Downsizing the jet: A forecast of economic effects of increased automation in aviation

Roman Zakharenko[1 2]
Alexander Luttmann[3]

Abstract

Although commercial aircraft are offered in a wide range of capacities, there has been a clear downward trend in aircraft size over time. We argue that this decline is primarily driven by a past reduction in the minimum flight crew requirement. In line with this argument, we develop a theory of optimal aircraft size, where the cost of the flight crew is the primary factor driving the use of larger aircraft, while passenger utility is primary factor driving the use of smaller aircraft. After fitting our model to U.S. data, we perform a counterfactual experiment where the minimum crew size requirement is further relaxed from two pilots to one, a policy currently being discussed by aviation experts. Implications are derived for the number of aircraft demanded and its size distribution, demand for pilots, passenger traffic, flight frequency, and where new nonstop service may be introduced.

Keywords: Aircraft size, non-scalable cost, aviation market equilibrium, single-pilot aircraft
JEL codes: L62, L93, O33, R41

[1]Corresponding author
[2]International College of Economics and Finance, HSE University, Moscow, Russia. Email: rzakharenko@hse.ru. Web: www.rzak.ru.
[3]MITRE Corporation. Email: aluttmann@mitre.org. Web: https://sites.google.com/a/uci.edu/aluttmann/.
The author's affiliation with the MITRE Corporation is for identification purposes only, and is not intended to convey or imply MITRE's concurrence with, or support for, the positions, opinions, or viewpoints expressed by the author.

January 31, 2022

1. Introduction

Commercial passenger aircraft are supplied in a wide range of capacities, from 40-seat regional jets to the 600+ seat Airbus A380. Over the past three decades, there has been a clear downward trend in aircraft size. For example, the average capacity of aircraft serving the New York (JFK) to Los Angeles (LAX) market has declined from 258 seats in 1990 to 164 seats in 2019, despite a 44% increase in passengers.[4] This downward trend has coincided with the demise of two of the largest aircraft types, the Boeing 747 and the Airbus A380. In particular, the Airbus A380 was in production for only 15 years and is now considered a commercial failure, with production stopped and many units already withdrawn from service.[5]

In this paper, we examine the factors contributing to the decline in average aircraft size. From an economist's perspective, optimal aircraft capacity is increased by non-scalable operating costs, i.e., components of costs that cannot be scaled up or down in proportion with aircraft size. To reduce these costs per passenger, larger aircraft are used. The downside of larger aircraft is a decrease in passenger utility, resulting from (i) less frequent service, (ii) more routes requiring the use of connecting flights, and (iii) a more cramped on-board experience. Since these disutility factors associated with larger aircraft have hardly changed in recent decades, the downward trend in aircraft size is most likely explained by technological changes in the industry that have reduced non-scalable flight costs.

A number of analysts (e.g., see footnote 5) have put forth what we call an "engine" theory of aircraft size. Forty years ago, three or more engines were required to fly further than 60 minutes from the nearest airport[6], which, according to the engine theory, required aircraft to

[4] These numbers are computed from the Bureau of Transportation Statistics' T100 flight segment database for July 1990 and July 2019.

[5] For example, see *"Why the Airbus A380 failed to take off"* by Michael Newell, The New Economy, June 28, 2019.

[6] Extended-range Twin-engine Operational Performance Standards, best known as ETOPS.

be large. Beginning in the late 1980s, these rules were gradually relaxed so that two-engine aircraft could fly between continents. According to the theory, fewer engines allowed aircraft to get smaller.

Using our above terminology, the engine theory views jet engines as non-scalable operating costs (i.e., more engines must be associated with greater capacity). However, the history of aviation has counterexamples to this proposed relationship. For instance, British Aerospace (BAE) 146, a commercially successful regional jet with only 70 seats in its smallest configuration, has four engines. In fact, its configuration is so similar to that of the seven-times larger Boeing-747 "Jumbo jet" that the former was nicknamed the "Jumbolino". This aircraft was an economic success as 387 units were produced from 1978 to 2001 and several major regional airlines operated the aircraft. Moreover, key engine specifications (e.g., weight/thrust, fuel consumption/thrust) of the BAE 146 "Jumbolino" are very similar to those of same-era Boeing 747-300. This counterexample demonstrates that the engine cost *per se* is indeed scalable. Hence, contrary to the engine theory, historical engine regulations did not necessarily increase non-scalable costs, and therefore did not directly constrain aircraft size.

In this paper, we propose an alternative "crew" theory where the primary non-scalable cost in airline operations is that of the flight crew. The data described below demonstrates that the cost of the flight crew is indeed significant and constitutes 10-20% of an airline's total operating cost. Today, almost all aircraft are operated by two crew members, the captain and the first officer, whose skills are close substitutes and who differ primarily in their experience. In the past, long-range aircraft with many engines also required a flight engineer, which inflated non-scalable costs and thus resulted in larger optimal aircraft capacity. Relaxation of engine number requirements coincided with the removal of the flight engineer, which, we argue, is the dominant cause of aircraft size reduction.

Figure 1 provides a visual test of the crew theory using a local discontinuity in crew

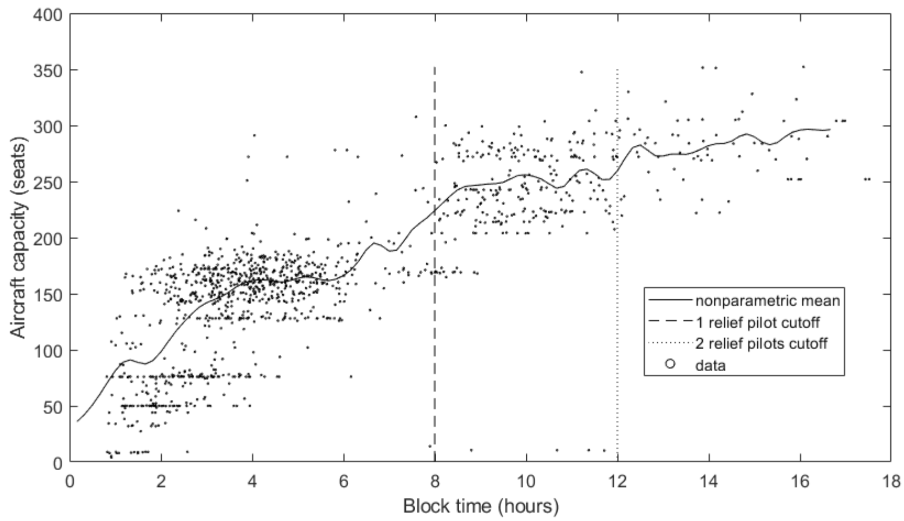

Figure 1: Block time vs. aircraft capacity, international flights in/out USA

size. Regulations limit the maximum work period of pilots to 8 or 9 hours, meaning that flights with ramp-to-ramp "block" time over 8 hours must have an additional relief pilot onboard, and flights with block time over 12 hours must have two relief pilots. Using the T100 flight segment data described in section 5.1, Figure 1 maps aircraft size against the 95th percentile of block time for a given origin-destination pair, which we assume is used by airlines when making staffing decisions. This figure illustrates only international flights departing or arriving into the U.S., because (i) there is an upper bound on duration of domestic flights, and (ii) domestic flights may be correlated with aircraft size by means of serving smaller airports without border service. In line with our theory, a spike in aircraft size around the 8-hour-flight cutoff is clearly visible: almost all flights under 8 hours use aircraft with less 200 seats, while almost all flights over 8 hours use aircraft with capacity over 200. The spike around 12 hours is also visible but less pronounced, which may be explained by the fact that the market for flights over 12 hours is too thin to create a dedicated aircraft type.

The goal of this paper is to develop a structural model of optimal aircraft capacity driven by non-scalable costs. The model is useful to analyze the potential effects of relaxing the current minimum crew size requirement from two pilots to one, a policy that is currently being discussed by aviation experts.[7] In particular, our model allows us to predict the effects of a reduction in crew size on the number of aircraft demanded and its size distribution, demand for pilot labor, total passenger traffic, flight frequency, and identify where new nonstop service may be introduced.

The rest of this paper is organized as follows. Section 2 summarizes previous literature with a particular emphasis on empirical studies of the airline industry. Section 3 presents econometric evidence on scalable and non-scalable airline costs. Section 4 describes our structural theoretical model. Section 5 describes the data used to estimate the parameters of the model. Section 6 details the generalized method of moments (GMM) procedure used to recover model parameters and presents parameter estimates. Section 7 presents results from our counterfactual experiment where the current minimum crew size requirement is relaxed from two pilots to one. Finally, Section 8 provides concluding remarks.

2. Previous Literature

Our paper contributes to four distinct research areas. Foremost, our paper complements studies in other disciplines that examine the feasibility of transitioning from two-pilot to single-pilot commercial aircraft. Second, our paper complements previous work on the existence of economies of density, scale, and scope in transportation and a variety of other industries. Third, our paper enhances the literature on the determinants of optimal aircraft size. Finally, we add to the growing number of empirical studies that employ structural models for policy evaluation.

[7] For example, see *"Some airlines want Boeing's new 797 to fly with just one pilot on board"* by David Reid, CNBC, May 20, 2019.

2.1. Single-pilot aircraft

Single-pilot aircraft are currently prevalent in air taxi, general aviation, and military operations. Given projected growth in air travel that is expected to result in a pilot shortage (ICAO, 2011; GAO, 2014)[8], aircraft manufacturers and government agencies have explored leveraging advances in technology to extend single-pilot operations to commercial aircraft. For example, the Federal Aviation Administration and the National Aeronautics and Space Administration (NASA) have conducted research aimed at developing standards for single-pilot aircraft (Warwick, 2013). In addition, Airbus and Boeing are examining how to redesign cockpits for single-pilot use (Freed and Hepher, 2018; Park, 2017).

If implemented, airlines are expected to follow a NASA-developed approach that would involve a single-pilot in the cockpit and a second ground-based pilot providing support to as many as five flights simultaneously (Bouchard and Baggioni, 2017). This ground-based pilot would have the ability to remotely control an aircraft in the event of an emergency (e.g., the death or incapacitation of the onboard pilot). Several recent studies have focused on the implementation and feasibility of such a system (Bailey et al., 2017; Bilimoria et al., 2014; Graham et al., 2014; Koltz et al., 2015; Lachter et al., 2014; Lim et al., 2017a,b; Liu et al., 2016; Myers and Starr, 2021; Stanton et al., 2014; Vu et al., 2018).

Although the widespread adoption of single-pilot passenger aircraft is likely over a decade away, it is not clear how such an adoption is expected to affect key aspects of airline operations such as flight frequency, network size, and pilot demand. For example, pilots may worry that removing the first officer will reduce the demand for pilots. However, it is also possible that cost savings from removing this crew member are passed through to passengers, stimulating demand to such an extent that aggregate demand for pilots increases. We explore this possibility in our counterfactual experiment (see Section 7).

[8]These forecasts were generated prior to the Covid-19 pandemic.

2.2. Economies of density, scale, and scope

Economies of density, scale, and scope occur in a variety of industries. For example, they occur in the air cargo (Lakew, 2014), container shipping (Cullinane and Khanna, 2000; Talley et al., 1986), electric power generation (Christensen and Greene, 1976; Roberts, 1986), manufacturing (Bain, 1954), public transportation (Farsi et al., 2007; Basso and Jara-Díaz, 2006; Savage, 1997; Gschwender et al., 2016), rail freight (Bitzan and Keeler, 2007; Braeutigam et al., 1984; Harris, 1977), service (Morikawa, 2011), and water supply (Kim and Clark, 1988; Nauges and Van den Berg, 2008) industries.

For the airline industry, Caves et al. (1984), Brueckner and Spiller (1994), and Jara-Díaz et al. (2013) provide evidence of substantial economies of traffic density. Following deregulation, the growth of hub-and-spoke networks allowed carriers to reduce their costs by funneling passengers through a hub airport. Our theoretical model of section 4 assumes the time spent at a connecting airport reduces passenger utility, meaning that passengers prefer to make connections at larger hubs with more frequent service.

2.3. Economics of aircraft size

To accommodate an increase in travel demand, airlines may either increase flight frequency, aircraft size, or both. However, frequency cannot always be increased due to runway capacity constraints.[9] Nevertheless, altering flight frequency and aircraft size have potential adverse affects, such as increasing congestion and exacerbating environmental externalities. For example, Givoni and Rietveld (2010) find that increasing aircraft size and decreasing flight frequency (i.e., to offer similar seating capacity) decreases the overall climate change impact resulting from aircraft emissions at the expense of increasing local pollution.

Considering that travel demand is expected to increase over time, several studies have

[9] When runway capacity is expanded, Takebayashi (2011) finds that airlines respond by using smaller aircraft at higher frequencies. In addition, Berster et al. (2015) finds that larger aircraft typically service congested airports, although this relationship does not hold at all airports.

focused on the factors that affect aircraft size and flight frequency. In particular, Pai (2010) finds that flight frequency and aircraft size increase with population, income, and runway length while an increase in distance between route endpoints leads to lower frequency with the use of larger planes. In addition, an increase in the proportion of managerial workers in the labor force or the proportion of population below the age of 25 results in greater frequency with the use of smaller planes.

