Spatial Competition in the Airline Industry*

Dennis B. McWeeny†

Job Market Paper

November 16, 2018

Most recent version here

Abstract

Airline passengers consider flights departing from airports in different cities to be substitutes and sometimes travel large distances to board a flight with lower fares. While this "airport leakage" phenomenon is a major concern for airport administrators, the industrial organization literature has assumed that flights departing from airports in different cities are in totally separate markets. This assumption rules out these substitution patterns altogether and could yield biased estimates of elasticities and markups. Using an airline passenger survey conducted annually at San Francisco International Airport, I estimate a structural model of air travel demand that allows consumers to choose among flights departing from airports in different cities. I then compare the results from my model to those from the conventional model that defines markets as origin-destination airport pairs. I find that leisure passengers are willing to travel up to 69 miles to save $100 on airfare. As a result, demand is 74 percent more elastic and markups are 41 percent lower when spatial competition is accounted for. These results suggest that airlines face substantial competition from flights departing from nearby airports and that the origin-destination airport pair definition of airline markets overstates market power.

*I am extremely grateful for the guidance and support provided by my advisor Alan Sorensen and committee members Kenneth Hendricks and Lorenzo Magnolfi. I also thank Jean-François Houde, Christopher Sullivan, Gary Baker, Diwakar Raisingh, and Elan Segarra for insightful comments as well as numerous seminar participants at the University of Wisconsin-Madison for excellent suggestions. This research was funded in part by a Juli Plant Grainger Summer Research Fellowship.

†Department of Economics, University of Wisconsin-Madison. Email: dmcweeny@wisc.edu;
Website: http://sites.google.com/site/dennismcweeny/

1 Introduction

Airline passengers view flights departing from airports in different cities as substitutes. Many passengers are willing to travel significant distances to board flights with lower fares than those offered at their local airport. For example, Figure 1 shows the distribution of home locations for a sample of domestic passengers departing from San Francisco International Airport (SFO). While a large fraction of passengers reside in the city of San Francisco itself, a significant share travel from as far away as San Jose and even Sacramento. San Francisco International is the closest airport for only 43 percent of these passengers, meaning a majority bypass airports that are closer to their home in favor of a flight out of SFO.

Airports are aware of these substitution patterns and often run marketing campaigns to prevent "airport leakage" by enticing consumers to fly from their local airport rather than a more distant hub. For example, Des Moines International Airport launched a "Do the Math" advertising campaign in 2005 to encourage consumers to avoid the long drive to Omaha or Kansas City and fly from Des Moines instead. Similarly, Milwaukee's General Mitchell International Airport used the slogan "Avoid the Chicago ORDeal" to convince passengers to fly from Milwaukee rather than Chicago's O'Hare International Airport. More recently, Dane County Regional Airport in Madison, Wisconsin launched a "Fly Local" ad campaign,[1] including an online trip cost calculator that compares Madison's fares, parking fees, and travel costs to those for Milwaukee's General Mitchell International Airport and Chicago's O'Hare International Airport.[2]

Despite this evidence of substitution between flights departing from different cities, the industrial organization literature has assumed that airline markets consist of either origin-destination city pairs, or more often, origin-destination *airport* pairs. This assumption entirely rules out the substitution patterns discussed above since flights originating from different airports are assumed to be in totally separate markets. Ignoring these substitution patterns and defining the market too narrowly could yield biased estimates of elasticities and markups and thus overstate the degree of market power.

In this paper, I use an airline passenger survey conducted annually at San Francisco International Airport to estimate a structural model of air travel demand that allows consumers to choose among flights departing from airports in different cities. The model accounts for the cost of ground travel between a passenger's home and the departure airport for her flight.

[1] Graduate students at the University of Wisconsin-Madison almost never fly from Dane County Regional (even for domestic flights), choosing to take a 2.5-hour bus ride to Chicago-O'Hare instead.

[2] See http://www.msnairport.com/flight_travel/cost. In my study area, Charles M. Schulz Sonoma County Airport in Santa Rosa, California provides a similar calculator; see http://sonomacountyairport.org/passengers/cost-comparison/.

