Bruno Henrique Castelo Branco

Aviation Technology and Air Traffic Networks

Dissertação de Mestrado

Thesis presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor : Prof. Pedro Carvalho Loureiro de Souza
Co-advisor: Prof. Leonardo Bandeira Rezende

Rio de Janeiro
July 2017

Bruno Henrique Castelo Branco

Aviation Technology and Air Traffic Networks

Thesis presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Economia. Approved by the undersigned Examination Committee.

Prof. Pedro Carvalho Loureiro de Souza
Advisor
Departamento de Economia – PUC-Rio

Prof. Leonardo Bandeira Rezende
Co-advisor
Departamento de Economia – PUC-Rio

Prof. Fabio Miessi Sanches
Departamento de Economia – PUC-Rio

Prof. João Paulo Cordeiro de Noronha Pessoa
Escola de Economia de São Paulo – EESP/FGV

Prof. Augusto Cesar Pinheiro da Silva
Vice Dean of Graduate Studies
Centro de Ciências Sociais – PUC-Rio

Rio de Janeiro, July the 28th, 2017

Bruno Henrique Castelo Branco

The author holds a Bachelor's degree in Economics from Fundação Getulio Vargas (EPGE/FGV).

Acknowledgments

To my advisors Pedro Souza and Leonardo Rezende, for the countless comments and suggestions from the planning to the execution of this thesis. Professors Fabio Miessi and Nathalie Gimenes also helped me with useful comments.

To PUC-Rio, for providing an exceptional Department of Economics, a vibrant academic environment, and a remarkably beautiful campus. I thank all the administrative staff of the Department of Economics, especially Maria das Graças and professor Gustavo Gonzaga.

I have also learned a lot from my colleagues from the master's program, who are amongst the most capable people I have ever met. I consider a privilege to have studied alongside them and I'm thankful for it. With a special mention to Pablo 'Pableti' Rocha.

I am also grateful to Lucas A. de Lima, for his friendship and all the help and advice provided since our undergraduate years.

Lastly, but most important of all, I thank my family, Margarida, and Gabriella.

Abstract

Castelo Branco, Bruno Henrique; Carvalho Loureiro de Souza, Pedro (Advisor); Bandeira Rezende, Leonardo (Co-Advisor). **Aviation Technology and Air Traffic Networks**. Rio de Janeiro, 2017. 61p. Dissertação de mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This paper studies to what extent the development of new aircraft shapes airlines' network structure. I argue that modern aircraft are more efficient and well suited to operate flights between smaller and less central cities, hence favoring the service of more markets in the periphery of the network. Using U.S. air traffic data, I employ a discrete choice framework to model airlines' entry decisions and the subsequent aircraft choice to each market. Counterfactual experiments show that had aircraft technology ceased to improve in 1999, the air traffic network as a whole would be more centralized, airlines would be operating more hub-centered networks, reaching fewer cities, and serving fewer markets.

Keywords

Airline industry;   Airline networks;   Technological impact;   Discrete Choice;   Counterfactual experiments.

Resumo

Castelo Branco, Bruno Henrique; Carvalho Loureiro de Souza, Pedro; Bandeira Rezende, Leonardo. **Tecnologia de Aviação e Redes de Tráfego Aéreo**. Rio de Janeiro, 2017. 61p. Dissertação de Mestrado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Esse estudo investiga em que medida o desenvolvimento e introdução de novas aeronaves moldam a estrutura da rede das companhias aéreas. Argumento que aeronaves modernas são mais eficientes e adequadas para operar voos entre cidades menores e menos centrais, favorecendo assim o serviço em mais mercados na periferia da rede. Com dados sobre o tráfego aéreo dos Estados Unidos, utilizo um arcabouço de escolha discreta para modelar as decisões de entrada das companhias e a subsequente escolha de aeronave em cada mercado. Experimentos contrafactuais mostram que, caso o desenvolvimento de tecnologia tivesse cessado em 1999, a rede de tráfego aéreo como um todo estaria mais centralizada, a maioria das companhias estariam operando redes mais centradas em torno de hubs, alcançando menos cidades e servindo menos mercados.

Palavras-chave

Indústria aeronáutica; Rede aéreas; Impacto tecnológico; Escolha Discreta; Experimentos contrafactuais;

Table of contents

List of figures

List of tables

1
Introduction

Aviation is highly technological and one of the most innovative industries in the world, aspects that may have consequences to firms decisions in several dimensions. In this study, I examine the impact of the introduction of new aircraft on airlines' decisions regarding market entry in the process of network formation. The rationale is twofold: first, a simple reduction of operating expenses, caused by the improvement in fuel economy in a passenger/mile basis; second, the adequacy of each aircraft to each market - features such as maximum range and number of seats make a difference when considering whether to operate each market, especially long and low-density ones. Due to its increased efficiency, newer aircraft allow a larger number of nonstop flights between small cities, operating markets that would not be considered cost-effective under older technology.

Using data on flights operated in the United States during the years of 1999 and 2015, I employ a discrete choice framework to model aircraft choice in each market, defined as an undirected city pair. Firms choose whether or not to enter a market, and in the positive case, the aircraft with which the flight will be operated. Firms are concerned with the revenue generated by market characteristics, operational characteristics, and the fuel expense, which I calibrate with fuel prices and fuel burn estimates from the European Environment Agency. By comparing estimates, I provide evidence that the importance of fuel efficiency to entry decisions increased over the period. I also perform the counterfactual experiment for the ceasing of technological improvements in aircraft in 1999, and show that most airlines would be reaching less cities and serving less markets in their networks.

Another relevant finding of this paper is that the development of new aircraft have an impact towards shifting firms' network structure to a more point-to-point system - rather than the typicall hub-and-spoke configuration. In a pure point-to-point network, every route is made of only one market, and passengers board at the origin airport and deplane only at the final destination. On the other hand, a pure hub-and-spoke configuration means that every route must pass through the hub airport, and routes may have one (in case the origin or destination is the hub itself) or two segments.

Figure 1.1: **Hub-and-spoke (left) and point-to-point (right) configuration.**
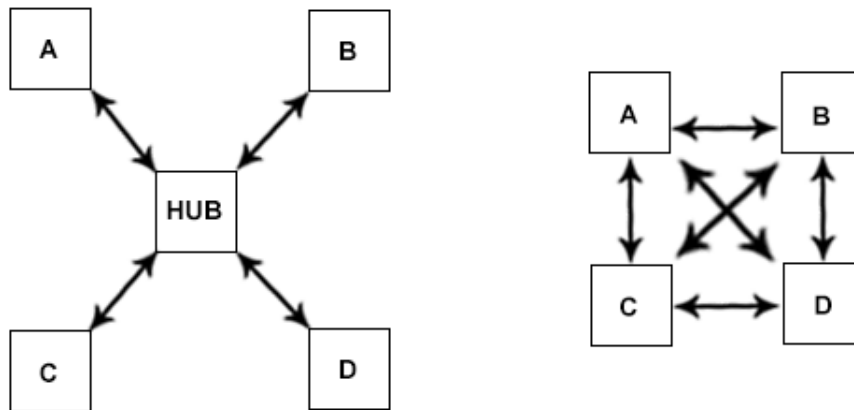
In a point-to-point network, routes are made only of nonstop flights, whereas in a hub-and-spoke configuration, passengers must make a connection in the hub airport before reaching the final destination.

In practice, airlines operate hybrid networks, varying its centralization, that is, how close they are to a pure hub-and-spoke architecture. This difference can be exemplified by the network structure of Southwest and Delta, depicted in Figure 1.2. Whereas Southwest is known for its point-to-point network, Delta's operations revolve around its hubs.

Figure 1.2: **Network representation for Delta and Southwest.**
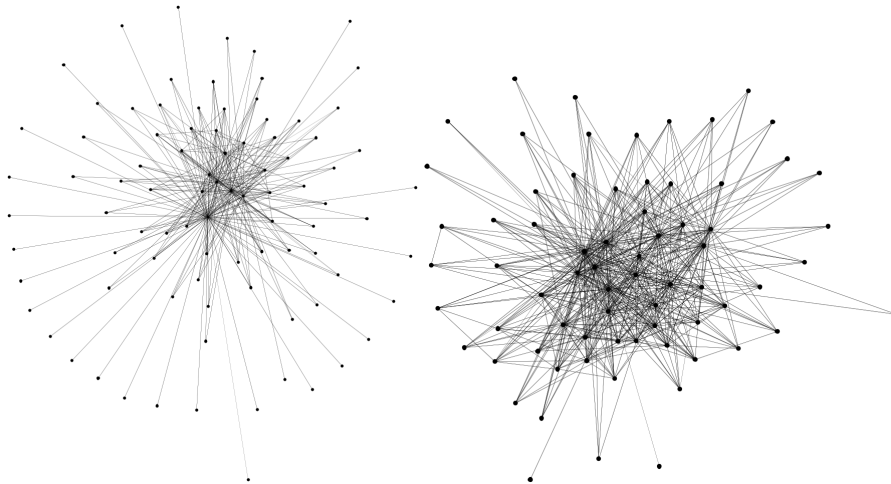
On the left, Delta's hub-and-spoke network structure resembles a star, whereas Southwest's (right) is more diffuse, closer to a point-to-point system.[1]

A shift in network structure is an interesting phenomenon because the idea of economies of scope and its advantages emerging as a result of hub size are well known in the literature. The idea is that by aggregating passengers

[1]The layout algorithm that generated these graphs disposes the nodes in such a manner that all of them are visible and also considering the distance between them.

from different origins and final destinations in the same flight, firms can increase the average number of passengers per flight and use larger aircraft, which tend to be more fuel efficient. Economies of density would lead to economies of scope among different itineraries with shared flights and induce a hub-and-spoke system.

Caves, Christensen, and Tretheway (1984) were the first to develop a model that differentiated between returns to density and returns to scale, and found evidence of substantial positive effects on profits in both cases. Berry, Carnall, and Spiller (2006) show that in 1985 a considerable share of the consumers was willing to pay an average premium of 20% for perks related to large hubs. In their model, the economies of density assumption is incorporated through a marginal cost function that is decreasing in the number of passengers.

Berry and Jia (2010) use a model based in BCS to study how the industry evolved over the period of 1999 to 2006, and show evidences of changes in demand towards a higher disutility of connecting flights, as passengers seem less likely to pay for hub-related features. This finding is consistent with those of Borenstein (2005) and Borenstein and Rose (2007), which pointed out a decline in hub premium over the years. According to them, the cost advantage of connecting flights through hub airports that existed in 1999 disappeared in 2006.

Whereas Berry and Jia (2010) call attention to a 'dehubbing' trend[2] and relate it to preference changes that favor nonstop flights, I explore the technological channel, which may have been overlooked in the past literature. It's interesting to observe that there is a possible endogeneity in this process, and both explanations would be complementing each other rather than competing: perhaps new aircraft made more nonstop flights available, which in turn may have had an effect on consumer preferences, and vice versa.

Flights between smaller cities usually have a relatively lower demand, making conventional jets too large in seat count to operate at profitable load factor levels. Besides, large jets need more ground structure than small ones, such as longer runways, bigger hangars, and more ground equipment that are typically only available at larger airports. In the recent past, smaller aircraft were mostly turboprops, which have a lower seat count, but lack flight range in most cases. For some markets, the problem could be solved by using turboprops to feed hubs, and jets to connect one hub to another. Nowadays, jets are

---

[2]They point that besides an increasing number of direct flights during the 1999-2006 period, most hubs were serving fewer connecting passengers. Refer also to Redondi, Malighetti, and Paleari (2012) for further discussion and examples.

tailored to meet the needs of a wider range of markets and thus are able to provide point-to-point service more frequently.

A remarkable example of improvement in aviation technology is the case of regional jets, whose popularity increased in the 1990s along with its main manufacturers Embraer and Bombardier. These jets have a lower seat count than conventional ones and comparable flight range, hence they are more appropriate to serve low-density markets. The decision of serving a market then becomes less dependent on city and hub sizes, and more links in the periphery of the network are created. When this happens, the number of links created between less central cities increase, the hubs have its importance diminished, and eventually a dehubbing event develops.

The present study is close to the work of Aguirregabiria and Ho (2012), which essentially extends the theoretical model of static duopoly in Hendricks et al (1999) to a dynamic framework with N firms and an empirical application. They treat network formation as endogenous, and the airlines are allowed to choose routes in a dynamic game, taking into account the economies of density and scale that reduce fixed and marginal costs as entry decisions are made and hub sizes increase. Their goal is to disentangle demand, costs, and strategic factors as influences to the adoption of a hub-and-spoke system.

This paper contributes to the economics of the airline industry literature by examining how central aircraft technology availability is to market entry decisions. To the extent of my knowledge, this is the first paper to investigate such matter and to use fuel burn data to model operating expenses in trying to evaluate the relevance of aircraft development and fuel efficiency to airline's network configuration.

2
Model and Estimation

Every period t, each firm i decides in which market (defined as a city pair) m it will operate nonstop flights by choosing an aircraft $j \in J_t$, where J_t is an aircraft menu, and a flight frequency $\mathcal{F}_{imjt}$. The list of aircraft available to choose from at a given period is the same for every firm and comprises all the aircraft used by at least one firm in that period. Among these options I include the outside alternative $j = 0$, which represents not operating in that market, and normalize its profit to 0.

From the perspective of the researcher, the profits collected by each combination of firm, market, period, and aircraft are composed by an observable and an unobservable portion, represented by $\overline{\Pi}_{imjt}(.)$ and ε_{imjt}, respectively. The observable part is a function of market characteristics, the aircraft's characteristics, and the flight frequency chosen by the airline.