In general, larger aircraft are more cost efficient (Wei and Hansen, 2003; Ryerson and Hansen, 2013). For example, Wei and Hansen (2003) find that for any given flight length there is an optimal aircraft size that increases with flight length. However, when economies of aircraft size exist, Zhang (2014) finds that airlines prefer to increase flight frequency but not aircraft size to accommodate traffic growth. Similarly, Wei and Hansen (2005) find that airlines obtain higher returns in market share from increasing frequency than from increasing aircraft size. In other words, airlines have an economic incentive to use aircraft smaller than the least-cost aircraft, since for the same capacity provided in the market, an increase in frequency attracts more passengers.

2.4. Structural Models of Supply and Demand

There are several examples of structural models being employed for policy evaluation in the airline industry. For example, Berry and Jia (2010) develop a structural model to evaluate how demand and supply shocks resulting from the dot-com bubble and the 9/11 terrorist attack affected air-travel demand. Other studies have employed structural models to perform merger analysis (Peters, 2006; Chen and Gayle, 2019), estimate the consumer welfare gain from Open Skies agreements (Winston and Yan, 2015), evaluate the welfare consequences of codesharing (Armantier and Richard, 2008; Gayle, 2007, 2008, 2013; Gayle and Brown, 2014; Gayle and Xie, 2018, 2019; Shen, 2017), value the non-price characteristics of airline networks (Gayle and Yimga, 2018; Israel et al., 2013), and determine if privatizing the San Francisco Bay area airports would improve passenger welfare (Yan and Winston, 2014). In Section 7, we use our structural model to predict how a reduction in the minimum crew size requirement is expected to affect aircraft size, pilot demand, passenger traffic, flight frequency, and the number of segments serviced.

3. Non-scalable aircraft cost: econometric evidence

Prior to developing a structural model of aircraft choice by airlines, we separate airline scalable and non-scalable costs econometrically, to verify that the latter costs are primarily related to the flight crew, and to assess their magnitude. The data we use are Air Carrier Financial Reports, Schedule P-5.2 "Aircraft operating expenses"("P-5.2 data" henceforth), which contains detailed information on operating costs that can be linked to a specific aircraft type, and which is reported by every major U.S. airline. The costs included into these data constitute 44% of all costs by reporting airlines. Each observation is a unique combination of airline, aircraft type, quarter of the year, and operating region (domestic/international). We use only observations on purely passenger aircraft in 2019, with positive flight costs and flight times, totaling 771 observations from 27 U.S. airlines across 42 different aircraft types.[10] The available variables include, besides various types of costs, the total hours of operation, which allows us to calculate costs per hour.

3.1. Flight crew cost

Our first exercise is to separate flight crew costs into scalable and non-scalable. We define flight crew cost as the sum of the following P-5.2 cost variables (all in section *"Flying Operations"*, i.e., excluding personnel unrelated to operation of aircraft): *Pilots and Copilots, Trainees and Instructors, Personnel Expenses, Employee Benefits and Pensions, Payroll*

[10] Reporting airlines include most regional carriers and all full-service (e.g., Alaska, American, Delta, Hawaiian, United) and low-cost carriers (e.g., Allegiant, Frontier, JetBlue, Southwest, Spirit, Sun Country).

Taxes. This cost constitutes 27% of all P-5.2 costs, or 12% of total airline operating costs. For each aircraft type a, airline k, and quarter-region combination t, we define variable $crcph_{akt}$ as flight crew cost per hour of aircraft operation, averaging $1034/h.

While all 42 aircraft types in the data are operated by two crew members, the number of crew members on board may vary due to relief pilot requirements for long-haul flights. To account for this, we introduce the number-of-crews cost factor ncr, as follows. For flights under 8 hours, we normalize $ncr = 1$. Flights between 8 and 12 hours should have two captains and one first officer (FO) on board; assuming the cost of a captain is double that of a FO,[11] we set $ncr = \frac{5}{3}$. For flights over 12 hours, two full crews are required, hence $ncr = 2$.

Because durations of individual flights are not reported in the P-5.2 data, we match this data to Schedule T100 of Air Carrier Traffic Statistics ("T100 data" henceforth), the detailed source of segment-level U.S. flight data. Each observation is a unique combination of airline, aircraft type, origin airport, destination airport, and month of the year; the data on the number of departures, available and used capacity, and flight-time is reported. The latter matches almost perfectly to that in the P-5.2 data, which allows us to infer the fraction of 8+ and 12+ hour flights in the total flight-time, and hence infer the average value of ncr_{akt} for every P-5.2 observation.

Empirically, flight crew on larger aircraft are paid more, i.e., part of the crew cost is scalable. This phenomenon partly offsets the incentives to use larger aircraft.[12] To account for this fact, we separate the cost of *standard* crew per hour into scalable and non-scalable, using the following regression:[13]

$$crcph_{akt}/ncr_{akt} = \gamma_f^{cr} + \gamma_S^{cr} S_a + \nu_k^{cr} + \eta_{akt}^{cr}. \quad (1)$$

[11]For example, see Table 2 in Cook and Tanner (2008) or https://pea.com/airline-pilot-salary/ for recent salary comparisons. Accessed on Sept.24, 2021.

[12]For example, see Wei and Hansen (2003) for additional discussion.

[13]Notation used in multiple sections is cataloged in Appendix A.

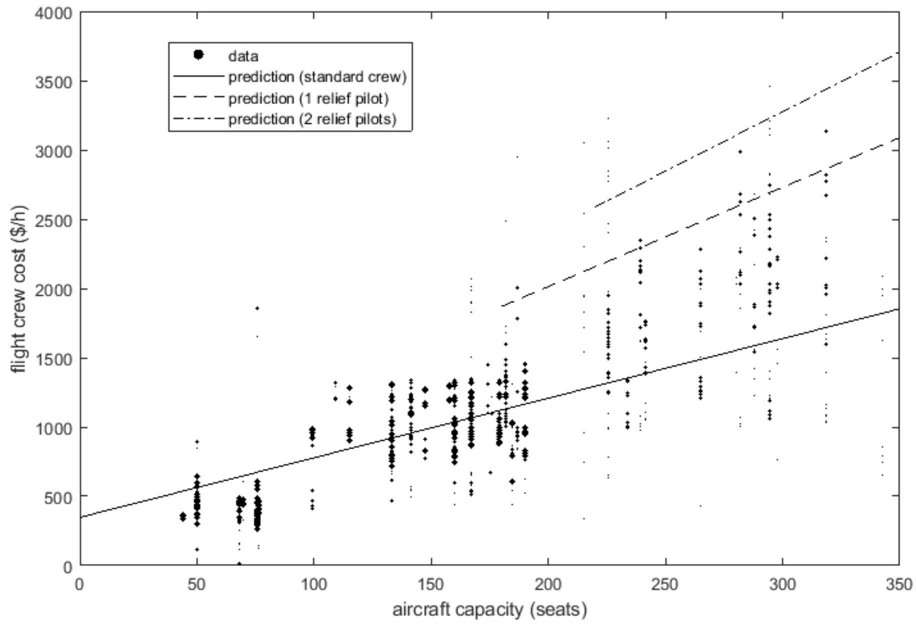

Figure 2: Aircraft capacity vs. flight crew cost per hour. Dot size indicates observation weight (number of departures).

Here, the capacity S_a of aircraft type a is inferred from T100 data; ν_k^{cr} is the airline random effect, while η_{akt}^{cr} is the error term. Quarter effects are not included because crew costs do not vary much during the course of the year. The number of departures (inferred from the T100 data) are used as observation weights.

The resulting estimates (standard errors) are $\gamma_f^{cr} = 348(27.0)$, $\gamma_S^{cr} = 4.30(0.194)$. Figure 2 visualizes flight crew cost *crcph*, and provides linear predictions for flights with various relief crew sizes. The above estimates imply that about one-third of flight crew cost is non-scalable.[14]

3.2. Other costs

We have also analyzed whether aircraft operating costs other than that of the flight crew have a non-scalable component. Using the P-5.2 data, we defined non-crew cost per

[14] $\$348/\$1034 = 0.3366$.

block hour, $ncrph_{akt}$, as total aircraft operating cost per hour minus $crcph_{akt}$, averaging \$2834/h. Using $ncrph_{akt}$ as the dependent variable in a regression similar to (1), we obtain the following estimates (standard errors): $\gamma_f^{other} = 54.4(75.5)$, $\gamma_S^{other} = 17.1(0.540)$, which implies the non-scalable component of non-crew cost constitutes only 1.92% of all *ncrph* and is not statistically different from zero.[15]

Besides aircraft operating costs, airline accounting includes other costs that constitute about 56% of the total. These costs are detailed in Schedule P-7 of the same database. Unfortunately, these costs are not linked to a specific aircraft type, rendering impossible regression analysis similar to that performed in section 3.1. However, most of these costs are theoretically unrelated to aircraft size. For example, in-flight passenger service, traffic service (i.e., that of passengers and their baggage at the airport), reservation and sales, advertising, administrative, and other transport-related costs (e.g., outsourcing fees to other airlines).

One potential candidate non-scalable cost are aircraft servicing fees levied by airport authorities. However, in the U.S., airport fees are proportional to aircraft weight and are thus fully scalable (Givoni and Rietveld, 2009).[16] Another theoretical candidate are air navigation fees, as the cost of navigating aircraft is likely unrelated to its size. Again, in the U.S., these fees are not levied directly but rather funded by the Federal Aviation Administration via fuel taxes and are thus fully scalable.

4. The model

This section introduces a model of passenger flight demand and supply, with a particular emphasis on factors affecting equilibrium aircraft size. On the supply side, the model features

[15] \$54.4/\$2834 = 0.0192.

[16] Some studies have focused on how takeoff and landing fees affect an airline's choice of aircraft size. For example, in a game-theoretic model applied to duopoly markets, Wei (2006) finds that higher landing fees force airlines to use larger aircraft at lower frequencies. However, because landing fees in the U.S. are often weight-based, these fees don't always provide airlines with an incentive to use larger aircraft at lower frequencies.

an explicit distinction of scalable and non-scalable costs. On the demand side, besides the usual effects of ticket price, we pay attention to the choice of flight routes (i.e., nonstop vs. connecting) and to the demand response to departure intervals.

Consider a model set up in continuous infinite time. The space consists of an exogenous set of *cities*, each having one or more *airports*. The airports are connected by an endogenous network $\mathcal{J}$ of nonstop flight *segments*. We consider only time-invariant steady states; in particular, the endogenous mean interval between departures on segment $j \in \mathcal{J}$ is denoted z_j. Each segment is also characterized by exogenous flight distance d_j and flight ramp-to-ramp "block" time h_j.

4.1. Travel demand

4.1.1. Overview

The model of flight demand is similar to the nested logit in Berry and Jia (2010), but was developed independently to target the research question of this paper and has a number of differences.[17] Flight demand is defined at the level of *markets*, i.e., bidirectional pairs of origin and destination cities, which may or may not be connected by a nonstop flight segment. For every market m, there is a continuous flow A_m of potential *passengers* traveling in each direction. Every passenger faces a menu $\mathcal{R}_m$ of *routes*, i.e., combinations of segments that connect an origin airport to a destination airport. For mathematical tractability, we consider only routes with at most two segments, i.e., either direct flights or those with one connection. By d_m, we denote the direct distance between endpoints of market m, labeled *market distance* henceforth.

Denote by $\mathcal{R}_1$ ($\mathcal{R}_2$) the set of all nonstop (2-segment) routes. For $r \in \mathcal{R}_2$, the flight interval z_r is assumed to equal the maximal interval among the two segments: $z_r = \max\{z_{j_1}, z_{j_2}\}$,

[17] Several other studies of the airline industry model consumer demand using a nested logit. For example, see Peters (2006); Chen and Gayle (2019); Gayle and Brown (2014); Gayle and Wu (2014); Gayle and Thomas (2016); Gayle and Yimga (2018); Shen (2017).

$\{j_1, j_2\} = r$. For example, if the first segment has shorter flight intervals $z_{j_1} < z_{j_2}$, the passenger will consider only departures that connect well with those on the second segment, effectively increasing the flight interval on the first segment to z_r.

To achieve mathematical tractability, previous theoretical models of passenger schedule delay (e.g., Brueckner (2004)) had to assume that passengers base their decision on expected rather than actual schedule delay, as if they had to purchase tickets before knowing their ideal departure time (IDT henceforth). Our model improves on previous studies by allowing each passenger i in market m to make their choice with full knowledge of their IDT. For each possible route $r \in \mathcal{R}_m$, such passenger considers two *flights*: the latest flight k *before* and the earliest k' *after* the IDT. Denote the schedule delay (i.e., the absolute difference between IDT and actual departure time, SD henceforth) of the earlier flight by t_{ik}; then the SD of the later flight is equal to $t_{ik'} = z_r - t_{ik}$. From the perspective of airlines, given constant flow of passengers, the distribution of t_{ik} across passengers is uniform on $[0, z_r]$ for any given route. We also make the following assumption:

Assumption 1. *The scheduling process is independent across routes, so t_{ik} for any departure k on route r is independent from $t_{ik'}$ for any departure k' on route $r' \neq r$, for any passenger i and any pair of routes $r, r' \in \mathcal{R}_m$.*

4.1.2. Utility

For a typical passenger i in market m, the utility of traveling from the origin to the destination by using flight k on route $r \in \mathcal{R}_m$ is given by

$$u_{ik} = -\lambda \frac{p_k + \alpha t_{ik} + \beta h_r}{\mu_m} + \lambda \xi_r + \epsilon_{ir}. \tag{2}$$

Here p_k is the monetary cost of travel using flight k,[18] μ_m is a market-specific *willingness-to-pay* parameter of utility, and λ is introduced below. α is the dollar-valued disutility of SD: with higher α, passengers become more sensitive to flight intervals. h_r is time loss from having to make a connection (if any), while β is the disutility of such time loss; higher β makes nonstop routes and shorter connections more attractive. ξ_r is a route-specific preference shock, common to all passengers; it picks up all unaccounted factors of route choice, such as heterogeneity in service quality across different routes, or variation in actual connection times for 2-segment routes. ϵ_{ir} is an idiosyncratic route preference shock of passenger i.

