C. Results Using Actual Flight Times and Delays as Dependent Variables

Actual Flight Times.—Table 8 shows the results for the analysis using the actual flight time as the dependent variable. We compute this measure as the time elapsed between the scheduled flight departure time and the actual arrival time. According to the results from column 1, an increase of 0.2 in Internet penetration, which roughly equals one standard deviation for this variable,¹³ is associated with an increase in actual flight time of between 1 and 2.9 minutes, depending on the level of competition. Accordingly, the overall effect of the Internet, which grew between 1997 and 2007 from 0.194 to 0.763 (see Table 1), amounts to between 2.6 and 8.2 minutes in actual flight times, depending on the level of competition.

Arrival Delays.— Our results suggest that both the scheduled and actual flight times increased as more passengers gained Internet access. Table 9 shows the estimation results for the relationship between the Internet and flight delays, measured as the (positive or negative) difference in minutes between a flight's actual time of arrival and its scheduled time of arrival at its destination. The estimates for the Internet coefficient are positive and significant in all specifications, suggesting that flight delays have increased as more airline passengers have gained access to the Internet. Notably, delays have increased even though airlines have been scheduling longer flight times. A possible interpretation of this result is as more passengers gained Internet access, airlines preferred to degrade the quality of the less transparent flight attribute (flight delay) rather than to excessively pad scheduled flight durations. Alternatively, airlines may have been looking for ways to reduce operational costs and allocated fewer resources to ensure that flights arrive on schedule.¹⁴ This incentive to reduce costs is arguably greater for flights that serve relatively many passengers with Internet access. The estimates for HHI coefficients are positive but are not always statistically significant at traditional significance levels. The positive competition coefficients suggest that

¹³ Most of the variation in the Internet variable comes from the longitudinal dimension. Regressing the Internet variable on segment and aircraft fixed effects, we get that the standard deviation of the residuals of the Internet penetration is 0.16. When we add date fixed effects to this regression, the standard deviation of the residuals of Internet penetration falls to 0.02.

¹⁴ We also explored whether the increased actual time is driven by longer departure delays, measured as the elapsed time between the actual gate pushback time and the scheduled departure time. This additional test is important because airlines are likely to have more control over delays occurring on the ground before departure compared to the flight airtime. The regression results, presented in the online appendix, suggest that departure delays are significantly higher for flights on segments in which larger shares of passengers have Internet access. This difference is greater in competitive markets.

although airlines maintained shorter scheduled elapsed times on more competitive segments they were nevertheless able to achieve shorter delays for these flights. The interaction term between the Internet variable and HHI is negative and highly significant. This finding suggests that delays increased more in more competitive environments because in these environments airlines responded more aggressively to the change in the mode of competition.

Our results imply that an increase of 0.2 in Internet penetration among the passengers on a given flight is associated with an increase of slightly more than 2 minutes in arrival delays. This represents a 24 percent increase in arrival delay time compared with its average level. For monopolistic segments, an increase of 0.2 in Internet penetration is associated with an increase of 0.8 minutes in arrival delays.

Our competition estimates also suggest that the net effect on competition on flight delays depends on passengers' access to the Internet. While in low Internet environment our estimates suggest that more competition is associated with fewer delays and better quality, this relationship may change as passengers gain more Internet access.

The coefficients on the numbers of operations at the origin and destination airports are positive and statistically significant, as expected. The coefficients for other variables (not reported in Table 9) are consistent with our expectations: a longer arrival delay for the previous flight on the same aircraft is associated with a longer delay for the given flight, while a longer scheduled time between a given flight's departure and the arrival of the previous flight on the same aircraft is associated with a shorter arrival delay for the given flight. Finally, a flight that takes place later in the day (as measured by the Flight in Day variable) is associated with a longer arrival delay. In column 2 we add the logarithm of a given airline's average fare on a given segment in a given quarter. The coefficient on average fare is negative and statistically significant, supporting theoretical models predicting that decreases in search costs could lead to decline in quality (e.g. Rogerson 1988; Dranove and Satterthwaite 1992; Fishman and Levy 2013). The results do not change qualitatively when we account for potential endogeneity of fares in column 3. The coefficient on the Internet penetration variable is still positive and significant. We obtain similar results in columns 4-6 when we include also the airport/carrier fixed effects.