Let X_j, Z_{imt} and $\mathcal{F}_{imj}$ be the vector of aircraft j characteristics, market m characteristics for firm i in period t, and the flight frequency chosen for that aircraft in that market and period, respectively. In the vector of market characteristics, Z_{imt}, I include the population, the hub size in each of the end cities, the distance between them (market distance) and the distance squared.

Then the profit Π_{imjt} collected by firm i when choosing aircraft j in market m and period t is:

$$\Pi_{imjt}(X_j, Z_{imt}, \mathcal{F}_{imjt}) = \overline{\Pi}_{imjt}(X_j, Z_{imt}, \mathcal{F}_{imjt}) + \varepsilon_{imjt}$$

The model is partially motivated by the data. Profits are observed on a quarterly basis, but the fuel burn data is reported by distance flown. Thus, we know the fuel expense of a single flight in a given market but we still need a flight frequency to obtain the quarterly fuel expense. To deal with this issue, I model the firm decision in two stages: first it chooses an aircraft, and then a flight frequency for that aircraft in that period.

Assume that for each aircraft in a given market and period there is an optimal flight frequency $\mathcal{F}_{imjt}^*$ that maximizes the profit for that aircraft:

$$\Pi_{imjt}(X_j, Z_{imt}, \mathcal{F}_{imjt}^*) \geq \Pi_{imjt}(X_j, Z_{imt}, \mathcal{F}_{imjt}) \ \forall \ \mathcal{F}_{imjt} \in \mathbb{R}^+$$

Then I solve this problem backwards. First, I find the optimal frequency for each (aircraft, market, period) and plug it back into the profit equation. This enables us to resolve the first stage, which consists in the aircraft choice (or outside alternative), considering that each one would be operated at its optimal flight frequency.

Consequently, the profit collected by firm i when operating flights in market m in period t with aircraft j is given by:

$$\Pi_{imjt} = \overline{\Pi}_{imjt}(X_j, Z_{imt}, \mathcal{F}^*_{imjt}) + \varepsilon_{imjt}$$

The observable part of the profit can be broken down in Operating Revenue (OR_{imjt}), Variable Cost (VC_{imjt}), and Fixed Cost (FC_{imt}):

$$\overline{\Pi}_{imjt}(X_j, Z_{imt}, \mathcal{F}_{imjt}) = OR_{imjt} - VC_{imjt} - FC_{imt}$$

Operating Revenue

The total number of passengers demanding a flight in each market depends not only on the end cities characteristics but also on the flight frequency. More frequent flights are more likely to fulfill the passengers' needs (easier to fit their schedules, for instance) and convert potential consumers into actual passengers.[1] As the frequency increases, this conversion rate also increases, but at diminishing rate of returns. Therefore, we'd expect the revenue to be concave in the frequency, and to make subsequent calculations easier, I model it as a function of the square root of the frequency in the market and period:

$$OR(Z_{imt}, \mathcal{F}_{imj}) \equiv Z_{imt}\beta + \delta\sqrt{\mathcal{F}_{imjt}}$$

Variable Costs

During the period 1999-2015, fuel expenses have always stood out as the largest share of operational expenses, ranging from 20% to 35%. For this reason, and to emphasize the interest in fuel efficiency, variable costs are summarized by the expense generated by the fuel cost of operating flights in a given market plus an aircraft fixed effect.

---

[1] see Hansen and Liu (2015) for a more detailed discussion on the effects of frequency on demand.

Motivated by the data,[2] the fuel burned is assumed to be a deterministic function of distance conditional on the aircraft type, considering that other factors that affect fuel consumption - such as payload weight, cruise altitude and speed, temperature, humidity and wind speed - are at the most frequently observed levels for each aircraft. I use p_{it} to represent the average fuel price payed by firm i in period t; FB_{jm} to represent the fuel burned when the aircraft j operates market m; and γ_j^a as a fixed effect for aircraft. The aircraft fixed effect captures the net effect of all aircraft specific aspects other than fuel efficiency. The variable cost is then defined as:

$$VC_{imjt} \equiv p_{it}\mathcal{F}_{imjt}FB_{jm} + \gamma_j^a$$

Note also that aircraft purchase prices could be significant when considering entry with different aircraft, but the prices remained roughly constant across the period. As an example, a Boeing 767-200ER could be purchased in 1999 by something around \$140,000,000 (in 2015 dollars) its price was in the ballpark of \$150,000,000. Another noteworthy point is that it is usual for airlines to lease aircraft instead of purchasing, and leasing rates have also remained more or less constant (when correct for inflation) in the period. Since there is no data on whether each aircraft was purchased or leased, and the in what conditions the deal was made, these details are not included in the model.

Fixed Costs

Fixed costs are summarized by fixed effects minus hub-size effects, which arise because of the economies of scale already discussed. It's important to observe that as the model is not dynamic, we cannot disentangle fixed and sunk entry costs, therefore the interpretation of the fixed effects must take this into account as they do not reflect solely the fixed costs, and are biased upwards. Firm and city fixed effects capture costs such as landing and takeoff slot prices.

Fixed and entry costs depend on the scale of operation of the airline. Fixed costs can be lowered by a large scale of operation as the airline is able to negotiate better slot prices and better contracts with ground handling providers such as cleaning, catering and maintenance. Similarly, starting

---

[2]Fuel burn per distance estimates based on flight movement data, available for download at the European Environment Agency website: https://www.eea.europa.eu/themes/air/emep-eea-air-pollutant-emission-inventory-guidebook

operations in a new airport is considerably more expensive than just adding one connection to a new destination, and entry costs in markets with large hubs are expected to be associated with lower costs.

Let γ_i^f be a fixed effect for firm; γ^t a fixed effect for period; γ_m^O and γ_m^D represent fixed effects for origin and destionation cities in the market; H_{imt}^O the hub size for firm i in the origin city in period t, and H_{imt}^D the same for the destination city. The purpose of the measures of hub size is to capture economies of scope that arise when the scale of operations of an airline increase in a given airport - and I also include the lags to take into account the dynamics of the entry process. I define the hub size for firm i in city c as the number of different cities this firm connects to city c with nonstop flights. Fixed costs are then defined as:

$$FC_{imt} \equiv \gamma_i^f + \gamma^t + \gamma_m^O + \gamma_m^D - \gamma_1 H_{imt}^O - \gamma_2 H_{imt}^D - \gamma_3 H_{im,t-1}^O - \gamma_4 H_{im,t-1}^D$$

Optimal Frequency

Bringing all together, the observable profit is given by:

$$\overline{\Pi}_{imjt} = Z_{imt}\beta + \delta\sqrt{\mathcal{F}_{imjt}} + \gamma_j^a - p_{it}\mathcal{F}_{imjt}FB_{jm}$$
$$- \left(\gamma_i^f + \gamma^t + \gamma_m^O + \gamma_m^D - \gamma_1 H_{it}^O - \gamma_2 H_{it}^D - \gamma_3 H_{im,t-1}^O - \gamma_4 H_{im,t-1}^D\right)$$

And as explained before in this section, I assume a profit maximizing behavior to find the optimal frequency. For the sake of simplicity I assume the frequency to be a continuous variable, and take the derivative:

$$\frac{\partial \Pi_{imjt}}{\partial \mathcal{F}_{imjt}} = \frac{\delta}{\sqrt{\mathcal{F}_{imjt}}} - p_{it}FB_{jm}$$

Setting it equal to zero and solving for frequency results in:

$$\mathcal{F}^*{}_{imjt} = \left(\frac{\delta}{2p_{it}FB_{jm}}\right)^2$$

Thus, using Π_{imjt}^* and $\overline{\Pi}_{imjt}^*$ to represent the profit and its observable part attained by firm i when operating flights in market m with aircraft j in period t at the aircraft's optimal frequency, the reduced form profit, which is relevant when comparing alternatives, is given by:

$$\Pi^*_{imjt} = \overline{\overline{\Pi}}_{imjt}\left(X_j, Z_{imt}, \mathcal{F}^*{}_{imjt}\right) + \varepsilon_{imjt}$$

$$\Pi^*_{imjt} = Z_{imt}\beta + \gamma^a_j + \frac{\delta^2}{4p_{it}FB_{jm}}$$
$$- \left(\gamma^f_i + \gamma^t + \gamma^O_m + \gamma^D_m - \gamma_1 H^O_{imt} - \gamma_2 H^D_{imt} - \gamma_3 H^O_{im,t-1} - \gamma_4 H^D_{im,t-1}\right) + \varepsilon_{imjt}$$

Finally, in order to enhance the identification of the parameters of interest, I subtract the market mean of the right-hand side of the profit equation whenever possible. Define for a generic variable $\Delta k = k - \overline{k}$, where $\overline{k}$ is the market mean. This procedure is required because due to the size of the network, it is computationally infeasible to introduce market dummies. We are left with:

$$\Pi^*_{imjt} = Z_{imt}\beta + \gamma^a_j + \frac{\delta^2}{4p_{it}FB_{jm}}$$
$$- (\gamma^f_i + \gamma^t + \gamma^O_m + \gamma^D_m - \gamma_1 \Delta H^O_{imt} - \gamma_2 \Delta H^D_{imt} - \gamma_3 \Delta H^O_{im,t-1} - \gamma_4 \Delta H^D_{im,t-1}) + \varepsilon_{imjt}$$

Estimation

Firm i chooses to operate nonstop flights in the market m in period t using the aircraft j if this choice results in the highest profit among all alternatives, including not operating in this market. Note that as a consequence of the profit maximizing behavior, if a firm chooses a certain aircraft to operate flights in a market, it will do so only at its optimal frequency. Therefore we need only to compare the profits generated by each choice evaluated at its optimal frequency, and aircraft j is chosen by firm i to operate market m in period t if, and only if:

$$\Pi^*_{imjt}(X_j, Z_{imt}, \mathcal{F}^*_{tmjt}) \geq \Pi^*_{imkt}(X_k, Z_{imt}, \mathcal{F}^*_{tmkt}), \ \forall k \neq j$$

Because of the unobservable portion, we can define the inequalities in probabilistic terms. Using y_{imt} to represent the aircraft (or outside alternative) choice of firm i in market m and period t, we can represent the probability of choosing aircraft j conditional on the market and its characteristics as:

$$\Pr\left(y_{imt} = j | Z_{imt}, X_j\right) = \Pr\left(\Pi^*_{imjt} > \Pi^*_{imkt} \ \forall k \neq j\right)$$
$$= \Pr\left(\overline{\Pi}^*_{imjt} + \varepsilon_{imjt} > \overline{\Pi}^*_{imkt} + \varepsilon_{imkt} \ \forall k \neq j\right)$$
$$= \Pr\left(\varepsilon_{imkt} < \overline{\Pi}^*_{imjt} - \overline{\Pi}^*_{imkt} + \varepsilon_{imjt} \ \forall k \neq j\right)$$

In order to build a tractable likelihood function, I make a set of assumptions similar to those in Aguirregabiria and Ho (2012). First, we need decentralization assumptions regarding the firm decision: (1) each firm has, for each market in the network, a manager (i, m) who decides whether the firm operates flights (and with which aircraft) in the market or not; (2) each manager has his own way of dealing with each aircraft, which is independent between aircraft and managers even when in the same firm. Albeit strong assumptions, they are needed due to the large dimension of the industry network. Without the decentralization assumption, for each period the likelihood would have to take into account every combination of aircraft choice for each market and firm given every possible network of each of the other firms, amounting to something around $30^{14*6328}$ possibilities.[3]

I also make distributional assumptions about the random shocks: (3) each of the random shocks follows Type I Extreme Value (or Gumbel) distribution, with density $f(\varepsilon_{imjt}) = e^{-\varepsilon_{imjt}} e^{-e^{-\varepsilon_{imjt}}}$ and cumulative $F(\varepsilon_{imjt}) = e^{-e^{-\varepsilon_{imjt}}}$. This assumption is standard in discrete choice models, as it is required by the conditional logit set up and it is also useful to explain choices deviating from what we would expect from the observable part; and (4) the random shocks are independently distributed over time.

Assuming that the random portion of the profits are independently distributed across firms, markets, and choices is convenient because it explains heterogeneity but stills avoids endogeneity problems. Given the set of assumptions made, we can define the conditional choice probability as:

$$\Pr\left(y_{imt} = j | Z_{imt}, X_j\right) = \int \left(\prod_{j \neq i} e^{-e^{-\left(\overline{\Pi}^*_{imjt} - \overline{\Pi}^*_{imkt} + \varepsilon_{imjt}\right)}}\right) e^{-\varepsilon_{imjt}} e^{-e^{-\varepsilon_{imjt}}} d\varepsilon_{imjt}$$

As shown in McFadden (1974), this integral has a closed form expression:

$$\Pr\left(y_{imt} = j | Z_{imt}, X_j\right) = \frac{\exp\left(\overline{\Pi}^*_{imjt}\right)}{\sum_{k=0}^{J} \exp\left(\overline{\Pi}^*_{imkt}\right)}$$

And finally, we can write the log-likelihood for firm i:

$$\ell_i(\theta; y_{imt} | Z_{imt}, X_j) = \sum_{t=1}^{T} \sum_{m=1}^{M} \sum_{j=1}^{J} 1\left[y_{imt} = j\right] \ln\left[Pr(y_{imt} = j | Z_{imt}, X_j)\right]$$

where $\theta \equiv (\delta, \gamma_j, \gamma_i, \gamma_m^O, \gamma_m^D \gamma_1, \gamma_2)$ is the vector of parameters.

[3]Considering 14 firms, 29 aircraft and 6328 markets in each period, which is the case in 2015. For comparison, the estimated number of atoms in the known universe is between 10^{78} and 10^{82}.

3
Data

The main database used in this study is the T-100 Domestic Segment, provided by the American Bureau of Transportation Statistics (BTS). It keeps monthly records of every scheduled nonstop flight in the United States, along with several information about each of them, such as the aircraft and airline that operated the flight, origin, destination, flight time, frequency, distance, available seats, and passengers. It amounts to 303,744 observations in 1999 and 354,368 in 2015, and allows us to observe the network structure in detail in each year.

