How much do Consumers really Value Air travel On-time Performance, and to what extent are Airlines Motivated to Improve their On-time Performance?

Philip G. Gayle* and Jules O. Yimga**

Forthcoming in ***Economics of Transportation***

This draft: December 1, 2017
First draft: August 13, 2015

Abstract

This paper estimates the value consumers place on air travel on-time performance (OTP), and the extent to which airlines are motivated to improve their OTP. We find robust evidence that consumers value OTP and are willing to pay to avoid delays. Airlines can invest to improve OTP, but would independently choose to do so only if on-time performance improvement leads to increases in profitability. Using a methodology that does not require having actual cost data to draw inference on cost changes associated with improvement in OTP, we analyze airlines' optimal OTP-improvement investment choice. The modeling framework allows us to provide estimates of OTP-related marginal investment costs per minute of improvement necessary to achieve specific percent reductions in arrival delay minutes from the current levels of arrival delay minutes observed in the data.

Keywords: Airline On-time Performance; Commercial Aviation

JEL classification codes: L93; L13

Acknowledgements: For very helpful comments and suggestions we thank editor Mogens Fosgerau, two anonymous referees, Yang-Ming Chang, Jin Wang, and Tian Xia. Any remaining errors are our own.

*Kansas State University, Department of Economics, 322 Waters Hall, Manhattan, KS 66506; Voice: (785) 532-4581; Fax: (785) 532-6919; email: gaylep@ksu.edu; **Corresponding author**.

** Embry-Riddle Aeronautical University, School of Business, 3700 Willow Creek Road, Prescott, AZ 86301; Voice: (928) 777-3315; Fax: (928) 777-3827; email: jules.yimga@erau.edu.

1. Introduction

Punctuality is certainly a key performance indicator in the airline industry, and carriers with excellent on-time performance (OTP) record use it as a marketing tool by prominently displaying it on their websites. Given increased competition that followed deregulation of the airline industry in 1978, many carriers have resorted to product quality differentiation as a key to long-term profitability. Although airlines generally compete based on pricing, flight OTP is a very important component of airline service quality, which drives customer satisfaction and loyalty. For example, in the 1990's American Airlines ran ads calling itself "The On-Time Machine."[1] Likewise, airlines that produce excessive flight delays receive a great deal of negative publicity.

In 1987, the U.S. Congress passed the flight on-time disclosure rule amidst chronic air traffic delays that stirred public outcry and media coverage. The disclosure rule made it mandatory for airlines with at least one percent of all domestic traffic to publish flight-by-flight delay data. Airlines are required to track and report five segments of travel time for each of their flights to the Federal Aviation Administration (FAA): *i)* departure delay, *ii)* taxi-out, *iii)* air time, *iv)* taxi-in, and *v)* arrival delay.

Remarkably, even with the flight on-time disclosure rule of 1987, the industry's OTP is still far below satisfactory levels. A report from the U.S. Department of Transportation's (DOT) Office of Aviation Enforcement and Proceedings[2] revealed that the most prevailing consumer air travel complaint in the year 2000, stems from flight problems namely cancellations, delays and missed connections. In fact, 1 out of 4 flights was either delayed, canceled or diverted (Rupp,

[1] Boozer et al. (1990)
[2] US Department of Transportation Office of Aviation Enforcement and Proceedings (USDTOAEP) Feb. 2001 p. 34

Owens, and Plumly, 2006). According to Mayer and Sinai (2003), in year 2000, flights that arrived at their destination within 15 minutes of their scheduled arrival time and without being canceled or diverted, accounted for less than 70 percent. Even more recently, the Bureau of Transportation statistics (BTS) showed that 23.02% of U.S. domestic flights were delayed[3] in year 2014, an increase from 14.69% in year 2012. The BTS maintains an archive of monthly and yearly OTP data that is also accessible through the Internet.[4] Thus, passengers' most common source of frustration is flight delay.

In the midst of these delay statistics, airlines often claim that air traffic delays are out of their control, placing the blame on adverse weather or air traffic control as the most common culprits.[5] A good portion of delay can be attributed to extreme weather, air traffic control and security checks (U.S. DOT, 2015). In June 2003, the Air Carrier On-Time Reporting Advisory Committee defined five broad categories for the cause of any flight delay:[6]

1. ***Air Carrier***: The cause of the cancellation or delay was due to circumstances within the airline's control (e.g. maintenance or crew problems, aircraft cleaning, baggage loading, fueling, etc.).
2. ***Extreme Weather:*** Significant meteorological conditions (actual or forecasted) that, in the judgment of the carrier, delays or prevents the operation of a flight (e.g. tornado, blizzard, hurricane, etc.). Weather delays are also included in the National Aviation System and late-arriving aircraft categories.

[3] A flight is considered delayed if it arrived at (or departed) the gate 15 minutes or more after the scheduled arrival (departure) time.
[4] The BTS archived data are located at http://www.transtats.bts.gov/homedrillchart.asp
[5] http://www.washingtonpost.com/lifestyle/travel/what-to-do-when-airlines-blame-flight-problems-on-circumstances-beyond-our-control/2015/02/12/7298b264-a57f-11e4-a7c2-03d37af98440_story.html
[6] http://www.rita.dot.gov/bts/help/aviation/html/understanding.html

3. ***National Aviation System (NAS)***: Delays and cancellations attributable to the National Aviation System that refer to a broad set of conditions—non-extreme weather conditions, airport operations, heavy traffic volume, air traffic control, etc.
4. ***Late-arriving Aircraft:*** A previous flight with same aircraft arrived late, causing the present flight to depart late, and
5. ***Security:*** Delays or cancellations caused by evacuation of a terminal or concourse, re-boarding of aircraft because of security breach, inoperative screening equipment and/or long lines in excess of 29 minutes at screening areas.

Although some of these factors are uncontrollable, airlines still have a substantial level of control over their OTP. An airline can schedule a longer flight time to absorb potential delays on the taxiways, or choose a longer layover on the ground to buffer against the risk of a late incoming aircraft (Mayer and Sinai, 2003). Figure 1 shows declining shares of flight delay caused by weather and air traffic control (NAS) over time, while over the same time period, the shares of delay caused by late-arriving aircraft and air carrier, continue to rise. Figures 1 and 2 suggest that OTP improvement potential within the reach of airlines is significant.

The objective of this paper is twofold. First, we examine the monetary value that consumers place on OTP. In order to make our case about consumers valuing improved OTP, we estimate a discrete choice demand model, which allows us to quantify the opportunity cost of delays to consumers. Thus, incorporating OTP into our demand model affords us the advantage of measuring how much OTP matters to consumers. In essence, we estimate how much consumers are willing to pay to avoid each minute of delay.

Figure 1: Causes of Delay by Percent Share of Total Delay Minutes

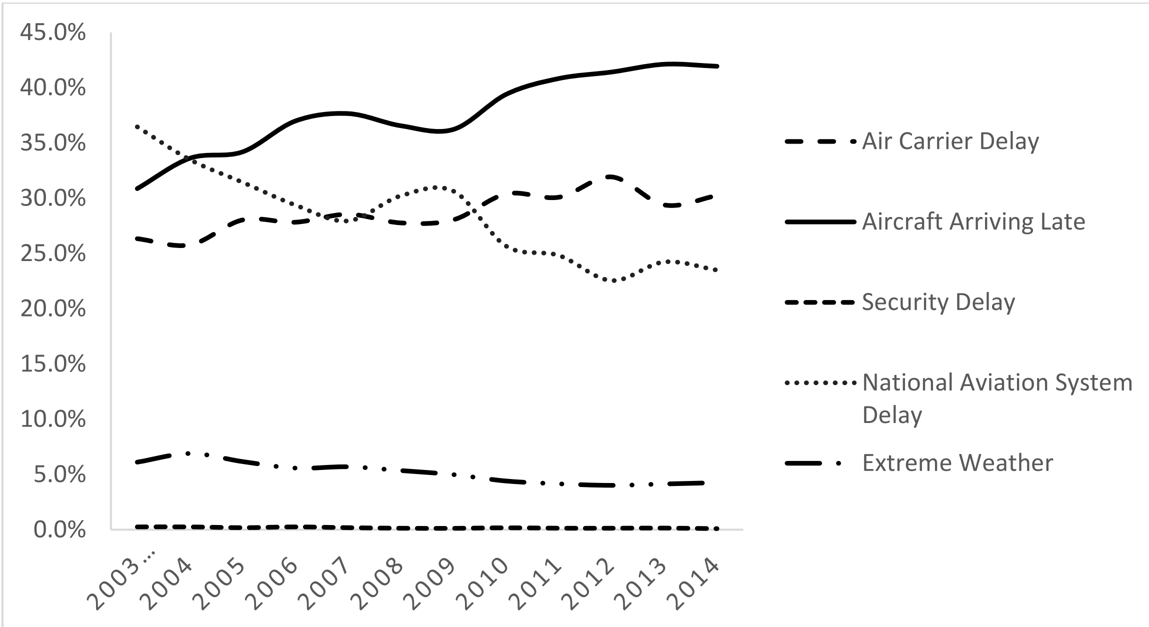
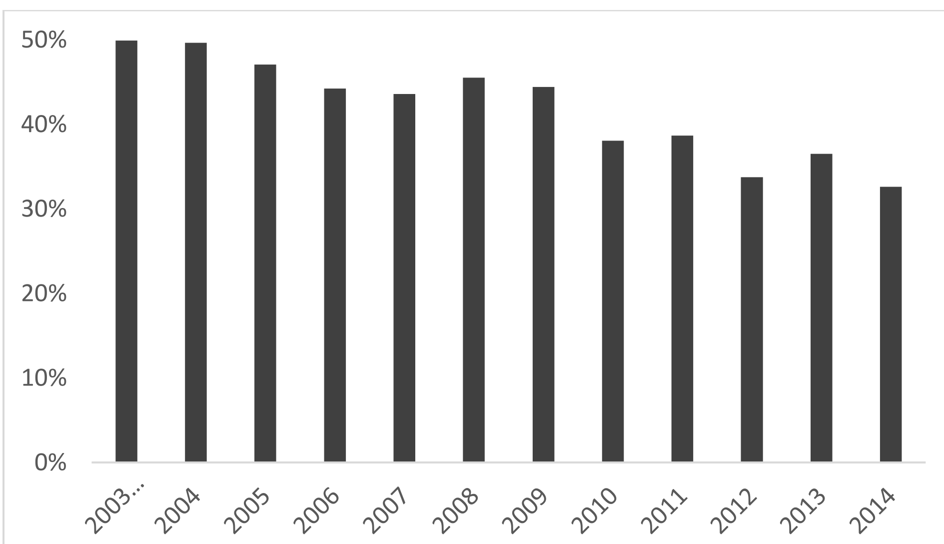