D. Robustness

To further examine the robustness of our findings, we performed several analyses. The results of these analyses and additional results mentioned in the text are reported in the online appendix. For instance, we experimented with several non-linear prices and verified that the results are unaffected. We also found that in most specifications, the Internet variable has a positive and statistically significant coefficient even if we exclude the interaction term between the

Internet and competition variable. We also reran our regressions to include all airlines that reported on-time performance statistics to the U.S. Department of Transportation (instead of including only airlines that reported on-time performance throughout the entire period studied). The results we obtained were qualitatively the same. We also ran the regressions separately for earlier and later time periods. These regressions suggest that most of the effect of the Internet on scheduled elapsed flight times took place between the years 1998 and 2007. We also verified that our results remained unchanged when we used flights from other days of the week or for flights performed between the largest 75 airports instead of the top 100 airports. In addition, to alleviate concerns that the results are driven by changes in the weights that are part of our measure of Internet access rather than by changes in Internet penetration at the MSA level, we fixed the weights based on the total number of passengers in each route between 1997 and 2007. We then reran the regressions with the fixed weights. The results remained qualitatively and quantitatively the same. An additional concern is that the HHI variable could be endogenous. Therefore, we also ran instrumental variable regressions where we used lagged HHI values on the same segment as instruments. Again, the regression results remained qualitatively the same. We also verified that our results are not sensitive to the exclusion of the late arrival and the scheduled buffer variables. One might be concerned that the scheduled buffer variable is under the control of the airline, and hence is potentially endogenous. Yet, excluding these variables does not qualitatively change our results. Finally, we verified that the significance of coefficients is not affected when we cluster the standard errors at alternative levels, such as at the origin/date level.

V. Alternative Explanations

Our findings uncover strong positive correlation between Internet access and various flight time measures. We argue that this association is driven by a change in the mode of competition in the airline industry, making the flight duration be less important means to sell tickets. Throughout the paper we already mentioned why we think other alternative explanations are less probable. In this section we reiterate the main alternative explanations.

A. The Effect of the Internet on Airport Congestion

An increase in Internet usage could lead to lower prices, which may result in more flights operating at the local airport. This additional air traffic may translate into congestion and longer flight times. Thus, our Internet estimates may capture the impact of the Internet on the number of aircraft operations and on airport congestion. We believe that this mechanism is unlikely to capture the impact of the Internet for three main reasons. As we explain above (Section III) we include several variables

that represent the scale of operations at the origin and destination airports. Our regression estimates indeed indicate that these variables have a sizeable effect on flight duration. The inclusion of these variables, however, implies that our Internet variables are capturing a mechanism which is not driven by the number of airport operation. Second, our identification strategy implies that we compare the effect of the Internet on flights departing from the same airport but carry passengers who purchased their tickets elsewhere, i.e., in areas with different levels of access to the Internet. Finally, it is unclear how this alternative explanation is consistent with our findings that flight times in more competitive markets and on flights with the shortest time have increased more than on other flights.

B. The Tradeoff Between Price and Performance

The Internet enabled air passengers to sort flights by price and consequently induced airlines to set lower prices. Airlines, facing lower margins per passenger, may choose to allocate fewer resources to ensure shorter flight times. We agree that this argument has merit and our findings support this argument. Nevertheless, the fact that we obtain strong positive correlation between Internet and scheduled flight duration even when we control the price variable indicates that other channels through which the Internet affected scheduled flight duration are also important. Including various non-linear price terms (see the online appendix) does not change these results. Furthermore, it is difficult to see how this alternative explanation relates to our other findings, for instance, that elapsed scheduled times in markets in which low-cost-carriers operate were affected less than in other markets.