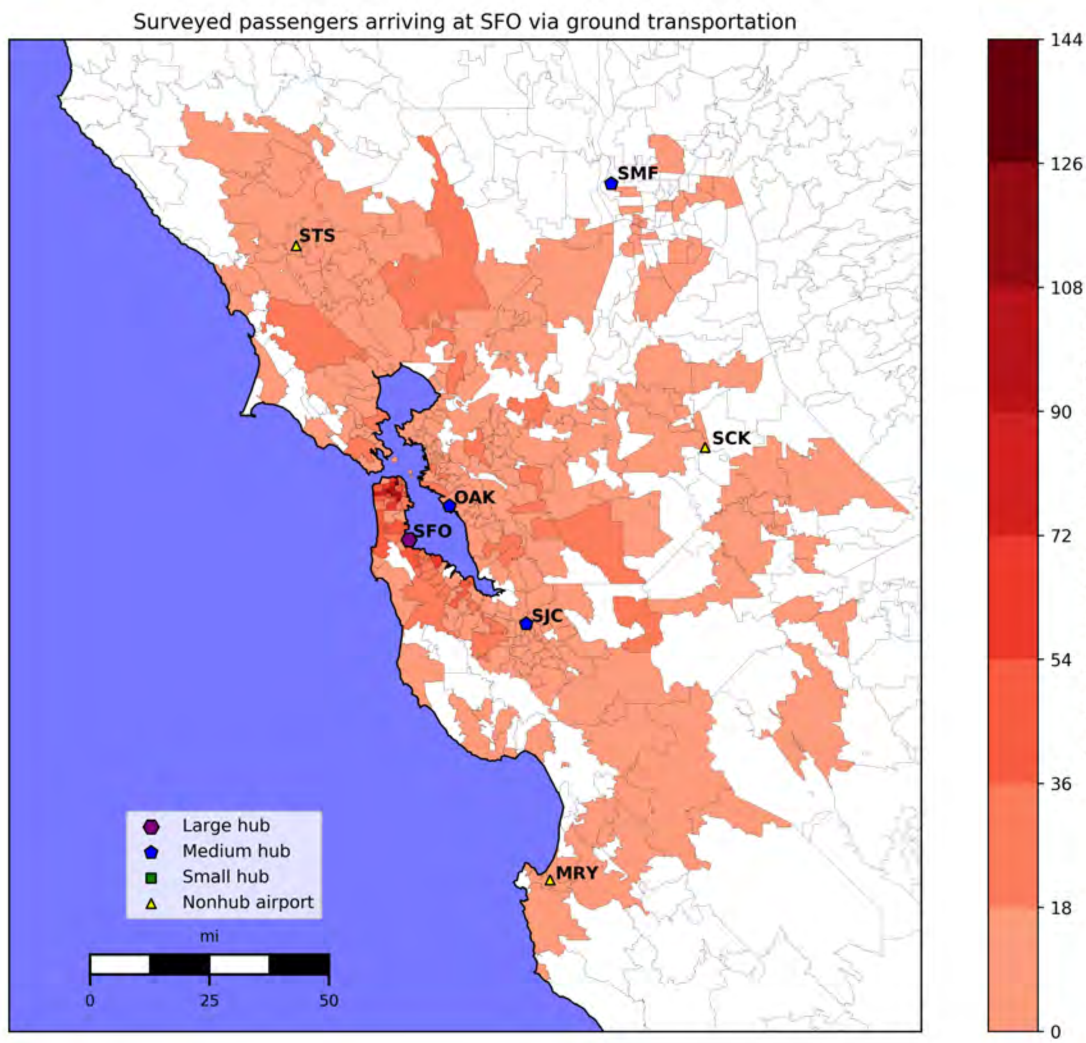

Figure 1: Spatial distribution of home locations for domestic passengers departing from San Francisco International Airport (SFO), 2010–2017

Since the passenger survey data contains information on not only the flight taken by each passenger but also the location of their home, I am able to precisely estimate the degree to which passengers trade off costly ground travel against lower fares. I then compare the results generated by my model to those generated by the conventional model that defines markets as origin-destination airport pairs.

I estimate the cost of ground travel for leisure passengers to be $1.45 per mile; these passengers are willing to travel up to 69 additional miles on the ground in order to save $100 on airfare. These costs are sufficiently low to allow significant substitution between airports. For example, my results show that United's flight from San Francisco to Denver is nearly as close a substitute for Southwest's flight from San Jose to Denver as it is for Southwest's flight from San Francisco to Denver. As a consequence, I find that demand is 74 percent more elastic and markups are 41 percent lower when spatial competition is accounted for. Taken together, these results indicate that airlines face significant competition from flights departing from nearby airports, and failing to account for this additional competition yields misleading estimates of price elasticities and markups. As a result, the narrow airport pair definition of airline markets overstates market power.

These results have significant implications for merger evaluation and antitrust enforcement, in which the definition of the relevant geographic market plays a crucial role. The United States Department of Justice and Federal Trade Commission use a simple "hypothetical monopolist test" to determine the relevant market. As the 2010 Horizontal Merger Guidelines state,

> "The hypothetical monopolist test requires that a hypothetical profit-maximizing firm that was the only present or future seller of the relevant product(s) to customers in the region would impose at least a SSNIP [small but significant and non-transitory increase in price] on some customers in that region. A region forms a relevant geographic market if this price increase would not be defeated by substitution away from the relevant product or by arbitrage, e.g., customers in the region travelling outside it to purchase the relevant product" (§4.2.2).

The antitrust authorities typically use city pairs as the definition of airline markets; if consumers do indeed substitute towards flights at airports in different cities in response to a price increase, then both the airport pair and city pair definitions of airline markets fail the hypothetical monopolist test and are therefore inappropriate market definitions.