Besides air travel, passenger i has an outside alternative, which may include ground transportation or no travel at all. Their utility from this option is $u_{i0} = \ln T_0 + \epsilon_{i0}$, for some $T_0 > 0$.

The idiosyncratic utility shocks are assumed to have the nested logit structure, such that all flight options are in one nest and the outside opportunity in the other. The cumulative distribution of utility shocks takes the form

$$\exp\left(\left(-\sum_{r\in\mathcal{R}_m} e^{-\frac{\epsilon_{ir}}{\lambda}}\right)^{\lambda} - e^{-\epsilon_{i0}}\right),$$

where $\lambda \in (0, 1]$ is the nested logit parameter, inversely related to the elasticity of substitution between flight routes.

[18] Unlike Berry and Jia (2010), we do not consider price heterogeneity on a given route and focus on average price only. Fares may vary with the class of service, season, and with the timing of ticket purchase. In addition, some passengers are willing to pay a premium to travel on a given airline due to brand loyalty (e.g., frequent flyer program benefits (De Jong et al., 2019)). We believe all of these dimensions are irrelevant for the choice of aircraft size by airlines. In particular, cabin class configuration can be adjusted without changing the aircraft type.

4.1.3. Optimal choice of route and flight

First, observe from (2) that the utility-maximizing flight k on a given route r must minimize $p_k + \alpha t_{ik}$. As elaborated in section 4.1.1, each passenger i considers two flights: k that precedes i's ideal departure time, and the next flight k'. The utility-maximizing choice between these two is determined by cutoff

$$\bar{t}_r(p_k, p_{k'}) \equiv \frac{z_r}{2} + \frac{p_{k'} - p_k}{2\alpha}, \tag{3}$$

such that i prefers flight k iff $t_{ik} \le \bar{t}_r(p_k, p_{k'})$.

Consider a flight k on some route r with off-equilibrium price p_k, which competes with earlier and later flights on the same route; assume the price for both earlier and later flights is at some equilibrium level p_r. Flight k will be considered only by passengers with ideal departure time within $[-z_r + \bar{t}_r(p_r, p_k), \bar{t}_r(p_k, p_r)] = [-\bar{t}_r(p_k, p_r), \bar{t}_r(p_k, p_r)]$ of departure k. Also note that in equilibrium ($p_k = p_r$) the cutoff SD is $\bar{t}_r(p_r, p_r) = \frac{z_r}{2}$. That is, all passengers on route r simply choose the nearest departure. Denote by $t_{ir} \in [0, \frac{z_r}{2})$ the SD of passenger i under equilibrium prices (i.e., when choosing the nearest departure on route r).

The fact that flight k on some route r is better for passenger i than other flights on the same route does not guarantee the choice of k, because i also considers flights on other routes. Consider a passenger i who compares route r to all alternative routes $r' \in \mathcal{R}_m \backslash r$. Suppose that on any given alternative route r', the ticket cost $p_{r'}$ is at equilibrium, i.e., the same for all flights, hence the utility-maximizing SD $t_{ir'}$ on route r' is distributed uniformly on $\left[0, \frac{z_{r'}}{2}\right]$. Then, from the nested logit properties[19] of the model described in section 4.1.2, the probability of passenger i choosing flight k on route r is given by

$$\Pr_{ik}(p_k, \mathbf{p}_m, \mathbf{t}_i) = \frac{U_{ir}(p_k, p_r, t_{ik}, h_r)}{T_{im}(p_k, \mathbf{p}_m, \mathbf{t}_i)} \frac{T_{im}(\cdot)^{\lambda}}{T_0 + T_{im}(\cdot)^{\lambda}}, \tag{4}$$

[19]For example, as described in section 4.2 of Train (2009).

where

$$U_{ir}(p_k, p_r, t_{ik}, h_r) = \begin{cases} \exp\left(-\frac{p_k+\alpha t_{ik}+\beta h_r}{\mu_m} + \xi_r\right), & t_{ik} < \bar{t}_r(p_k, p_r) \\ 0, & t_{ik} \geq \bar{t}_r(p_k, p_r) \end{cases}; \tag{5}$$

$$T_{im}(p_k, \mathbf{p}_m, \mathbf{t}_i) = U_{ir}(p_k, p_r, t_{ik}, h_r) + \sum_{r' \in \mathrm{R}_m \backslash r} U_{ir'}(p_{r'}, p_{r'}, t_{ir'}, h_{r'}); \tag{6}$$

$\mathbf{p}_m \equiv \{p_{r'}\}$ is the vector of equilibrium prices; $\mathbf{t}_i = \{t_{ik}, t_{ir'}\}$ are schedule delays for flights preferred by i, $\forall r' \in \mathcal{R}_m$.

In (4), the first ratio is the probability that passenger i chooses route r and departure k, conditional on the choice of travel by air. The second ratio is the probability that such a passenger chooses air travel.

The fact that passengers taking the outside opportunity are unobserved allows us to make the following simplifying assumption:

Assumption 2. *The utility of flight is low enough that almost all passengers in each market m end up choosing the outside opportunity: $T_{im}(\cdot)^\lambda << T_0, \forall m$.*

Given this assumption, the probability of travel by air (second ratio in (4)) is approximated by $\frac{T_m(\cdot)^\lambda}{T_0}$, which allows us to simplify (4) to

$$\Pr_{ik}(p_k, \mathbf{p}_m, \mathbf{t}_i) = \frac{U_{ir}(p_k, p_r, t_{ik}, h_r)}{T_{im}(p_k, \mathbf{p}_m, \mathbf{t}_i)^{1-\lambda} T_0}. \tag{7}$$

Calculation of the number of passengers boarding flight k involves finding the expectation of (7) over $\mathbf{t}_i$. Under assumption 1, the general solution exists but is cumbersome and includes infinite rows; instead we use a much simpler approximate solution. The approximation is as follows: we replace $E_{\mathbf{t}_i} \frac{U_{ir}(\cdot)}{T_{im}(\cdot)^{1-\lambda} T_0}$ in (7) by $\frac{E_{\mathbf{t}_i}(U_{ir}(\cdot))}{(E_{\mathbf{t}_i} T_{im}(\cdot))^{1-\lambda} T_0}$. This approximation is accurate if the denominator of (7) varies little with $\mathbf{t}_t$; we believe this is the case for the following reasons:

(i) Routes taken by few passengers constitute a small fraction of $T_{im}(\cdot)$ and thus contribute little to its variance.

(ii) Routes r' taken by many passengers tend to have smaller intervals $z_{r'}$, meaning less variance in $t_{ir'}$ and in corresponding elements of $T_{im}(\cdot)$.

(iii) The fact that $\lambda > 0$ also reduces the variance of the denominator; $\lambda = 1$ results in a constant denominator.

The density of $t_{ir'}$ is $\frac{2}{z_{r'}}$, where coefficient 2 is due to the fact that the SD can be both backward and forward in time. Then, the expectation $E_{t_{ik}} U_{ir}(p_k, p_r, t_{ik}, h_r)$ is given by

$$\bar{U}_r(p_k, p_r, z_r, h_r) = \frac{2\mu_m}{\alpha z_r} \exp\left(-\frac{p_k + \beta h_r}{\mu_m} + \xi_r\right)\left(1 - \exp\left(-\frac{\alpha \bar{t}_r(p_k, p_r)}{\mu_m}\right)\right), \tag{8}$$

while

$$\bar{T}_m(p_k, \mathbf{p}_m, \mathbf{z}) = E_{\mathbf{t}_i} T_{im}(p_k, \mathbf{p}_m, \mathbf{t}_i) = \bar{U}_r(p_k, p_r, z_r, h_r) + \sum_{r' \in \mathrm{R}_m \setminus r} \bar{U}_{r'}(p_{r'}, p_{r'}, z_{r'}, h_{r'}),$$

and $\mathbf{z}$ is the vector of all flight intervals.

For a nonstop route $r \in \mathcal{R}_1$, denote by $\bar{D}_r(p_k, \mathbf{p}_m, \mathbf{z})$ the demand for off-equilibrium flight k, i.e., the number of passengers expected to take route r using flight k. It is the product of three multipliers: (i) the total flow of passengers considering travel A_m, (ii) the probability of choosing route r, averaged across schedule delays, $E_{\mathbf{t}_i} \Pr_{ik}(\cdot)$, and (iii) the flight interval z_r. Given above discussion, flight demand is approximated by

$$\bar{D}_r(p_k, \mathbf{p}_m, \mathbf{z}) = A_m z_r \frac{\bar{U}_r(p_k, p_r, z_r, h_r)}{\bar{T}_m(p_k, \mathbf{p}_m, \mathbf{z})^{1-\lambda} T_0}. \tag{9}$$

In equilibrium $p_k = p_r$, the average flow of passengers taking route r is given by $pax_r = \frac{\bar{D}_r(\cdot)}{z_r}$. As flight interval z_r rises from zero to infinity, this quantity decreases from some fraction of A_m to zero. In words, flights departing every minute maximize the flow of passengers taking the route, but still do not guarantee the 100% market share, due to other factors of choice such as price and taste shocks. As flight intervals increase, passengers switch to other routes or to the outside option, hence traffic on this route decreases.

At the same time, the equilibrium flight demand $\bar{D}_r(p_r, \mathbf{p}_m, \mathbf{z})$ increases from zero to a finite maximum as z_r rises from zero to infinity. As $z_r = \infty$ corresponds to a stand-alone flight, the upper bound on $\bar{D}_r$ corresponds to the number of passengers using such flight. With more frequent flights (lower z_r), the number of passengers per flight $\bar{D}_r$ is smaller due to rising competition between flights on route r.

For 2-segment routes $r = \{j_1, j_2\} \in \mathcal{R}_2$, assuming $z_{j_1} \geq z_{j_2}$, the number of passengers expected to take the first (less frequent) segment, per departure, is also given by (9); this quantity is labeled as segment-j_1 flight demand from route r. For the more frequent segment j_2, the flight demand is $\bar{D}_r \frac{z_{j_2}}{z_{j_1}}$. Under equilibrium prices $p_k = p_r$, we can define average route traffic per minute as (cf.(9)) $pax_r = A_m \frac{\bar{U}_r(\cdot)}{\bar{T}_m(\cdot)^{1-\lambda} T_0}$; then the segment-$j$ flight demand by passengers from route r is $pax_r z_j$.

4.2. Flight supply and equilibrium

4.2.1. Aircraft size and load factors

Flights are provided by competing airlines that can choose capacity of their aircraft. Assume for simplicity that a continuum of capacities S is available. Empirically, aircraft capacity for a given flight segment is slightly higher than the average number of passengers D traveling on that segment, causing the typical *load factor* $L \equiv \frac{D}{S}$ to be less than unity. Two effects contribute to lower load factors: (i) flight demand imbalances (i.e., the number of passengers traveling from A to B may be higher than that from B to A, causing seats to remain vacant on the return flight) and (ii) flight demand stochasticity (arrival of passengers is a random process that varies from one day to another, causing variation in the number of occupied seats).

Theoretically, the first effect reduces load factors regardless of aircraft size, while the second effect is primarily relevant for small aircraft. For large aircraft, stochasticity is reduced by the law of large numbers, allowing airlines to choose capacity closer to average demand and thus make the load factor closer to unity. In the 2019 T100 flight segment data, we indeed observe that very small aircraft with 15 or fewer seats have an average load factor of $L = 0.55$ on commercial flights. For aircraft with more than 40 seats that serve over 99% of commercial passengers, average load factors hardly vary with aircraft size, ranging from 79% for 40-60 seat aircraft to 85% for aircraft with over 250 seats. This observation allows us to make a simplifying assumption that the empirically observed segment-level load factor $L_j = \frac{D_j}{S_j}$ is exogenous and does not change in counterfactual experiments with changing aircraft size. For segments currently not served, an average value of L can be assumed.

4.2.2. Flight costs

Airline costs are divided into scalable and non-scalable. The latter are defined on the segment level, as f_j per minute of flight. Scalable costs may contain multiple elements, some of which are attributable to a particular flight segment (e.g., fuel cost, in-flight service, etc.) while others are not (e.g., advertising, customer service). We denote by c_r the scalable cost per passenger on route r.

4.2.3. Optimal pricing

Calculation of profit-maximizing prices is complicated by the fact that most routes are two-segment, meaning that they may be provided by two different airlines. We assume that prices are set by such airlines cooperatively, so that the joint profit is maximized.