The T-100 database provides information on passengers transported, but not the price paid for the tickets. For this reason, I use the DB1B database to calculate average fares charged by each firm in each market and quarter. The DB1B database, also known as the Airline Origin and Destination Survey, is a 10% random sample of all issued airline tickets - which corresponds to 17,116,437 tickets in 1999 and 24,836,077 tickets in 2015 - and provides information such as the fare charged, origin, destination, and which carrier operated the flight, but not the aircraft type.

Using the tuple (origin, destination, airline, quarter) available in both data sets, I can calculate and match the average fares from the DB1B to the network information supplied by the T-100. Data on city population is available in the U.S. Census Bureau website, and the GDP per capita for each metropolitan area is provided by the U.S. Bureau of Economic Analysis.

Working sample

In order to make inference about network centralization, I hold fixed the nodes to examine the same network in different moments. If we let the nodes vary, then the resulting change in centrality could be just in response to a new network configuration, not necessarily meaning an increase or decrease in concentration. Therefore, I select all flights between cities with a population of at least 100,000 in 1999 as the network to be studied, which amounted to 224 cities or nodes in this network. Some of these cities are not in the continental U.S., and I exclude them. Also, there are cities without primary

airports - 10,000 enplanements or more per year - and cities that belong to the same metropolitan areas or share the same airports, in which case I add their population and consider them as a single node in the network (as if they were a single "city"). Finally, there are 113 cities left, the nodes of the network considered in the estimation afterwards. I exclude firms with less than 1% of market share or that ceased operations during each given year - for instance, US Airways was acquired by American and ceased to report flight data in the last quarter of 2015.

The costs involved in the operation of large jet aircraft are considerably different from those of small, piston-engine aircraft (e.g. a Boeing 777 vs Cessna 172) and with this in mind, I exclude flights operated by aircraft that does not meet a minimum of 30 seats. I consider a firm to be active in a market in a given quarter if it operates roundtrip flights in this market at a minimum frequency of 16 flights (per quarter, which roughly corresponds to 2 flights a week), and serving at least 80 passengers a month. The restriction on minimum passengers per quarter is also useful to avoid flights whose primary purpose is not to transport passengers, be it aircraft relocation due to maintenance or any other reasons. Other criteria to define whether a firm is active were tested without great changes in the resulting set of aircraft or the proportion of choices of each aircraft, and I decided to choose the less restrictive one while still being consistent with past literature.

In the construction of the fares data set, I exclude tickets marked as bulk fare and with more than 1 coupon (a segment or market composing the route). I also follow previous literature in excluding the outliers - the top 5% more expensive or cheapest tickets - because they may represent coding errors. I consider the price of a flight in a given quarter as the average fare charged by that firm in that market and quarter. Some airlines operate the same market with more than one aircraft type. In those cases, I consider that the chosen aircraft for that market and quarter is the one in which the largest number of passengers were transported. I experimented different 'tiebreaker' criteria such as highest frequency, but again there were no significant changes in the resulting set of chosen aircraft or the proportion of choices of each aircraft.

In order to make a more accurate description of the hub relevance for each firm, I take into account codeshare agreements[1] and aggregate those firms which are not independent. Air Wisconsin, for example, operated as a feeder for United Airlines in 1999, and in 2015 it operated under contract to American.

---

[1]These are agreements in which more than one airline share the same flight, even though consumers may purchase seats in this flight from different firms. This is a common and advantageous business scheme because it allows firms to offer routes out of their operation network.

Firms operating as feeders usually have at its disposal all the structure and ground support of the parent carrier, and for this reason it is important to aggregate them to assess the importance of hubs. Following Aguirregabiria and Ho (2012), I assume that the reporting carrier is the firm that pays the cost and and receives the revenues for providing the service.

The DB1B database reports for each ticket the operating carrier and reporting carrier. The first one is the airline that actually operated the flight (i.e. provided the aircraft and staff), whereas the second one is the firm that submitted the ticket information to the Bureau of Transportation. Most of the time the operating carrier and reporting carrier are the same airline, but when that is not the case I assume that the reporting carrier is the firm which pays for the expenses and collect the revenues of the flight. I observe 12 firms in 1999 and 14 in 2015, in 4 different quarters in each year, making entry decisions in 6,328 markets. This results in a total of 303,744 decisions in 1999 and 354,368 in 2015.

Descriptive statistics

First, it is interesting to observe that the number of passengers traveling in nonstop flights increased considerably in relation to connecting flights. Table 3.1 shows the percentage change in the number of total passengers by the number of markets in the route they flew. There is a considerable decrease in the percentage of passengers in routes with 3 or more markets and pronounced increase the percentage of passengers in nonstop flights (1 market), which means that airlines are, on average, operating a more point-to-point network.

Table 3.1: **Passengers by number of markets in the route flown.**

Markets	1999	2015	% change
1	4,708,298	8,805,558	+87.02
2	5,844,916	7,097,796	+21.43
3	432,736	358,784	-17.08
4	56,333	26,709	-52.58
5+	5,336	2,281	-57.25
Total	11,047,619	16,291,128	+47.46

In 1999, 42.6% of all passengers were in nonstop flights, and this group represented the second largest share of all transported passengers. In 2015, the group of passengers in nonstop flights came to represent the largest share, rising to 54.0% of the total.

There is also a milder increase in the percentage of passengers in routes with 2 markets, probably reflecting a natural growth of the network, as more

cities are reached from the existing hubs. Observe that the total number of passengers transported in this network increased 47.46%, while the growth in passangers transported in nonstop flights was almost twice that in percentage points.

A market is a undirected city pair, and with 113 nodes in the network there are $\frac{113*112}{2} = 6,328$ markets or possible links. The number of active firms operating at least one flight between the cities of the selected network increased from 12 to 14; the average number of seats of aircraft used went down from 132 to 121, possibly reflecting a more intense use of regional jets - an observation also made by Berry and Jia (2010) when comparing the period of 1999 to 2006. Table 3.2 presents the number of markets by active airlines and shows that there are 150 new markets being served in 2015, an increase of 12.2%. It also shows an increase in competitiveness: markets with 5 or more firms increased from 21 to 237, with 4 firms increased from 33 to 153, with 3 firms increased from 112 to 231, while the number of monopolies decreased from 742 to 446.

Table 3.2: **Number of markets by active airlines.**

Active airlines	1999	2015
1	742	446
2	314	305
3	112	231
4	33	153
5 or more	21	237
Total	1,222	1,372

Tables 3.3 and 3.4 show the most active markets[2] by percentage of passengers transported in a given year. Observe that this is not a ranking by demand, it shows the total count of passengers transported between two cities for any purpose. It means that the passenger count for the market New York - Chicago includes connecting flights whose origin or final destination is neither New York nor Chicago, hence if the a passenger travels through a 2-market route, this person will be counted twice, one time for each market of the route.

[2]Order does not matter: the city pair Atlanta-Chicago is the same as Chicago-Atlanta.

Table 3.3: **Most flown markets in 1999.**

Market	Passengers	% of total	Airlines
Chicago, IL-New York, NY	4,236,515	0.86	6
New York, NY-Atlanta, GA	4,115,325	0.83	3
Dallas, TX-Houston, TX	4,024,962	0.81	4
New York, NY-Los Angeles, CA	3,832,615	0.77	6
Los Angeles, CA-Las Vegas, NV	3,607,327	0.73	9
Chicago, IL-Los Angeles, CA	3,522,521	0.71	3
New York, NY-Boston, MA	3,337,179	0.67	7
Washington, DC-New York, NY	3,090,660	0.62	7
San Francisco, CA-Los Angeles, CA	3,034,144	0.61	4
Los Angeles, CA-Phoenix,AZ	2,968,635	0.60	3

Table 3.4: **Most flown markets in 2015.**

Market	Passengers	% of total	Airlines
Chicago, IL-New York, NY	5,896,280	1.03	10
New York, NY-Los Angeles, CA	4,925,070	0.86	5
San Francisco, Ca-Los Angeles, CA	4,580,542	0.80	8
Dallas, TX-Los Angeles, CA	3,975,225	0.70	7
New York, NY-San Francisco, CA	3,894,413	0.68	5
New York, NY-Atlanta, GA	3,866,111	0.68	8
New York, NY-Orlando, FL	3,848,381	0.67	5
Chicago, IL-Los Angeles, CA	3,789,660	0.66	6
New York, NY-Miami, FL	3,653,488	0.64	6
New York, NY-Fort Lauderdale, FL	3,425,244	0.60	6

In a network, the degree of a node is defined as the number of links that connect it to different nodes. In terms of air traffic, the nodes are airports and links are nonstop flights, hence the degree of an airport is the number of nonstop flights connecting it to different airports. I define the largest hub of an airline as the airport with the highest degree, and in Tables 5 and 6 I list the firms by market share, along with their largest hubs and its degree.

Table 3.5: **Airlines, market-shares, largest hub and hub degree in 1999.**

Airline	Share (%)	Largest Hub	Degree
Delta (DL)[a]	15.57	Atlanta, GA	79
Southwest (WN)	14.85	Las Vegas, NV	29
United (UA)[b]	12.88	Chicago, IL	66
US Airways (US)[c]	10.78	Charlotte, NC	46
American (AA)[d]	10.92	Dallas, TX	77
Northwest (NW)[e]	7.51	Detroit, MI	60
Continental (CO)[f]	6.19	Houston, TX	61
Trans World (TW)	3.67	St. Louis, MO	57
America West (HP)	3.22	Phoenix, AZ	41
Alaska (AS)[g]	2.93	Seattle, WA	15
AirTran (FL)	1.33	Atlanta, GA	21
ATA (TZ)	1.22	Chicago, IL	12

[a]Delta + Atlantic Southwest; [b]United + Air Wisconsin; [c]US Airways + USAir Shuttle; [d]American + Envoy; [e]Northwest + Mesaba; [f]Continental + ExpressJet; [g]Alaska + Horizon.

Table 3.6: **Airlines, market-shares, largest hub and hub degree in 2015.**

Airline	Share (%)	Largest Hub	Degree
Southwest (WN)	24.76	Chicago, IL	56
Delta (DL)[a]	14.24	Atlanta, GA	79
American (AA)[b]	14.10	Dallas, TX	85
United (UA)[c]	9.22	Chicago, IL	46
SkyWest (OO)[d]	6.2	Chicago, IL	66
JetBlue (B6)	5.56	Boston, MA	31
Alaska (AS)[e]	4.27	Seattle, WA	36
Spirit (NK)	3.19	Dallas, TX	21
Allegiant (G4)	3.06	Las Vegas, NV	20
Frontier (F9)	2.47	Denver, CO	37
Republic (YX)	1.81	Chicago, IL	27
Virgin America (VX)	1.41	San Francisco, CA	9
Mesa (YV)	1.34	Dallas, TX	34
Shuttle America (S5)	1.02	New York, NY	43

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

Fuel burn

Fuel is expensive, heavy, and voluminous, aspects that can severely curtail flight range and profits, consequently making fuel efficiency a major concern to the airlines, and a demand presented to manufacturers. Each new generation of aircraft may reach up to double-digit fuel efficiency improvements when compared to the previous one. The gains come not only from engine-related performance but also from several other features ranging from the use of lighter-weight composite materials to more efficient control systems. Aerodynamic efficiency alone has increased in the vicinity of 15% from 1959 to 2000.[3]

Figure 3.1: **Average fuel burn for new commercial jet aircraft across time.**
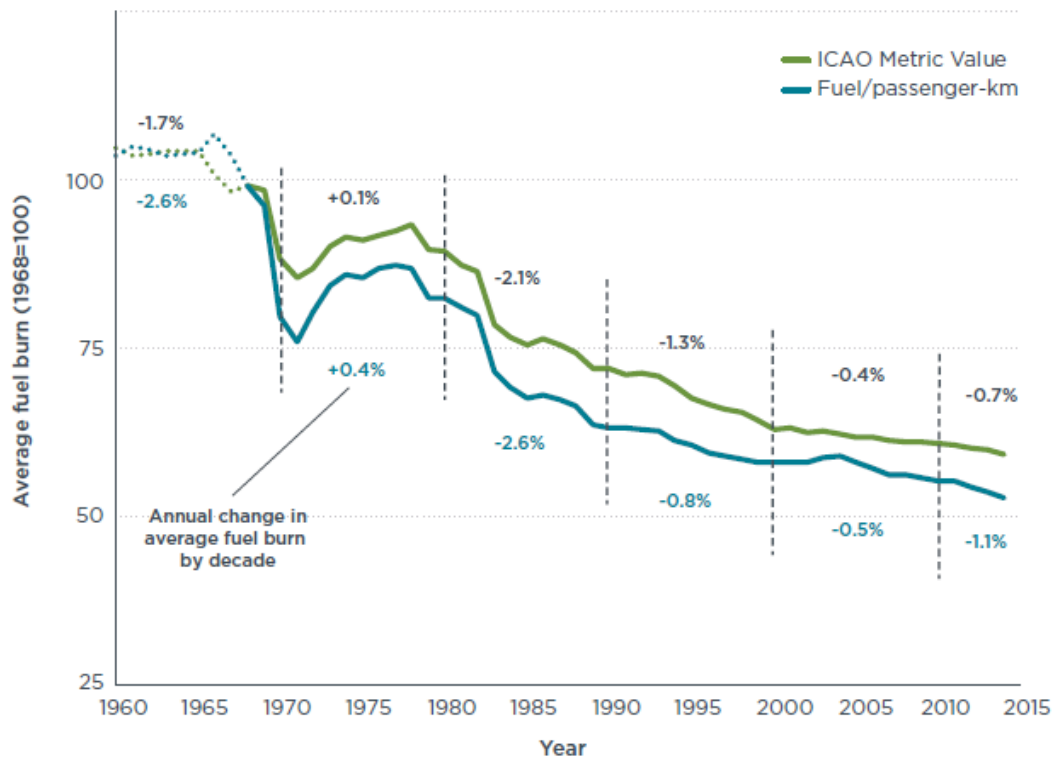
The figure is reproduced from *Fuel efficiency trends for new commercial jet aircraft: 1960 to 2014* by the International Council for Clean Transportation. The ICAO metric value is based on carbon dioxide emissions. 1968 is the benchmark (1968 = 100).