To quantify the implications of spatial competition for antitrust policy, I use both models to simulate the price effects of two recent mergers: the 2015 merger between American Airlines and US Airways and the 2018 merger between Alaska Airlines and Virgin America. Though these results are preliminary, I find that predicted postmerger price increases are about one percent in the home address model but between 2.5 and 3.5 percent in the airport pair model for nonstop routes in which the merging parties competed directly. These results again indicate the necessity of accounting for spatial competition.

The remainder of the paper is organized as follows. Section 2 discusses the related literature. I introduce my home address model as well as the conventional airport pair model of air travel demand in Section 3. Section 4 discusses the data, including the customer survey data from San Francisco International Airport. Section 5 outlines the estimation procedure and discusses identification. Section 6 shows the estimation results, including the price elasticities and markups implied by each model. I discuss the simulations of the American Airlines–US Airways and Alaska Airlines–Virgin America mergers in Section 7. Finally, Section 8 concludes.

2 Related Literature

This paper adds to the growing literature on spatial competition in markets with differentiated products. Several papers have expanded the model of product differentiation in Berry, Levinsohn and Pakes (1995) to allow for spatial differentiation in the spirit of the Hotelling model. In particular, these models account for the cost of travel between the consumer's home address and the location of the product to be purchased, generating a trade-off between lower prices and the costs of ground travel. Most of these papers study local retail establishments such as gasoline stations (Manuszak, 2001; Manuszak and Moul, 2009), grocery stores (Smith, 2004), fast food restaurants (Thomadsen, 2005), movie theaters (Davis, 2006), coffee shops (McManus, 2007), bank branches (Ho and Ishii, 2011), and liquor stores (Seim and Waldfogel, 2013). Houde (2012) also studies spatially differentiated gasoline stations but models travel costs as a function of deviations from a consumer's regular commuting path; this model nests the home address model as a special case. Since air travel is less frequent and more expensive than retail purchases, consumers are likely to consider airline products that are much further from their home than in these retail markets (e.g., see Figure 1). As a result, the set of competing products may also be more geographically dispersed than in typical retail markets.

There is also a vast literature in industrial organization on the U.S. airline industry; here I restrict attention to studies that estimate differentiated product models of air travel demand. Berry, Carnall and Spiller (2006) was the first study to apply the methods of Berry, Levinsohn and Pakes (1995) to the airline industry; they study the effects of hubs on both costs and demand. Peters (2006) estimates nested logit and generalized extreme value models of demand and compares post-merger prices predicted by the model to observed post-merger prices for four mergers in the 1980s. Armantier and Richard (2008) use a similar approach to study the welfare implications of airline alliances and codeshare agreements. Berry and Jia (2010) study the evolution of the airline industry after the wave of mergers and bankruptcies in the early 2000s. Ciliberto and Williams (2014) study the link between multimarket contact and collusive behavior and relax the usual assumption of Bertrand-Nash pricing behavior. Two recent papers, Ciliberto, Murry and Tamer (2016) and Li et al. (2018), jointly model demand, marginal costs, and entry in airline markets and simulate the effects of recent mergers.

Most of these papers define airline markets solely as origin-destination airport pairs; the only exceptions are Berry, Carnall and Spiller (2006), Peters (2006), and Berry and Jia (2010).[3] Peters (2006) recognizes not only that flights from different airports within the same

[3]As Peters (2006) remarks, "It is surprisingly rare in the airline economics literature to allow for imperfect

city likely compete with each other, but also that flights departing from or arriving at the same airport are likely to be closer competitors. He uses a generalized extreme value model that groups products departing from and arriving at the same airports into a nest.[4] Though their primary specification uses an airport pair definition, Berry and Jia (2010) recognize that this definition might be too restrictive in cities with multiple airports. As a robustness check, they allow six groups of geographically adjacent airports to be in the same market[5] and find that demand is about 20 percent more elastic when the broader market definition is used. Though these papers address potential substitution among airports in the same city, none have recognized that passengers may travel from far outside the city to board a flight, creating competition among distant airports.

Brueckner, Lee and Singer (2014) address the airline market definition problem directly; they use regression analysis to relate route-level fare changes at a metro area's primary airport to the number of competing airlines at other airports in the area. Based on their results, they suggest that flights from San Francisco International and Oakland International be grouped in the same market, but not flights from San Jose International. In contrast, the SFO Customer Survey data shows that 32 percent of consumers flying from SFO have also chosen to fly from San Jose in the past two years, indicating that the appropriate market definition in the Bay Area should also include San Jose International. My estimation results show that flights from San Jose are nearly as close substitutes for flights from San Francisco as are flights from Oakland.