Define the *gross profit* for a given route $r \in \mathcal{R}_m$ the revenue minus scalable cost, per departure k on the less frequent segment. Gross profit is then given by

$$\phi_r(p_k, \mathbf{p}_m, \mathbf{z}) = (p_k - c_r)\bar{D}_r(p_k, \mathbf{p}_m, \mathbf{z}). \tag{10}$$

The term in parentheses is the *price markup* (i.e., revenue minus scalable cost). The profit-maximizing price markup is as follows:

$$p_k - c_r = -\frac{\bar{D}_r(\cdot)}{\frac{\mathrm{d}\bar{D}_r(\cdot)}{\mathrm{d}p_k}} = -\frac{1}{\frac{\mathrm{d}\ln \bar{D}_r(\cdot)}{\mathrm{d}p_k}}. \tag{11}$$

In equilibrium, where all flights on route r cost the same ($p_k = p_r$), the profit-maximizing markup takes the form (cf.(3,8,9))

$$mkup_r(\mathbf{p}_m, \mathbf{z}) = p_r - c_r = \mu_m \frac{1 - \exp\left(-\frac{\alpha z_r}{2\mu_m}\right)}{1 - \frac{1}{2}\exp\left(-\frac{\alpha z_r}{2\mu_m}\right)} \frac{1}{1 - (1-\lambda)\frac{\bar{U}_r(p_r,p_r,z_r,h_r)}{\bar{T}_m(p_r,\mathbf{p}_m,\mathbf{z})}}, \tag{12}$$

which defines the optimal price p_r. Here, the first ratio is the effect of route-r flight interval on markup: as z_r rises from zero to infinity, it rises from zero to unity. The second ratio is the effect of route-r market share, $\frac{\bar{U}_r(\cdot)}{\bar{T}_m(\cdot)}$, on markup; it rises from unity when the route has zero market share, to $\frac{1}{\lambda}$ for the monopoly route.

4.2.4. Free entry

For 2-segment routes $r = \{j_1, j_2\}$, the markup (11) has to be divided between the two segments; we assume the share of segment j_1, $B_{j_1 r}$, is proportional to flight time: $B_{j_1 r} = \frac{h_{j_1}}{h_{j_1} + h_{j_2}}$. For nonstop routes $r = \{j\}$, naturally $B_{jr} = 1$. Denote $\mathcal{R}_j$ the set of all routes that use segment j; the optimal airline gross profit on segment j, per flight, is given by

$$\phi_j(\mathbf{p}, \mathbf{z}) = \sum_{r \in \mathcal{R}_j} \phi_r(p_r, \mathbf{p}_m, \mathbf{z}) B_{jr} \frac{z_j}{z_r}, \tag{13}$$

where $\mathbf{p}$ is the vector of all prices. The flight interval ratio for 2-segment routes is explained at bottom of section 4.1.3.

We assume free entry of airlines into every particular segment. This imposes a constraint that gross profit cannot exceed the non-scalable cost:

$$\phi_j(\mathbf{p}, \mathbf{z}) \leq f_j h_j, \tag{14}$$

with equality for segments served in equilibrium (i.e. those with $z_j < \infty$). This defines the equilibrium flight intervals $\mathbf{z}$.

4.3. Connection time

The connection time loss h_r (cf.(2)) is obviously zero for nonstop routes, and positive for 2-segment routes. While calculation of h_r is theoretically possible from available schedule data, it is complicated because (i) our definition of routes pools all airlines and both directions of flight, making the exact value of h_r unclear; (ii) our model is designed for analysis of a counterfactual equilibrium, in which precise departure schedules are not computed. For this reason, we replace the actual h_r by its expectation, as follows. First, we assume that the maximum connection time $\bar{h}_r$ is equal to the minimal flight interval of the two routes: $\bar{h}_r = \min\{z_{j_1}, z_{j_2}\}, \{j_1, j_2\} = r$. Even if the second segment has greater interval than the first, $z_{j_1} < z_{j_2}$, strategic choice of the first departure will ensure that the connection time does not exceed z_{j_1}. Second, by assumption 1, departures on the two segments $\{j_1, j_2\}$ are independent from each other, hence the connection time h_r is distributed uniformly on $[0, \bar{h}_r]$. But then, we can calculate the expectation of (8) over h_r for 2-segment routes, at optimal price p_r, as follows:

$$U_r(p_r, \mathbf{z}) = \frac{2\mu_m^2}{\alpha\beta z_r \bar{h}_r} \exp\left(-\frac{p_r}{\mu_m} + \xi_r\right)\left(1 - \exp\left(-\frac{\alpha z_r}{2\mu_m}\right)\right)\left(1 - \exp\left(-\frac{\beta\bar{h}_r}{\mu_m}\right)\right), \tag{15}$$

where $z_r = \max\{z_{j_1}, z_{j_2}\}$ and $\bar{h}_r = \min\{z_{j_1}, z_{j_2}\}$. For non-stop routes, we can redefine U_r from (8) as $U_r(p_r, \mathbf{z}) = \bar{U}_r(p_r, p_r, z_r, 0)$, which also equals the limit of the right-hand side of (15) when $\bar{h}_r \to 0$.

Note that the connection time factor of utility (2) creates airport agglomeration effects: a higher β increases comparative advantage of 2-segment routes with a connection at larger airports with shorter departure intervals (lower $\bar{h}_r$), relative to other 2-segment routes.

We also define the equilibrium flight demand as (cf.(9)):

$$D_r(\mathbf{p}_m, \mathbf{z}) = \bar{D}_r(p_r, \mathbf{p}_m, \mathbf{z}). \tag{16}$$

4.4. An illustration of equilibrium

To illustrate an equilibrium in this model, consider the most simple setting with only two airports and a single market/segment/route connecting them. Figure 3 illustrates flight profits $\pi = (p - c)D(p, z) - fh$ as functions of passenger price p, for various flight intervals z. The upward-sloping dashed line highlights combinations of profit-maximizing price, given by (12), and corresponding profit; the optimal price ranges from c when $z = 0$ to $c + \frac{\mu}{\lambda}$ for stand-alone flights ($z = \infty$). The profit curves do not intersect because, for given price, flight profit strictly increases with flight interval. The free-entry condition (14) pins down the equilibrium flight interval, the one for which the maximal profit is zero.

As evident from figure 3, there is a unique equilibrium in a "toy" model with non-stop flights only. The general model with 2-segment flights is more complex, due to agglomeration effects explained in section 4.3. These agglomeration effects may lead to multiple equilibria in terms of location of connection hubs. In our counterfactual exercise of section 7, we search for an equilibrium using the observed segment network as the starting point.

5. Data

This section describes the data used to estimate parameters of the model. We focus on the 2019 market of U.S. passenger flights.

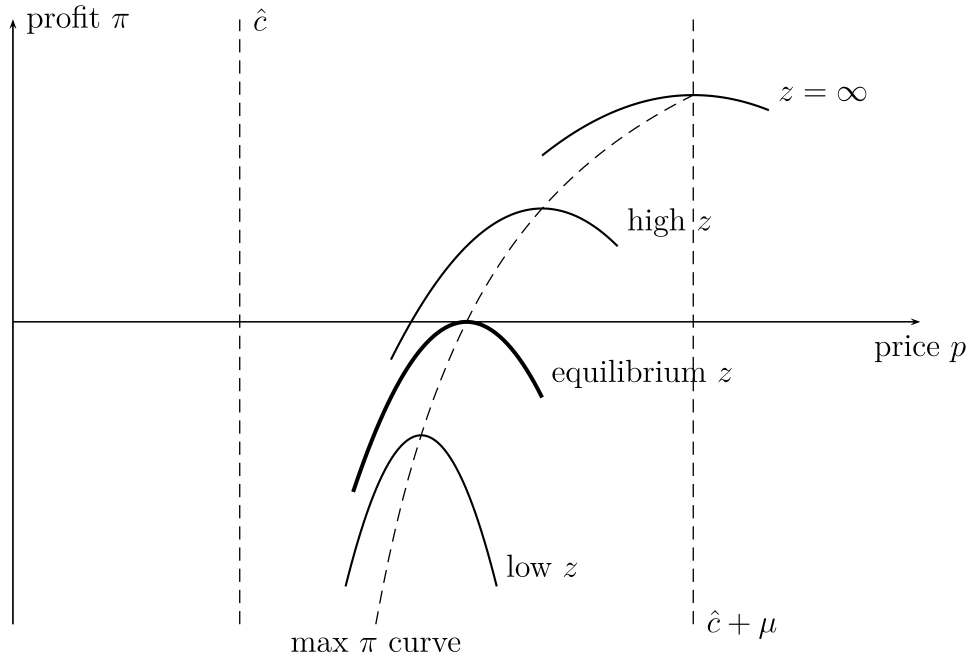

Figure 3: Flight profit per departure vs. price. z is flight interval.

5.1. Segment data

The primary source of information on flight segments is the T100 data (see section 3.1 for description). We keep only data on purely passenger aircraft (i.e., cargo and mixed-type aircraft are dropped). We also drop observations on aircraft with under 25 seats, which primarily perform non-scheduled flights or serve very small and remote destinations. We also drop observations with the same origin and destination (primarily sightseeing flights).

Then, we pool observations from different months of 2019, different airlines and aircraft types; we also pool the two directions of the same origin-destination pair. The resultant observations are bidirectional *segments*, identified by origin and destination airports; for each segment j, we observe the number of flights, average aircraft capacity S_j, average number of passengers per flight D_j, and average block time h_j. The flight interval z_j is the inverse of the number of flights.

Next, we drop segments with an average load factor under 25%. We end up with the set of 11,324 bidirectional flight segments with at least one endpoint inside the U.S., which we refer to as *T100 segments*. The resultant data includes 1.06 billion passenger-segments flown, or 99.78% of traffic observed in the original T100 data.

While we focus on U.S. market, some passenger itineraries to/from the U.S. include segments fully outside the U.S., not covered by the T100 data. For example, there were an estimated 90K passengers flying to/from Bangkok (Thailand) to Narita airport (Tokyo, Japan) for further connection to/from U.S. airports. For these segments, we use data on the number of flights in 2019 and average aircraft capacity from Innovata, a database of global flight schedules. The T100 and Innovata segments jointly constitute the set of *empirical segments*.

5.2. Route data

We derive information on 2019 passenger flight demand from the Airline Origin and Destination Survey (DB1B), another dataset provided by the U.S. Bureau of Transportation Statistics. These data are a 10% random sample of all tickets sold by U.S. airlines, and has information on the flight route, price, class of service, airline, and quarter of the year. We pool all passengers taking the same route (in both directions); for every such bidirectional route r, we use data on the number of observed passengers $paxsam_r$ and average price p_r. We drop the following routes: those with 3+ segments, with one-way ticket price under \$.02 and over \$3 per mile, as well as over \$5,000 total, with missing segment data (cf. section 5.1), routes including ground transportation between flight segments, those with both endpoints abroad (connecting via the U.S.), with both endpoints in the same city, and 2-segment routes where both flight intervals exceed 48 hours. The resultant set $\mathcal{R}_e$ of 225K bidirectional routes represents 96.49% of all passengers recorded in the 2019 DB1B data; we refer to $\mathcal{R}_e$ as the set of *DB1B routes*.

Next, we impute the total number of passengers pax_r on every $r \in \mathcal{R}_e$. For each T100 segment j, denote $q_j = \frac{\sum_{r \in \mathcal{R}_j} paxsam_r}{pax_j}$, where pax_j is the number of passengers using segment j as recorded in the T100 data. Due to DB1B sampling design, q_j is an asymptotically normal variable with mean 0.1 and variance $\frac{0.1\times 0.9}{pax_j}$. However, while the T100 passenger count is comprehensive, our DB1B count omits some passengers: those purchasing tickets from foreign airlines (which are not included in the DB1B), and also those taking routes that were dropped as described above. To address this issue, we impose a lower bound on q_j, equal to lower bound of the theoretical 99% confidence interval for q_j: $q_j \geq 0.1 - 2.5\left(\frac{0.1\times 0.9}{pax_j}\right)^{\frac{1}{2}}$.

For 2-segment routes $r = \{j_1, j_2\}$, we impute total passengers as $pax_r = \frac{paxsam_r}{\max\{q_{j_1}, q_{j_2}\}}$. For nonstop routes $r = \{j\}$, imputed passengers are T100 segment passengers minus all 2-segment imputed passengers using segment j, but no greater than $\frac{paxsam_r}{q_j}$:

$$pax_r = \min\left\{pax_j - \sum_{r' \in \mathcal{R}_j \cap \mathcal{R}_2} pax_{r'}, \frac{paxsam_r}{q_j}\right\}.$$

The upper bound implies that for some segments (primarily operated by foreign airlines), the sum of imputed route passengers $\sum_{r \in \mathcal{R}_j} pax_r$ will be lower than the T100 segment passenger count pax_j. We will treat these missing passengers as "exogenous" in the analysis that follows.

Define the set $\mathcal{J}_e$ as all T100 segments where less than 70% of passengers are exogenous. In the counterfactual exercise below, we will treat intervals $z_j, j \in \mathcal{J}_e$ as endogenous. Out of all T100 segments defined in section 5.1, only 4886 (43.2%) are included into $\mathcal{J}_e$, i.e. match well with the route data, but they account for 91.4% of passenger traffic in T100 data. The remaining empirical segments will be treated as exogenous, i.e., their flight intervals and other equilibrium parameters will be assumed unchanged in counterfactual analysis.

5.3. Counterfactual segments and routes

As one of objectives of this paper is to predict new flight destinations, in this section we specify which segments and routes will be added in the counterfactual analysis. We define the set $\mathcal{J}_c$ of *new segments* as all pairs *od* of airports not included in our list of empirical segments, and having 50+ DB1B passengers traveling between them: $\sum_{r \in \mathcal{R}_{od}} paxsam_r \geq 50$, where $\mathcal{R}_{od} \subset \mathcal{R}_e$ are all DB1B routes connecting the *od* airport pair. A total of 22K new segments are introduced. The block time h_j for the new segments, as well as for some empirical segments with missing data, is predicted (in minutes) by its estimated relationship with flight distance (in miles) as $h_j = 45.1 + 0.114 d_j$.