Figure 2: Weather's Share of Total Delay Minutes



Second, if consumers do value OTP, to what extent may airlines benefit from improving their OTP? How does improved OTP affect airlines' variable profits in an oligopoly market structure, a strategic environment where few firms are competing with each other? One way to answer these questions is by examining how airlines' price-cost markup and demand for air travel respond to changes in OTP. To facilitate this part of the analysis we specify a supply-side of the model assuming that multiproduct airlines set prices for their differentiated products according to a static Nash equilibrium. The supply-side of the model is first used to generate estimates of product-level price-cost markups and airline-level variable profits, and then used to conduct counterfactual exercises to assess the extent to which improvements in arrival delay (i.e. improvements in OTP) influence variable profits of airlines. The variations in airlines' variable profit due to counterfactual improvements in OTP are used to assess the incentive a given airline has to improve OTP.

The rationale for using variable profit changes as measures of airlines' incentive to improve OTP is that airlines care about their bottom line, and if improving OTP leads to increases in variable profits, then improving OTP might be a worthwhile proposition for airlines. In addition, because variable profit is a function of price-cost markup and demand level, we are able to decompose the changes in variable profit into changes in its components and examine how these components drive the changes in variable profits.

Over the last three decades, empirical studies on air travel have neglected to explicitly incorporate OTP measures of service quality into air travel demand estimation. De Vany (1975) is among the first to incorporate service quality, proxied by flight frequency, in a demand model. Anderson and Kraus (1981), Ippolito (1981), Abrahams (1983) and De Vany (1975) estimated air travel demand models with schedule delay[7] as a measure of service quality. Our present paper contributes to this literature.

A novel feature of this study is that we model demand based on passenger origin-destination[8] markets rather than flight segments within a broader origin-destination market. Previous demand studies that are based on origin-destination markets have not incorporated flight delay,[9] and studies that have incorporated delay (Abrahams, 1983; Anderson and Kraus, 1981; De Vany, 1975; Douglas and Miller, 1974; Ippolito, 1981), model demand based on flight segments rather than passenger origin-destination markets. Since much of air travel from passengers' origin to their destination use several flight segments rather than a single non-stop flight, it is reasonable to model demand within an origin-destination framework, which captures the imperfect substitutability between non-stop and intermediate-stop products within an origin-destination market. In fact, our dataset shows that only 17.6 percent of itineraries are non-stop flights from the relevant passengers' origin to destination. Travelers typically demand air transportation between a directional origin and destination pair rather than segment-by-segment flights. Our focus on origin-destination markets not only helps to predict passengers' behavioral intentions, but provides a more realistic structure for the measurement of consumer welfare effects of flight delay.

Several conclusions emerge from the empirical analysis. First, other things equal, consumers value OTP and are willing to pay for it. Our demand estimates show that consumers are willing to pay $1.56 per minute late to avoid delay. Second, we find that, a 10% reduction in arrival delay minutes (improved OTP) results in an increase in variable profit by a mean 3.95

[7] Defined as the sum of frequency delay and stochastic delay. Frequency delay is the gap between one's desired and the nearest offered departure time while stochastic delay is time lost due to the nearest offered departure being unavailable.
[8] Tickets are issued for the entire itinerary, which may include intermediate airport(s).
[9] Origin-destination passenger data contain no information on routings' on-time performance.

percent. Furthermore, we find that it is the increase in demand levels, as compared to increase in markup, that accounts for most of the increase in variable profits.

Given the finding that reductions in arrival delay minutes (improved OTP) yield increases in variable profits, we are able to use the model to recover estimates of the cost per minute of delay improvement that rationalizes the level of delay minutes we observe in the data. The model predicts that a 2.39% increase in OTP-related marginal investment cost per minute of improvement is necessary to achieve a 10% reduction in arrival delay minutes below their current levels..

The remainder of the paper is organized as follows. The next section provides a review of the literature. Section 3 describes the data used for analysis. Section 4 describes variables used for estimation. Section 5 discusses the research methodology and estimation procedure used to analyze the OTP effects. Results are presented and discussed in Section 6, while concluding remarks are gathered in Section 7.

2. Literature Review

Researchers have written extensively on airline flight delays. The literature on flight delays abounds in both operations management and economics. The operations management literature uses models that attempt to explain flight delays from an operational standpoint of running an airline. Shumsky (1995) contributed to the literature of airline scheduling performance analysis by examining U.S. air carriers' response to the on-time disclosure rule of 1987. The rule creates incentives for carriers to improve their OTP by either reducing the amount of time to complete a flight or lengthening the amount of time scheduled for a flight. Shumsky (1995) shows that although actual flight times have fluctuated, scheduled flight times have increased significantly. Ramdas and Williams (2006) investigate the tradeoff between aircraft utilization and OTP using queuing theory and found that flight delays increase with increasing aircraft utilization.

In the economics literature, researchers have tried to explain variations in flight delays by estimating how flight delay relates to airline hub size and airport concentration (Mayer and Sinai, 2003), competition (Mayer and Sinai, 2003; Mazzeo, 2003; Rupp et al., 2006), multimarket contact (Prince and Simon, 2009), prices (Forbes, 2008) and entry or threat of entry (Prince and Simon, 2014), among others. Mayer and Sinai (2003) found that as origin (destination) airport concentration increases, flight delays originating (arriving) from (to) that airport decrease. On the other hand, for both origin and destination airports, flight delays increase with increasing airport hub size. Mazzeo (2003) found that the prevalence and duration of flight delays are significantly greater on routes where only one airline provides direct service. Rupp and Holmes (2006) examined the determinants of flight cancellations such as revenue, competition, aircraft utilization, and airline network. Prince and Simon (2009) tested the mutual forbearance hypothesis (Edwards, 1955) using different measures of OTP. This hypothesis suggests that firms that meet in multiple markets compete less aggressively because they recognize that a competitive attack in any one market may call for response(s) in all jointly contested markets. They conclude that multimarket contact increases delays and that the effect is substantially larger in less competitive markets.

Forbes (2008) examines the effect of air traffic delays on airline fares and found that prices fall by $1.42 on average for each additional minute of flight delay, and that the price response is substantially larger the more competitive the markets are. Prince and Simon (2014) examine whether entry and entry threats by Southwest Airlines cause incumbent airlines to improve their OTP as a way to protect their market share. Surprisingly, their results show that incumbents' delays increase with entry and entry threats by Southwest Airlines. They provide two possible explanations for their findings: 1) incumbents worsen OTP in an effort to cut costs in order to compete against Southwest's low costs/prices; or 2) incumbents worsen on-time performance to differentiate away from Southwest, a top-performing airline in on-time performance.

3. Definitions and Dataset Construction

3.1. Definitions

A *market* is defined as a directional, round-trip between an origin and destination city during a specific time period. By directional, we mean that a round-trip air travel from Chicago to Boston is a distinct market from a round-trip air travel from Boston to Chicago. This directional definition of a market controls for heterogeneity in demographics across origin cities that may affect air travel demand (Berry, Carnall and Spiller, 2006; Gayle, 2007).

An itinerary is a planned route from an origin city to a destination city. It entails one or more flights, each flight typically representing point-to-point travel between two airports.

There are two types of carriers in the data—ticketing carrier and operating carrier. The ticketing carrier is the airline that issues the flight reservation or ticket to consumers. The operating carrier is the airline that operates the aircraft, i.e., the airline that actually transports the passengers.

An air travel *product* is defined as an itinerary-carrier combination during a particular time period. We focus on products that use a single airline as both ticketing and operating carrier for all segments of the trip. Table 1 reports the names and associated code of the carriers in our sample.