Finally, several studies in the transportation management literature have used a 1995 San Francisco Bay Area travel survey to estimate multinomial logit models of airport and airline choice. The most closely related of these is Ishii, Jun and Van Dender (2009), who focus on travel from the Bay Area to one of the four airports in the Los Angeles area. Though they examine how passengers trade off various airport and airline characteristics, they do not allow substitution to an outside option (e.g., not flying), do not account for unobserved product-level characteristics (lumping these into the logit error), and do not properly address the endogeneity of fares.[6] Furthermore, they do not report own- or cross-price elasticities and do not examine market power.

substitution between airports. More commonly, airports within the same metropolitan area are treated as identical or as completely separate markets" (p. 632, footnote 15).

[4]This method is a particularly attractive alternative to the home address model when data on ground travel distances is not available. I plan to estimate a similar cross-nested logit model and compare the results.

[5]One of these groups of airports consists of the three major airports in the Bay Area: San Francisco International, Oakland International, and San Jose International.

[6]They use a large set of control variables but do not instrument for fares. As a result, their fare coefficients are close to zero and statistically insignificant in most specifications, indicating they have not identified demand.

3 Model

This section discusses my model of demand and supply in the airline industry. My modeling approach closely resembles Berry, Carnall and Spiller (2006) and Berry and Jia (2010), who use random coefficient discrete choice models à la Berry, Levinsohn and Pakes (1995) to model demand. I augment these models by explicitly modeling consumer substitution between departure airports. In particular, I account for the cost of ground travel prior to reaching the desired airport by using a "home address" model of the kind used in other studies of spatial competition with differentiated products.

3.1 Demand

The primary modeling objective is to relax the assumption that airline markets consist of origin-destination city pairs. To do so, I define a market as travel from the San Francisco Bay Area to another large city in a given quarter. The set of products in each market consists of all flights from any airport in the Bay Area to that city in that quarter, including both nonstop and connecting service. Thus a product is defined as a combination of an origin airport, destination airport, airline, and nonstop or connecting service type.

The set of potential consumers in each market consists of all people within 100 miles of San Francisco. Consumers reside at the population centroid of each ZIP code in the Bay Area (i.e., their "home address"). In each quarter and in each market, all consumers decide whether to fly to that destination city or not, and conditional on choosing to fly, they choose which product to purchase. Consumers differ in their valuations for each product for three reasons. First, consumers living in different ZIP codes must travel different distances to reach the departure airport for a given flight; thus ground travel distances vary across consumers. Second, consumers are distinguished by their traveler type, defined as business and leisure travelers. This distinction has been shown to be important in the airline industry and allows correlation in preferences for certain product characteristics. Third, consumers differ in their idiosyncratic preferences for each product.

More formally, the indirect utility of consumer i of traveler type r living in location l from taking flight j to destination city m in quarter t is given by

$$u^r_{iljmt} = \delta_{jmt} + \eta^r_{iljmt}$$

where

$$\begin{aligned} \delta_{jmt} &= \boldsymbol{x}'_{jmt}\boldsymbol{\beta} + \alpha p_{jmt} + \xi_{jmt} \\ \eta^r_{iljmt} &= \beta^r + \alpha^r p_{jmt} + \tau d_{lj} + \zeta_{imt}(\sigma) + (1-\sigma)\epsilon_{ijmt}. \end{aligned}$$

Thus indirect utility can be decomposed into two components. The first component, δ_{jmt}, is product-specific but does not vary across consumers, while the second component, η^r_{iljmt}, depends on both product and consumer characteristics. Within δ_{jmt}, $\boldsymbol{x}_{jmt}$ is a vector of observed product characteristics (e.g., flight distance), p_{jmt} is the fare for product j, and ξ_{jmt} captures product characteristics that are unobserved by the econometrician. The consumer-specific component of indirect utility consists of a traveler type-specific term $(\beta^r + \alpha^r p_{jmt})$, a location-specific term (τd_{lj}), and an idiosyncratic term $(\zeta_{imt}(\sigma) + (1-\sigma)\epsilon_{ijmt})$. The type-specific constant β^r allows flying to be more attractive for some consumers (e.g., leisure travelers may be willing to take a road trip while business travelers face time constraints that make this infeasible). Similarly, the coefficient α^r allows business travelers to be less price-sensitive than leisure travelers (as business travelers are typically reimbursed for travel expenses through their employer). Utility also depends on d_{lj}, the great circle distance in miles[7] between the consumer's home and the departure airport for product j, which captures spatial differentiation among departure airports for a given consumer. The parameter τ is thus the marginal disutility of ground travel.

Finally, ϵ_{ijmt} is a Type 1 Extreme Value random variable that captures consumer i's idiosyncratic preferences for product j. The term $\zeta_{imt}(\sigma)$ follows a Cardell distribution and differentiates air travel from the outside option (i.e, not flying).[8] Under these assumptions, the idiosyncratic component of indirect utility $[\zeta_{imt}(\sigma) + (1-\sigma)\epsilon_{ijmt}]$ follows a Type 1 Extreme Value distribution for any $\sigma \in [0, 1)$. The indirect utility of the outside option is normalized to $u^r_{il0mt} = \epsilon_{i0mt}$, where ϵ_{i0mt} is also a Type 1 Extreme Value random variable. These distributional assumptions imply that choice probabilities have a nested logit form, where one nest consists of all air travel products and the other nest contains only the outside option. The nesting parameter σ measures the degree of differentiation between air travel and the outside option. As σ approaches zero, the choice probabilities take the usual multinomial logit form and there is no distinction between nests; as σ approaches one, products within the air travel nest become closer substitutes.