The set $\mathcal{R}_c$ of *new routes* includes nonstop routes that coincide with new segments, as well as all 2-segment routes such that (i) one segment is new and the other has empirical flight interval under 48h, (ii) the endpoints of the route are connected by a DB1B route (i.e., at least one DB1B passenger has traveled between them), and (iii) the sum of segment distances $d_{j_1} + d_{j_2}$ does not exceed 1.5 times the market distance d_m. A total of 689K new routes are considered.

6. Fitting the model

Some consumer preference parameters are borrowed from previously published studies. Morrison et al. (1989) estimate the cost of schedule delay in 1983 at \$2.98/h, or \$7.67/h in 2019 dollars. We assume $\alpha = 7.67/60 = 0.1278$ \$/min. The cost of connection time is set equal to the cost of travel time estimated by Federal Aviation Administration,[20] $\beta = 47.10/60 = 0.785$ \$/min.[21] The value of the logit parameter is borrowed from Berry and Jia (2010), at $\lambda = 0.72$.

The expected non-scalable flight cost per flight minute is calibrated using estimates from

[20] Although the value of travel time may increase with trip duration (e.g., see Fosgerau and Engelson (2011)), we use the average hourly value from Table 4 of the Department of Transportation's 2016 Revised Value of Travel Time Guidance instead of making additional assumptions about how the value of time changes for longer trips.

[21] This per-minute disutility estimate is consistent with the estimate in Luttmann (2019). In that paper, passengers taking a connecting flight were found to be compensated with a fare that is \$0.71 to \$0.79 cheaper per minute (i.e., \$42.74 to \$47.60 per hour) of connecting time.

section 3:

$$f_j^e = \gamma_f^{cr} * ncr_j + \gamma_f^{other}. \tag{17}$$

The consumer willingness-to-pay parameter (cf.(2)) in market m is assumed to take the form

$$\mu_m = \mu_0 \frac{\mu_1 d_m}{\mu_0 + \mu_1 d_m}. \tag{18}$$

This functional form implies that μ_m, as function of d_m, increases from 0 to μ_0, and has slope of μ_1 at zero. The unknown parameters $M \equiv \{\mu_0, \mu_1\}$ are estimated using a GMM procedure described later in this section. The moments used in the procedure aim to minimize the discrepancies between the model (section 4) and the data (section 5) at the level of segments. We now proceed to describe calculation of route preference shocks ξ and segment-level errors for a given value of M. Because most demand-side parameters are borrowed from the literature, route shocks ξ are not used directly in the GMM procedure; they are needed to evaluate the market size A_m which is used to predict counterfactual demand.

In this section, all relevant variables are presented as functions of model parameters M and ξ.

6.1. Scalable costs

For DB1B routes $r \in \mathcal{R}_e$, scalable costs are computed simply as $c_r(M) = p_r - mkup_r(\mathbf{p}_m, \mathbf{z}, \mu_m(M))$, where p_r is the observed ticket price and $mkup_r(\mathbf{p}_m, \mathbf{z}, \mu_m(M))$ is given by (12). Since the new route price is not observed, we extrapolate c_r from DB1B routes onto new routes. Specifically, we first estimate the regression

$$c_r = c_0 + \sum_{j \in r} X_j^c \delta_c + c_m + \eta_r^c, r \in \mathcal{R}_e.$$

Here X_j^c are regressors of costs related to segment j, which include: dummies for segments fully inside the U.S., segments crossing the U.S. border, and segments fully abroad; flight time; flight time for segments to/from Essential Air Service (EAS) airports;[22] individual dummies for airports with 50M+ traffic (to account for possible congestion charges), and a dummy for EAS airports. c_0 and δ_c are parameters estimated by least squares, c_m are market random effects, and η_r^c are route-level residuals. We use the number of observed DB1B passengers as regression weights. We then use estimates c_0, δ_c, c_m from this regression to predict scalable costs c_r on new routes $r \in \mathcal{R}_c$.

6.2. Selection bias

DB1B routes, by definition, are ones taken by at least one passenger sampled by the DB1B database. By construction, such data is subject to selection bias: for thin routes with low expected traffic, a positive preference shock ξ_r increases the probability that a passenger chooses that route, and thereby increases the probability that the route is recorded in the DB1B. Therefore, for thin routes, expectation $E\xi_r$ is positive rather than zero. While other studies using DB1B data ignore this problem, we develop a procedure to explicitly evaluate $E\xi_r$ for every DB1B route, as follows.

First, observe that the probability of passenger i in market m to use route r is (cf.(7,15)) $\frac{U_r(p_r,\mathbf{z},M,\xi_r)}{T_m(\mathbf{p}_m,\mathbf{z},M,\xi)^{1-\lambda}T_0}$; the probability that the passenger is also sampled by the DB1B data is $\frac{1}{10}$ of that quantity. With A_m passengers considering flight, the probability that none are recorded in DB1B data is then

$$\Pr(paxsam_r = 0) = \left(1 - \frac{U_r(p_r, \mathbf{z}, M, \xi)}{10T_m(\mathbf{p}_m, \mathbf{z}, M, \xi)^{1-\lambda}T_0}\right)^{A_m}. \tag{19}$$

By assumption 2, T_0 is a large quantity; for (9) to match empirical aircraft size, A_m should

[22] Essential Air Service is a U.S. federal program that subsidizes flights to/from certain "disadvantaged" communities.

be large, too. But then, probability (19) can be approximated by

$$\Pr(paxsam_r = 0) \approx \exp\left(-\frac{A_m U_r(p_r, \mathbf{z}, M, \xi)}{10 T_m(\mathbf{p}_m, \mathbf{z}, M, \xi)^{1-\lambda} T_0}\right). \tag{20}$$

Under the "true" value of preference shocks ξ_r, the ratio in (20) is the expectation of the number of passengers sampled by the DB1B, $paxsam_r$, and can therefore be approximated by the latter. For an arbitrary $\tilde{\xi}_r$, the approximate probability (20) is given by (assuming for simplicity the denominator in (20) is constant) $\exp\left(-paxsam_r e^{\tilde{\xi}_r - \xi_r}\right)$. But then, the expectation of $\tilde{\xi}_r$, conditional on the fact that at least one passenger was observed, is given by

$$E\xi_r = \frac{\int x\left(1 - \exp\left(-paxsam_r e^{x-\xi_r}\right)\right) \mathrm{d}F(x)}{\int 1 - \exp\left(-paxsam_r e^{x-\xi_r}\right) \mathrm{d}F(x)}, \forall r \in \mathcal{R}_e, \tag{21}$$

where $F(\cdot)$ is the c.d.f. of preference shocks ξ. $F(\cdot)$ is assumed to be normal with empirically evaluated variance. The latter, along with values of ξ_r in (21), are estimated using the methodology of section 6.3, under a preliminary assumption of $E\xi_r = 0$.

6.3. Preference shocks and market size

Preference shocks ξ_r are assumed to consist of two components: expectation $E\xi_r$ and demeaned preference shocks ξ_r^d that add up to zero for each market. Calculation of $E\xi_r$ is explained in section (6.2); we now detail ξ_r^d. Unlike earlier papers (e.g., Berry and Jia (2010)) that had to use a numerical multi-iterative contraction-mapping procedure to find preference shocks, out method allows to find them explicitly in a single step.

From (9,16), equilibrium passenger traffic on route r can be expressed as

$$pax_r \equiv \frac{D_r}{z_r} = C_m U_r(p_r, \mathbf{z}, M, \xi_r) = U_r(p_r, \mathbf{z}, M, \xi_r + \ln(C_m)), \tag{22}$$

where C_m is some market-specific constant. Given this observation, we can compute the "biased" value $\tilde{\xi}_r$ of preference shocks that solve $pax_r = U_r(p_r, \mathbf{z}, M, E\xi_r + \tilde{\xi}_r)$. We then subtract the mean $\tilde{\xi}_r$ across DB1B routes for each market to obtain the demeaned shocks: $\xi_r^d = \tilde{\xi}_r - \frac{\sum_{r' \in \mathcal{R}_e \cap \mathcal{R}_m} \tilde{\xi}_{r'}}{|\mathcal{R}_e \cap \mathcal{R}_m|}$. Note that for a market with a single DB1B route, ξ_r^d for that route is zero by construction.

The market size A_m can be inferred from the data on the total passenger traffic on market m (cf.(4,15))

$$pax_m \equiv \sum_{r \in \mathcal{R}_m} pax_r = A_m \frac{\sum_{r \in \mathcal{R}_m} U_r(p_r, \mathbf{z}, M, \xi)}{T_m(\mathbf{p}_m, \mathbf{z}, M, \xi)^{1-\lambda}(T_0 + T_m(\cdot))^\lambda} = \frac{A_m T_m(\mathbf{p}_m, \mathbf{z}, M, E\xi_r + \xi_r^d)^\lambda}{(T_0 + T_m(\cdot))^\lambda}. \tag{23}$$

We do not investigate the determinants of A_m, but rather extrapolate it as given into the counterfactual environment.

The total preference shock is $\xi_r = E\xi_r + \xi_r^d$. For new routes $r \in \mathcal{R}_c$, the first component is zero while the second is unknown. We simulate ξ_r^d in our counterfactual exercise presented in section 7; for the purposes of parameter estimation, we make a preliminary assumption $\xi_r^d = 0, r \in \mathcal{R}_c$.

6.4. Moment conditions

We use two types of moment conditions to identify M: empirical segment moments, and a new segment moment.

Given the values of unknown parameters M, inequality (14) identifies the value of non-scalable cost f_j for empirical endogenous segment $j \in \mathcal{J}_e$, and the lower bound for f_j for counterfactual segments $j \in \mathcal{J}_c$. We now describe the procedure for computing these.

First, we compute the price markup $mkup_r(\mathbf{p}_m, \mathbf{z}, \mu_m(M))$ from (12), for all routes. The second step is to compute route gross profit $\phi_r(\cdot, M)$ from (10), which requires knowledge of flight demand $D_r = pax_r z_r$. For empirical routes $r \in \mathcal{R}_e$, this quantity is known from the data.

One moment condition described below is based on new segment profits, for which $D_r, r \in \mathcal{R}_c$ (the number of passengers who would take a stand-alone flight on a counterfactual segment) has to be computed from (16). For that purpose, we use estimates of counterfactual scalable cost c_r from section 6.1, market size A_m defined in section 6.3, and assumed zero preference shocks.

For given segment j, $\phi_j(\mathbf{p}, \mathbf{z}, M)$ (as defined in (13)) is segment gross profit from *endogenous* passengers; we also need to account for exogenous passengers (i.e., those not matched to DB1B data) described in section 5.2. We assume the contribution of each exogenous passenger to gross profit is the same as that of an average endogenous passenger on the same segment. Denote by $shex_j$ the share of exogenous passengers on segment j; the above assumption then modifies (14) by multiplying its right-hand side by $(1 - shex_j)$.

To achieve a match between gross profits and non-scalable costs, we assume the latter is stochastic and takes the form $f_j = f_j^e e^{\zeta_j}$, where f_j^e is defined by (17) and ζ_j are *cost shocks*. For empirical endogenous segments, cost shocks are then given by (cf.14)

$$\zeta_j(M) = \ln \phi_j(\mathbf{p}, \mathbf{z}, M) - \ln(1 - shex_j) - \ln h_j - \ln f_j^e, j \in \mathcal{J}_e. \tag{24}$$

The zero expectation of cost shocks defines the first moment function: $g_j^{\mathcal{J}}(M) = \zeta_j(M), j \in \mathcal{J}_e$. The nature of ζ_j (unexplained part of airline non-scalable cost) also allows us to specify additional moments, that ζ_j is uncorrelated with aircraft capacity S_j and departure interval z_j. Because capacity and departure intervals are endogenous, we instrument for them using predicted values z_j^{ins} (S_j^{ins}) from the regression of $\ln z_j$ ($\ln S_j$) on log passenger traffic of endpoint airports, log distance, and individual dummies for airports with 50M+ passengers. Then, two additional moment functions are $g_j^S(M) = \zeta_j(M) \ln S_j^{ins}$ and $g_j^z(M) = \zeta_j(M) \ln z_j^{ins}, j \in \mathcal{J}_e$.

For new segments, net profits are unknown, cost shocks are not identifiable and are assumed to equal zero. (14) implies the log of gross profit does not exceed the log of non-scalable cost. This observation allows us to specify the counterfactual moment function as $g_j^c(M) \equiv \max\{\ln \phi_j(\mathbf{p}, \mathbf{z}, M) - \ln h_j - \ln f_j^e, 0\}, \forall j \in \mathcal{J}_c$. This counterfactual moment identifies M by capping profits of stand-alone counterfactual flights on new segments.

6.5. GMM estimation

To summarize section 6.4, we have specified four moment conditions of two types:

- three empirical segment moments $\hat{g}^k(M) \equiv \frac{1}{|\mathcal{J}_e|}\sum_{j\in\mathcal{J}_e} g_j^k(M)$, with $k \in \{\mathcal{J}, S, z\}$;
- one new segment moment $\hat{g}^c(M) \equiv \frac{1}{|\mathcal{J}_c|}\sum_{j\in\mathcal{J}_c} g_j^c(M)$.