C. Visibility of On-Time Performance Statistics

As Internet penetration across the US has increased, statistics on flight delays and on-time performance have become more readily available to passengers through various sources, such as online media sources, online travel agencies and the U.S. Department of Transportation website. Arguably, airlines may pad their schedules in order to improve their on-time performance statistics, especially for passengers who likely purchase their tickets online. We do not think that this alternative explanation poses a real threat for our argument for few reasons. First, to improve on-time performance statistics airlines should also reduce, or at least maintain, their flights' actual flight times. However, as our estimates show, actual flight times have increased, especially for flights with higher share of passengers who purchased their tickets online. Consistent with this view, in a separate analysis reported in the online appendix, we measure delays using the definition used by the Department of Transportation to compute airlines' on-time performance statistics. In other words, we use as a dependent variable a dummy variable which equals one if the

flight is late more than 15 minutes and zero otherwise. In this analysis we do not find that the Internet is associated with fewer delays. Second, this alternative explanation cannot explain the different magnitudes of the relationship between Internet access and the shortest flight on a given market. Similarly, it is unclear how such an explanation is related to our finding on business flights. Third, on-time performance statistics are measured at the carrier level. Hence, carriers that try to improve their statistics can do so at any of their network flights, i.e., they need not focus on the flights that serve passengers with Internet access. The fact that these passengers have better access to on-time performance statistics does not mean that improvements in on-time performance need to be achieved for the flights that they take.

To sum up, we acknowledge that the Internet led to reduced usage of traditional travel agents, to better airfare information, and to better information about flight delays. While separately quantifying these three effects is not possible, our findings imply that the Internet led to increased flight times through the first two channels, i.e. the exit of traditional travel agents and increased price competition.

VI. Discussion and Concluding Remarks

Empirical research and competition policy tend to focus almost exclusively on price setting, making relatively little progress studying the effect of competition and information on non-price attributes. In this paper, we address this gap by examining the effect of the Internet on performance and product quality in the airline industry, using scheduled elapsed flight times as our main measure for flight performance. Our empirical analysis provides evidence that increased Internet usage led to longer scheduled elapsed flight times. This increase was most pronounced in flight segments in which legacy carriers competed with one another and among the flights with the shortest scheduled times. We argue that these findings can best be explained by a fundamental shift in the mode of competition in the airline industry – which rendered the use of scheduled elapsed flight times as a primary means to sell airfares. We also find the actual flight times and flight delays have increased as more passengers gained Internet access.

Following the deregulation of the U.S. airline industry in 1978, airlines began competing on price at the expense of flight quality. More than twenty years later, the US airline industry experienced another dramatic change as passengers began using the Internet to search for and easily compare prices across airlines. In the search for ways to reduce costs and address the lower margins, legacy carriers considered abandoning their traditional business model in which passengers paid more in return for better service. Our findings suggest that airlines' response involved not only price cuts

but also reductions in the quality provided to consumers and worse performance, especially in more competitive environments. Thus, our results suggest that the advent of the Internet, like the 1978 airline deregulation, led to lower prices and at the same time to worse performance and lower flight quality. The magnitudes of effects we find are in line with findings of other studies that examined the relationship between market structure and airline performance. Mayer and Sinai (2003) look at the relationship between market structure and flight times. They find that flights from and to concentrated hub airports are 4.5 to 7.2 minutes longer. Forbes and Lederman (2010) look at the relationship between vertical integration and departure delays. They find that departure delays are 2.3 minutes shorter on normal-weather days and 5.6 minutes shorter on rainy days for integrated airlines than for non-integrated airlines.

This paper makes two main contributions to the literature. First, to the best of our knowledge, it is the first paper that examines whether increased access to information has a predominantly positive or negative effect on performance and product quality. Economic theory has long emphasized the importance of information for the efficient functioning of markets. Nevertheless, previous papers have typically explored how access to information affects prices, and generally have not investigated other ways in which firms respond to these changes. Our findings – that the Internet had a negative effect on performance and product quality – provide one such explanation. More generally, our findings that quality fell, together with previous findings that prices have fallen as Internet usage has increased, could suggest that the effect of the Internet on consumers in the airline industry might be negative. Alternatively, it is quite possible that in the pre-Internet era airlines were overproviding short flight durations. As passengers gained Internet access, they could choose flights that are better aligned with the preferences. i.e., lower prices in exchange to longer flights. In other words, the Internet enabled passengers to directly buy fares, thereby mitigating potential agency problems created by the pre-Internet market institutional details and travel agents' incentives. Second, our paper contributes to the growing literature that investigates the effect of information and communication technologies on productivity and performance. Not surprisingly, the existing literature focuses on the large benefits associated with the adoption of information and communication technologies. Our findings, however, highlight a potentially negative effect of online distribution channels, which have indirectly reduced the incentives of airlines to compete on product quality. The results of this study are also related to other industries. In particular, in the context of online search, being on top of the consumers' search results is considered important for online vendors as it translates into higher revenues. Our paper offers direct evidence on firms' response to changes in the criterion used to rank products.