[7] The great circle distance may be a poor approximation for a few reasons. First, the presence of San Francisco Bay means that consumers often must take circuitous routes to get to their destination, either around the bay or across one of the five bridges in the area. Second, consumers can take many different forms of ground transportation (car, bus, BART, etc.), each of which have different costs. Third, travel times can vary a lot depending on traffic patterns and the time of the day. In the future, I plan to use average travel times rather than great circle distance in order to obtain a better approximation of travel costs.

[8] Note that the outside option does not distinguish other forms of transportation (such as car, bus, or train) from not traveling to the destination at all. Though this is an important distinction, due to insufficient data it is impossible to separately identify the share of consumers who travel by some other mode from those who do not travel at all.

The probability that a consumer of type r in home location l purchases product j is

$$s^r_{ljmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta}) = \exp\left(\frac{\delta_{jmt} + \beta^r + \alpha^r p_{jmt} + \tau d_{lj}}{1-\sigma}\right) \cdot \frac{1}{D^r_{lmt} + (D^r_{lmt})^\sigma} \tag{1}$$

where

$$D^r_{lmt} = \sum_{k=1}^{J_{mt}} \exp\left(\frac{\delta_{kmt} + \beta^r + \alpha^r p_{kmt} + \tau d_{lk}}{1-\sigma}\right).$$

Market shares for a given product are then calculated by aggregating the choice probabilities over traveler types and home locations:

$$s_{jmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta}) = \sum_r \sum_l \mu^r_l \cdot s^r_{ljmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta})$$

where μ^r_l is the probability mass of consumers of traveler type r in location l. The model's prediction of the number of passengers who choose product j is then given by $q_{jmt} = M_{mt} \cdot s_{jmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta})$, where M_{mt} is a measure of market size (i.e., the number of potential consumers). I will discuss market size in detail in Section 4.2.

I model substitution between flights departing from different airports in a very flexible manner. In principle, one could apply the same technique to model substitution among flights *arriving* at different airports as well. I focus on substitution between departure airports for two reasons. First, customer survey data from San Francisco International Airport allows me to directly observe the ground distance traveled prior the the passenger's arrival at the departure airport and aids in the identification of travel costs. Analogous information on the distance between the arrival airport and the passenger's ultimate destination is not available. Second, ground travel at a passenger's destination is typically more onerous than it is near their home; passengers can drive their own vehicles to their departure airport but must either rent a car or rely on taxi service or public transportation at their destination. Thus, it is likely that the degree of substitution between distant arrival airports is small relative to the degree of substitution between departure airports. I therefore allow products arriving at different airports within the same city to be in the same market but exclude flights landing in other cities.

3.2 Supply

I assume that airlines choose prices according to a static Bertrand-Nash equilibrium.[9] Airline a's variable profits in market m in quarter t are

$$\Pi_{amt} = \sum_{j \in \mathcal{J}_{amt}} (p_{jmt} - mc_{jmt}) \cdot M_{mt} \cdot s_{jmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta})$$

where $\mathcal{J}_{amt}$ is the set of products offered by airline a, mc_{jmt} is the per-passenger marginal cost for flight j, and M_{mt} is the market size.

The first order condition in p_{jmt} is

$$\frac{\partial \Pi_{amt}}{\partial p_{jmt}} = s_{jmt}(\boldsymbol{\delta}_{mt}, \boldsymbol{p}_{mt}; \boldsymbol{\theta}) + \sum_{k \in \mathcal{J}_{amt}} (p_{kmt} - mc_{kmt}) \cdot \frac{\partial s_{kmt}(\cdot)}{\partial p_{jmt}} = 0$$

where

$$\frac{\partial s_{jmt}(\cdot)}{\partial p_{kmt}} = \begin{cases} \sum_r (\alpha + \alpha^r) \sum_l \mu_l^r \cdot s_{ljmt}^r \cdot \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} s_{ljmt|F}^r - s_{ljmt}^r \right) & \text{if } j = k \\ -\sum_r (\alpha + \alpha^r) \sum_l \mu_l^r \cdot s_{ljmt}^r \cdot \left(-\frac{\sigma}{1-\sigma} s_{lkmt|F}^r - s_{lkmt}^r \right) & \text{if } j \neq k \end{cases}$$

where $s_{ljmt|F}^r$ is the probability a consumer of type r in location l chooses product j, *conditional on choosing to fly.*

In each m and t, the first order conditions for all firms can be rewritten in matrix notation as

$$\boldsymbol{s}_{mt} + \boldsymbol{\Omega}_{mt}(\boldsymbol{p}_{mt} - \boldsymbol{mc}_{mt}) = \boldsymbol{0}. \tag{2}$$