For an efficient GMM estimate, moment conditions should be weighted by the inverse matrix of moment covariance. Covariance within each of the two groups is evaluated from the data; covariance across empirical and new segments is assumed zero.

The GMM estimates (standard errors) of M (cf.(18)) are as follows: $\mu_0 = 150.1(6.1)$, $\mu_1 = 0.1049(0.0064)$. For example $\mu_m = \$17.6$ for the Boston-New York market or \$126 for the Los Angeles-Sydney market. This variation implies, among other things, that passengers on the latter market are seven times more patient, allowing less frequent service and larger aircraft. Empirically, nonstop flights on the former market used aircraft with 102 seats on average and departed every 23 min, compared to 336 seats and 311 min on the latter market.[23]

To further evaluate model fit, we compare the route-level own-price elasticity generated from our model to that of other studies. In our model, this elasticity is route-level demand response to a change in price for all departures on that route (cf.(15,16)): $p_r \frac{\mathrm{d}\ln D_r(\mathbf{p_m}, \mathbf{z})}{\mathrm{d}p_r} = -\frac{p_r}{\mu_m}\left(1 - (1-\lambda)\frac{U_r(p_r, \mathbf{z})}{T_m(\cdot)}\right)$; the average elasticity (weighted by DB1B passengers) is -3.29.

[23]These aircraft capacity and departure interval metrics for the Boston-New York and Los Angeles-Sydney markets are computed from 2019 T100 data.

This estimate matches that from previous studies using a nested logit to model air travel demand. For example, own-price elasticities range from -3.2 to -3.6 in Peters (2006) and from -2.0 to -3.5 in Israel et al. (2013).

7. Counterfactual experiment

7.1. Counterfactual exogenous parameters

This section predicts the effects of reduced crew requirements. Currently, all commercial aircraft (except very small aircraft with 10-15 seats) are operated by two crew members. We investigate the effects of new technology allowing aircraft to be operated by a single Captain. By assumption of section 3.1, the cost of the Captain is $\frac{2}{3}$ of standard crew cost ncr. Therefore, the counterfactual "effective" number of crews (i.e., the crew cost relative to that of the standard 2-pilot crew) is $ncr_j^c = \frac{2}{3}$ for segments with block time under 8 hours, and $ncr_j^c = \frac{4}{3}$ for segments over 8 hours, where a relief captain must be present. The expected non-scalable cost per min f_j^e, defined by (17), is then replaced by $f_j^c \equiv \gamma_f^{cr} * ncr_j^c + \gamma_f^{other} < f_j^e$.

Section 3.1 points out that part of crew cost is scalable (i.e., proportional to aircraft size). We argue that this part of crew cost is the effect of pilot experience rather than aircraft size *per se*: larger aircraft are typically operated by more experienced pilots. In the counterfactual equilibrium, the distribution of aircraft size may change, but pilot experience will probably not; this means that aircraft of the same size will be matched to more experienced (and thus more highly paid) pilots. This effect is likely to offset the effect of reduced crew size on scalable cost; we will assume no change of c_r from the estimates of section 6.1. In figure 2, the counterfactual change is then a parallel downward shift of the cost lines.

While all parameters of empirical segments and routes are identified by observed prices, traffic, and zero-profit conditions, the same parameters for new segments and routes are only partially identified by the non-positive profit condition. Negative profit on a given segment $j \in \mathcal{J}_c$ may be due to low preference shock $\xi_r, r \in \mathcal{R}_j$, or high cost shock ζ_j, or both.[24] To keep the model credible and tractable at the same time, we make the following simplifying assumptions. To make unobserved heterogeneity uni-dimensional, we assume the cost shocks are the same for all new segments and equal to mean empirical cost shock: $\zeta_j = \frac{1}{|\mathcal{J}_e|} \sum_{j' \in \mathcal{J}_e} \zeta_{j'}, \forall j \in \mathcal{J}_c$. The GMM procedure aims to make this quantity equal to zero (cf. section 6.4), but it is generally non-zero because the GMM model is under-identified.

Given the above assumption, any variation in profit is driven exclusively by preference shocks ξ_r^d. Furthermore, because ξ_r^d are defined at the route level while airline profit at the segment level, we make an additional assumption that ξ_r^d are the same for all new routes that include a given new segment: $\xi_r^d = \xi_j^d, \forall r \in \mathcal{R}_j \cap \mathcal{R}_c$.

Given these assumptions, profit condition (14) identifies the upper bound $\bar{\xi}_j^d$ on preference shocks. To calculate $\bar{\xi}_j^d, j \in \mathcal{J}_c$, first observe that T_m in (9) does not depend on counterfactual shocks, only empirical ones, in empirical (observed) equilibrium. But then, (15), flight demand (16), route gross profit (10), and segment gross profit (13) are all multiplied by factor $\exp(\xi_j^d)$ as the counterfactual preference shock for routes $r \in \mathcal{R}_j$ increases from zero to ξ_j^d. Then, (cf.(13)) $\ln \phi_j(\mathbf{p}, \mathbf{z}, M, \xi) = \ln \phi_j(\mathbf{p}, \mathbf{z}, M, 0) + \xi_j^d$; the upper bound is given by $\bar{\xi}_j^d = \ln h_j + \ln f_j^e - \ln \phi_j(\mathbf{p}, \mathbf{z}, M, 0)$. The simulated values of ξ_j^d are drawn from normal distribution with zero mean and variance equal to that of $\xi_r^d, r \in \mathcal{R}_e$, conditional on $\xi_j^d < \bar{\xi}_j^d$.

7.2. Calculation of counterfactual equilibrium

The counterfactual equilibrium is given by vector of flight intervals $z_j, j \in \mathcal{J}_e \cup \mathcal{J}_c$, such that (14) hold for all $j \in \mathcal{J}_e \cup \mathcal{J}_c$ under counterfactual exogenous parameters, with equality if $z_j < \infty$. As mentioned in section 4.4, the equilibrium is not guaranteed to be unique due to agglomeration effects of 2-segment connection hubs; moreover, equilibria may differ from one simulation of $\xi_j^d, j \in \mathcal{J}_c$ to another. We perform 100 independent simulations of the

[24]Unobserved heterogeneity in scalable cost c_r and thus in ticket price p_r is also possible; from (2), its effect is very similar to that of heterogeneity of ξ_r.

latter to provide probabilistic rather than deterministic properties of the equilibrium.

In the remainder of the paper, we present all relevant endogenous variables as functions of the vector of flight intervals z; the parameters M and ξ are at their estimated or simulated value.

To find the equilibrium vector z, we exploit the fact that the majority of passengers on a typical segment are nonstop passengers, and therefore airline profit on segment j usually does not depend too much on flight intervals $z_{j'}$ on a connecting segment j'. Given this feature, we seek the equilibrium in a multi-iterative process; on each iteration a for each segment j, we seek to find $z_j^{[a]}$ that meets (14), while treating $z_{j'}^{[a-1]}, j' \neq j$ as given. We do so using the Newton's method. Each iteration a of our algorithm thus updates flight interval from previous iteration $z_j^{[a-1]}, j \in \mathcal{J}_e \cup \mathcal{J}_c$, as follows:

$$\ln z_j^{[a]} = \ln z_j^{[a-1]} - \frac{\phi_j(\mathbf{p}^{[a-1]}, \mathbf{z}^{[a-1]}) - f_j^c h_j}{z_j^{[a-1]} \frac{\partial \phi_j(\mathbf{p}^{[a-1]}, \mathbf{z}^{[a-1]})}{\partial z_j}}.$$

For segments without service ($z_j^{[a-1]} = \infty$) but positive counterfactual profit ($\phi_j(\mathbf{p}^{[a-1]}, \mathbf{z}^{[a-1]}) > f_j^c h_j$), we set $z_j^{[a]}$ equal to a certain initial value (i.e., "introduce" service). For segments with $z_j^{[a-1]}$ above the a certain high value and $\phi_j(\mathbf{p}^{[a-1]}, \mathbf{z}^{[a-1]}) < f_j^c h_j$, we set $z_j^{[a]} = \infty$ (i.e., "cancel" service). We also make sure z_j does not change too much in one iteration.

This algorithm shows excellent convergence properties. We define "convergence" by net profit per departure being within interval $[-\$10, \$10]$ for all segments with finite interval z; this goal is typically achieved within less than 30 iterations. Moreover, the profit is typically within interval $[-\$1, \$1]$ for 99.5% of segments.

	endog. segments	counterfactual	
	2019 data	2019 segments	new segments
Traffic (10^6 pax-segments)	968	1263	31
Mean aircraft capacity (weighted by flight-hours)	156	131	109
# of departures (10^6)	9.03	14.2	0.416
Seat-miles flown (10^9)	1427	1753	65.4
Cost of flight crew (10^9 \$)	24.14	28.06	1.03

Table 1: Counterfactual equilibrium: segment-level parameters

7.3. Results

7.3.1. Segments

Among 4886 empirical endogenous segments $\mathcal{J}_e$ defined in section 5.2, 4877 have an increased expected (across the 100 simulations) number of departures. The number of segments with at least weekly service increases from 3773 to 5132. Table 1 compares the values of some segment-level aggregate parameters in 2019 data, expected counterfactual values for segments with at least one departure in 2019, and expected counterfactual values for new segments.

The overall traffic increases by about 27%. The number of seat-miles flown is roughly proportional to fuel consumption and emissions. Thus, automation, by increasing traffic, will also increase emissions by the aviation industry unless complemented by greener engines. The effect on aggregate emissions however may be lower, because some of the new passengers may be switching from personal automobiles. As air travel is about two times more energy-efficient than a car,[25] the net emissions effect will be zero if one-half of new passengers are previous automobile users.

As expected, aircraft are becoming smaller, with a 16% reduction in capacity on existing

[25] The fuel efficiency of Airbus A320 is estimated at 97 seat-miles per gallon (Carbon Offsetting & Air Travel, Anja Kollmuss and Jessica Lane, 2008), or 78 passenger-miles per gallon under 80% load factor. In comparison, the average U.S. automobile of the 2019 model year had efficiency of 24.9 mpg (Highlights of the Automotive Trends Report by the U.S. Environmental Protection Agency), or 37 passenger-miles per gallon under the assumption of average occupancy of 1.5.

	DB1B data	counterfactual
Total passengers (10^6)	701	898
nonstop passengers (10^6)	501	593
Mean price per mile flown, $	0.222	0.218
Mean flight interval, min (all routes of given market)	89.4	60.1

Table 2: Counterfactual equilibrium: demand-side parameters. Exogenous passengers excluded.

segments. Section 7.3.4 provides more detail about counterfactual aircraft demand.

Despite reduced spending on flight crew on a given flight, the aggregate flight crew expenses increase by 20.5%, due to (i) more passengers and (ii) higher departures/passengers ratio. Our calculations suggest that pilots should view automation as a positive rather than negative shock to their profession.

7.3.2. Passengers

Table 2 illustrates aggregate demand-side parameters in a typical simulation. The passenger count in table 1 is higher than in 2 because (i) the former includes some "exogenous" passengers, excluded from the latter table, and (ii) passengers using 2-segment routes are counted twice in the former table. Contrary to our expectation, the share of nonstop passengers does not increase, but slightly decreases from 501/701=71.5% to 593/883=66.0%. This result however is not the same for all markets. Figure 4 compares empirical and counterfactual share of nonstop passengers (smoothed non-parametrically) as a function of initial market size. In smaller markets with initial traffic under $\exp(11.3) \approx 80$K passengers, the share of non-stop traffic increases by about 5 percentage points; the aggregate decrease is mostly coming from larger markets. We interpret these findings as follows. The share of non-stop travel on small markets increases due to more flights by small aircraft in/out small airports. For larger markets, while the absolute number of non-stop passengers may be increasing, the number of 2-segment passengers increases faster due to increased attractiveness of smaller airports as connection hubs. The next section has more on connection airports.

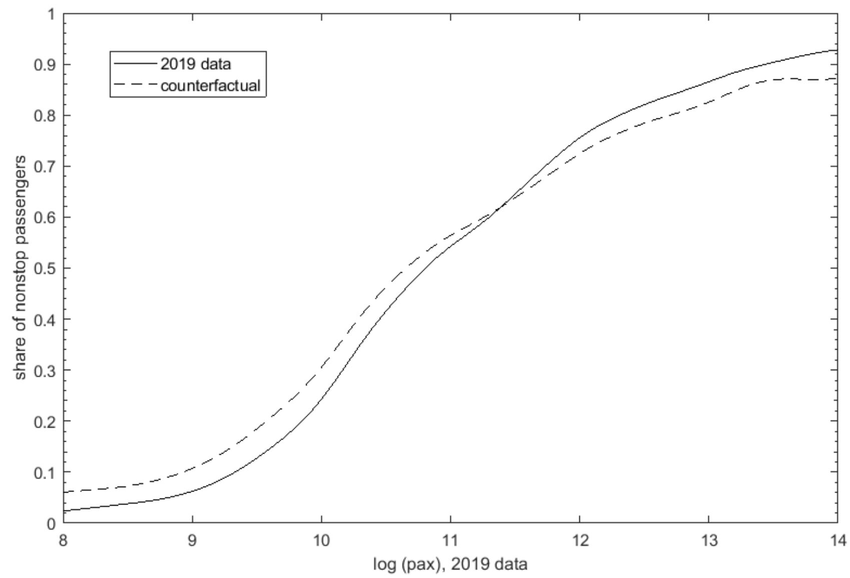

Figure 4: 2019 market size vs. share of nonstop passengers

7.3.3. Airports

Figure 5 maps, for the U.S. airports with endogenous passengers, the initial airport traffic against the counterfactual increase in traffic. It is evident that smaller communities gain much more in terms of traffic (and therefore in terms of transport accessibility). While traffic at the biggest airports increases by only about 10%, the median increase in log traffic is 0.49, or $\exp(0.49) - 1 = 63\%$ increase in passengers.