3.2. Dataset Construction

We construct our dataset using data from two sources that span from the first quarter of 2002 to the fourth quarter of 2012. First, we use data from the Airline Origin and Destination Survey (DB1B) collected by the Office of Airline Information of the Bureau of Transportation Statistics. The data are quarterly and represent a 10 percent sample of airline tickets from reporting carriers. Each record in this database contains the following information; (i) the identities of origin, destination, and intermediate stop(s) airports on an itinerary; (ii) the identities of ticketing and operating carriers on the itinerary; (iii) the price of the ticket; (iv) the number of passengers who bought the ticket at that price during the relevant quarter; (v) total itinerary distance flown from origin to destination; and (vi) the nonstop distance between the origin and destination. Regrettably, passenger-specific information, that would facilitate the estimation of a richer demand model than the one we use, is not available. Information on ticket restrictions such as advance-purchase and length-of-stay requirements are unavailable as well.

Table 1: Airlines in Sample

Code	Airline
AA	American Airlines
AQ	Aloha Airlines
AS	Alaska Airlines
B6	JetBlue Airways
CO	Continental Air Lines
DH	Independence Air
DL	Delta Air Lines
F9	Frontier Airlines
FL	AirTran Airways
HA	Hawaiian Airlines
HP	America West Airlines
NW	Northwest Airlines
OO	SkyWest
TZ	ATA Airlines
UA	United Air Lines
US	US Airways
VX	Virgin America Inc.
WN	Southwest Airlines
XE	ExpressJet Airlines
YX	Midwest Airlines

Second, we also use the U.S. Department of Transportation (DOT) Bureau of Transportation Statistics (BTS) On-Time Performance database to construct our on-time performance measure of product quality. All U.S. domestic carriers with revenues from domestic passenger flights of at least one percent of total industry revenues must report flight OTP data. The data frequency is monthly and covers scheduled-service flights between points within the United States. So, a record in this survey represents a flight, where the given flight may contain passengers who are using very different routing itineraries that share this flight on a segment of each passengers' multi-segment itinerary. Each record or fight[10] contains information on the operating carrier, the origin and destination airports for the flight, miles flown, flight times, and departure/arrival delay information.

Perhaps owing to the challenge of matching the two distinct databases described above, previous demand studies that are based on origin-destination data have not incorporated OTP data. A reason for the matching challenge is that the origin-destination database gathers information at the travel ticket level but contains no information at the operating flight level, while the OTP database gathers information at the operating flight level but contains no information at the travel ticket level.

To construct a product quality variable from the OTP data, we compute the average departure (arrival) delay across flights for each carrier at any given origin (destination) airport in a quarter for a given year. This results in an aggregated on-time performance data in which a single observation is now at the carrier-airport-quarter level instead of being at the initial less aggregated level of operating flight-airport-month. This aggregated OTP data is then matched to

[10] Some flights could be segments of itineraries with intermediate stop(s)

the DB1B origin-destination dataset. The matching process is done at all airports of the passengers' itineraries.

In this study, we focus on carriers' OTP at the itinerary's final destination airport. In order to construct our data set, we place some restrictions on the raw data:

(i) We confine our analysis to U.S. domestic flights operated by U.S. domestic carriers.

(ii) Our analysis focuses on performed flights arriving at their destination and do not include cancellations or diversions since they do not accurately depict OTP.

(iii) We focus on passengers purchasing round-trip, coach class tickets.

(iv) We exclude real airfares less than \$25 or greater than \$2,000. Dropping real airfares that are too low gets rid of discounted airfares from passengers using their frequent-flyer miles to offset the full price of the trip or employee travel tickets. Likewise, excluding real airfares that are too high gets rid of first-class or business-class tickets.

(v) Our analysis is limited to air travel products possessing at least 9 passengers per quarter to exclude products that are not part of the regular offerings by an airline.

(vi) Our analysis focuses on itineraries: (1) within the 48 states in U.S. mainland; (2) no more than one intermediate stop; and (3) with a single and the same ticketing and operating carrier.

(vii) Following Aguirregabiria & Ho (2012), markets selection focuses on air travel amongst the 63 largest U.S. cities. City size is based on the Census Bureau's Population Estimates Program (PEP), which publishes estimates of U.S. population. Data are drawn from the category "Cities and Towns". We use the size of population in the origin city as a proxy for potential market size. Unlike Aguirregabiria & Ho (2012), we do not group cities that belong to the same metropolitan areas and share the same airport since airport grouping will lessen the heterogeneity in OTP.

(viii) Given that there are often multiple records for the same itinerary because different passengers paid different prices, we construct the price and quantity variables by averaging the airfares and aggregating the number of passengers respectively based on our product definition and then collapse the data by product. So, in the collapsed data that we use for analyses, a product appears only once during a given time period.

Our final working dataset includes a total of 63 airports representing 1,281,413 air travel products bought across 156,743 different directional city-pair markets.

4. Variables used for Analysis

4.1 On-Time Performance Measure

Delay-based measures are obtained using on-time performance from the DOT BTS' dataset. According to the U.S. DOT, flights that do not arrive at (depart from) the gate within 15 minutes of scheduled arrival (departure) time are late arrivals (departures). This represents performance measured against airlines' published schedules. For example, if your flight is scheduled to arrive at 3:30 p.m. and does not get in until 3:44 p.m., it is not late according to the U.S. DOT's definition of flight lateness. With this measurement standard, 81.9 percent of flights arrived on time in April 2015.[11] However, if we count all flights that arrive after their scheduled arrival time, including when they are one minute late, the industry's "true" on-time performance drops to about 60 percent.

The OTP measure used for analysis is "arrival minutes late" at the destination airport. There are other measures of OTP reported by the DOT BTS such as the percentage(s) of flights arriving at least 15 (or 30) minutes late that have been explored.[12] However, a potential limitation of the percentage(s) of flights arriving at least 15 (or 30) minutes late is that it only

[11] US Department of Transportation (2015)
[12] See Barnett et al. (2001), Rupp et al. (2006), Forbes (2008) and Prince and Simon (2009) amongst others

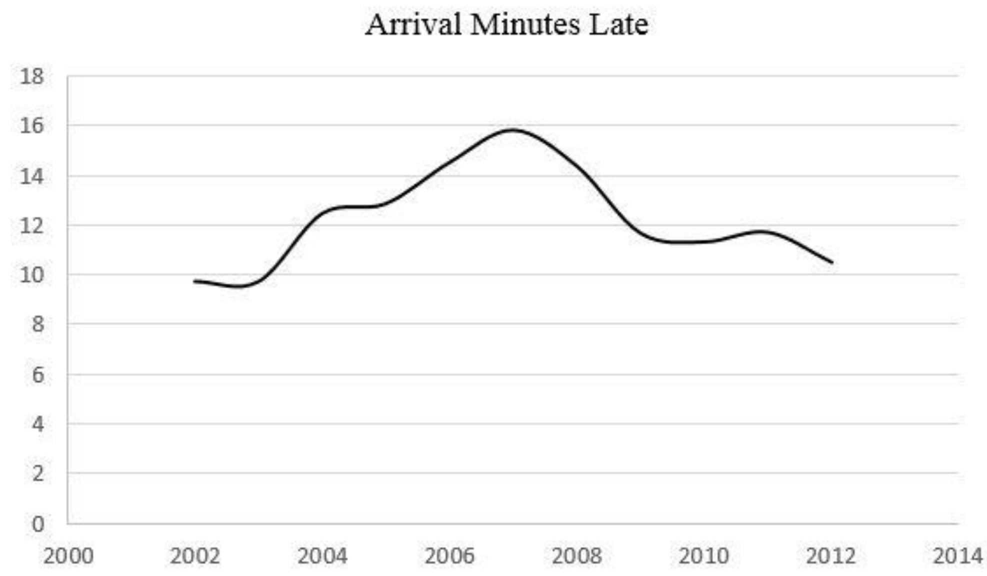
captures delay beyond a certain threshold. "Minutes late," on the other hand, is a more relevant measure given that it is a continuous variable.

Table 2 summarizes OTP by carrier over the 2002 to 2012 time span. Hawaiian Airlines (HA) tops all carriers with the best arrival on-time performance. Figure 3 shows that over the 2002 to 2012 time span, airlines performed the worst in 2007.