The matrix $\boldsymbol{\Omega}_{mt}$ is determined by elementwise multiplication of the Jacobian of the demand system and an "ownership matrix" indicating whether products j and k are operated by the same airline. More formally,

$$\Omega_{jkmt} = \begin{cases} \frac{\partial s_{jmt}(\cdot)}{\partial p_{kmt}} & \text{if } \exists a : \{j, k\} \subseteq \mathcal{J}_{amt} \\ 0 & \text{otherwise} \end{cases}.$$

[9] I recognize that this is a strong assumption. In reality, airlines use complex price discrimination and revenue management schemes to set fares (Lazarev, 2013; Sweeting, Roberts and Gedge, 2018; Williams, 2018). Furthermore, there is some evidence of tacit collusion in the industry (Evans and Kessides, 1994; Ciliberto and Williams, 2014; Ciliberto, Watkins and Williams, 2018). However, the static Bertrand-Nash assumption is standard in the literature.

A simple rearrangement of equation (2) yields

$$\boldsymbol{mc}_{mt} = \boldsymbol{p}_{mt} + \boldsymbol{\Omega}_{mt}^{-1}\boldsymbol{s}_{mt}.$$

This means that although marginal costs are not directly observed, they can be obtained from observed prices, market shares, and estimates of the demand parameters. Thus I first estimate demand and then recover marginal costs.[10]

3.3 Airport Pair Model

In addition to the home address model described above, I also estimate a conventional model that defines airline markets as origin-destination airport pairs; I refer to this model as the airport pair model. The structure of the airport pair model is exactly the same as the model described above, except for two important features. First, following the previous literature, there is no spatial heterogeneity among consumers in the airport pair model; ground travel costs do not enter the utility function. However, I still allow preferences to vary across business and leisure travelers; the overall fraction of each type of traveler is the same as in the home address model. Second, the set of competing products is different in each model. In the home address model, all flights departing from any airport in the San Francisco Bay region and arriving at the same city are considered to be competitors, while in the airport pair model the set of competing products is restricted to flights that share the same origin and destination airports.[11] As a result of these fundamentally different market definitions, market shares are calculated differently in each model.

4 Data

This section describes the data used in this study. I combine data from several sources, the most novel of which is a survey of airline passengers conducted annually at San Francisco International Airport (SFO). I also use data from the Airline Origin and Destination Survey (DB1B) from the Department of Transportation, a source frequently used in other studies

[10]In the future I plan to estimate both demand and marginal costs jointly. Joint estimation somewhat complicates the estimation algorithm, as $\boldsymbol{\Omega}_{mt}$ must be calculated for every candidate value of the parameters.

[11]Recall that the set of competing products at the destination end of the market in the home address model is also more broadly defined than in the airport pair model. Though I believe that my results are overwhelmingly driven by the broader market definition at the origin end of the trip, I plan to conduct two robustness checks that modify the airport pair model to allow (1) substitution between flights arriving at different airports within the same city, and (2) substitution between flights at different airports at both the origin and destination cities, which would define markets as city pairs. The results shown in the current draft of the paper show only the opposite ends of the spectrum of plausible market definitions.

of the airline industry. Finally, I obtain data on the population of each ZIP code and metropolitan statistical area (MSA) from the Census Bureau and data on the distribution of household income by ZIP code from the Internal Revenue Service.

4.1 San Francisco International Airport Customer Survey

Each year, San Francisco International Airport (SFO) conducts a comprehensive survey of passengers flying on every route and airline departing from the airport. The survey is administered at the boarding gate prior to flight departure; though response is voluntary, the sample is representative of all departing passengers.[12] The survey is typically conducted over the span of a few weeks in May, though two of the surveys were conducted in September.[13] I use the surveys from 2010 through 2017.

The survey asks a wide variety of questions, not only about the passenger's experience as a customer at SFO (e.g., whether the bathrooms are clean and whether airline staff are friendly), but also about their travel choices and demographic characteristics. These questions include the purpose of their current trip (e.g., business, vacation, visiting relatives, etc.) as well as their gender, age, and household income. Most importantly for my purposes, the survey also asks how the passenger arrived at the airport, what county the passenger departed from prior to arriving at the airport, and the passenger's home ZIP code. These data allow me to observe how far each passenger traveled on the ground prior to boarding their flight and identifies the cost of ground travel in the model.

Despite the wealth of information contained in the survey, there are two important pieces of information that the survey does not contain. The first of these is the passenger's final destination. Though I observe the exact flight number and thus destination of the flight departing from San Francisco, I do not observe whether the passenger made any additional connections after the conclusion of that flight. According to data from the Department of Transportation's Airline Origin and Destination Survey, over 92 percent of all domestic passengers departing from SFO in the second quarter of 2017 were nonstop passengers. Because this fraction is so high, I assume that all surveyed passengers are nonstop passengers.[14] How-

[12] The response rate was 53 percent in 2017. According to a conversation with one of the airport executives, the response rate in prior years was similar. Though non-response bias is a legitimate concern, I have found that the characteristics of respondents broadly match the characteristics of airline passengers obtained from other sources. For example, the fraction of respondents who are business travelers approximately matches the fraction found by Borenstein (2010), who used data from the 1995 American Travel Survey.