Information on fuel burned by distance for each aircraft is reported by the Air Pollutant Emission Inventory Guidebook issued by the European Environment Agency. These are estimates based on fuel sales and the information reported by the ICAO Aircraft Engine Emissions Databank, which compiles data provided by the manufacturers on emissions of aircraft engines. To the

[3]Lee et al (2000).

best of my knowledge, this is the first study to model aircraft choice and consequently the first one to use a database of the kind to model fuel expenditure.

The data on fuel burned by distance put together with seat count for each aircraft allow us to make efficiency comparisons as in Figure 3.2. The Boeing 737-300 was the most popular aircraft operating in domestic routes in 1999, with jet engines and capacity for 145 passengers; the ATR-42 is a twin turboprop, which is a category of engine typically more fuel efficient than jets for short routes, it has capacity for 46 passengers; and the Embraer 195, a regional jet with capacity for 124 passengers. Note that the Embraer 195 is more fuel efficient than the Boeing 737-300 for every distance for which data is available, and that even though the ATR 42 is the most efficient for very short routes, it is left behind at around 200 nautical miles (~230 miles).

Figure 3.2: **Efficiency by distance: Boeing 737-300, ATR-42, and Embraer 195.**
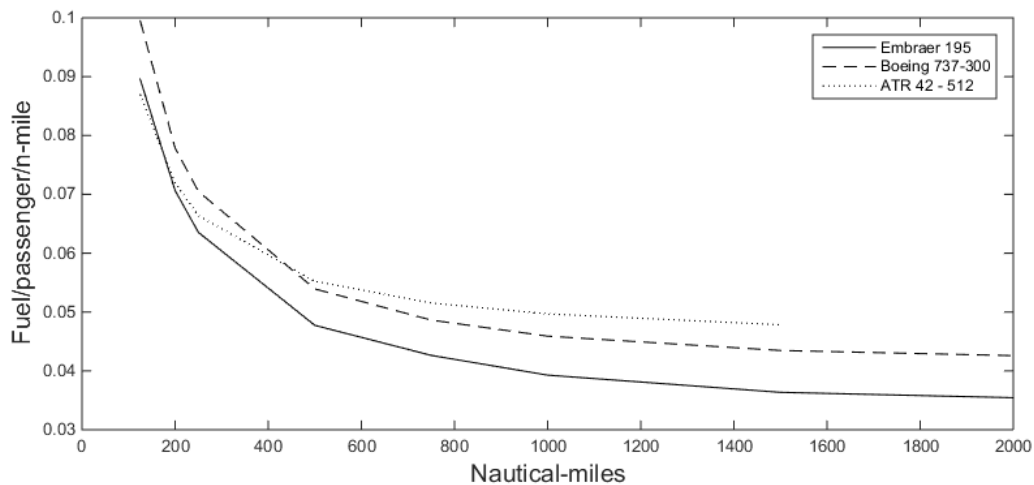
Although important, economy on fuel per passenger basis is one among many features of an aircraft, therefore firms don't simply choose the most fuel efficient. There is a myriad of factors to be considered when building up a fleet, such as maintenance costs, leasing conditions, and airport slot prices.

There may be a considerable lag of up to 20 years between the release of a new aircraft and the adoption by airlines, but the new models are substantially more efficient than the average aircraft in the carrier fleet and we can expect the overall stock of aircraft to become more efficient at a steady pace, with an average annual rate of performance boost equivalent or slightly above the rate of the improvement of new aircraft models.[4]

---

[4]IEA/OECD (2009) Transport, Energy, and CO2: Moving toward sustainability.

4
Results

Estimation results

I rewrite the estimated equation below and present the results in Table 4.1. I have also made a robustness check for sample selection. For conciseness, it was omitted from the table and left to the Appendix, along with the fixed effects estimates.

$$\Pi^*_{imjt} = Z_{imt}\beta + \gamma^a_j + \frac{\delta^2}{4p_{it}FB_{jm}}$$
$$- \left(\gamma^f_i + \gamma^t + \gamma^O_m + \gamma^D_m - \gamma_1\Delta H^O_{imt} - \gamma_2\Delta H^D_{imt} - \gamma_3\Delta H^O_{im,t-1} - \gamma_4\Delta H^D_{im,t-1}\right) + \varepsilon_{imjt}$$

Table 4.1: **Estimation results.**

	1999	**2015**
$\frac{1}{\textbf{p.FB}}$	0.104 (0.047)	1.010 (0.281)
Origin hub	0.105 (0.009)	0.191 (0.006)
Destination hub	0.100 (0.007)	0.183 (0.006)
Lag of Origin hub	-0.0001 (0.008)	-0.012 (0.006)
Lag of Destination hub	0.002 (0.006)	-0.009 (0.006)
Avg. population	2.275 (0.229)	2.183 (0.248)
Distance	-0.755 (0.644)	1.143 (0.878)
Distance^2	-0.055 (0.241)	-0.231 (0.439)

Population is measured in millions, distance in 1,000 miles and the hub size of an airport is defined as the number of nonstop flights connecting it to different airports . Standard errors were clustered by firms and are in parentheses.

Measured by the coefficient of the term $\frac{1}{p.FB}$, the importance of fuel economy to entry decision and aircraft choice increased considerably from 1999 to 2015. A caveat to the interpretation of this parameter is that it could also capture other effects, such as competitiveness, that for the sake of simplicity, the model leaves unaccounted for. All coefficients have the expected sign, and

from 1999 to 2015, the importance of hub size in both ends of the market decreased slightly.

Some variables, such as the hub size, for example, have an impact on both the revenue and costs, affecting the profits positively in both cases. Others, such as market distance, may have a positive impact on revenue while also possibly increasing the costs, and what we observe in the profits is the net effect. For this reason, the coefficients estimates don't have a clear interpretation. This should not be a concern since point estimation is not the purpose of this study. The interest of this paper lies in the counterfactuals, and regarding the coefficients as control variables seems more pertinent.

Counterfactuals and network structure

In this section, I describe and perform counterfactual experiments to examine how the development of new aircraft and fuel economy impact network structure - as measured by number of markets and destinations reached, hubbing concentration, and nodes centrality - through changes in the entry pattern.

Experiment I: In this experiment, I seek to understand how firms would have configured its networks if aviation technology had ceased to improve in 1999. I remove from the aircraft list every aircraft that wasn't available in 1999 - that is, the new aircraft list for 2015, J_{c1}, is such that $J_{c1} = J_{15} \cap J_{99}$.

Experiment II: The goal of this experiment is to uncover the impact caused in the network solely by improvements in aviation technology. Similarly to experiment I, any aircraft introduced later than 1999 is removed from the list, but now I also set the fuel price paid by firms to the average price in 1999 (\$0.71 a gallon). If prices had not increased, it is possible that the impact of new aircraft would be reduced.

Experiment III: To get a better understanding of how the process of network formation behaves in response to a change in aircraft fuel efficiency, I simulate the choices in 1999 considering an increase of 10%, 20% and 40% in aircraft fuel economy. This experiment helps to separete the effect of the introduction of new aircraft from the effect of improvements in fuel efficiency.

Experiment IV: Similarly to experiment III, I simulate choices considering different levels of increase in fuel efficiency in 2015. Motivated by the recent improvements in aviation technology, I increase the fuel efficiency of all aircraft in 10%, 20% and 40%, being the first two in line with the efficiency claimed by manufacturers for new generation aircraft such as the Bombardier CS100, Embraer 190-E2, Mitsubishi MRJ90LR and Airbus A330neo.

To allow each firm to examine the profitability of each market in the network and make an entry decision, we need the distance between every 2 points in the network. Every flight scheduled in US is reported in the T100 data base, along with the distance between the two cities of origin and destination. Using all the available data I can recover 6,099 of 7,503 distances, or 81.3% of the complete network. Therefore, there is distance information only to those markets in which at least one flight was scheduled since 1990.

The of goal of the experiments described is to assess to what extent the introduction of new aircraft and higher fuel efficiency facilitates entry. Therefore, in the computation of counterfactuals, we must bear in mind that there is a bias towards not finding an effect, since some of the possible markets are not even being considered because we lack information on its distance. To help in the examination of the impact, in addition to the results after each experiment, I also report the results of a simple simulation using the parameter estimates.

To mitigate the effect of the random shock, the statistics reported are an average of those obtained after 20 simulations for each exercise.

Markets served and destinations reached

In Table 4.2, 4.3 and 4.4, I report the observed number of markets served for each firm in 2015, the predicted number after simple choices simulation with the estimated parameters and the result after counterfactuals described for experiments I and II.

Table 4.2: **Markets served in 2015: observed, simulation with the estimated parameters, experiment I, and experiment II.**

Airline	Observed	Sim.	Experiment I	Experiment II
Southwest (WN)	571	682.25	576.6	605.2
Delta (DL)[a]	386	447.35	404.05	420.8
American (AA)[b]	431	526.95	477.25	497.2
United (UA)[c]	243	215.25	167.9	180.4
SkyWest (OO)[d]	574	574.3	498.35	523
JetBlue (B6)	105	171.55	128.7	141.6
Alaska (AS)[e]	77	110.8	78.65	85.2
Spirit (NK)	116	261.35	180.85	207.2
Allegiant (G4)	98	131.4	84.7	103.2
Frontier (F9)	116	165.7	119.05	139.4
Republic (YX)	185	174	126.7	143
Virgin America (VX)	28	62.45	38.35	45.8
Mesa (YV)	169	223.75	162.35	175
Shuttle America (S5)	232	198.9	147.3	163.4

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

Table 4.3: **Experiment III: markets served for various levels of fuel efficiency in 1999.**

Airline	Observed.	Sim.	+10%	+20%	+40%
Delta (DL)[a]	287	219.3	221.3	221.05	222.2
Southwest (WN)	270	164.7	164.95	164.85	166.4
United (UA)[b]	202	216.1	216.65	217	218
US Airways (US)[c]	221	188.2	188.4	188	191.05
American (AA)[d]	232	181.3	182.55	183.25	185.2
Northwest (NW)[e]	170	128.6	129.7	129.6	130.25
Continental (CO)[f]	184	106.3	107.15	107.2	109.15
Trans World (TW)	82	74.8	74.3	74.8	75.75
America West (HP)	86	97.2	97.15	97.8	98.8
Alaska (AS)[g]	35	43.2	44.3	44.8	45.6
AirTran (FL)	28	11.5	11.4	11.55	11.4
ATA (TZ)	24	13.6	14.2	14.4	14.05

[a]Delta + Atlantic Southwest; [b]United + Air Wisconsin; [c]US Airways + USAir Shuttle; [d]American + Envoy; [e]Northwest + Mesaba; [f]Continental + ExpressJet; [g]Alaska + Horizon.

First, as we can see from the "observed" columns in tables 8 and 9, firms were on average serving more markets in 2015 than in 1999. The average

number of different nonstop flights operated rose from 151.7 in 1999 to 238 in 2015, an increase of 56.7%. In practice, there are 150 new markets being served in 2015.

Table 4.4: **Experiment IV: markets served for increases in efficiency in 2015.**

Airline	Observed	Sim.	+10%	+20%	+40%
Southwest (WN)	571	682.25	691	700	785
Delta (DL)[a]	386	447.35	450.5	453.5	494.5
American (AA)[b]	431	526.95	530.5	534.5	587.5
United (UA)[c]	243	215.25	215	224	274.5
SkyWest (OO)[d]	574	574.3	574.5	580	633
JetBlue (B6)	105	171.55	170.5	174.5	213
Alaska (AS)[e]	77	110.8	111.5	114	141
Spirit (NK)	116	261.35	261.5	270.5	355
Allegiant (G4)	98	131.4	139.5	150	207
Frontier (F9)	116	165.7	163	172	238.5
Republic (YX)	185	174	175	179	232.5
Virgin America (VX)	28	62.45	70	72	101.5
Mesa (YV)	169	223.75	227	233	285.5
Shuttle America (S5)	232	198.9	204.5	209	253.5

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

An increase in both measures is desirable from the point of view of the consumer since it means more convenience. As the number of nodes and links in each airline's network increase, this translates into an increase in the number of destinations and differents paths to them, more travel options available to passengers. The removal of newer aircraft caused a strong reduction in the number of markets served for every firm, but as shown by experiments III and IV. It also caused a reduction of 16.8 in the average number of cities reached (comparing the average for the simple simulation of 77.6 to 60.8 in experiment I). This is an evidence that new aricraft are more adequate to reach smaller cities, but do note that the effect of changes in efficiency are relatively weak in both years, especially regarding the number of destinations reached.

Table 4.5: **Destinations reached in 2015: observed, simulation with the estimated parameters.**

Airline	Observed	Sim.	Experiment I	Experiment II
Southwest (WN)	69	84.4	78.5	79.6
Delta (DL)[a]	94	99.1	99.5	99.8
American (AA)[b]	104	101.1	100.7	101.0
United (UA)[c]	62	64.4	56.4	59.2
SkyWest (OO)[d]	106	93.3	91.3	92.4
JetBlue (B6)	40	70.3	62.5	65.6
Alaska (AS)[e]	41	62.0	52.3	54.8
Spirit (NK)	23	61.2	55.7	58.8
Allegiant (G4)	48	57.4	50.3	55.0
Frontier (F9)	42	64.3	57.0	59.0
Republic (YX)	53	59.3	52.6	55.8
Virgin America (VX)	14	44.5	32.9	36.6
Mesa (YV)	74	70.4	61.5	63.6
Shuttle America (S5)	57	69.0	62.2	64.2

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon. Observed number of destinations reached for each firm in 2015, the predicted number after simple choices simulation with the estimated parameters and the result after counterfactuals described for experiments I and II.