Table 2: Airlines' Mean Arrival Delay

Code	Airline	Mean Minutes Late
HA	Hawaiian Airlines	4.75
AQ	Aloha Airlines	9.23
VX	Virgin America Inc.	9.90
WN	Southwest Airlines	9.93
AS	Alaska Airlines	10.50
HP	America West Airlines	10.60
US	US Airways	10.93
F9	Frontier Airlines	12.02
DL	Delta Air Lines	12.16
CO	Continental Air Lines	13.01
NW	Northwest Airlines	13.08
TZ	ATA Airlines	13.13
FL	AirTran Airways	13.22
OO	SkyWest	13.42
XE	ExpressJet Airlines	13.48
UA	United Air Lines	13.51
AA	American Airlines	13.73
YX	Midwest Airlines	14.48
B6	JetBlue Airways	15.14
DH	Independence Air	15.49
Overall Mean		12.31

Figure 3: Overall Airline On-Time Performance (2002:Q1—2012:Q4)



4.2 Routing Quality Measure

We constructed and include the distance-based measure, *Routing Quality,* into our analysis following the literature.[13] *Routing Quality* is defined as the ratio of nonstop fight distance to the product's itinerary fight distance used to get passengers from the origin to destination. Based on our routing quality measure, a nonstop flight between the origin and destination will have the shortest itinerary flight distance. Hence, air travel products that require intermediate airport stop(s) that are not on a straight path between the origin and destination, will have an itinerary flight distance that is longer than the nonstop flight distance. Our rationale for considering this measure is that the longer the itinerary flight distance of an intermediate-stop product relative to the nonstop flight distance, the lower the routing quality of the intermediate-stop product.

[13] See Reiss and Spiller (1989); Borenstein (1989); Ito and Lee (2007); Fare, Grosskopf and Sickles (2007); Gayle (2007 and 2013); Chen and Gayle (2013) and Gayle and Yimga (2015).

4.3 Creation of Other Variables

In the collapsed and matched dataset, we create more variables to include in the empirical model. The observed product share variable is created by dividing product quantity sold by the market size. Measured non-price product characteristic variables include: *Nonstop* and *Origin Presence*. *Nonstop* is a zero-one indicator variable that takes the value one only if a product has no intermediate stop. This variable constitutes one measure of the travel inconvenience embodied in a product's itinerary since, assuming all other product characteristics equal, passengers likely prefer a non-stop product to one with intermediate stop(s). The *Origin Presence* variable counts the number of different cities that an airline provides service to via a nonstop flight from the origin airport of the market.

We include zero-one dummy variables for quarter, year, origin, destination, and carrier to capture product characteristics unobserved to us that vary across time periods, origins, destinations, and carriers. Table 3 reports summary statistics for variables used in the analysis.

Table 3: Summary Statistics

Variables	*Mean*	*Std. Dev.*	*Min*	*Max*
Price[a]	171.810	57.4635	51.579	1588.565
Quantity (Number of passengers that purchase the product)	166.841	511.468	9	11266
Observed Product Share[b]	0.00028	0.001	1.07e-06	0.095
Origin presence	21.162	24.103	0	142
Nonstop (dummy variable)	0.176	0.380	0	1
Itinerary distance flown (miles)[c]	1556.71	694.619	47	3982
Nonstop flight distance (miles)	1379.813	646.842	47	2724
Routing Quality[d]	0.889	0.129	0.338	1
Arrival On-Time Performance Variable:				
Minutes Late	12.31	4.88	0	68.43
Number of Products	1,281,413			
Number of Markets [e]	156,743			

[a] Adjusted for inflation

[b] Observed Product Share is computed as product quantity sold divided by the population of origin city

[c] Reported as "market miles flown" in the DB1B database

[d] Defined as the ratio of non-stop distance to itinerary distance

[e] A market is an origin-destination-time period combination

5. The Model

5.1 Demand

The nested logit model is used to specify air travel demand. A typical passenger *i* may either buy one of *J* products (air travel products), *j=1,…,J,* or otherwise choose the outside option (*j=0*), where the outside option includes driving or using other modes of transportation. Thus, passenger *i* makes a choice among $J_{mt}+1$ alternatives in market *m* during time period *t*. The nested logit model classifies products into *G* groups, and one additional group for the outside good. Therefore, products are organized into *G+1* mutually exclusive groups. We group products by carriers. A passenger solves the following utility maximization problem:

$$\underset{j\epsilon\{0,1,\ldots J_{mt}\}}{\text{Max}} U_{ijmt} = \delta_{jmt} + \sigma\varsigma_{imtg} + (1-\sigma)\varepsilon_{ijmt} \tag{1}$$

$$\delta_{jmt} = x_{jmt}\beta + \alpha p_{jmt} + \eta_j + \upsilon_t + origin_m + dest_m + \xi_{jmt} \tag{1.1}$$

where:

- U_{ijmt}: passenger i's utility from choosing product j.
- δ_{jmt} : mean level of utility across passengers that choose product j.
- ς_{imtg}: a random component of utility common across all products within the same group.
- ε_{ijmt}: an independently and identically distributed (across products, consumers, markets and time) random error term assumed to have an extreme value distribution.
- x_{jmt} : vector of explicit product quality measures and other non-price product characteristics described below
- p_{jmt}: price
- η_j: airline-specific fixed effects
- υ_t: time period fixed effects
- $origin_m$: origin city fixed effects
- $dest_m$: destination city fixed effects
- ξ_{jmt} : unobserved (by the researcher) component of product characteristics that affects consumer utility.

Vector x_{jmt} includes explicit product quality measures (OTP measure and *Routing Quality*), a measure of the size of an airline's airport presence (*Origin Presence*), and a zero-one indicator variable, *Nonstop,* that equals to unity only if the product uses a nonstop flight to get passengers from the origin to destination.

The vector β measures the passenger's marginal utilities associated with non-price product characteristics. The parameter α captures the marginal utility of price. The parameter σ lies between 0 and 1 and measures the correlation of consumer utility across products belonging to the same airline. The correlation of preferences increases as σ approaches 1. In the case where σ is 0, the model collapses to the standard logit model where products compete symmetrically. For notational convenience, we drop the market and time subscripts to complete the derivation of the model.

Let there be G_g products in group g. If product j is in group g then the conditional probability of choosing product j given that group g is chosen is given by:

$$s_{j/g} = \frac{e^{(1-\sigma)^{-1}\delta_j}}{D_g} \text{ where } D_g = \sum_{j\epsilon G_g} e^{(1-\sigma)^{-1}\delta_j} \tag{2}$$

The probability of choosing group g or group g 's predicted share is given by:

$$s_g = \frac{D_g^{1-\sigma}}{D_0^{1-\sigma} + \sum_{g=1}^{G} D_g^{1-\sigma}} \tag{3}$$

The outside good is the only good in group 0. Therefore, $D_0^{1-\sigma} = e^{\delta_0}$. We normalize the mean utility of the outside good to zero. This implies $D_0^{1-\sigma} = 1$. Equation (3) can be rewritten as:

$$s_g = \frac{D_g^{1-\sigma}}{1 + \sum_{g=1}^{G} D_g^{1-\sigma}} \tag{4}$$

The unconditional probability of choosing product j, or the market share of product j is:

$$s_j = s_{j/g} * s_g = \frac{e^{(1-\sigma)^{-1}\delta_j}}{D_g} \frac{D_g^{1-\sigma}}{1+\sum_{g=1}^{G} D_g^{1-\sigma}} \text{ or equivalently } s_j = \frac{e^{(1-\sigma)^{-1}\delta_j}}{D_g^{\sigma}\left[1+\sum_{g=1}^{G} D_g^{1-\sigma}\right]} \quad (5)$$

Therefore, the demand for product *j* is given by:

$$d_j = M * s_j(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \alpha, \beta, \sigma) \quad (6)$$

where M is a measure of market size—the population in the origin city. The market share of product j predicted by the model is represented compactly as $s_j(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \alpha, \beta, \sigma)$, where $\mathbf{x}$, $\mathbf{p}$ and $\boldsymbol{\xi}$ are vectors of observed non-price product characteristics, price, and product characteristics unobserved to us the researchers, respectively. α, β and σ are parameters to be estimated. The specific functional form for $s_j(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \alpha, \beta, \sigma)$ is given above in equation (5).

5.2 Variable Profits and Product Markups

We assume that carriers simultaneously choose prices as in a static Bertrand-Nash model of differentiated products. Let each carrier *f* offer for sale a set F_{fm} of products in market *m*. Firm *f*'s variable profit in market *m* is given by:

$$VP_{fm} = \sum_{j \epsilon F_{fm}} (p_{jm} - mc_{jm}) q_{jm}, \quad (7)$$

where $q_{jm} = d_{jm}(\mathbf{p})$ in equilibrium, q_{jm} is the quantity of travel tickets for product *j* sold in market *m*, $d_{jm}(\mathbf{p})$ is the market demand for product *j* in equation (6), $\mathbf{p}$ is a vector of prices for the *J* products in market *m*, and mc_{jm} is the marginal cost to provide product *j* in market *m*.

The corresponding first-order conditions are:

$$\sum_{r \epsilon F_f} (p_r - mc_r) \frac{\partial s_r}{\partial p_j} + s_j = 0 \text{ for all } j = 1, \dots, J \quad (8)$$

where the market subscript *m* is suppressed only for notational convenience. Equation (8) can be rewritten in matrix notation as:

$$(\Omega * \Delta) \times (\mathbf{p} - \mathbf{mc}) + s(\mathbf{p}) = 0 \quad (9)$$

where $\mathbf{p}$, $\mathbf{mc}$, and s(·) are J×1 vectors of product prices, marginal costs, and predicted product shares respectively, while $\Omega * \Delta$ is an element-by-element multiplication of two matrices. Δ is a J×J matrix of first-order derivatives of model predicted product market shares with respect to prices, where element $\Delta_{jr} = \frac{\partial s_r(\cdot)}{\partial p_j}$. Ω is a J×J matrix of appropriately positioned zeros and ones which describes carriers' ownership structure of the J products. For example, let Ω_{jr} denote an element in Ω, where

$$\Omega_{jr} = \begin{cases} 1 \text{ if there exists } f: \{j, r\} \subset F_f \\ 0 \text{ otherwise} \end{cases}$$

That is, $\Omega_{jr} = 1$ if products *j* and *r* are offered for sale by the same carrier, otherwise $\Omega_{jr} = 0$.