[13] The September surveys were taken in 2010 and 2014. These delays were a result of contracting issues in 2010 and major runway maintenance in 2014. According to a conversation with an airport executive, SFO's passenger traffic patterns in May and September tend to be very similar, which is why September was chosen in those years.

[14] One possible alternative is to "integrate" over the distribution of final destinations conditional on reaching the endpoint of the flight departing from SFO. This would relax the nonstop assumption but significantly

ever, the model still allows the choice sets for these passengers to contain connecting flights. The second important omission is the actual fare paid by the passenger; the survey does not ask any questions regarding fares or ancillary fees. I instead use information on the average fare for each route from the Airline Origin and Destination Survey.

In all, approximately 3,000 passengers are surveyed each year. To obtain a sample of passengers to use during estimation, I drop passengers who arrived at SFO via a connecting flight, passengers flying outside the contiguous United States, and passengers flying to a destination that is not in the sample of 75 cities discussed in the next subsection. I also drop passengers who do not report their home ZIP code or county of departure since this information is crucial for my analysis. I also drop any passengers departing from a county other than their home county, as the model assumes that all passengers depart from their home prior to arriving at SFO.[15] Finally, I drop passengers whose home ZIP code is more than 100 miles from SFO, as it is highly likely that they misunderstood one of the survey questions. Specifically, the question asking for the passenger's mode of arrival is stated as "How did you get to *the airport* today?" (italics mine). The home locations for these outlying passengers tend to be clustered around smaller airports far to the north and south of San Francisco.[16] Since this question is also how connecting status is determined, it is possible that these passengers actually arrived at SFO via a connecting flight from their local airport and interpreted "*the airport*" as their local airport rather than SFO.[17] Overall, this cleaning process yields a final sample of 3,802 passengers, a little less than 500 per year.

As discussed in the introduction, Figure 1 shows the spatial distribution of home locations for passengers in my sample. Though a large number of passengers live in downtown San Francisco, many passengers travel from as far away as Monterey, Stockton, Santa Rosa, and even Sacramento prior to boarding a flight at SFO.[18] Figure 2 shows the distribution of ground travel distances for passengers in the sample. The average ground travel distance is 21 miles, and the median is 15 miles. Again, there is a large mass of consumers living 15 miles away because the densely populated area of downtown San Francisco is located

complicate the estimation procedure.

[15]It might be the case that this step disproportionately drops business travelers, as some may depart from their office or stay overnight in a hotel rather than depart directly from their home. I plan to conduct a robustness check in which such passengers are retained in the sample.

[16]These include California Redwood Coast–Humboldt County Airport (250 miles northwest of SFO), Fresno Yosemite International Airport (158 miles southeast of SFO), Redding Municipal Airport (199 miles north of SFO), and San Luis Obispo County Airport (191 miles southeast of SFO).

[17]These passengers account for less than 1.5 percent of all remaining passengers. As a robustness check, I plan to re-estimate the model with these passengers included.

[18]Figure 7 in Appendix A shows the spatial distribution of business and leisure travelers departing from SFO. As one might expect, leisure travelers tend to come from further away than business travelers, but this difference is slight.