Table 4.6: **Experiment III: destinations reached for increases in fuel economy in 1999.**

Airline	Observed	Sim.	+10%	+20%	+40%
Delta (DL)[a]	76	84.6	84.9	84.5	85.0
Southwest (WN)	45	59.2	59.6	59.6	59.9
United (UA)[b]	73	85.5	85.9	86.0	85.9
US Airways (US)[c]	51	66.9	66.7	66.8	67.0
American (AA)[d]	81	85.3	85.5	85.4	85.8
Northwest (NW)[e]	63	68.3	68.7	68.6	69.0
Continental (CO)[f]	54	60.4	60.8	61.0	61.4
Trans World (TW)	47	59	58.8	59.1	59.6
America West (HP)	44	42.0	56.7	56.7	57.6
Alaska (AS)[g]	16	38.9	38.9	39.2	39.7
AirTran (FL)	16	18	17.6	17.8	17.5
ATA (TZ)	13	20.8	21.5	21.8	20.9

[a]Delta + Atlantic Southwest; [b]United + Air Wisconsin; [c]US Airways + USAir Shuttle; [d]American + Envoy; [e]Northwest + Mesaba; [f]Continental + ExpressJet; [g]Alaska + Horizon.

Table 4.7: **Experiment IV: destination reached for increases in efficiency in 2015.**

Airline	Observed	Sim.	+10%	+20%	+40%
Southwest (WN)	69	84.4	84.5	85.0	89.0
Delta (DL)[a]	94	99.1	98.5	98.5	99.5
American (AA)[b]	104	101.1	101.0	101.5	101.5
United (UA)[c]	62	64.4	65.0	66.5	74.0
SkyWest (OO)[d]	106	93.3	93.0	93.5	95.5
JetBlue (B6)	40	70.3	70.5	71.0	74.0
Alaska (AS)[e]	41	62.0	63.0	63.5	67.0
Spirit (NK)	23	61.2	60.0	60.5	70.5
Allegiant (G4)	48	57.4	58.5	60.0	67.5
Frontier (F9)	42	64.3	64.0	64.5	75.0
Republic (YX)	53	59.3	60.5	60.5	69.0
Virgin America (VX)	14	44.5	49.0	50.5	56.0
Mesa (YV)	74	70.4	71.5	73.0	79.5
Shuttle America (S5)	57	69.0	68.5	69.8	72.0

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

Hubbing concentration

To gauge the propensity of an airline to operate in a hub-and-spoke fashion, I follow Aguirregabiria and Ho (2012) and calculate the *Hubbing Concentration Ratio* (HCR), defined as the ratio between the number of nonstop connections operated by an airline that include its largest hub over the total number of nonstop connections operated by the same airline. Note that an airline operating a pure hub-and-spoke system would have HCR = 1, and operating a pure point-to-point system would result in HCR = 0.

More formally, let $\mathbb{C}_m$ be set of cities in pair that define the market m, and $h_i^{(1)}$ be the largest hub of firm i, defined as the airport connected with the largest number of cities by nonstop flights. Define $a_{im} \in \{0,1\}$ as the binary indicator for the presence of airline i in market m, then we can write the HCR for firm i:

$$HCR_i \equiv \frac{\sum_{m=1}^{M} a_{im} 1\left[h_i^{(1)} \in \mathbb{C}_m\right]}{\sum_{m=1}^{M} a_{im}}$$

The impact on the Hubbing Concentration Ratio of each airline after the counterfactual experiments I and II are shown in Table 4.8, and the results for

experiment III and IV in Table 4.9 and Table 4.10.

Table 4.8: **Hubbing Concentration Ratios in 2015: observed, simulation with the estimated parameters, experiment I, and experiment II.**

Airline	Observed	Sim.	Experiment I	Experiment II
Southwest (WN)	5.5	5.8	6.2	6.0
Delta (DL)[a]	12.1	12.6	14.0	13.5
American (AA)[b]	11.9	12.2	13.5	12.9
United (UA)[c]	12.1	12.2	13.2	12.4
SkyWest (OO)[d]	7.5	8.8	9.6	9.3
JetBlue (B6)	17.5	22.4	25.7	24.3
Alaska (AS)[e]	28.8	30.3	36.7	34.8
Spirit (NK)	10.5	9.9	11.6	10.5
Allegiant (G4)	14.8	13.9	16.3	15.4
Frontier (F9)	19.7	22.8	28.5	25.6
Republic (YX)	10.6	14.0	15.2	14.7
Virgin America (VX)	18	14.2	17.8	16.5
Mesa (YV)	14.7	15.9	18.7	18.6
Shuttle America (S5)	13.6	20.0	23.8	22.9

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

Table 4.9: **Experiment III: Hubbing Concentration Ratios for various levels of fuel efficiency in 1999.**

Airline	Observed	Sim.	+10%	+20%	+40%
Delta (DL)[a]	19.5	25.2	24.9	24.8	24.8
Southwest (WN)	8.3	8.6	8.5	8.6	8.5
United (UA)[b]	18.7	21.3	21.3	21.2	21.2
US Airways (US)[c]	12.3	12.2	12.3	12.5	12.5
American (AA)[d]	22.5	27.8	27.8	27.7	27.5
Northwest (NW)[e]	23.3	25.8	25.6	25.5	25.6
Continental (CO)[f]	21.9	24.6	24.9	25.0	24.7
Trans World (TW)	36.0	43.9	44.0	44.3	43.9
America West (HP)	31.4	24.6	24.3	24.3	23.9
Alaska (AS)[g]	26.3	16.1	16.1	15.5	15.5
AirTran (FL)	47.6	46.8	46.2	46.3	46.4
ATA (TZ)	40.0	45.1	44.8	45.0	45.0

[a]Delta + Atlantic Southwest; [b]United + Air Wisconsin; [c]US Airways + USAir Shuttle; [d]American + Envoy; [e]Northwest + Mesaba; [f]Continental + ExpressJet; [g]Alaska + Horizon.

On average, there is a decrease in the hub dependency, as revealed by a comparison between the average HCR in each year.[1] If we only consider firms operating in both years, we also note a reduction in the HCR for most of them: Delta's HCR went from 19.5 to 12.1, American's from 22.5 to 11.9, United's from 18.7 to 12.1. The exception is Alaska Airlines, probably because it had most of its operations in the state of Alaska in 1999 - whose cities were excluded from the sample - and expanded operations in U.S. mainland in 2015.

Overall, the changes in HCR in the experiments occurred as expected: removing aircraft from the list induced higher Hubbing Concentration Ratios. When prices were set back to 1999 levels, the impact diminishes, but it still exists (experiment II). Experiments III and IV show that even though the introduction of new aircraft has an effect in the HCR, it is not driven by improvements in fuel economy.

Table 4.10: **Experiment IV: Hubbing Concentration Ratios for increases in efficiency in 2015.**

Airline	Observed	Sim.	+10%	+20%	+40%
Southwest (WN)	5.5	5.8	5.7	5.6	5.3
Delta (DL)[a]	12.1	12.6	12.5	12.4	11.5
American (AA)[b]	11.9	12.2	12.2	12.1	10.9
United (UA)[c]	12.1	12.2	11.7	11.6	10.8
SkyWest (OO)[d]	7.5	8.8	8.8	8.8	8.1
JetBlue (B6)	17.5	22.4	23.2	23.0	19.9
Alaska (AS)[e]	28.8	30.3	29.8	29.2	26.1
Spirit (NK)	10.5	9.9	10.5	10.4	8.8
Allegiant (G4)	14.8	13.9	13.4	13.3	11.7
Frontier (F9)	19.7	22.8	21.7	20.9	17.8
Republic (YX)	10.6	14.0	13.5	13.5	11.3
Virgin America (VX)	18	14.2	10.4	10.8	10.0
Mesa (YV)	14.7	15.9	16.9	16.6	14.4
Shuttle America (S5)	13.6	20.0	20.1	19.7	16.7

[a]Delta + Endeavor; [b]American + Envoy + Air Wisconsin; [c]United + Commutair; [d]SkyWest + ExpressJet; [e]Alaska + Horizon.

Betweenness centrality

Hubs are relative to firms, and for this reason, the HCR is a measure of operational centralization of each firm, but not the industry as a whole - i.e.

---

[1]I have also calculated the HCRs without firm aggregation, the reduction in the average HCR is stronger.

the aggregate network of all firms. For this reason, to make inference about the industry's network I use the *Betweenness Centrality (BC)*. In a given network, a node with a higher BC is more important when it comes to moving from one point to another using the fewest links possible.

In order to define this centralization measure, first we need to define two concepts: (1) a *path* between two nodes in a network is a sequence of links such that each node between the links of the sequence is different, and (2) the *shortest path* between two nodes is the path connecting these nodes with the fewest number of links possible. The Betweenness Centrality then is defined as the fraction of shortest paths between any pair of nodes in the network that passes through the node of interest.

Formally, let $BC(v)$ represent the Betweenness Centrality of node v, and we can define:

$$BC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the number of shortest paths between nodes s and t, and $\sigma_{st}(v)$ is the number of those that pass through node v. In tables 4.11, 4.12, and 4.13, I report the top 10 cities in Betweenness Centrality and the changes caused by each experiment.

Table 4.11: **Betweenness Centrality in 2015: observed, simulation with the estimated parameters, experiment I, and experiment II.**

Observed	Sim.	Experiment I	Experiment II
17.4	16.9	19.4	18.4
11.2	13.0	15.2	14.5
9.9	7.7	7.9	7.8
7.0	6.0	6.1	6.1
6.3	5.2	5.1	5.1
4.8	4.4	4.2	4.2
4.0	3.4	3.3	3.5
3.3	3.2	2.9	3.2
3.0	2.3	2.3	2.3
2.5	2.1	2.1	2.1

Table 4.12: **Experiment III: Betweenness Centrality for various levels of fuel efficiency in 1999.**

Observed	Simulation	+10%	+20%	+40%
20.3	18.2	18.1	18.0	17.8
17.7	16.7	16.7	16.6	16.8
10.4	12.8	12.8	13.1	12.5
4.9	7.2	7.2	7.1	7.2
4.8	5.4	5.4	5.5	5.5
4.2	4.3	4.3	4.3	4.3
4.0	3.5	3.5	3.5	3.5
3.2	3.3	3.3	3.3	3.3
3.2	2.9	2.9	2.9	2.8
2.6	2.3	2.2	2.2	2.2

Table 4.13: **Experiment IV: Betweenness Centrality for increases in efficiency in 2015.**

Observed	Sim.	+10%	+20%	+40%
17.4	16.9	17.5	17.2	15.1
11.2	13.0	12.5	12.2	11.1
9.9	7.7	7.6	7.4	7.1
7.0	6.0	6.0	5.9	5.9
6.3	5.2	5.2	5.3	5.2
4.8	4.4	4.5	4.5	4.5
4.0	3.4	3.2	3.3	3.2
3.3	3.2	3.2	3.2	3.0
3.0	2.3	2.3	2.3	2.4
2.5	2.1	2.2	2.2	2.2

In every experiment the strongest effect is on the most central node, as expected: the expected impact on the other nodes is less clear since a network reconfiguration could either decrease all the centralities or cause an increase for those which came to occupy a more central position. The removal of newer aircraft (experiment I) caused the most significant impact, and as in the previous measures, changes in fuel economy had a weak effect.

Discussion

In this model, firms decide whether to enter each market (or link), and it is through the aggregate of these decisions that effects on the network level unravel. The mechanism underlying the expansion of links formed is well illustrated by the new Boeing 787 Dreamliner. According to the executive director of Boeing's airline network and fleet planning, Alex Heiter[2], in 2015 the new 787 operated over 350 routes, and about one in six of those were entirely new. "We have talked about for many years at Boeing this concept of network fragmentation and how airlines appeal to passenger preferences by offering more services to more cities nonstop, and we are seeing the Boeing 787 doing just that" he said, citing British Airways' London – Austin and United's San Francisco – Chengdu routes as perfect examples.

Much like the Dreamliner, many modern aircraft are designed to operate longer nonstop flights, without much reliance on its density.[3] When the new, more efficient, technology is off the table, the decision of whether to provide flights in a given market becomes more dependent on hub and city sizes, and links between smaller cities are undone. This, in turn, impacts not only the number of cities reached and markets served but also causes the network structure to become more centralized around hubs and large cities, as shown by the counterfactuals.

Even though the availability of aircraft seems to impact network structure, the effect of fuel efficiency is surprisingly weak, especially in 2015. A possible explanation is that firms started to see fuel efficiency as a concern relatively long ago, due to the rise and instability of oil prices of the 1970s, for example. Firms are concerned with the total quantity of gallons burned, and if fuel prices started to be a problem in the early 1970s, it is plausible that most of the gains in terms of gallons burned - and consequently, in monetary terms - occurred in the following years. If this is the case, most of the effect would have been captured comparing these early decades. In other words, it could be that in 1999 aircraft were already efficient enough so that the an additional improvement of 20% or even 40% would not translate into a considerable economy in terms of gallons of fuel.