REFERENCES

- Arbatskaya, M. 2007. "Ordered Search." *Rand Journal of Economics* 38: 119-126.
- Ater, I. 2012. "Internalization of Congestion at U.S. Hub Airports." *Journal of Urban Economics*, 72: 196–209.
- Berry, S., and P. Jia. 2010. "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics* 2: 1–43.
- Bloom, N., R. Sadun, and J. V. Reenen. 2012. "Americans Do I.T. Better: US Multinationals and the Productivity Miracle." *American Economic Review* 102:167–201.
- Borenstein, S. 1992. "The Evolution of U.S. Airline Competition." *Journal of Economic Perspectives* 6: 45–73.
- Bowen, B. D., and D. E. Headley, "Airline Quality Rating", W. Frank Barton School of Business, Wichita, Kansas, 2011. Available at: <http://www.airlineinfo.com/public/2011aqr.pdf>
- Brown, J., and A. Goolsbee. 2002. "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *Journal of Political Economy* 110: 481–507.
- Dana, J. D., and E. Orlov. 2013. "Internet Penetration and Capacity Utilization in the US Airline Industry." Working paper.
- Dranove, D., and M. Satterthwaite. 1992. "Monopolistic Competition When Price and Quality are Imperfectly Observable" *Rand Journal of Economics* 23: 518–534.
- Ellison, G., and S. F. Ellison. 2009. "Search, Obfuscation, and Price Elasticities on the Internet." *Econometrica* 77: 427–452.
- Fishman, A., and N. Levy. 2013. "Search Costs and Risky Investment in Quality." Working paper
- Forbes, S. 2008. "The Effect of Service Quality and Expectations on Product Quality on Customer Complaints." *Journal of Industrial Economics* 56: 190–213.
- Forbes, S., and M. Lederman. 2010. "Does Vertical Integration Affect Performance? Evidence from the Airline Industry." *Rand Journal of Economics* 41: 765–790.
- Forbes, S., M. Lederman, and T. Tombe. 2012. "Do Firms Game Quality Ratings? Evidence from the Mandatory Disclosure of Airline On-Time Performance." Working paper.
- GAO. 1995. "International Aviation – Airline Alliances Produce Benefits, but Effect on Competition is Uncertain." U.S. General Accounting Office Report.
- Gerardi, K. S., and A. H. Shapiro. 2009. "Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry." *Journal of Political Economy* 117: 1–37.
- Goldmanis, M., A. Hortacsu, C. Syverson, and O. Emre. 2010. "E-Commerce and the Market Structure of Retail Industries." *The Economic Journal* 120: 651–682.