Based on equation (9), product-level markups can be obtained as:

$$Mkup(\mathbf{x}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma}) = \mathbf{p} - \mathbf{mc} = (\Omega * \Delta)^{-1} \times s(\mathbf{p}) \quad (10)$$

where $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}$ are estimates of the demand parameters. Let $markup_j(\mathbf{x}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma})$ be an element in $Mkup(\mathbf{x}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma})$. Note that $markup_j(\mathbf{x}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma})$ is a product markup function which depends exclusively on demand-side variables and parameter estimates.

With computed product markups in hand, variable profit can be computed by:

$$VP_{fm} = \sum_{j \epsilon F_{fm}} markup_{jm}(\mathbf{x}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma}) \times d_{jm}(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma}) \quad (11)$$

where the demand function $d_{jm}(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \hat{\alpha}, \hat{\beta}, \hat{\sigma})$ is based on equations (5) and (6).

5.3 Estimation of Demand

The estimation strategy of the demand parameters (α, β, σ) is such that the observed market shares $\mathbf{S}_{jmt}$ are equal to the market shares predicted by the model s_{jmt}. A well-known result in empirical industrial organization is that the demand model presented above results in the following linear equation:

$$\ln(\mathbf{S}_{\mathrm{jmt}}) - \ln(\mathbf{S}_{\mathrm{0mt}}) = x_{jmt}\beta + \alpha p_{jmt} + \sigma\ln(\mathbf{S}_{\mathrm{jmt/g}}) + \eta_j + \upsilon_t$$

$$+origin_m + dest_m + \xi_{jmt}. \tag{12}$$

$\mathbf{S}_{jmt}$ is the observed share of product j computed from the data by $\mathbf{S}_{jmt} = \frac{q_{jmt}}{M}$, where q_{jmt} is the quantity of air travel product j sold, and M is the population of the origin city. $\mathbf{S}_{0mt} = 1 - \sum_{j \in J_m} S_{jmt}$ is the observed share of the outside good. $\mathbf{S}_{\mathrm{jmt/g}}$ is the observed within-group share of product j, and the other variables are as previously described in equation (1.1). Equation (12) can be estimated using Two Stage Least Squares (2SLS) since price p_{jmt} and $\ln(\mathbf{S}_{\mathrm{jmt/g}})$ are endogenous.

5.4 Instruments for Endogenous Variables in Demand Equation

To construct instruments for price and within group product shares we exploit the fact that the set of product choices offered by airlines in a market is predetermined at the time of exogenous shocks to demand, and the non-price characteristics of the menu of products offered are primarily determined by the pre-existing route network structure of the airline.[14]

The instrument variables we use for the 2SLS estimation are: (1) number of competing products offered by other airlines that a given product faces, where the competing products each have the same number of intermediate stops as the given product; (2) the product's itinerary distance flown; and (3) the deviation of a product's itinerary flying distance-based routing quality measure from the mean routing quality measure across the set of products offered by the

[14] Unlike price and within group product share, airline route network structure is fixed in the short run, which mitigates the influence of demand shocks on the menu of products offered and their associated non-price characteristics (Gayle and Thomas, 2016).

carrier.[15] The rationale for using these instruments is discussed in Gayle and Thomas (2016). (1) and (2) instrument for price and are motivated by supply theory which predicts that a product's price is affected by changes in its markup and marginal cost. By rearranging Equation (10), it can be shown that price is a function of markup and marginal cost. Instrument (1) captures the degree of competition facing a product in a market, which in turn affects the size of a product's markup. Instrument (2) follows from the notion that flying distance is likely to be correlated with the marginal cost of providing the product. (3) instruments for the within group product share. The intuition for instrument (3) is as follows. Recall that we group products by airline in the nested logit demand model. So, instrument (3) is likely to be correlated with the within group product share because, all else equal, among the products offered by an airline in a given origin-destination market consumers are likely to prefer the product with the most direct routing, i.e., the product with the highest routing quality measure (Chen and Gayle, 2013 and Gayle and Thomas, 2016).

6. Empirical Results

6.1 Demand Results

We estimate the demand equation using both Ordinary Least Square (OLS) and Two-stage Least Squares (2SLS). Table 4 shows the demand regression results. As stated in Section 5.3, price p_{jmt} and within-group product shares $\mathbf{S}_{\mathrm{jmt/g}}$ are endogenous variables in the demand equation. Thus, OLS estimation produces biased and inconsistent estimates of the price coefficient and σ. A Hausman test confirms by rejecting the exogeneity of price and within-group product share at conventional levels of statistical significance.

[15] For instances where the routing quality of a given product is equal to the mean routing quality of all products offered by the carrier in a market, the deviation of routing quality instrument variable is constructed to take the maximum value of the routing quality measure of 1.

To confirm the validity of instruments used in the 2SLS regression, we estimate first-stage reduced-form regressions for each of the endogenous variables. First-stage reduced-form regressions where we regress p_{jmt} and $\ln(\mathbf{S}_{\mathrm{jmt/g}})$ against the instruments suggest that the instruments explain variations in the endogenous variables. R^2 measures for the regressions of price and within-group product share against the instruments are 0.1210 and 0.2281 respectively. Since the use of instruments is justified, we only discuss the 2SLS estimates. We also performed the Stock and Yogo's (2005) test for weak instrument. Here, we can reject the null hypothesis that the instruments are weak, because the test statistic of 2147.17 exceeds the critical values of conventional rejection rates of a nominal 5% Wald test. On the basis of this test, we do not have a weak-instrument problem.

The coefficient estimate on the price variable has the expected negative sign. All else equal, an increase in the product's price reduces the probability that a typical passenger will choose the product. The coefficient estimate on $\ln(\mathbf{S}_{\mathrm{jmt/g}})$, which is an estimate of σ should lie between zero and one and measures the correlation of consumers' preferences for products offered for sale by the same airline. Given that we nest products by airlines and σ is statistically significant, this suggests that passenger choice behavior shows some level of brand-loyalty to airlines. However, since the estimate of σ is closer to zero than it is to one, evidence of brand-loyal behavior is not very strong. Even though airlines use customer loyalty programs to strengthen relationships with their customers, such programs do not provide exceptional advantages when potential gain can be quickly eroded by competitive forces (Dowling and Uncles, 1997).

The coefficient estimate on *Origin Presence* is positive. This result is consistent with our expectations and suggests that travelers prefer to fly with airlines, ceteris paribus, that offer services to more destinations from the travelers' origin city. Chen and Gayle (2013), Gayle and Le (2013) and Berry (1990) among others, obtained similar findings.

Table 4: Demand Estimation Results

	OLS	2SLS
	(1)	(2)
Price	0.0003***	-0.0070***
	(1.58e-05)	(0.0001)
ln(Sj/g)	0.4372***	0.1409***
	(0.0006)	(0.0017)
Origin Presence	0.0143***	0.0148***
	(4.52e-5)	(0.0001)
Nonstop	1.0206***	1.0835***
	(0.0025)	(0.0032)
Routing Quality	1.8849***	1.9991***
	(0.0072)	(0.0083)
Arrival Minutes Late	-0.0110***	-0.0109***
	(0.0002)	(0.0002)
Constant	-10.6880***	-9.7242***
	(0.0118)	(0.0156)
Carrier Fixed Effects	✓	✓
Quarter and Year fixed effects	✓	✓
Market Origin fixed effects	✓	✓
Market Destination fixed effects	✓	✓
No. of Obs.	1,281,413	1,281,413
Endogeneity Test. *H_0*: Price and $\ln(S_{jmt\|g})$ are exogenous. Wu-Hausman:		F(2, 1281250)= 36437.2*** (p = 0.0000)

Note: Standard errors are in parentheses.
***p < 0.01; **p < 0.05; *p < 0.10

The positive coefficient estimate on the *Nonstop* variable suggests that direct flights are associated with higher levels of utility compared to connecting flights. Since we only consider nonstop products and products with one intermediate stop, passengers prefer products with nonstop flight itineraries to those with one intermediate stop when traveling from origin to destination. In fact, the ratio of the coefficient estimates on price and *Nonstop* suggest consumers are willing to pay up to $155 extra[16], on average, to obtain a product with a nonstop itinerary in order to avoid products with intermediate stop.

The positive coefficient estimate on *Routing Quality* suggests that passengers prefer the most direct route to the destination. Consumers show preference for products with itinerary flight distances as close as possible to the nonstop flight distance between the origin and destination. So, consumer choice behavior is consistent with the premise that better routing quality is associated with a more passenger-desirable itinerary. In fact, consumers are willing to pay up to $286 extra[17], on average, for each percentage point increase that the nonstop flight distance is of the actual itinerary flight distance.