Figure 6 shows similar airport-level results, but for connecting passengers only. It confirms the increasing role of smaller airports as connecting hubs. This result is given for a single typical simulation and may be non-robust for individual airports, but the overall trend is clear.

7.3.4. Aircraft

Figure 7 illustrates the change in demand for aircraft by airlines. Three range categories are analyzed separately, as aircraft size is known to vary with range. Aircraft are clustered by size into 20-seat categories, from 20-39 to 360-379 seat aircraft. The height of each bar is the

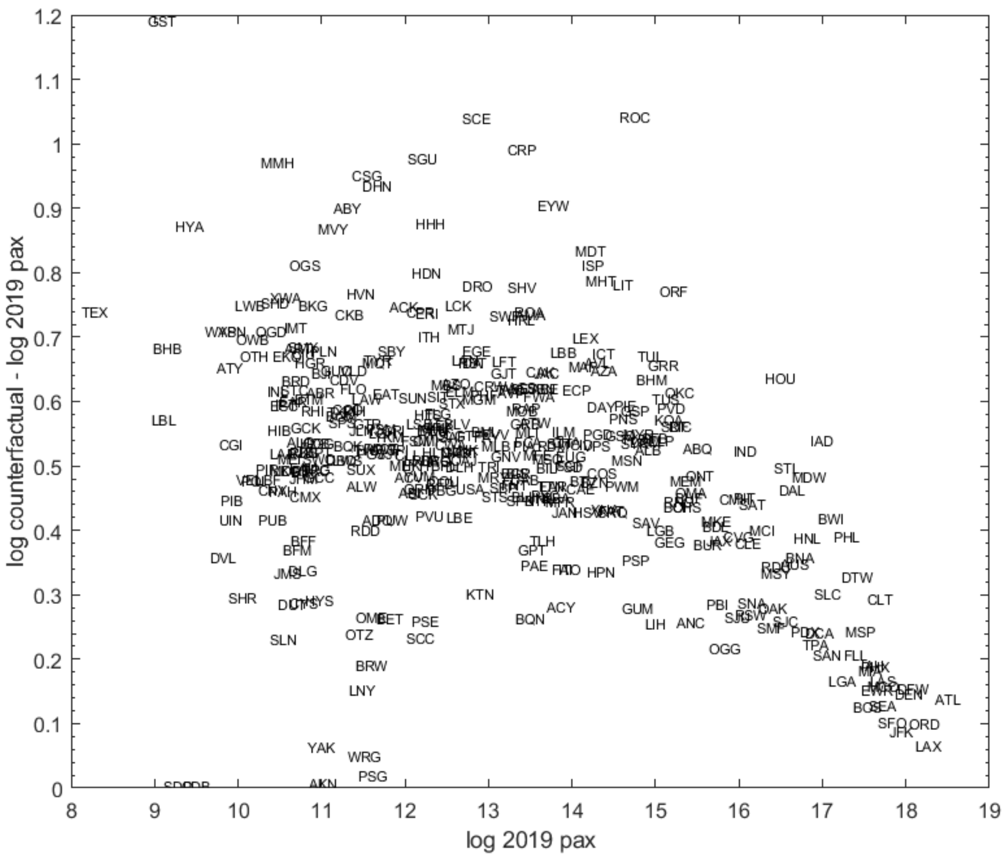

Figure 5: 2019 airport traffic (passenger enplanements+deplanements) vs. counterfactual change.

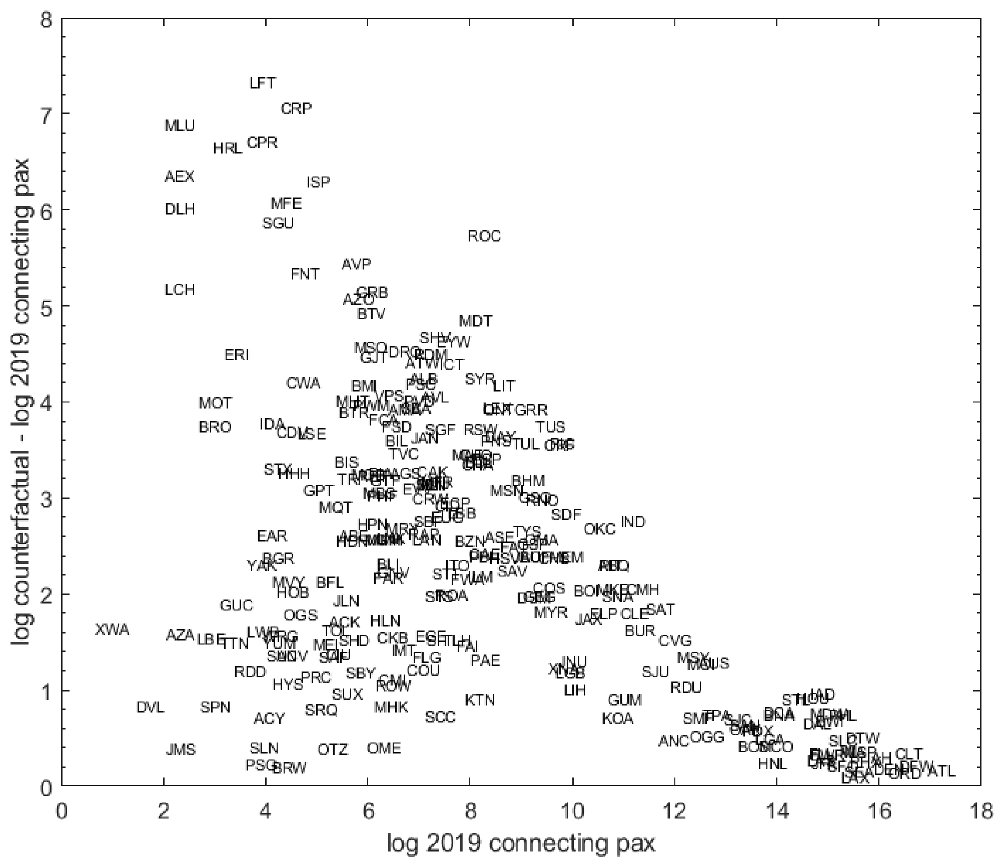

Figure 6: 2019 airport connecting traffic vs. counterfactual change.

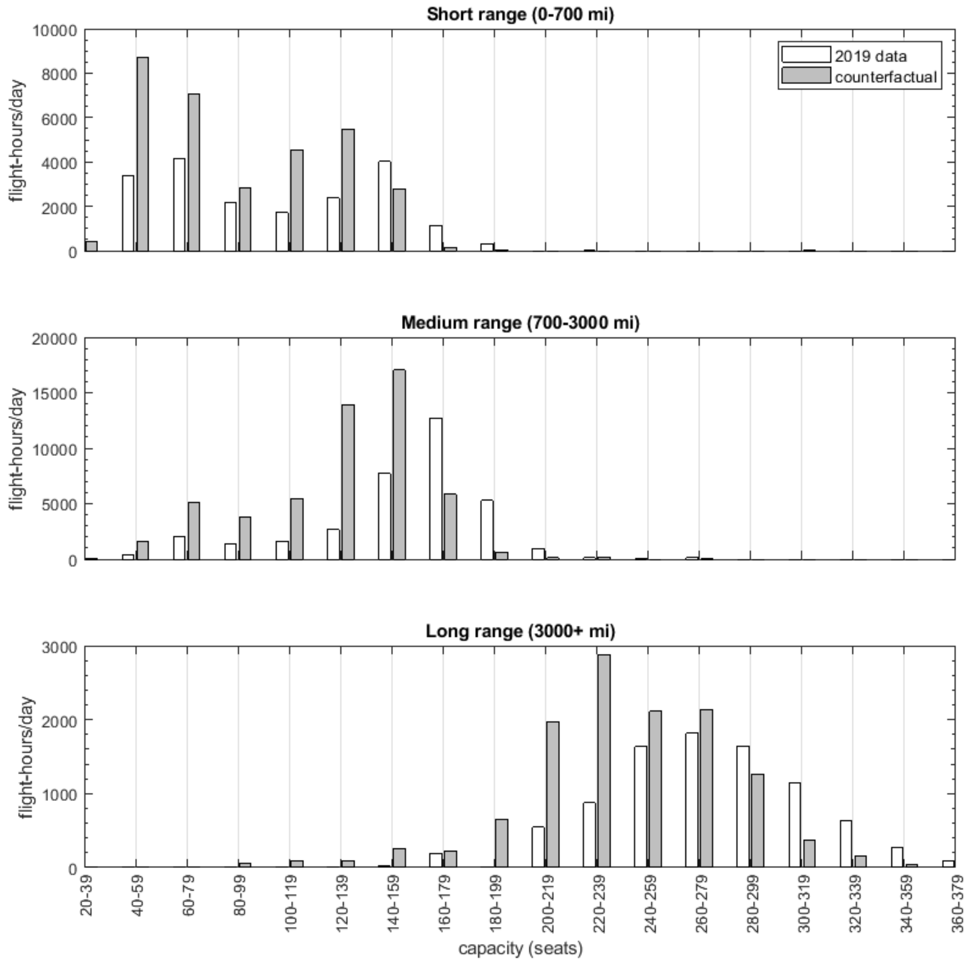

Figure 7: Demand for aircraft by size and range: 2019 data vs. counterfactual

total number of flight-hours demanded by U.S. airlines (excluding “exogenous” segments), per day, within each aircraft size and range category. This quantity is roughly proportional to the number of aircraft units demanded.

For short-range aircraft, there are two peaks in size distribution, both of which shift towards smaller size. The market for 60-79 seat aircraft, such as Embraer ERJ-175, is increasing but is dwarfed by that for 40-59 seat aircraft, such as Embraer-145. The second peak at 140-159 seats (e.g., Airbus A320) decreases for short-range markets, to be replaced by demand for aircraft with 120-139 seats (e.g., Airbus A319).

In the mid-range market, the demand for now-popular 160-179 seat aircraft (e.g., Boeing 737-800) decreases dramatically; at the same time, there is a powerful rise in demand for 120-159 seat models (e.g., Boeing 737-700). And in the long-range market, peak demand shifts from models with 260-279 seats (e.g., Boeing 777-200) to 220-239 seat aircraft, (e.g., the smallest variants of the Boeing 787 Dreamliner).

8. Conclusion

Although commercial aircraft are offered in a range of capacities, average aircraft size has declined over time. We argue that this decline is primarily driven by a past reduction in the minimum flight crew requirement. Consistent with this argument, we develop a novel structural model of aircraft choice that combines data on passenger traffic, flight frequency, and airline costs. In our model, the cost of the flight crew is the primary factor driving the use of larger aircraft (i.e., non-scalable costs), while passenger utility is the primary factor that drives the use of smaller aircraft. After recovering the parameters of our structural model, we perform a counterfactual experiment where the minimum crew requirement is further relaxed from two pilots to one, a policy being considered as advances in automation increase safety by reducing the amount of flight time requiring manual pilot intervention.

Our counterfactual simulations reveal several important insights with respect to aircraft size, pilot demand, passenger traffic, flight frequency, and the number of markets served. If such a policy were implemented, our results indicate that passenger traffic would increase by over 25% and flight frequency would increase on virtually all segments. As expected, average aircraft size decreases by 16%. However, despite a reduction in the cost of the flight crew per departure, aggregate flight crew expenses rise by 20% due to an increase in the number of flights. Accordingly, advances in automation are expected to *increase* aggregate pilot income.

Overall, our counterfactual experiment indicates that consumer utility increases with the relaxation of the minimum crew requirement, primarily due to more frequent departures and shorter connection times. While passenger traffic increases, most of this increase occurs in smaller non-congested airports, meaning that (i) most welfare gains accrue to smaller communities, and (ii) additional congestion in already congested airports is modest.

Appendix A. Notational glossary

Notation used within a single section is omitted.