Take the Boeing 737-300 operating in a market with a distance of 1,000 miles as an example. According to the data, this would require approximately 5,950 gallons of fuel, which at the average fuel price in 2015 of \$1.82 means a fuel expense of \$10,829. If we could boost its efficiency by 20% - which is a

[2]In the World Route Strategy Summit, 2015.
[3]see Boeing's *Current Market Outlook 2015.*

strong improvement whereas still realistic - it would burn 4,760 gallons, at a cost of $8,663, amounting to an economy of $2,166 per flight. Aguirregabiria and Ho (2012) estimate an average fixed cost of almost $120,000 per quarter. Therefore, in a given quarter, if aircraft operational costs will correspond to an important share of the total expenses of serving a market depends on the flight frequency. It could be that at the present state of technology and distribution of flight frequencies, improvements in fuel efficiency would not be that relevant to entry decisions.

An illustrative case could be the supersonic jet Concorde, whose project date back to 1955. Produced by a consortium between British Aircraft Corporation and Aerospatiale, the Concorde made its debut flight in 1969, and was clearly not projected regarding fuel efficiency as a primary concern: when full, the Concorde achieved roughly 0.06 gallons of fuel per passenger-mile, against 0.043 achieved by the 737-300 (whose first flight was in 1984). Despite the fact that it is one of the most famous aircraft ever made, it became economically impracticable to operate after the oil shocks in the 1970s. Merely 14 units were sold for commercial use, and it was produced only until 1979.

5
Conclusion

In order to to assess how the introduction of new technologies impact firms entry decisions and network archictecture, I developed and estimated a structural model of aircraft choice upon market entry. With data on U.S. air traffic network and fuel burn by distance for each aircraft, I present a framework in which firms make entry decisions and maximize profits by choosing an aircraft to operate each market. Results show an impact stemming from the development of new aircraft on several dimensions of network structure, namely increasing the number of cities reached, the number of markets served, and also decreasing hub centralization (as defined either by the hub concentration ratio or betweenness centrality).

From the firm perspective, these results could have implications to the optimization of fleet renewal and hub investment planning. From a policy perspective, it is known that airport congestion is a common problem caused by large hubs, but taking notice that there are structural changes taking place in this industry reduces the appeal for public intervention. Aircraft purchases often happen years before it is actually delivered, therefore leaving time for forecasting network and airport congestion. It is also worth noting that network traffic is related to airport development, and as shown by Sheard (2017), airports have a positive impact on city GDP, rate of employment and other economic outcomes of public interest.

Finally, I acknowledge that important considerations such as forward-looking firms and strategic interactions - when competing in prices or benefiting from economies of scale in existing hubs - are limitations of the model and ideally should also be taken into account. Future work should also look into the tractability of strategic network formation.

References

LEE, J.; LUKACHKO, S.; WAITZ, I. ; SCHAFER, A.. **Historical and future trends in aircraft performance, cost, and emissions**. Annual Review of Energy and the Environment, 2001.

REDONDI, R.; MALIGHETTI, P. ; PALEARI, S.. **De-hubbing of airports and their recovery patterns**. Journal of Air Transport Management, 18(1):1–4, 2012.

HENDRICKS, K.; PICCIONE, M. ; TAN, G.. **Equilibria in networks**. Econometrica, 67(6):1407–1434, 1999.

HENDRICKS, K.; PICCIONE, M. ; TAN, G.. **The economics of hubs: The case of monopoly**. The Review of Economic Studies, 62(1):83–99, 1995.

BERRY, S.; JIA, P.. **Tracing the woes: An empirical analysis of the airline industry**. American Economic Journal, 2(1-43), 2010.

BERRY, S.; CARNALL, M. ; SPILLER, P.. **Airline hubs: Costs, markups and the implications of customer heterogeneity**. Advances in Airline Economics: Competition Policy and Antitrust, 1:183–214, 2006.

BERRY, S.. **Estimation of a model of entry in the airline industry**. Econometrica, 60(4):889–917, 1992.

AGUIRREGABIRIA, V.; HO, C.. **A dynamic oligopoly game of the us airline industry: Estimation and policy experiments**. Journal of Econometrics, 168:156–173, 2012.

MCFADDEN, D.. **Economic choices**. The American Economic Review, 91(3):351–378, 2001.

MCFADDEN, D.; TALVITIE, A.; REID, F. ; JOHNSON, M.. **Urban Travel Demand Forecasting Project**, volumen I. Institute of Transportation Studies, University of California, Berkeley, 6 1977.

MCFADDEN, D.. **Conditional logit analysis of qualitative choice behavior, in p. Zaremkba, Frontiers in Econometrics**. Academic Press, 1974.

BORENSTEIN, S.. **Hubs and high fares: Airport dominance and market power in the u.s. airline industry**. The RAND Journal of Economics, 20(3):344–365, 1989.

BORENSTEIN, S.; ROSE, N.. **Economic Regulation and Its Reform: What Have We Learned?**, chapter How airline markets work... or do they? Regulatory reform in the airline industry. University of Chicago Press, 2014.

BOREINSTEIN, S.. **U.S. domestic airline pricing, 1995-2004**. Competition Policy Center Working Paper, (CP05):05–48, 2005.

CAVES, D.; CHRISTENSEN, L. ; TRETHEWAY, M.. **Economies of density versus economies of scale: why trunk and local airline cost differ**. RAND Journal of Economics, 15:471–489, 1984.

SHEARD, N.. **Airport size and urban growth**. Working Paper, 2017.

HANSEN, M.; LIU, Y.. **Airline competition and market frequency: a comparison of the s-curve and schedule delay models**. Transportation Research Part B: Methodological, 78:301–317, 2015.

A
Derivation of logit probabilities

Reproduced from Kenneth Train's *Discrete Choice Methods with Simulation*:[1]

The utility that the decision maker obtains from alternative j is decomposed into (1) a part labeled V_{nj} that is known by the researcher up to some parameters, and (2) an unknown part ϵ_{nj} that is treated by the researcher as random:

$$U_{ij} = V_{ij} + \varepsilon_{ij} \forall j$$

The probability that decision marker i chooses alternative j is P_{ij}.

$$P_{ij} = Prob\left(V_{ij} + \varepsilon_{ij} > V_{ik} + \varepsilon_{ik} \forall k \neq i\right)$$
$$= Prob\left(\varepsilon_{ik} < \varepsilon_{ij} + V_{ij} - V_{ik} \forall k \neq i\right)$$

The choice probability is the integral over all values of ε_{ij} weighted by its density, $f(\varepsilon_{ik}) = e^{-\varepsilon_{ik}} e^{-e^{\varepsilon_{ik}}}$:

$$P_{ij} = \int \left(\prod_{k \neq i} e^{-e^{-\left(\varepsilon_{ij} + V_{ij} - V_{ik}\right)}}\right) e^{-\varepsilon_{ik}} e^{-\varepsilon_{ij}} d\varepsilon_{ij}$$

where s is ε_{ij}. Note that $V_{ij} - V_{ij} = 0$ and collect the terms in the exponent of e:

$$P_{ij} = \int_{s=-\infty}^{\infty} \left(\prod_{k \neq i} e^{-e^{-\left(s + V_{ij} - V_{ik}\right)}}\right) e^{-s} ds$$
$$= \int_{s=-\infty}^{\infty} \exp\left(-\sum_k e^{-(s + V_{ij} - V_{ik})}\right) e^{-s} ds$$
$$= \int_{s=-\infty}^{\infty} \exp\left(-e^{-s} \sum_k e^{-(V_{ij} - V_{ik})}\right) e^{-s} ds$$

[1] Available online at https://eml.berkeley.edu/books/choice2.html

Define $t = \exp(-s)$ such that $-\exp(-s)ds = dt$. Note that as s approaches infinity, t approaches zero, and as s approaches negative infinity, t becomes infinitely large.

$$
\begin{aligned}
P_{ni} &= \int_{\infty}^{0} \exp\left(-t \sum_{j} e^{-(V_{ij}-V_{ik})}\right)(-dt) \\
&= \int_{0}^{\infty} \exp\left(-t \sum_{j} e^{-(V_{ij}-V_{ik})}\right) dt \\
&= \left. \frac{\exp\left(-t \sum_{j} e^{-(V_{ij}-V_{ik})}\right)}{-\sum_{j} e^{-(V_{ij}-V_{ik})}} \right|_{0}^{\infty} \\
&= \frac{1}{\sum_{k} e^{(V_{ij}-V_{ik})}} = \frac{e^{V_{ij}}}{\sum_{k} e^{V_{ik}}}
\end{aligned}
$$

B
Fixed effect estimates for 1999

Table B.1: **Aircraft list, year of debut flight, and fixed effects for 1999.**

Boeing 737-500 (1989)	0.290	(0.360)
Fokker 100 (1986)	0.187	(0.521)
McDonnell Douglas DC9 Super 80/MD81/82/83/88 (1979)	2.211	(0.287)
Boeing 727-200/231A (1963)	1.668	(0.318)
Boeing 737-300 (1984)	2.457	(0.275)
Boeing 757-200 (1967)	1.917	(0.280)
Aerospatiale/Aeritalia ATR-72 (1988)	-7.621	(0.588)
Saab-Fairchild 340/B (1983)	-0.349	(0.438)
McDonnell Douglas MD-90 (1993)	-1.784	(0.952)
Embraer EMB-120 Brasilia (1983)	-2.225	(0.582)
Boeing 767-300/300ER (1986)	0.145	(0.550)
Boeing 737-100/200 (1967)	1.179	(0.323)
McDonnell Douglas DC-9-30 (1972)	0.864	(0.335)
Airbus Industrie A320-100/200 (1987)	0.858	(0.387)
Boeing 737-400 (1988)	-0.010	(0.643)
Airbus Industrie A319 (1995)	-0.766	(0.367)
British Aerospace BAe-146-300 (1988)	-1.669	(0.445)

Table B.2: **Aircraft list, year of debut flight, and fixed effects for 1999.**

Lockheed L-1011-500 Tristar (1978)	-39.879	(1.053)
Lockheed L-1011-1/100/200 (1970)	-1.117	(0.881)
Canadair RJ-200ER /RJ-440 (1991)	-20.372	(0.843)
Boeing 767-200/ER/EM (1981)	-0.596	(0.913)
McDonnell Douglas DC-9-50 (1975)	-1.283	(0.692)
De Havilland DHC8-100 Dash-8 (1983)	-2.825	(0.957)
Fokker F28-4000/6000 Fellowship (1967)	-1.645	(0.524)
McDonnell Douglas DC-10-10 (1970)	-2.166	(1.014)
Embraer-145 (1995)	-4.226	(0.816)
Boeing 737-800 (1997)	-19.235	(1.039)
Aerospatiale/Aeritalia ATR-42 (1984)	-0.791	(0.557)
Boeing 777-200ER/200LR/233LR (1994)	-3.064	(1.033)
Airbus Industrie A300-600/R/CF/RCF (1983)	-2.290	(1.034)
Avroliner RJ85 (1993)	-6.986	(0.843)
McDonnell Douglas DC-9-10 (1970)	-13.878	(1.007)
McDonnell Douglas DC-10-40 (1972)	-14.518	(1.031)
Boeing 747-100 (1969)	-6.143	(0.926)
Boeing 747-400 (1988)	-2.203	(0.795)
McDonnell Douglas DC-9-40 (1968)	-47.141	(1.088)

Table B.3: **City, firm, and quarter fixed effects estimates for 1999.**

Allentown, PA	-2.991	(0.468)	Corpus Christi, TX	-2.790	(0.902)
Albuquerque, NM	-1.882	(0.970)	Columbus, GA	-6.891	(1.002)
Waco, TX	-5.263	(0.851)	Cincinnati, OH	-1.837	(0.772)
Dallas, TX[a]	-2.293	(1.023)	Dayton, OH	-2.941	(1.016)
Colorado Springs, CO	-3.132	(0.454)	Detroit, MI	-1.858	(0.625)
Pittsburgh, PA	-1.198	(0.856)	Des Moines, IA	-3.238	(0.962)
Akron, OH	-12.193	(0.849)	Houston, TX	-1.981	(0.526)
Huntsville, AL	-2.753	(1.067)	Orlando, FL	0.671	(0.356)
Amarillo, TX	-4.572	(0.909)	Erie, PA	-35.956	(1.057)
Denver, CO[b]	-0.899	(0.927)	Eugene, OR	-2.224	(0.927)
Atlanta, GA	-0.910	(0.554)	Evansville, IN	-11.036	(1.063)
Austin, TX	-1.734	(0.611)	Fresno, CA	-2.513	(1.021)
Tucson, AZ	-2.040	(0.754)	Fayetteville, NC	-2.738	(1.004)
Phoenix, AZ[c]	-1.957	(0.752)	Minneapolis, MN[f]	-1.501	(0.657)
Shreveport, LA	-4.976	(0.450)	Fort Lauderdale, FL	0.514	(0.534)
Seattle-Tacoma, WA[d]	0.421	(0.943)	New York, NY[g]	-6.687	(0.884)
Bakersfield, CA	-18.348	(0.395)	Flint, MI	-5.713	(1.057)
Mobile, AL	-25.086	(1.052)	Sioux Falls, SD	-6.752	(0.987)
Birmingham, AL	-2.931	(1.051)	Fort Wayne, IN	-6.131	(0.980)

[a]Dallas-Fort Worth-Arlington-Plano, TX; [b]Denver-Aurora-Lakewood, CO;
[c]Phoenix-Tempe-Mesa-Glendale-Chandler, AZ; [d]Seattle-Tacoma, WA; [e]Beaumont-Port Arhutr, TX;
[f]Minneapolis-St. Paul , MN; [g]New York-Jersey-Newark, NY; [h]Los Angeles-Santa Ana-Anaheim-Fullerton,
CA; [h]Kansas, MO-Kansas, KS.