- Goolsbee, A., and C. Syverson. 2008. "How Do Incumbents Respond to the Threat of Entry? Evidence from Major Airlines." *Quarterly Journal of Economics* 123: 1611–1633.
- Guerin-Calvert, M. 1992. "Vertical Integration as a Threat to Competition: Airline Computer Reservations Systems," In *The Antitrust Revolution: The Role of Economics*, edited by J. Kwoka and L. White. New York: Oxford University Press.
- Hubbard, T. 2003. "Information, Decisions, and Productivity: On-Board Computers and Capacity Utilization in Trucking." *American Economic Review*, 93: 1328–53.
- Matsa, D. A. 2011. "Competition and Product Quality in the Supermarket Industry." *Quarterly Journal of Economics* 126: 1539–1591.
- Mayer, C., and T. Sinai. 2003. "Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Created Evil." *American Economic Review* 93: 1194–1215.
- Mazzeo, M. 2003. "Competition and Service Quality in the U.S. Airline Industry." *Review of Industrial Organization* 22: 275–296.
- Nevo, A. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69: 307–342.
- Oestreicher-Singer, G. and A. Sundararajan. 2012. "The Visible Hand? Demand Effects of Recommendation Networks in Electronic Markets." *Management Science*, forthcoming.
- Orlov, E. 2011. "How Does the Internet Influence Price Dispersion? Evidence from the Airline Industry." *Journal of Industrial Economics* 59: 21–37.
- Prince, J., and D. Simon. 2009. "Multimarket Contact and Service Quality: Evidence from On-Time Performance in the U.S. Airline Industry." *Academy of Management Journal* 52: 336–354.
- Rogerson, W. 1988. "Price Advertising and the Deterioration of Product Quality." *Review of Economic Studies* 55: 215–229.
- Salant, Y. 2011. "Procedural Analysis of Choice Rules with Applications to Bounded Rationality." *American Economic Review* 101: 724–748.
- SITA Airline IT Trends Survey. <http://www.sita.aero/node/1069> (accessed 18 February, 2013.)
- Syverson, C. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49: 326–365.

TABLE 1—INTERNET PENETRATION ACROSS METROPOLITAN STATISTICAL AREAS
(*N*= 243)

Year	Mean	Std. Dev.	Min	Max
1997	0.194	0.074	0.043	0.489
1998	0.411	0.118	0.094	0.723
2000	0.555	0.110	0.207	0.836
2001	0.662	0.102	0.225	0.911
2003	0.698	0.100	0.323	0.923
2007	0.763	0.085	0.415	0.970

Source: Computer Use and Ownership Supplement to the Consumer Population Survey.

TABLE 2—DIFFERENCES ACROSS AIRLINES

	Average Scheduled Flight Time (mins)	Average Arrival Delay (mins)	Average Departure Delay (mins)
Alaska Airlines	137.0	9.8	9.2
America West	138.3	9.8	9.4
American Airlines	162.4	6.4	8.3
Continental Airlines	159.9	11.3	10.7
Delta Airlines	133.1	9.3	9.1
Northwest Airlines	135.6	8.7	8.5
Southwest Airlines	97.4	7.1	10.6
United Airlines	150.5	11.7	12.2
US Airways	116.7	9.9	9.2

Notes: Each cell contains average values over each airline's directional segments in the sample over 24 days (one Thursday in each quarter in the years 1997, 1998, 2000, 2001, 2003, and 2007). US Airways merged with America West in 2005.

TABLE 3—DESCRIPTIVE STATISTICS (286,004 OBSERVATIONS)

Variable	Mean	Std. Dev.	Min	Max
INTERNET	0.6	0.2	0.1	0.9
HHI (Segment)	0.701	0.255	0.164	1.000
SCHEDULED FLIGHT TIME	135.1	71.2	22.0	660.0
ARRIVAL DELAY	9.1	31.3	-65.0	358.0
DEPARTURE DELAY	9.7	28.0	-50.0	360.0
TOTAL ACTUAL TIME	144.2	77.5	3.0	832.0
<i>Congestion Variables:</i>				
LOAD FACTOR	0.693	0.117	0.075	0.980
AVG. FARE	138.5	69.7	1.9	658.0
# DEPARTURES PER HOUR, ORIGIN	33.4	24.0	1.0	122.0
# DEPARTURES PER HOUR, DEST	28.7	22.8	0.0	122.0
# ARRIVALS PER HOUR, DEST	32.3	24.4	1.0	171.0
# ARRIVALS PER HOUR, ORIGIN	29.0	22.6	0.0	1.0
HHI (Airport)	0.39	0.193	0.1	1
<i>Demographic Variables:</i>				
LOG (INCOME PER CAPITA)	10.4	0.1	9.9	10.9
LOG (POPULATION)	15.1	0.6	12.7	16.7
EMPLOYMENT, %	0.60	0.03	0.38	0.73
<i>Other Variables:</i>				
LATE ARRIVAL	5.8	26.0	-65.0	797.0
SCHEDULED BUFFER	100.8	93.7	-397.0	1406.0
FLIGHT IN DAY	3.2	2.0	1.0	14.0