The negative coefficient estimates on *Arrival Minutes Late* indicate that consumer choice behavior is consistent with our expectations that products with longer arrival delays at the destination airport are less desirable. The ratio of coefficient estimates of "Minutes late at destination" and price in column 2 of Table 4 suggests that consumers are willing to pay $1.56 on average for each additional minute of flight arrival delay to avoid delay. This implies substantial welfare effects knowing that on average a product is purchased by 167 passengers, is 12 minutes late and that our dataset consists of 1,281,413 products. So, extrapolating yields an estimated consumer welfare cost due to arrival minutes late of approximately $4 billion.[18]

This extrapolation is very conservative since it only accounts for delay at the final destination. In reality, costs borne by passengers may include: potential loss of business due to

---

[16] This is obtained by dividing the coefficient estimate on the *Nonstop* dummy variable by the coefficient estimate on *Price* from column 2 of Table 4.

[17] This is obtained by dividing the coefficient estimate on the *Routing Quality* variable by the coefficient estimate on *Price* from column 2 of Table 4.

[18] Welfare costs to consumers = $1.56 × 12 minutes × 1,281,413 products × 167 passengers

late arrival at a meeting; partial loss of social activity (Cook, Tanner, Williams and Meise, 2009); disrupted ground travel plans; forgone pre-paid hotel accommodations; and missed vacation times (Schumer and Maloney, 2008).

Studies that have examined consumers' reactions to product problems (Curren and Folkes, 1987; Folkes, 1984) show that passengers would be less willing to fly an airline again when delays are perceived to be controllable (e.g. caused by poor management) than when they are perceived to be uncontrollable (e.g. caused by bad weather). Also, even when passengers think that a delay may have arisen from an uncontrollable mechanical failure, they still nevertheless believe that the airline could take action to solve the problem (e.g., substitute another plane), and so refuse to fly that airline again (Folkes, Koletsky, and Graham, 1987).

Using the estimated demand model we can compute the own, as well as cross, demand elasticity of arrival delay minutes. To the best of our knowledge this is the first formal attempt in the economics literature to compute estimates of own and cross demand elasticities of any OTP measure in commercial aviation. The own elasticity of arrival delay minutes captures the extent to which consumers are willing to decrease their demand for a given air travel product when its arrival delay minutes increases. The mean own elasticity of arrival delay minutes is -0.256, which implies that for each percentage point increase in arrival delay minutes for a given product, consumers decrease their demand for the said product by 0.256%. This own elasticity estimate also implies that a 10% increase in a product's arrival delay minutes causes consumers to decrease their demand for the product by 2.56%. Since in our sample on average a product has 12 minutes of arrival delay and is purchased by 167 passengers per period, a 10% increase in arrival delay minutes corresponds to an additional delay of 1 minute and 12 seconds, which according to our own elasticity estimate of -0.256 causes 4 (= 0.025 × 167) fewer passengers per period to choose the product.

The cross elasticity of arrival delay minutes captures the extent to which consumers are willing to switch to substitute products when arrival delay increases for a given product. The mean cross elasticity of arrival minutes late is 0.004, which implies that for each percentage point increase in arrival delay minutes for a given product, consumers increase their demand for substitute products by 0.004%. In other words, an airline's worsening OTP will cause some of its customers to switch to competing airlines, ceteris paribus.

6.2 Counterfactual Analysis

We conduct counterfactual experiments to assess the extent to which improvements in arrival delay (i.e. improvement in OTP) influence variable profits of airlines. Furthermore, since variable profit is a function of product markup and demand level as shown in equation (7), we are able to decompose the change in variable profit and examine how these components drive the change in variable profits.

The counterfactual experiments are done by assuming each carrier experiences a reduction in arrival delay minutes (improvement in OTP) in each sample market. Assuming that the previously estimated preference parameters are unchanged, we use the supply-side of the model to solve for new product-level markups and predicted demand levels after each airline experiences: (1) a counterfactual 10% reduction in arrival delay minutes; (2) a counterfactual 25% reduction in arrival delay minutes; (3) a counterfactual 50% reduction in arrival delay minutes; (4) a counterfactual 75% reduction in arrival delay minutes; and (5) a counterfactual 100% reduction in arrival delay minutes. A comparison of the model's predicted product-level markup and demand level before and after counterfactual reductions in arrival delay minutes, reveals the extent to which OTP improvement influences product-level markup, demand level and ultimately variable profits.

6.2.1 Predicted Percent Change in Product-level Markup, Demand Levels and Variable Profits

Table 5 reports the predicted percent changes in markup, demand and variable profit assuming counterfactual reductions in arrival delay minutes. The model predicts that a 10% reduction in arrival delay minutes (improved OTP) results in an increase in product markup by a mean 1.51 percent, an increase in demand levels by a mean 2.39 percent, and an increase in variable profit by a mean 3.95 percent. Results in Table 5 reveal that as the percentage reductions in arrival delay minutes get larger, the percentage increase in product markup remains relatively stable at approximately 1.5 percent, but percentage increases in demand levels and ultimately variable profit become larger. For example, a 50% reduction in arrival delay minutes results in product markup increasing by approximately 1.52 percent, but an increase in demand levels by 12.64 percent, and an increase in variable profit by 14.43 percent. However, a 75% reduction in arrival delay minutes still results in an increase in product markup by 1.53 percent, but more substantial increases in demand levels and variable profits of 19.66 percent and 21.52 percent, respectively.

In summary, the pattern of results in Table 5 suggests that even though increases in markup and demand levels jointly drive increases in variable profits, the variable profit increases resulting from improved OTP are primarily driven by increases in product demand levels. This pattern of results is still evident when broken out by airlines, as revealed in Tables A1 and A2, which are located in the Appendix.[19]

[19] We thank an anonymous referee for this suggestion.

Table 5: Predicted Percent Change in Product Markup, Demand, and Variable Profit Assuming Counterfactual Reductions in Arrival Delay Minutes

Assumed Counterfactual Reduction in Arrival Delay Minutes	Mean Percent Change in **Markup** (Std. error of mean)	Mean Percent Change in **Demand** (Std. error of mean)	Mean Percent Change in **Variable Profit** (Std. error of mean)
10%	**1.51***** (0.000014)	**2.39***** (0.0012)	**3.95***** (0.0013)
25%	**1.51***** (0.000035)	**6.10***** (0.0025)	**7.739***** (0.0034)
50%	**1.52***** (0.000075)	**12.64***** (0.0053)	**14.43***** (0.0073)
75%	**1.53***** (0.00012)	**19.66***** (0.0085)	**21.52***** (0.0118)
100%	**1.53***** (0.00017)	**27.21***** (0.0122)	**29.35***** (0.0171)

Note: Standard errors are in parentheses.
*** Indicates statistical significance at the 1% level.

6.2.2 Recovering an Estimate of the Cost per minute of Delay Improvement

The natural question to raise at this point is: If improvements in OTP results in higher levels of variable profits for airlines, why don't airlines undertake the investments necessary to substantially mitigate airline-related delays? A reasonable answer to this question is that the increase in variable profits induced by OTP improvements may not be sufficiently high to cover additional cost the airline will need to incur to achieve improved OTP. In other words, it is reasonable to assume that airlines are profit maximizing in their decision-making, and that this behavior extends to their decisions on the extent of flight delays under their control to tolerate. As such, airlines incur costs to reduce each minute of delay, and this cost per minute of reduced delay rationalizes the level of delay minutes we observe in the data. A key objective in this subsection is to use our model to recover an estimate of the cost per minute of delay improvement that airlines face.

Let the profit of airline *f* in market *m* be given by:

$$\Pi_{fm} = VP_{fm} - FC_{fm}, \tag{13}$$

where specification of the variable profit function VP_{fm} is given in equation (11), while FC_{fm} represents the sum of recurrent fixed and sunk costs that airline *f* incurs in market *m*. We assume that FC_{fm} can be decomposed into two components, where one component is a function of OTP-related investment activities for reducing arrival delay minutes (improved OTP), while the other component is not a function of such investment activities. Let such decomposition of FC_{fm} be specified as follows:

$$\mathrm{FC}_{fm} = \Gamma_{fm} + \sum_{j \epsilon F_{fm}} \mathrm{g}_{jm}(I_{jm}), \tag{14}$$

where I_{jm} is a measure of the level of OTP-related investment activities attributed to product *j* in market *m*, and $g_{jm}(I_{jm})$ is the function that translates OTP-related investment activities into economic cost (explicit cost plus opportunity cost). We assume that function $g_{jm}(\cdot)$ has the following properties: $\frac{\partial g_{jm}}{\partial I_{jm}} > 0$ and $\frac{\partial^2 g_{jm}}{\partial I_{jm}^2} > 0$. Therefore, $g_{jm}(\cdot)$ is increasing and convex in I_{jm}. Γ_{fm} represents the portion of FC_{fm} that is invariant to OTP-related investment activities.