Table B.4: **City, firm, and quarter fixed effects estimates for 1999.**

Birmingham, AL	-2.931	(1.051)	Fort Wayne, IN	-6.131	(0.980)
El Paso, TX	-2,684	(0.935)	Spokane, WA	-1.570	(0.978)
Cleveland, OH	-0.509	(0.925)	Grand Rapids, MI	-2.680	(0.936)
Nashville, TN	-1.105	(0.772)	Greensboro, NC	-2.060	(0.915)
Boise, ID	-0.972	(0.715)	Las Vegas, NV	-0.420	(0.517)
Boston, MA	0.211	(0.950)	Indianapolis, IN	-1.422	(0.787)
Beaumont, TX[e]	-18.565	(0.430)	Little Rock, AR	-2.941	(0.976)
Brownsville, TX	-3.131	(1.037)	Lansing, MI	-5.929	(1.056)
Baton Rouge, LA	-3.331	(1.017)	Los Angeles, CA[h]	-3.146	(0.543)
Buffalo, NY	-2.605	(0.908)	Lubbock, TX	-3.510	(0.985)
Baltimore, MD	-0.311	(0.894)	Lexington, KY	-2.654	(0.949)
Washington, DC	0.116	(0.713)	Lafayette, LA	-4.344	(0.952)
Columbia, SC	-6.846	(0.368)	Long Beach, CA	-2.892	(0.999)
Wichita, KS	-3.754	(1.106)	Lincoln, NE	-3.598	(0.949)
Chicago, IL	-2.605	(0.897)	Laredo, TX	-3.031	(0.918)
Chattanooga, TN	-4.088	(0.411)	Louisville, KY	-1.909	(0.862)
Cedar Rapids, IA	-3.237	(0.991)	Macon, GA	-3.403	(1.013)
Charlotte, NC	-1.742	(0.959)	Sacramento, CA	-1.104	(0.867)
Columbus, OH	-1.247	(0.875)	Tampa, FL	0.501	(0.622)
St. Louis, MO	-0.872	(0.898)	Kansas, MO	-1.306	(0.831)
Jacksonville, FL	-2.157	(0.864)	San Antonio, TX	-2.179	(0.880)

[a]Dallas-Fort Worth-Arlington-Plano, TX; [b]Denver-Aurora-Lakewood, CO; [c]Phoenix-Tempe-Mesa-Glendale-Chandler, AZ; [d]Seattle-Tacoma, WA; [e]Beaumont-Port Arhutr, TX; [f]Minneapolis-St. Paul , MN; [g]New York-Jersey-Newark, NY; [h]Los Angeles-Santa Ana-Anaheim-Fullerton, CA; [h]Kansas, MO-Kansas, KS.

Table B.5: **City, firm, and quarter fixed effects estimates for 1999.**

Memphis, TN	-1.517	(0.761)	Savannah, GA	-4.039	(0.923)
McAllen, TX	-3.775	(0.897)	San Francisco, CA	0.256	(0.540)
Montgomery, AL	-7.205	(1.044)	Springfield, MO	-3.932	(0.856)
Manchester, NH	-1.823	(0.961)	San Jose, CA	-0.900	(0.836)
Miami, FL	0.128	(0.575)	Springfield, IL	-8.291	(1.039)
Omaha, NE	-2.865	(0.970)	Syracuse, NY	-1.595	(0.985)
Milwaukee, WI	-2.481	(0.965)	Toledo, OH	-4.252	(0.969)
Madison, WI	-3.962	(0.801)	Tallahassee, FL	-7.602	(1.018)
New Orleans, LA	-0.908	(0.716)	Knoxville, TN	-3.316	(0.997)
San Diego, CA	-1.097	(0.743)	Southwest (WN)	0.674	(0.360)
Norfolk, VA	-1.953	(0.875)	Delta (DL)[b]	0.151	(0.521)
Oakland, CA	-0.837	(0.658)	American (AA)[c]	-0.165	(0.287)
Oklahoma , OK	-2.319	(0.895)	United (UA)[d]	0.516	(0.318)
Ontario, CA	-1.650	(0.748)	US Airways (US)[e]	0.615	(0.275)

[a]Raleigh-Durham;[b]Delta + Atlantic Southwest;[c]American + Envoy; [d]United + Air Wisconsin; [e]US Airways + USAir Shuttle; [f]Alaska + Horizon;[g]Continental + ExpressJet;[h]Northwest + Mesaba.

Table B.6: **City, firm, and quarter fixed effects estimates for 1999.**

Memphis, TN	-1.517	(0.761)	Savannah, GA	-4.039	(0.923)
McAllen, TX	-3.775	(0.897)	San Francisco, CA	0.256	(0.540)
Montgomery, AL	-7.205	(1.044)	Springfield, MO	-3.932	(0.856)
Manchester, NH	-1.823	(0.961)	San Jose, CA	-0.900	(0.836)
Miami, FL	0.128	(0.575)	Springfield, IL	-8.291	(1.039)
Omaha, NE	-2.865	(0.970)	Syracuse, NY	-1.595	(0.985)
Milwaukee, WI	-2.481	(0.965)	Toledo, OH	-4.252	(0.969)
Madison, WI	-3.962	(0.801)	Tallahassee, FL	-7.602	(1.018)
New Orleans, LA	-0.908	(0.716)	Knoxville, TN	-3.316	(0.997)
San Diego, CA	-1.097	(0.743)	Southwest (WN)	0.674	(0.360)
Norfolk, VA	-1.953	(0.875)	Delta (DL)[b]	0.151	(0.521)
Oakland, CA	-0.837	(0.658)	American (AA)[c]	-0.165	(0.287)
Oklahoma , OK	-2.319	(0.895)	United (UA)[d]	0.516	(0.318)
Ontario, CA	-1.650	(0.748)	US Airways (US)[e]	0.615	(0.275)
Portland, OR	0.065	(0.726)	Alaska (AS)[f]	0.755	(0.280)
Philadelphia, PA	-1.267	(0.538)	Continental (CO)[g]	-1.524	(0.588)
Peoria, IL	-8.127	(0.980)	AirTran (FL)	-0.335	(0.438)
St. Petersburg, FL	-1.608	(0.934)	America West (HP)	0.841	(0.952)
Providence, RI	-1.449	(0.883)	Northwest (NW)[h]	0.226	(0.582)
Raleigh, NC[a]	-1.475	(0.887)	Trans World (TW)	0.148	(0.550)
Rockford, IL	-7.127	(1.005)	ATA (TZ)	0.011	(0.323)
Richmond, VA	-2.494	(0.883)	Quarter 1	-8.503	(0.335)
Reno, NV	-1.484	(0.852)	Quarter 2	-8.501	(0.387)
Rochester, NY	-2.283	(0.953)	Quarter 3	-8.454	(0.643)
Salt Lake, UT	0.068	(0.626)	Quarter 4	-8.531	(0.367)
Tulsa, OK	-3.016	(0.903)			

[a]Raleigh-Durham;[b]Delta + Atlantic Southwest;[c]American + Envoy; [d]United + Air Wisconsin; [e]US Airways + USAir Shuttle; [f]Alaska + Horizon;[g]Continental + ExpressJet;[h]Northwest + Mesaba.

C
Fixed effect estimates for 2015

Table C.1: **Aircraft list, year of debut flight, and fixed effects for 2015.**

Boeing 737-900ER (2006)	19.123 (0.431)
Embraer 190 (2004)	17.536 (0.392)
Embraer ERJ-175 (2003)	20.296 (0.367)
Embraer 170 (2001)	19.351 (0.423)
Boeing 717-200 (1999)	19.746 (0.385)
Canadair CRJ 900 (1999)	18.987 (0.380)
Canadair RJ-700 (1999)	19.619 (0.369)
Boeing 757-300 (1999)	20.549 (0.370)
De Havilland DHC8-400 Dash-8 (1998)	19.497 (0.382)
Boeing 737-700/700LR (1997)	19.775 (0.361)
Boeing 737-800 (1997)	18.177 (0.430)
Boeing 737-900 (1997)	20.062 (0.359)
Embraer-145 (1995)	18.870 (0.400)

This aircraft list comprises every aircraft used by at least one firm in the working sample for 2015. When estimating and simulating choices, I assume that every aircraft available in 2015 is in this list.

Table C.2: **Aircraft list, year of debut flight, and fixed effects for 2015.**

Airbus Industrie A319 (1995)	17.607 (0.468)
Embraer-140 (1995)	17.322 (0.389)
De Havilland DHC8-200Q Dash-8 (1995)	15.051 (1.021)
Airbus Industrie A321 (1994)	18.294 (0.400)
McDonnell Douglas MD-90 (1993)	18.804 (0.367)
Canadair RJ-200ER /RJ-440 (1992)	18.293 (0.398)
Boeing 737-500 (1989)	9.710 (0.858)
De Havilland DHC8-300 Dash 8 (1989)	17.648 (0.369)
Boeing 737-400 (1988)	17.259 (0.421)
Airbus Industrie A320-100/200 (1987)	2.614 (0.938)
Boeing 767-300/300ER (1986)	-20.111 (1.065)
Boeing 737-300 (1984)	0.035 (0.894)
De Havilland DHC8-100 Dash-8 (1983)	-42.387 (1.100)
Embraer EMB-120 Brasilia (1983)	6.136 (0.946)
McDonnell Douglas DC9 Super 80/MD81/82/83/88 (1979)	-42.230 (1.184)
Boeing 757-200 (1967)	-44.248 (1.158)

This aircraft list comprises every aircraft used by at least one firm in the working sample for 2015. When estimating and simulating choices, I assume that every aircraft available in 2015 is in this list.

Table C.3: **City, firm, and quarter fixed effects estimates for 2015.**

Allentown, PA	-9.683	(1.031)	Corpus Christi, TX	-7.409	(1.028)
Albuquerque, NM	-2.442	(0.978)	Columbus, GA	-47.375	(1.068)
Waco, TX	-47.153	(1.058)	Cincinnati, OH	-1.501	(0.773)
Dallas, TX[a]	5.156	(0.410)	Dayton, OH	-6.584	(1.003)
Colorado Springs, CO	-9.190	(0.927)	Detroit, MI	-0.013	(0.342)
Pittsburgh, PA	-1.900	(0.864)	Des Moines, IA	-4.541	(0.910)
Akron, OH	-4.922	(0.964)	Houston, TX	4.301	(0.358)
Huntsville, AL	-7.316	(0.949)	Orlando, FL	0.039	(0.641)
Amarillo, TX	-8.571	(1.021)	Erie, PA	-16.891	(1.046)
Denver, CO[b]	0.997	(0.311)	Eugene, OR	-3.533	(1.019)
Atlanta, GA	0.699	(0.469)	Evansville, IN	-6.882	(1.024)
Austin, TX	0.272	(0.772)	Fresno, CA	-5.738	(1.020)
Tucson, AZ	-5.778	(0.962)	Fayetteville, NC	-33.417	(1.040)
Phoenix, AZ[c]	4.646	(0.368)	Minneapolis, MN[f]	1.208	(0.452)
Shreveport, LA	-5.901	(1.020)	Fort Lauderdale, FL	-0.582	(0.380)
Seattle-Tacoma, WA[d]	1.683	(0.642)	New York, NY[g]	16.915	(1.010)
Bakersfield, CA	-3.609	(1.066)	Flint, MI	-7.844	(0.985)
Mobile, AL	-7.908	(1.028)	Sioux Falls, SD	-6.405	(0.971)
Birmingham, AL	-3.162	(0.940)	Fort Wayne, IN	-5.781	(1.019)

[a]Dallas-Fort Worth-Arlington-Plano, TX; [b]Denver-Aurora-Lakewood, CO;
[c]Phoenix-Tempe-Mesa-Glendale-Chandler, AZ; [d]Seattle-Tacoma, WA; [e]Beaumont-Port Arhutr, TX;
[f]Minneapolis-St. Paul , MN; [g]New York-Jersey-Newark, NY; [h]Los Angeles-Santa Ana-Anaheim-Fullerton,
CA; [h]Kansas, MO-Kansas, KS.

Table C.4: **City, firm, and quarter fixed effects estimates for 2015.**

Birmingham, AL	-3.162	(0.940)	Fort Wayne, IN	-5.781	(1.019)
El Paso, TX	-4.976	(0.966)	Spokane, WA	-3.733	(0.928)
Cleveland, OH	-0.719	(0.782)	Grand Rapids, MI	-2.002	(0.869)
Nashville, TN	-0.884	(0.851)	Greensboro, NC	-3.919	(0.972)
Boise, ID	-1.417	(0.844)	Las Vegas, NV	0.599	(0.353)
Boston, MA	1.174	(0.649)	Indianapolis, IN	-1.223	(0.843)
Beaumont, TX[e]	-36.916	(1.039)	Little Rock, AR	-6.162	(0.995)
Brownsville, TX	-14.112	(1.047)	Lansing, MI	-23.551	(1.048)
Baton Rouge, LA	-6.431	(1.024)	Los Angeles, CA[h]	9.006	(0.583)
Buffalo, NY	-4.041	(0.781)	Lubbock, TX	-10.347	(0.965)
Baltimore, MD	-0.158	(0.812)	Lexington, KY	-2.982	(0.934)
Washington, DC	1.299	(0.332)	Lafayette, LA	-5.219	(1.031)
Columbia, SC	-7.264	(0.992)	Long Beach, CA	-4.904	(0.977)
Wichita, KS	-5.008	(1.006)	Lincoln, NE	-6.175	(1.034)
Chicago, IL	4.419	(0.339)	Laredo, TX	-10.139	(1.050)
Chattanooga, TN	-7.462	(1.025)	Louisville, KY	-2.107	(0.986)
Cedar Rapids, IA	-5.533	(0.980)	Macon, GA	-47.543	(1.059)
Charlotte, NC	-0.368	(0.827)	Sacramento, CA	-1.422	(0.875)
Columbus, OH	-0.871	(0.864)	Tampa, FL	-0.412	(0.669)
St. Louis, MO	-2.132	(0.791)	Kansas, MO	0.165	(0.562)
Jacksonville, FL	-0.940	(0.883)	San Antonio, TX	0.404	(0.901)

[a]Dallas-Fort Worth-Arlington-Plano, TX; [b]Denver-Aurora-Lakewood, CO; [c]Phoenix-Tempe-Mesa-Glendale-Chandler, AZ; [d]Seattle-Tacoma, WA; [e]Beaumont-Port Arhutr, TX; [f]Minneapolis-St. Paul , MN; [g]New York-Jersey-Newark, NY; [h]Los Angeles-Santa Ana-Anaheim-Fullerton, CA; [h]Kansas, MO-Kansas, KS.