TABLE 4—REGRESSION RESULTS: SCHEDULED FLIGHT TIME

Dependent Variable:	SCHEDULED ELAPSED FLIGHT TIME, mins					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	OLS	OLS	IV
INTERNET	4.209*** (1.300)	3.865*** (1.281)	2.701* (1.383)	3.297*** (1.223)	3.114** (1.214)	2.687** (1.258)
HHI	1.288*** (.488)	1.775*** (.490)	3.422*** (.710)	1.546*** (.480)	1.905*** (.480)	2.746*** (.618)
INTERNET * HHI	-3.115*** (.665)	-3.296*** (.658)	-3.910*** (.721)	-3.085*** (.659)	-3.228*** (.651)	-3.562*** (.673)
# DEPARTURES ORIGIN AT FLIGHT'S HOUR	.060*** (.002)	.060*** (.002)	.060*** (.002)	.060*** (.002)	.060*** (.002)	.060*** (.002)
# DEPARTURES DESTINATION AT FLIGHT'S HOUR	.02*** (.002)	.026*** (.002)	.027*** (.002)	.025*** (.002)	.025*** (.002)	.026*** (.002)
# ARRIVALS DESTINATION AT FLIGHT'S HOUR	.029*** (.003)	.029*** (.003)	.028*** (.003)	.030*** (.003)	.030*** (.003)	.029*** (.003)
# ARRIVALS ORIGIN AT FLIGHT'S HOUR	.003 (.002)	.004* (.002)	.005** (.002)	.006*** (.002)	.007*** (.002)	.007*** (.002)
LOG (AVG. FARE)		-1.705*** (.262)	-7.476*** (1.576)		-1.358*** (.238)	-4.559*** (1.451)
Segment Fixed Effects	Y	Y	Y	Y	Y	Y
Aircraft Fixed Effects	Y	Y	Y	Y	Y	Y
Demographic Controls (X ^D)	Y	Y	Y	Y	Y	Y
Addl' Congestion Controls (X ^C)	Y	Y	Y	Y	Y	Y
Carrier/Day Fixed Effects	Y	Y	Y	Y	Y	Y
Origin/Carrier Fixed Effects	N	N	N	Y	Y	Y
Observations	286,004	286,004	286,004	286,004	286,004	286,004

Notes: Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

TABLE 5—REGRESSION RESULTS: SEGMENTS WITH AND WITHOUT LCC PRESENCE

Dependent Variable:	SCHEDULED ELAPSED FLIGHT TIME, mins		
	(1)	(2)	(3)
	OLS	OLS	IV
INTERNET	4.852 ^{***} (1.482)	4.594 ^{***} (1.466)	4.549 ^{***} (1.461)
HHI	2.835 ^{***} (.511)	2.963 ^{***} (.511)	2.967 ^{***} (.516)
INTERNET * HHI	-4.642 ^{***} (.731)	-4.630 ^{***} (.725)	-4.605 ^{***} (.721)
LCC	2.384 ^{***} (.338)	2.181 ^{***} (.343)	2.141 ^{***} (.342)
INTERNET * LCC	-3.520 ^{***} (.718)	-3.263 ^{***} (.722)	-3.226 ^{***} (.720)
INTERNET * HHI * LCC	.060 (0.997)	-.011 (0.998)	.001 (0.993)
LOG (AVG. FARE)		-.902 ^{***} (.267)	-1.015 ^{***} (.286)
Segment Fixed Effects	Y	Y	Y
Aircraft Fixed Effects	Y	Y	Y
Demographic Controls (X^D)	Y	Y	Y
Congestion Controls (X^C)	Y	Y	Y
Carrier/Day Fixed Effects	Y	Y	Y
Origin/Carrier Fixed Effects	N	N	N
Observations	243,113	243,113	243,113

Notes: Sample excludes Southwest Airlines. Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