Let $h_{jm}(I_{jm})$ be a function that characterizes the relationship between OTP-related investment activities and arrival delay minutes, i.e.:

$$L_{jm} = h_{jm}(I_{jm}), \tag{15}$$

where L_{jm} measures arrival delay minutes of product *j* in market *m*. We assume that $\frac{\partial L_{jm}}{\partial I_{jm}} = \frac{\partial h_{jm}}{\partial I_{jm}} < 0$ and $\frac{\partial^2 L_{jm}}{\partial I_{jm}^2} = \frac{\partial^2 h_{jm}}{\partial I_{jm}^2} > 0$, i.e. $h_{jm}(\cdot)$ is decreasing and convex in I_{jm} suggesting that increasing OTP-related investment activities can lower arrival delay minutes, but the marginal impact diminishes at higher levels of OTP-related investment activities.

Substituting the right-hand-sides of equation (14) into equation (13) yields:

$$\Pi_{fm} = VP_{fm}\big(\boldsymbol{L}(\boldsymbol{I})\big) - \Gamma_{fm} - \sum_{j\epsilon F_{fm}} \mathrm{g}_{jm}\big(I_{jm}\big), \qquad (16)$$

where $\boldsymbol{L}$ and $\boldsymbol{I}$ are a J ×1 vectors of arrival delay minutes and OTP-related investment activity levels associated with the J products in market m, respectively. Assume each airline chooses levels of OTP-related investment activities across its menu of products in a market to maximize its profit. A Nash equilibrium in OTP-related investment activities must satisfy the following first-order conditions implied by maximization of equation (16):

$$\frac{\partial VP_{fm}(\boldsymbol{L})}{\partial L_j}\frac{\partial L_j}{\partial I_j} - \frac{\partial g_{jm}}{\partial I_{jm}} = 0 \text{ for all } j = 1, \dots, J \qquad (17)$$

Equation (17) implies that:

$$\frac{\partial VP_{fm}(\boldsymbol{L})}{\partial L_j} = \frac{\partial g_{jm}}{\partial I_{jm}}\Big/\frac{\partial L_j}{\partial I_j} \text{ for all } j = 1, \dots, J \qquad (18)$$

The right-hand-side of equation (18) measures OTP-related marginal investment cost per minute of improvement in arrival delay minutes, while the left-hand-side of equation (18) measures the marginal increase in variable profit generated from improvements/reduction in arrival delay minutes.

Note that we the researchers do not have information on OTP-related investment activities, I_{jm}, and have not assumed any functional forms for $g_{jm}(\cdot)$ and $h_{jm}(\cdot)$, respectively, which imply that we cannot directly compute the right-hand-side of equation (18). Fortunately, we do have a functional form for $VP_{fm}(\boldsymbol{L})$ from our previously specified structural model, and we do observe arrival delay minutes, $\boldsymbol{L}$, for all products. Therefore, we can directly compute

$\frac{\partial VP_{fm}(\boldsymbol{L})}{\partial L_j}$ for all products, which according to equation (18) reveals the value of $\frac{\partial g_{jm}}{\partial I_{jm}} \Big/ \frac{\partial L_j}{\partial I_j}$ in equilibrium. In other words, we can obtain estimates of OTP-related marginal investment costs per minute of improvement in arrival delay minutes by computing $\frac{\partial VP_{fm}(\boldsymbol{L})}{\partial L_j}$ from our structural model. From equation (11) we know that $VP_{fm} = \sum_{j \epsilon F_{fm}} markup_{jm}(\mathbf{L}) \times d_{jm}(\mathbf{L})$, which implies that:

$$\frac{\partial VP_{fm}(\boldsymbol{L})}{\partial L_j} = \sum_{r \epsilon F_f} \frac{\partial d_r(\boldsymbol{L})}{\partial L_j} markup_r(\boldsymbol{L}) + \sum_{r \epsilon F_f} \frac{\partial markup_r(\boldsymbol{L})}{\partial L_j} d_r(\boldsymbol{L}) \qquad (19)$$

Based on equations (18) and (19), we find that the mean OTP-related marginal investment cost per minute of improvement in arrival delay minutes is \$568.13. In other words, the current levels of arrival delay minutes observed in our data are rationalized by a mean OTP-related marginal investment cost per minute of improvement equal to \$568.13. Achieving lower levels of arrival delay minutes would require larger OTP-related investment activities and larger marginal investment costs.

We further use the model to predict OTP-related marginal investment costs per minute of improvement necessary to achieve a ten percent reduction in arrival delay minutes from the current levels of arrival delay minutes observed in the data. As we previously reported, on average, a product has arrival delay of approximately 12 minutes. So, on average, a 10% improvement in OTP corresponds to a 1 minute and 12 seconds reduction in arrival delay. While a mean marginal investment cost per minute of improvement of \$568.13 sustains the current levels of arrival delay minutes observed in the data, a higher mean marginal investment cost per minute of improvement of \$581.68 is required to sustain arrival delay minutes that are 10% less than their current levels. Put another way, a 2.39% increase in OTP-related marginal investment cost per minute of improvement (from $568.13 to $581.68) is necessary to achieve a 10% reduction in arrival delay minutes below their current levels. The intuition is that reductions, i.e. improvements, in arrival delay minutes will require higher levels of OTP-related investment activities, which in turn pushes up the marginal investment cost per minute of improvement.

The model prediction above regarding the increase in OTP-related marginal investment cost required to achieve a reduction in arrival delay minutes can also be interpreted from a policy perspective. The estimates suggest that for airlines to achieve a 10% reduction in their arrival delay minutes, they will need to increase OTP-related investment activities, which will push the marginal investment cost per minute of improvement from $568.13 to $581.68. Knowing such information, policymakers could incentivize airlines to undertake the increase in OTP-related investment activities by offering an OTP-related investment credit to each airline of $13.55 (= $581.68 - $568.13) for each minute of improvement in arrival delay. However, the feasibility of such a policy relies on the ability of policymakers to verify with relative ease each airline's "true" OTP improvement.

7. Conclusion

Researchers have long been interested in explaining why airlines are late. To answer this question, most researchers have resorted to a reduced-form estimation approach where they explain variations in on-time performance (OTP) through a set of explanatory variables. This approach yields a set of parameters that describes the marginal impact of an explanatory variable on on-time arrival performance. In contrast, we use a structural estimation approach.

First, using a demand model, we measure the welfare cost of delay borne by consumers in terms of how much monetary value they are willing to pay to avoid delay. We find that consumers are willing to pay $1.56 per minute late to avoid arrival delay, which after extrapolation amounts to consumer welfare cost of $4 billion. Second, with consumers having a preference for flights that arrive at destination on time, we measure the incentive for airlines to provide on-time arrivals using counterfactual experiments. We find that, a 10% reduction in arrival delay minutes (improved OTP) results in an increase in variable profit by a mean 3.95 percent. Furthermore, we find that it is the increase in demand levels, as compared to increase in markup, that accounts for most of the increase in variable profits.

Given the finding that reductions in arrival delay minutes (improved OTP) yield increases in variable profits, we are able to use the model to recover estimates of the cost per minute of delay improvement that rationalizes the level of delay minutes we observe in the data. The model predicts that a 2.39% increase in OTP-related marginal investment cost per minute of improvement is necessary to achieve a 10% reduction in arrival delay minutes below their current levels.

Stronger conclusions may be drawn from future work about the underlying mechanisms through which product quality may impact product markup, demand and variable profit. OTP is one among other product quality dimensions such as mishandled baggage, oversold flights, in-flight amenities etc. Examining changes in these other quality dimensions along with OTP may provide insights about how airlines engage in overall quality differentiation in a strategic environment where few firms are competing with each other.

Appendix

Table A1: Predicted Percent Change in Product Markup, Demand, and Variable Profit Assuming a **10%** Counterfactual Reduction in Arrival Delay Minutes – By Airline

Code	Airline	Mean Percent Change in **Markup**	Standard Error of mean	Mean Percent Change in **Demand**	Standard Error of mean	Mean Percent Change in **Variable Profit**	Standard Error of mean
AA	American Airlines	1.507***	0.000017	2.718***	0.002	4.273***	0.003
AQ	Aloha Airlines	1.508***	0.000151	1.544***	0.090	3.084***	0.182
AS	Alaska Airlines	1.509***	0.000108	1.933***	0.015	3.509***	0.017
B6	JetBlue Airways	1.511***	0.000264	2.838***	0.010	4.418***	0.013
CO	Continental Air Lines	1.507***	0.000040	2.424***	0.003	3.943***	0.004
DH	Independence Air	1.511***	0.000793	3.030***	0.039	4.652***	0.044
DL	Delta Air Lines	1.509***	0.000063	2.322***	0.003	3.790***	0.003
F9	Frontier Airlines	1.506***	0.000074	2.270***	0.008	3.799***	0.009
FL	AirTran Airways	1.506***	0.000031	2.550***	0.006	4.108***	0.007
HA	Hawaiian Airlines	1.504***	0.000002	0.706***	0.043	2.221***	0.044
HP	America West Airlines	1.505***	0.000013	2.027***	0.005	3.533***	0.007
NW	Northwest Airlines	1.506***	0.000014	2.580***	0.004	4.128***	0.005
OO	SkyWest	1.505***	0.000240	2.609***	0.568	4.153***	0.577
TZ	ATA Airlines	1.506***	0.000073	2.429***	0.020	3.982***	0.024
UA	United Air Lines	1.507***	0.000016	2.603***	0.003	4.189***	0.004
US	US Airways	1.505***	0.000008	2.150***	0.003	3.662***	0.004
VX	Virgin America Inc.	1.507***	0.000179	1.642***	0.038	3.151***	0.045
WN	Southwest Airlines	1.506***	0.000010	1.968***	0.004	3.522***	0.002
XE	ExpressJet Airlines	1.504***	0.000012	2.751***	0.243	4.470***	0.576
YX	Midwest Airlines	1.505***	0.000043	2.723***	0.060	4.268***	0.064