Table C.5: **City, firm, and quarter fixed effects estimates for 2015.**

City	Value	SE	City/Firm	Value	SE
Memphis, TN	-2.786	(0.869)	San Francisco, CA	2.569	(0.469)
McAllen, TX	-8.019	(1.030)	Springfield, MO	-5.183	(1.002)
Montgomery, AL	-34.012	(1.039)	San Jose, CA	0.009	(0.933)
Manchester, NH	-5.940	(1.012)	Springfield, IL	-25.505	(1.042)
Miami, FL	-0.798	(0.764)	Syracuse, NY	-3.434	(1.009)
Omaha, NE	-2.334	(0.940)	Toledo, OH	-45.630	(1.050)
Milwaukee, WI	-3.888	(0.954)	Tallahassee, FL	-5.234	(1.039)
Madison, WI	-4.979	(0.945)	Knoxville, TN	-5.543	(1.002)
New Orleans, LA	-1.180	(0.669)	Southwest (WN)	-0.929	(0.410)
San Diego, CA	1.543	(0.910)	Delta (DL)[b]	-0.342	(0.507)
Norfolk, VA	-3.991	(1.009)	American (AA)[c]	-0.154	(0.517)
Oakland, CA	-2.301	(0.741)	United (UA)[d]	-0.663	(0.388)
Oklahoma , OK	-1.358	(0.879)	JetBlue (B6)	1.505	(0.516)
Ontario, CA	-5.489	(0.989)	Alaska (AS)[e]	1.196	(0.591)
Portland, OR	0.154	(0.617)	Frontier (F9)	1.902	(0.447)
Philadelphia, PA	1.851	(0.531)	Allegiant (G4)	2.299	(0.554)
Peoria, IL	-4.880	(1.003)	Spirit (NK)	2.378	(0.469)
St. Petersburg, FL	-3.975	(0.994)	SkyWest (OO)[f]	-0.860	(0.401)
Providence, RI	-3.835	(0.953)	Shuttle America (S5)	0.653	(0.399)
Raleigh, NC[a]	-0.218	(0.619)	Virgin America (VX)	1.695	(0.861)
Rockford, IL	-30.923	(1.040)	Mesa (YV)	1.486	(0.473)
Richmond, VA	-4.072	(0.980)	Republic (YX)	1.234	(0.457)
Reno, NV	-3.167	(0.947)	Quarter 1	-23.495	(0.655)
Rochester, NY	-4.890	(0.973)	Quarter 2	-23.372	(0.630)
Salt Lake, UT	0.191	(0.405)	Quarter 3	-22.910	(0.636)
Tulsa, OK	-4.268	(0.910)	Quarter 4	-23.794	(0.635)
Savannah , GA	-6.347	(1.007)			

[a]Raleigh-Durham; [b]Delta + Endeavor; [c]American + Envoy + Air Wisconsin; [d]United + Commutair; [e]Alaska + Horizon; [f]SkyWest + ExpressJet.

D
Robustness check for main parameter estimates

As a rubustness test, I re-estimate the model excluding only aircraft with less than 15 seats (instead of 30) and requiring only 40 passengers transported per month (instead of 80) in order to consider a firm active in a market.

Table D.1: **Estimation results for rubustness check.**

	1999	**2015**
$\frac{1}{\textbf{p.FB}}$	0.095	1.135
	(0.030)	(0.281)
Origin hub	0.113	0.151
	(0.009)	(0.006)
Destination hub	0.110	0.143
	(0.008)	(0.006)
Lag of Origin hub	-0.0003	-0.003
	(0.008)	(0.006)
Lag of Destination hub	0.002	0.0008
	(0.006)	(0.006)
Avg. population	2.105	2.183
	(0.242)	(0.247)
Distance	-0.468	1.143
	(0.287)	(0.841)
Distance^2	-0.080	-1.405
	(0.221)	(0.413)

Population is measured in millions, distance in 1,000 miles and the hub size of an airport is defined as the number of nonstop flights connecting it to different airports . Standard errors were clustered by firms and are in parentheses.

E
Figures

Figure E.1: Next generation aircraft.
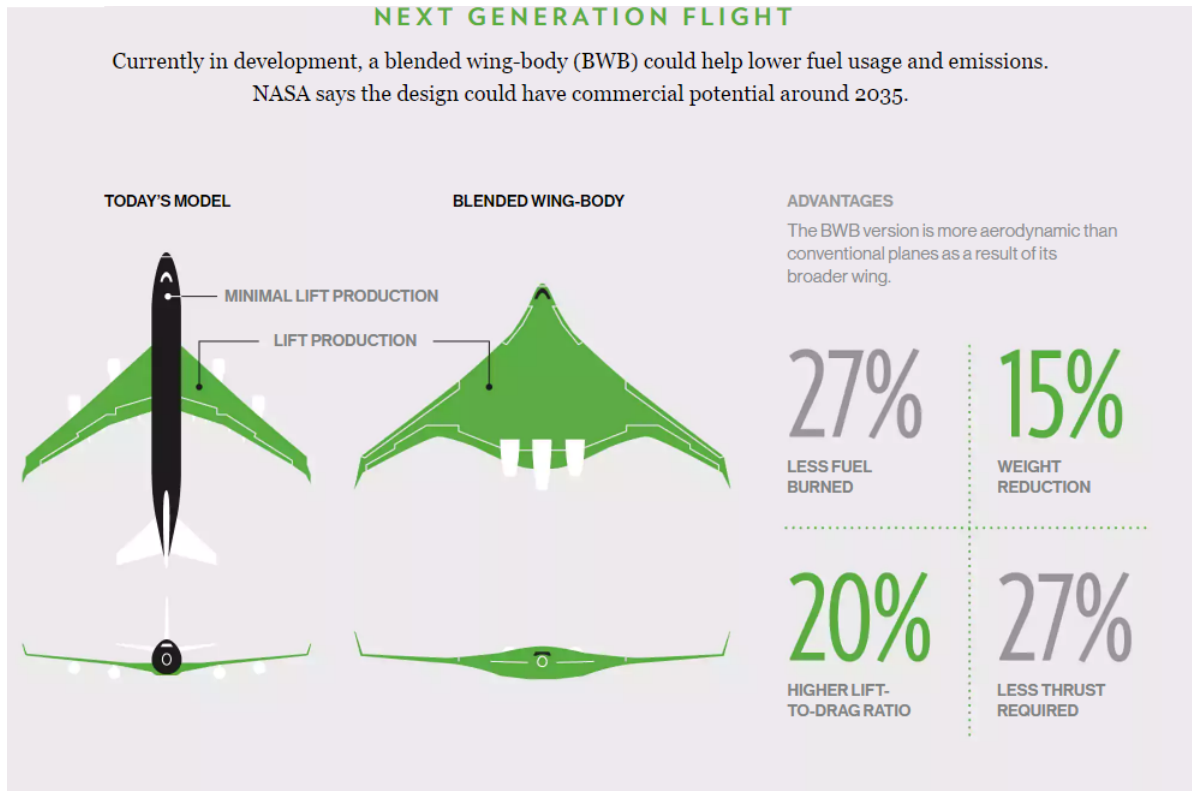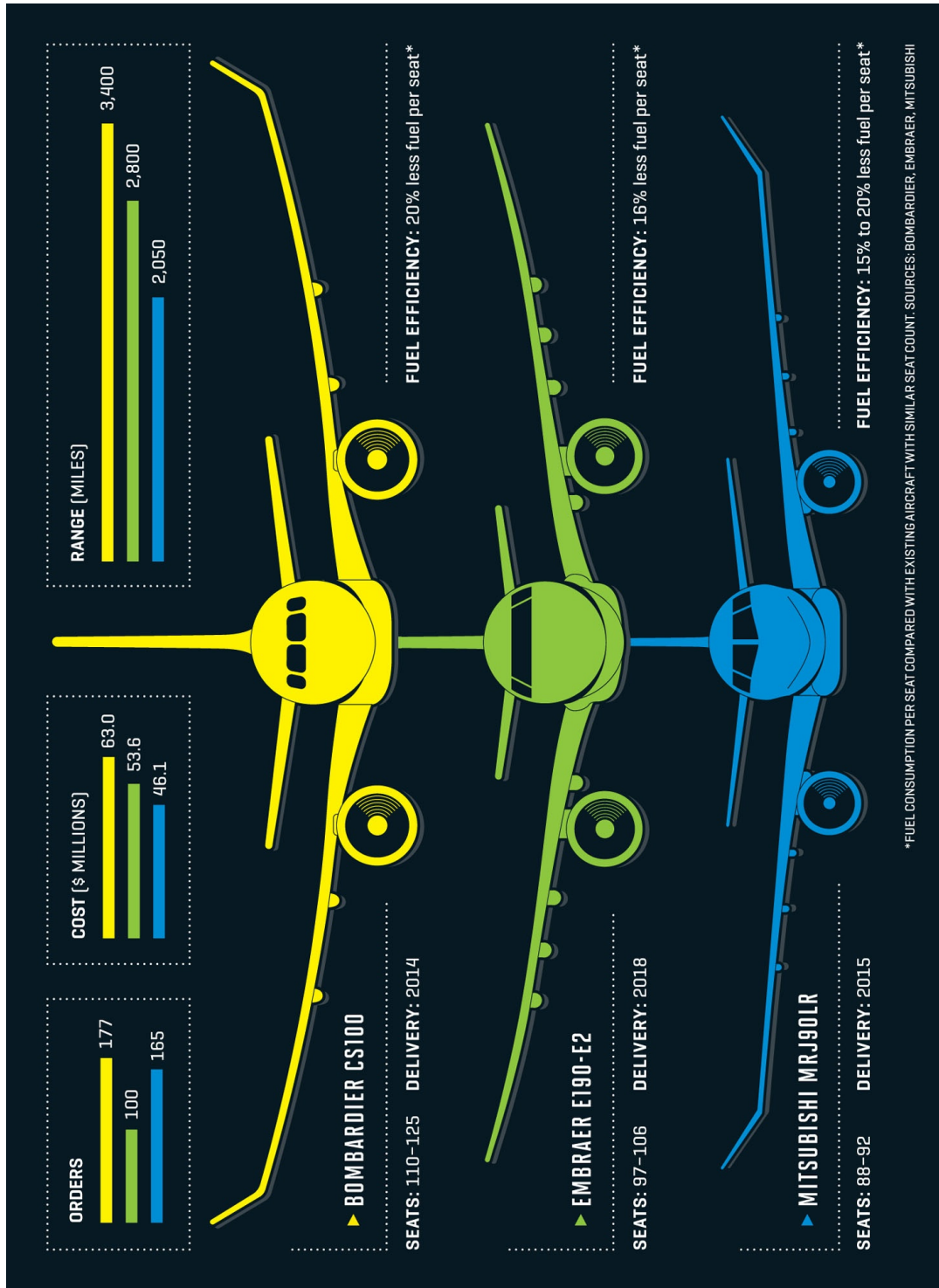
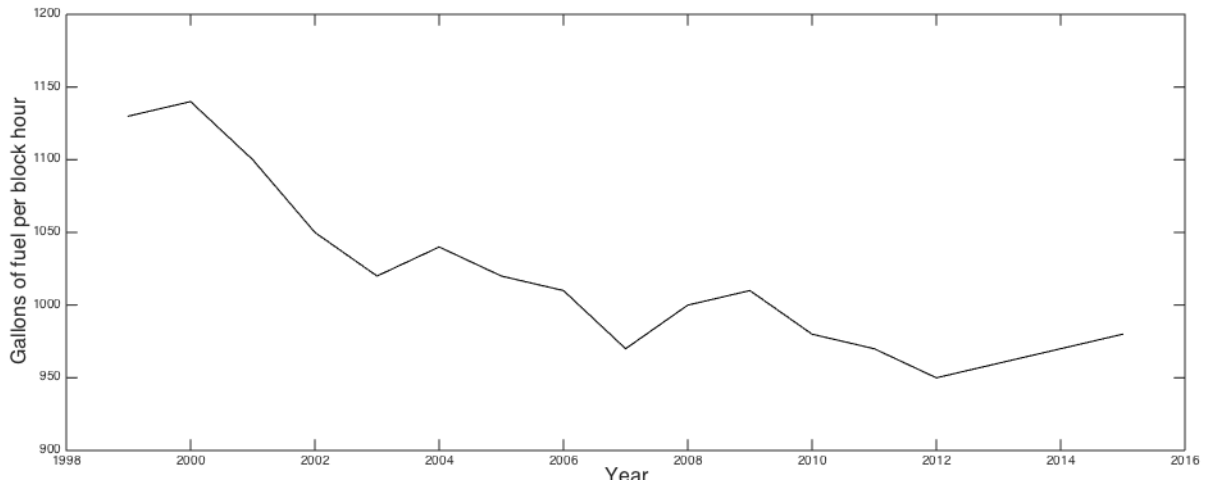Figure E.2: Next generation aircraft.

Figure E.3: Gallons of fuel burned by block hour.

Figure E.4: Fuel price and fuel expense, percentage change.