TABLE 6—REGRESSION RESULTS: BUSINESS VS. NON-BUSINESS FLIGHTS

Dependent Variable:	SCHEDULED ELAPSED TIME, mins		
	(1)	(2)	(3)
	OLS	OLS	IV
INTERNET	3.255** (1.301)	2.875** (1.292)	2.513* (1.360)
HHI	1.075** (.496)	1.179** (.521)	3.436*** (.759)
INTERNET * HHI	-2.614*** (.692)	-2.721*** (.692)	-3.869*** (.787)
MORNING	-1.314*** (.142)	-1.293*** (.142)	-0.954*** (.113)
INTERNET * MORNING	1.249*** (.211)	1.233*** (.212)	.676*** (.123)
INTERNET * LARGE	1.002*** (.372)	.756** (.364)	.035 (.417)
INTERNET * LARGE * MORNING	.093 (.253)	.089 (.253)	.07 (.250)
LOG (AVG. FARE)		-1.687*** (.261)	-7.418*** (1.599)
Segment Fixed Effects	Y	Y	Y
Aircraft Fixed Effects	Y	Y	Y
Demographic Controls (X^D)	Y	Y	Y
Congestion Controls (X^C)	Y	Y	Y
Carrier/Day Fixed Effects	Y	Y	Y
Origin/Carrier Fixed Effects	N	N	N
Observations	286,004	286,004	286,004

Notes: Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

TABLE 7—REGRESSION RESULTS: SHORTEST SCHEDULED ELAPSED TIME

	(1)	(2)	(3)	(4)
	OLS	OLS	OLS	OLS
INTERNET	7.574*** (.122)	11.90*** (.396)	3.533*** (0.979)	4.079*** (0.995)
HHI		5.909*** (.381)	6.064*** (.376)	6.883*** (.377)
INTERNET * HHI		-5.336*** (.498)	-5.259*** (.489)	-5.483*** (.488)
# DEPARTURES ORIGIN AT FLIGHT'S HOUR				.051*** (.009)
# DEPARTURES DESTINATION AT FLIGHT'S HOUR				-.041*** (.009)
# ARRIVALS DESTINATION AT FLIGHT'S HOUR				.047*** (.009)
# ARRIVALS ORIGIN AT FLIGHT'S HOUR				-.044*** (.009)
Segment Fixed Effects	Y	Y	Y	Y
Demographic Controls (X ^D)	N	N	N	Y
Day Fixed Effects	N	N	Y	Y
Observations	48,169	48,169	48,169	48,169

Notes: Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

TABLE 8—REGRESSION RESULTS: ACTUAL ELAPSED TIME

Dependent Variable:	ACTUAL ELAPSED TIME, mins					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	OLS	OLS	IV
INTERNET	14.546*** (4.673)	13.571*** (4.591)	11.297** (4.585)	13.027*** (4.654)	12.513*** (4.598)	11.962*** (4.586)
HHI	2.758 (1.722)	4.163** (1.739)	7.782*** (2.111)	3.530** (1.710)	4.682*** (1.719)	7.472*** (2.086)
INTERNET * HHI	-9.516*** (2.436)	-10.008*** (2.427)	-11.106*** (2.534)	-9.923*** (2.428)	-10.356*** (2.415)	-11.260*** (2.488)
# DEPARTURES ORIGIN AT FLIGHT'S HOUR	.092*** (.005)	.092*** (.005)	.092*** (.005)	.090*** (.005)	.090*** (.005)	.091*** (.005)
# DEPARTURES DESTINATION AT FLIGHT'S HOUR	.036*** (.005)	.037*** (.005)	.037*** (.005)	.035*** (.005)	.035*** (.005)	.036*** (.005)
# ARRIVALS DESTINATION AT FLIGHT'S HOUR	.050*** (.007)	.049*** (.007)	.046*** (.007)	.052*** (.007)	.051*** (.007)	.048*** (.007)
# ARRIVALS ORIGIN AT FLIGHT'S HOUR	.012** (.006)	.014** (.005)	.019*** (.006)	.016*** (.006)	.016*** (.006)	.018*** (.006)
LOG (AVG. FARE)		-4.734*** (.816)	-14.830*** (3.624)		-4.223*** (.824)	-12.803*** (3.989)
Segment Fixed Effects	Y	Y	Y	Y	Y	Y
Aircraft Fixed Effects	Y	Y	Y	Y	Y	Y
Demographic Controls (X^D)	Y	Y	Y	Y	Y	Y
Addl' Congestion Controls (X^C)	Y	Y	Y	Y	Y	Y
Carrier/Day Fixed Effects	Y	Y	Y	Y	Y	Y
Origin/Carrier Fixed Effects	N	N	N	Y	Y	Y
Observations	286,004	286,004	286,004	286,004	286,004	286,004