Notation	Description	Unit	Section of 1st use
α	Disutility of schedule delay	\$/min	4.1.2
β	Disutility of (max) connection time	\$/min	4.1.2
γ	Predictors of airline costs		3.1
ζ	Non-scalable cost shocks		6.4
λ	Nested logit elasticity parameter		4.1.2
μ_m	Consumer willingness to pay, market m	\$	4.1.2
ξ	Route preference shock		4.1.2
ϕ	Route gross profit per departure	\$	4.2.3
A_m	Size of market m	pax/min	4.1.1
B_{jr}	Markup share of segment j on route r		4.2.3
c_r	Scalable cost per passenger, route r	\$	4.2.2
D	Flight demand per departure	pax	4.1.3
d_j (d_m)	Segment (market) distance	mi	4 (4.1.1)
f_j	Non-scalable airline cost	\$/min	4.2.2
h_j	Ramp-to-ramp time, segment j	min	4
$\bar{h}_r$	Max connection time, route r	min	4.1.2
$\mathcal{J}_e$ ($\mathcal{J}_c$)	Set of endogenous empirical (new) segments		5.2 (5.3)
L_j	Load factor, segment j		4.2.1
$M = \{\mu_0, \mu_1\}$	Vector of determinants of willingness to pay		6
p	Passenger price	\$	4.1.2
$\mathcal{R}_1$ ($\mathcal{R}_2$)	Set of nonstop (2-segment) routes		4.1.1
$\mathcal{R}_j$	Set of routes using segment j		4.2.4
$\mathcal{R}_e$ ($\mathcal{R}_c$)	Set of DB1B (new) routes		5.2 (5.3)
$\mathcal{R}_m$	Set of routes for market m		4.1.1
S	Aircraft capacity	seats	3.1
T_m	Logit param.: value of all routes, market m		4.1.3
t	Schedule delay	min	4.1.2
U_r	Logit param.: value of route r		4.1.3
u	Consumer utility		4.1.2
z	Departure interval	min	4

References

Armantier, O., Richard, O., 2008. Domestic airline alliances and consumer welfare. The RAND Journal of Economics 39, 875–904.

Bailey, R.E., Kramer, L.J., Kennedy, K.D., Stephens, C.L., Etherington, T.J., 2017. An assessment of reduced crew and single pilot operations in commercial transport aircraft operations, in: 2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC), IEEE. pp. 1–15.

Bain, J.S., 1954. Economies of scale, concentration, and the condition of entry in twenty manufacturing industries. The American Economic Review 44, 15–39.

Basso, L.J., Jara-Díaz, S.R., 2006. Distinguishing multiproduct economies of scale from economies of density on a fixed-size transport network. Networks and Spatial Economics 6, 149–162.

Berry, S., Jia, P., 2010. Tracing the woes: An empirical analysis of the airline industry. American Economic Journal: Microeconomics 2, 1–43.

Berster, P., Gelhausen, M.C., Wilken, D., 2015. Is increasing aircraft size common practice of airlines at congested airports? Journal of Air Transport Management 46, 40–48.

Bilimoria, K.D., Johnson, W.W., Schutte, P.C., 2014. Conceptual framework for single pilot operations, in: Proceedings of the International Conference on Human-Computer Interaction in Aerospace, pp. 1–8.

Bitzan, J.D., Keeler, T.E., 2007. Economies of density and regulatory change in the US railroad freight industry. The Journal of Law and Economics 50, 157–180.

Bouchard, J., Baggioni, N., 2017. As airlines aim for autonomous flight, near-term revolution will be going single pilot. Forbes

URL: https://www.forbes.com/sites/oliverwyman/2017/10/25/single-pilot-commercial-flights-are-not-far-off-even-if-fully-autonomous-flight-is/?sh=5f8a64b13b17.

Braeutigam, R.R., Daughety, A.F., Turnquist, M.A., 1984. A firm specific analysis of economies of density in the US railroad industry. The Journal of Industrial Economics , 3–20.

Brueckner, J.K., 2004. Network structure and airline scheduling. The Journal of Industrial Economics 52, 291–312.

Brueckner, J.K., Spiller, P.T., 1994. Economies of traffic density in the deregulated airline industry. The Journal of Law and Economics 37, 379–415.

Caves, D.W., Christensen, L.R., Tretheway, M.W., 1984. Economies of density versus economies of scale: Why trunk and local service airline costs differ. The RAND Journal of Economics , 471–489.

Chen, Y., Gayle, P.G., 2019. Mergers and product quality: Evidence from the airline industry. International Journal of Industrial Organization 62, 96–135.

Christensen, L.R., Greene, W.H., 1976. Economies of scale in US electric power generation. Journal of Political Economy 84, 655–676.

Cook, A., Tanner, G., 2008. Innovative Cooperative Actions of R&D in EUROCONTROL Programme CARE INO III: Dynamic Cost Indexing: Aircraft crewing–marginal delay costs. Technical Report. Transport Studies Group, University of Westminster.

Cullinane, K., Khanna, M., 2000. Economies of scale in large containerships: Optimal size and geographical implications. Journal of Transport Geography 8, 181–195.

DARPA, 2016. Aircrew labor in-cockpit automation system (alias). Defense Advanced Research Projects Agency URL: https://www.darpa.mil/program/aircrew-labor-in-cockpit-automation-system.

De Jong, G., Behrens, C., Van Ommeren, J., 2019. Airline loyalty (programs) across borders: A geographic discontinuity approach. International Journal of Industrial Organization 62, 251–272.

Farsi, M., Fetz, A., Filippini, M., 2007. Economies of scale and scope in local public transportation. Journal of Transport Economics and Policy 41, 345–361.

Fosgerau, M., Engelson, L., 2011. The value of travel time variance. Transportation Research Part B: Methodological 45, 1–8.

Freed, J., Hepher, T., 2018. Freight first as jetmakers study single-pilot airplanes. Reuters URL: https://www.reuters.com/article/us-singapore-airshow-pilotless/freight-first-as-jetmakers-study-single-pilot-airplanes-idUSKBN1FS0SV.

GAO, 2014. Aviation workforce: Current and future availability of airline pilots.

Gayle, P.G., 2007. Is virtual codesharing a market segmenting mechanism employed by airlines? Economics Letters 95, 17–24.

Gayle, P.G., 2008. An empirical analysis of the competitive effects of the delta/continental/northwest code-share alliance. The Journal of Law and Economics 51, 743–766.

Gayle, P.G., 2013. On the efficiency of codeshare contracts between airlines: Is double marginalization eliminated? American Economic Journal: Microeconomics 5, 244–73.

Gayle, P.G., Brown, D., 2014. Airline strategic alliances in overlapping markets: Should policymakers be concerned? Economics of Transportation 3, 243–256.

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, 178–194.

Gayle, P.G., Wu, C.Y., 2014. On the extent to which the presence of intermediate-stop (s) air travel products influences the pricing of nonstop air travel products. Review of Network Economics 13, 355–395.

Gayle, P.G., Xie, X., 2018. Entry deterrence and strategic alliances. Economic Inquiry 56, 1898–1924.

Gayle, P.G., Xie, X., 2019. Firms' markup, cost, and price changes when policymakers permit collusion: Does antitrust immunity matter? Journal of Economic Behavior & Organization 157, 680–707.

Gayle, P.G., Yimga, J.O., 2018. How much do consumers really value air travel on-time performance, and to what extent are airlines motivated to improve their on-time performance? Economics of Transportation 14, 31–41.

Givoni, M., Rietveld, P., 2009. Airline's choice of aircraft size–Explanations and implications. Transportation Research Part A: Policy and Practice 43, 500–510.

Givoni, M., Rietveld, P., 2010. The environmental implications of airlines' choice of aircraft size. Journal of Air Transport Management 16, 159–167.

Graham, J., Hopkins, C., Loeber, A., Trivedi, S., 2014. Design of a single pilot cockpit for airline operations, in: 2014 Systems and Information Engineering Design Symposium (SIEDS), IEEE. pp. 210–215.

Gschwender, A., Jara-Díaz, S., Bravo, C., 2016. Feeder-trunk or direct lines? Economies of density, transfer costs and transit structure in an urban context. Transportation Research Part A: Policy and Practice 88, 209–222.

Harris, R.G., 1977. Economies of traffic density in the rail freight industry. The Bell Journal of Economics , 556–564.

ICAO, 2011. Global and Regional 20-year Forecasts: Pilots, Maintenance Personnel, Air Traffic Controllers. International Civil Aviation Organization.

Israel, M., Keating, B., Rubinfeld, D.L., Willig, B., 2013. Airline network effects and consumer welfare. Review of Network Economics 12, 287–322.

Jara-Díaz, S.R., Cortés, C.E., Morales, G.A., 2013. Explaining changes and trends in the airline industry: Economies of density, multiproduct scale, and spatial scope. Transportation Research Part E: Logistics and Transportation Review 60, 13–26.

Kim, H.Y., Clark, R.M., 1988. Economies of scale and scope in water supply. Regional Science and Urban Economics 18, 479–502.

Koltz, M.T., Roberts, Z.S., Sweet, J., Battiste, H., Cunningham, J., Battiste, V., Vu, K.P.L., Strybel, T.Z., 2015. An investigation of the harbor pilot concept for single pilot operations. Procedia Manufacturing 3, 2937–2944.

Lachter, J., Brandt, S.L., Battiste, V., Ligda, S.V., Matessa, M., Johnson, W.W., 2014. Toward single pilot operations: Developing a ground station, in: Proceedings of the International Conference on Human-Computer Interaction in Aerospace, pp. 1–8.

Lakew, P.A., 2014. Economies of traffic density and scale in the integrated air cargo industry: The cost structures of Fedex Express and UPS airlines. Journal of Air Transport Management 35, 29–38.

Lim, Y., Bassien-Capsa, V., Ramasamy, S., Liu, J., Sabatini, R., 2017a. Commercial airline single-pilot operations: System design and pathways to certification. IEEE Aerospace and Electronic Systems Magazine 32, 4–21.

Lim, Y., Ramasamy, S., Gardi, A., Sabatini, R., et al., 2017b. A virtual pilot assistant system for single pilot operations of commercial transport aircraft, in: Proceedings of the 17th Australian International Aerospace Congress (AIAC 2017), pp. 26–28.

Liu, J., Gardi, A., Ramasamy, S., Lim, Y., Sabatini, R., 2016. Cognitive pilot-aircraft interface for single-pilot operations. Knowledge-Based Systems 112, 37–53.

Luttmann, A., 2019. Are passengers compensated for incurring an airport layover? Estimating the value of layover time in the US airline industry. Economics of Transportation 17, 1–13.

Morikawa, M., 2011. Economies of density and productivity in service industries: An analysis of personal service industries based on establishment-level data. The Review of Economics and Statistics 93, 179–192.

Morrison, S.A., Winston, C., Bailey, E.E., Kahn, A.E., 1989. Enhancing the performance of the deregulated air transportation system. Brookings Papers on Economic Activity. Microeconomics 1989, 61–123.

Myers, P.L., Starr, A.W., 2021. Single pilot operations in commercial cockpits: Background, challenges, and options. Journal of Intelligent & Robotic Systems 102, 1–15.

Nauges, C., Van den Berg, C., 2008. Economies of density, scale and scope in the water supply and sewerage sector: A study of four developing and transition economies. Journal of Regulatory Economics 34, 144–163.

Pai, V., 2010. On the factors that affect airline flight frequency and aircraft size. Journal of Air Transport Management 16, 169–177.

Park, K., 2017. Airbus looking forward to a pilotless future. Bloomberg URL: `https://www.bloomberg.com/news/articles/2017-11-21/airbus-exploring-single-pilot-autonomous-aircraft-in-tech-race`.

Peters, C., 2006. Evaluating the performance of merger simulation: Evidence from the US airline industry. The Journal of law and economics 49, 627–649.

Reid, D., 2019. Some airlines want boeing's new "797" to fly with just one pilot on board. CNBC URL: `https://www.cnbc.com/2019/05/20/boeings-new-797-could-be-built-to-fly-with-just-one-pilot-on-board.html`.

Roberts, M.J., 1986. Economies of density and size in the production and delivery of electric power. Land Economics 62, 378–387.

Ryerson, M.S., Hansen, M., 2013. Capturing the impact of fuel price on jet aircraft operating costs with Leontief technology and econometric models. Transportation Research Part C: Emerging Technologies 33, 282–296.

Savage, I., 1997. Scale economies in United States rail transit systems. Transportation Research Part A: Policy and Practice 31, 459–473.

Shen, C., 2017. The effects of major US domestic airline code sharing and profit sharing rule. Journal of Economics & Management Strategy 26, 590–609.

Stanton, N.A., Harris, D., Starr, A., 2014. Modelling and analysis of single pilot operations in commercial aviation, in: Proceedings of the International Conference on Human-Computer Interaction in Aerospace, pp. 1–8.

Takebayashi, M., 2011. The runway capacity constraint and airlines' behavior: Choice of aircraft size and network design. Transportation Research Part E: Logistics and Transportation Review 47, 390–400.

Talley, W.K., Agarwal, V.B., Breakfield, J.W., 1986. Economies of density of ocean tanker ships. Journal of Transport Economics and Policy , 91–99.

Train, K.E., 2009. Discrete choice methods with simulation. Cambridge university press.

Vu, K.P.L., Lachter, J., Battiste, V., Strybel, T.Z., 2018. Single pilot operations in domestic commercial aviation. Human Factors 60, 755–762.

Warwick, G., 2013. NASA outlines civil-aviation autonomy research plans. Aviation Week & Space Technology , 48.

Wei, W., 2006. Impact of landing fees on airlines' choice of aircraft size and service frequency in duopoly markets. Journal of Air Transport Management 12, 288–292.

Wei, W., Hansen, M., 2003. Cost economics of aircraft size. Journal of Transport Economics and Policy 37, 279–296.

Wei, W., Hansen, M., 2005. Impact of aircraft size and seat availability on airlines' demand and market share in duopoly markets. Transportation Research Part E: Logistics and Transportation Review 41, 315–327.

Winston, C., Yan, J., 2015. Open skies: Estimating travelers' benefits from free trade in airline services. American Economic Journal: Economic Policy 7, 370–414.

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

Zhang, Y., 2014. The puzzle of aircraft size and traffic growth. Journal of Transport Economics and Policy 48, 465–482.