*** Indicates statistical significance at the 1% level.

Table A2: Predicted Percent Change in Product Markup, Demand, and Variable Profit Assuming a **25%** Counterfactual Reduction in Arrival Delay Minutes – By Airline

Code	Airline	Mean Percent Change in **Markup**	Standard Error of mean	Mean Percent Change in **Demand**	Standard Error of mean	Mean Percent Change in **Variable Profit**	Standard Error of mean
AA	American Airlines	1.511***	0.00004	6.956***	0.006	8.588***	0.008
AQ	Aloha Airlines	1.513***	0.00038	3.927***	0.234	5.518***	0.471
AS	Alaska Airlines	1.517***	0.00028	4.923***	0.038	6.613***	0.044
B6	JetBlue Airways	1.521***	0.00068	7.266***	0.026	8.969***	0.035
CO	Continental Air Lines	1.511***	0.00010	6.194***	0.008	7.734***	0.010
DH	Independence Air	1.520***	0.00202	7.776***	0.102	9.590***	0.115
DL	Delta Air Lines	1.515***	0.00016	5.925***	0.006	7.331***	0.007
F9	Frontier Airlines	1.509***	0.00019	5.795***	0.022	7.361***	0.023
FL	AirTran Airways	1.509***	0.00008	6.520***	0.016	8.165***	0.019
HA	Hawaiian Airlines	1.504***	1.73E-06	1.774***	0.109	3.305***	0.111
HP	America West Airlines	1.507***	0.00003	5.158***	0.014	6.668***	0.018
NW	Northwest Airlines	1.508***	0.00004	6.596***	0.009	8.214***	0.012
OO	SkyWest	1.505***	0.00062	6.700***	1.496	8.306***	1.520
TZ	ATA Airlines	1.510***	0.00018	6.205***	0.053	7.836***	0.064
UA	United Air Lines	1.510***	0.00004	6.657***	0.007	8.373***	0.010
US	US Airways	1.507***	0.00002	5.485***	0.009	7.006***	0.011
VX	Virgin America Inc.	1.510***	0.00046	4.166***	0.097	5.682***	0.116
WN	Southwest Airlines	1.510***	0.00002	4.994***	0.006	6.632***	0.006
XE	ExpressJet Airlines	1.505***	0.00003	7.044***	0.644	9.138***	1.528
YX	Midwest Airlines	1.506***	0.00011	6.986***	0.160	8.596***	0.170

*** Indicates statistical significance at the 1% level.

References

Abrahams, M. (1983). A service quality model of air travel demand: An empirical study. *Transportation Research Part A: General, 17*(5), 385-393.

Aguirregabiria, V., & Ho, C. Y. (2012). A dynamic oligopoly game of the US airline industry: Estimation and policy experiments. *Journal of Econometrics*, *168*(1), 156-173.

Anderson, J. E., & Kraus, M. (1981). Quality of service and the demand for air travel. *The Review of Economics and Statistics,* 533-540.

Barnett, A., Shumsky, R., Hansen, M., Odoni, A., & Gosling, G. (2001). Safe at home? An experiment in domestic airline security. *Operations Research*, *49*(2), 181-195.

Berry, S. T. (1990). Airport presence as product differentiation. *The American Economic Review,* 394-399.

Berry, S., Carnall, M., & Spiller, P. T. (2006). Airline hubs: Costs, markups and the implications of customer heterogeneity. *Competition Policy and Antitrust*

Boozer, R. W., Wyld, D. C., & Grant, J. (1990). Using metaphor to create more effective sales messages. *Journal of Services Marketing*, *4*(3), 63-71.

Borenstein, S. (1989). Hubs and high fares: dominance and market power in the US airline industry. *The RAND Journal of Economics*, 344-365.

Chen, Y., & Gayle, P. G. (2013). *Mergers and Product Quality: Evidence from the Airline Industry* (No. 51238). University Library of Munich, Germany.

Cook, A., Tanner, G., Williams, V., & Meise, G. (2009). Dynamic cost indexing–Managing airline delay costs. *Journal of Air Transport Management*, *15*(1), 26-35.

Curren, M. T., & Folkes, V. S. (1987). Attributional influences on consumers' desires to communicate about products. *Psychology & Marketing*, *4*(1), 31-45.

De Vany, A. S. (1975). The effect of price and entry regulation on airline output, capacity and efficiency. *The Bell Journal of Economics,* , 327-345.

Douglas, G. W., & Miller, J. C. (1974). Quality competition, industry equilibrium, and efficiency in the price-constrained airline market. *The American Economic Review,* , 657-669.

Dowling, G. R., & Uncles, M. (1997). Do customer loyalty programs really work? *Research Brief, 1*

Edwards, C. D. (1955). Conglomerate bigness as a source of power. *Business concentration and price policy* (pp. 331-359) Princeton University Press.

Färe, R., Grosskopf, S., & Sickles, R. C. (2007). Productivity? Of US airlines after deregulation. *Journal of Transport Economics and Policy*, 93-112.

Folkes, V. S. (1984). Consumer reactions to product failure: An attributional approach. *Journal of Consumer Research,* , 398-409.

Folkes, V. S., Koletsky, S., & Graham, J. L. (1987). A field study of causal inferences and consumer reaction: The view from the airport. *Journal of Consumer Research,* , 534-539.

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

Gayle, P. G. (2007). Airline code-share alliances and their competitive effects. *JL & Econ.*, *50*, 781.

Gayle, P. G. (2013). On the efficiency of codeshare contracts between airlines: Is double marginalization eliminated? *American Economic Journal: Microeconomics, 5*(4), 244-273.

Gayle, P., & Le, H. (2013). Airline alliances and their effects on costs. *Unpublished Manuscript*

Gayle, P. G., & Thomas, T. (2016). Assessing Firm Behavior in Carve-out Markets: Evidence on the Impact of Carve-out Policy. *Journal of Economic Behavior & Organization*, 128 (2016), 178-194.

Gayle, P., & Yimga, J. (2015). Airline Alliance and Product Quality: The case of the US Domestic Airline Industry. *Unpublished Manuscript*

Ippolito, R. A. (1981). Estimating airline demand with quality of service variables. *Journal of Transport Economics and Policy*, , 7-15.

Ito, H., & Lee, D. (2007). Domestic code sharing, alliances, and airfares in the US airline industry. *Journal of Law and Economics, 50*(2), 355-380.

Mayer, C., & Sinai, T. (2003). Why do airline schedules systematically underestimate travel time? *Unpublished Paper, Wharton School of Business, University of Pennsylvania*

Mazzeo, M. J. (2003). Competition and service quality in the US airline industry. *Review of industrial Organization*, *22*(4), 275-296.

Prince, J. T., & Simon, D. H. (2009). Multimarket contact and service quality: Evidence from on-time performance in the US airline industry. *Academy of Management Journal, 52*(2), 336-354.

Prince, J. T., & Simon, D. H. (2014). Do incumbents improve service quality in response to entry? evidence from airlines' on-time performance. *Management Science,*

Ramdas, K., & Williams, J. (2006). An empirical investigation into the tradeoffs that impact on-time performance in the airline industry. *Washington Post*

Reiss, P. C., & Spiller, P. T. (1989). Competition and entry in small airline markets. *The Journal of Law & Economics, 32*(2), S179-S202.

Rupp, N. G., & Holmes, G. M. (2006). An investigation into the determinants of flight cancellations. *Economica, 73*(292), 749-783.

Rupp, N., Owens, D., & Plumly, L. (2006). Does competition influence airline on-time performance. *Advances in Airline Economics: Competition Policy and Antitrust, 1*, 251-272.

Schumer, C., & Maloney, C. B. (2008). Your flight has been delayed again: Flight delays cost passengers, airlines, and the US economy billions. *The US Senate Joint Economic Committee,*

Shumsky, R. A. (1995). *Dynamic statistical models for the prediction of aircraft take-off times* (Doctoral dissertation, Massachusetts Institute of Technology).

Stock, James H. and Motohiro Yogo (2005), "Testing for Weak Instruments in Linear IV Regression." Ch. 5 in J.H. Stock and D.W.K. Andrews (eds), Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg, Cambridge University Press. Originally published 2001 as NBER Technical Working Paper No. 284; newer version (2004)

US Department of Transportation (2015). *Air Travel Consumer Report*, Bureau of Transportation Statistics. various issues

US Department Of Transportation Office of Aviation Enforcement and Proceedings. (2001). *Air Travel Consumer Report* (January 2001). Washington, DC: US Government Printing Office