Notes: Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

TABLE 9—REGRESSION RESULTS: ARRIVAL DELAYS

Dependent Variable:	ARRIVAL DELAY, mins					
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	OLS	OLS	IV
INTERNET	10.372** (4.231)	9.807** (4.183)	8.876** (4.185)	9.674** (4.285)	9.366** (4.249)	9.154** (4.234)
HHI	1.414 (1.576)	2.228 (1.59)	3.851** (1.877)	1.952 (1.578)	2.644* (1.596)	4.210* (1.917)
INTERNET * HHI	-6.416*** (2.239)	-6.701*** (2.237)	-7.195*** (2.311)	-6.856*** (2.245)	-7.116*** (2.242)	-7.642*** (2.281)
# DEPARTURES ORIGIN AT FLIGHT'S HOUR	0.032*** (0.005)	0.032*** (0.005)	0.032*** (0.005)	0.031*** (0.005)	0.031*** (0.005)	0.032*** (0.005)
# DEPARTURES DESTINATION AT FLIGHT'S HOUR	0.012** (0.005)	0.012** (0.005)	0.012** (0.005)	0.011** (0.005)	0.011** (0.005)	0.012** (0.005)
# ARRIVALS DESTINATION AT FLIGHT'S HOUR	0.019*** (0.006)	0.018*** (0.006)	0.017*** (0.006)	0.021*** (0.006)	0.020*** (0.006)	0.018*** (0.006)
# ARRIVALS ORIGIN AT FLIGHT'S HOUR	0.009* (0.005)	0.010** (0.005)	0.013** (0.005)	0.010** (0.005)	0.011** (0.005)	0.012** (0.005)
LOG (AVG. FARE)		-2.741*** (0.741)	-6.648** (3.242)		-2.536*** (0.771)	-6.997* (3.816)
Segment Fixed Effects	Y	Y	Y	Y	Y	Y
Aircraft Fixed Effects	Y	Y	Y	Y	Y	Y
Demographic Controls (X^D)	Y	Y	Y	Y	Y	Y
Addl' Congestion Controls (X^C)	Y	Y	Y	Y	Y	Y
Carrier/Day Fixed Effects	Y	Y	Y	Y	Y	Y
Origin/Carrier Fixed Effects	N	N	N	Y	Y	Y
Observations	286,004	286,004	286,004	286,004	286,004	286,004

Notes: Standard errors are in parentheses. Asterisks denote the significance level of coefficients: *** - 1 percent, ** - 5 percent, * - 10 percent. Standard errors are clustered within a segment. Additional control variables are described in the text.

FIGURE 1. AVERAGE SCHEDULED FLIGHT TIME 1990 - 2007

Notes: Average is calculated across directional non-stop segments that had at least five daily flights. Data source: On-Time performance database.

FIGURE 2. AVERAGE SHORTEST SCHEDULED FLIGHT TIME 1997 - 2007

Notes: Average is calculated across directional non-stop segments. Data source: OAG

FIGURE 3. RELATIVE CHANGE IN SHORTEST VS. MEAN SCHEDULED FLYING TIME

Notes: The figure shows the ratio between the shortest and average scheduled flight times. The sample includes directional non-stop segments that had at least five daily flights. The calculation of the average flight time excludes flights with the fastest scheduled times. The figure illustrates that changes in the scheduled flight times were larger for the flights with the shortest scheduled times.
Data source: OAG database.