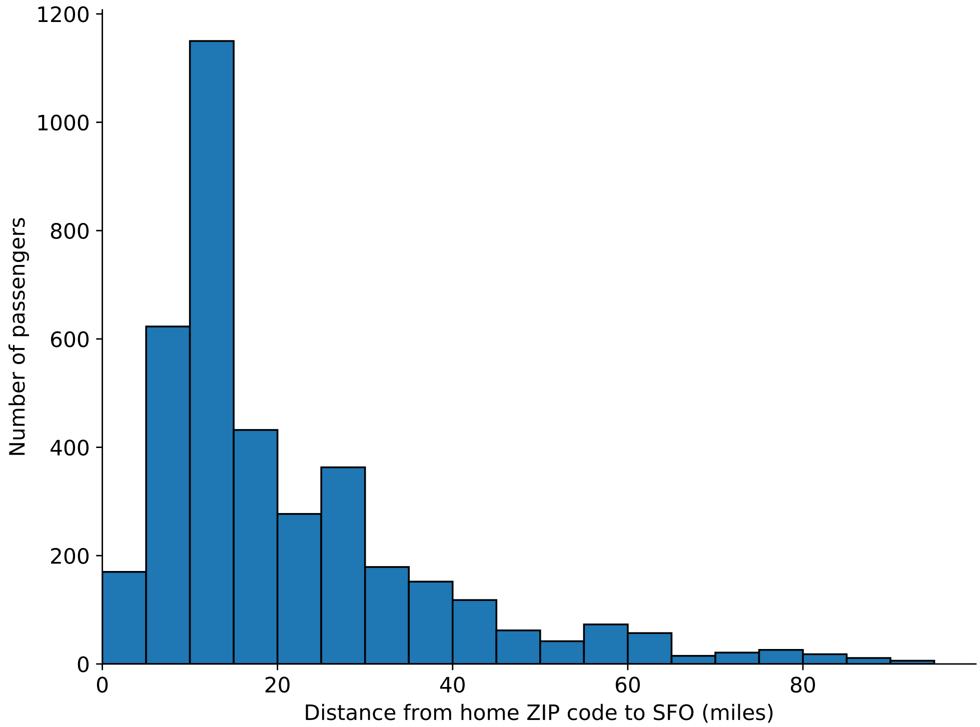

Figure 2: Distribution of ground travel distances for SFO passengers

approximately 15 miles north of SFO. However, there is also a long right tail of passengers traveling more than 30 miles to reach SFO.

Table 1 shows key summary statistics for the sample of passengers. Business passengers make up 32 percent of the sample, which closely matches the fraction of business travelers found in other studies of the industry (see Borenstein, 2010; Berry and Jia, 2010; Ciliberto and Williams, 2014). SFO is the closest airport for only 43 percent of passengers, meaning a majority of passengers bypass an airport that is closer to their home and instead choose to board a flight at SFO. Furthermore, 48 percent of the surveyed passengers also used Oakland International Airport (OAK) within the past two years, 32 percent used San Jose International Airport (SJC) in the same time period, and 17 percent used both airports. These facts suggest that passengers do not simply choose the flight departing from the nearest airport but instead shop for the best flight from several of the airports in the region. Regarding the mode of ground travel, 77 percent of passengers arrived at SFO via some kind of car, either driving themselves, being dropped off by a friend or relative, or taking a taxi, Uber, or Lyft. About 22 percent of passengers drove and parked their own car at the airport. Overall, the differences between business and leisure travelers along these dimensions are small, though business travelers are more likely to drive their own car to their airport and are more likely to have used OAK and SJC in the past two years (probably

Table 1: Summary statistics for surveyed passengers (SFO Customer Survey)

	Business passengers		Leisure passengers		All passengers	
	Mean	(Std. Dev.)	Mean	(Std. Dev.)	Mean	(Std. Dev.)
Ground travel distance (miles)	20.05	(14.82)	21.72	(17.25)	21.19	(16.52)
SFO is closest airport?	0.45	(0.50)	0.42	(0.49)	0.43	(0.50)
Used OAK in past 2 years?*	0.56	(0.50)	0.45	(0.50)	0.48	(0.50)
Used SJC in past 2 years?*	0.38	(0.49)	0.29	(0.45)	0.32	(0.47)
Used both OAK & SJC in past 2 years?*	0.22	(0.42)	0.14	(0.35)	0.17	(0.37)
Arrived via car?	0.81	(0.39)	0.75	(0.43)	0.77	(0.42)
Drove own car?	0.30	(0.46)	0.18	(0.39)	0.22	(0.41)
Number of passengers	1,220 (32.1%)		2,582 (67.9%)		3,802	

*Question asked 2013–2017

because they fly more frequently overall).

4.2 Airline Origin and Destination Survey (DB1B)

The second key data source for this study is the Airline Origin and Destination Survey (DB1B) from the Department of Transportation. The DB1B is a quarterly 10 percent sample of all domestic passenger itineraries from all large commercial airlines in the United States.[19] It contains information on the origin and destination airports, intermediate connections (if any), ticketing and operating carriers, airfare, and number of passengers on each itinerary. In each year from 2010 through 2017, I use the DB1B data for the quarter spanning the collection of each SFO Customer Survey: the third quarter in 2010 and 2014 and the second quarter in all other years.

To select a sample of destination city markets, I begin with the 100 largest metropolitan statistical areas (MSAs) according to their 2010 population. Some of theses MSAs had no flights from airports in the Bay Area during the time period and thus do not appear in the sample. I also combine two pairs of MSAs due to the close proximity of an airport to both MSAs.[20] This results in a sample of 75 potential destination city markets in each quarter.[21]

From the domestic DB1B, I keep itineraries for passengers flying to one of these 75 cities from any of the seven commercial airports within 100 miles of San Francisco.[22] The origin airports included in this sample are San Francisco International (SFO), Oakland International

[19]Some very small regional carriers are not required to report.

[20]These are the Washington-Arlington-Alexandria, DC-VA-MD-WV and Baltimore-Columbia-Towson, MD MSAs (associated with Baltimore-Washington International Thurgood Marshall Airport) and the Hartford, CT and Springfield, MA MSAs (associated with Bradley International Airport).

[21]Not all cities appear in every quarter due to changes in route networks. For example, a flight to a relatively small city may require one connection in some quarters but two connections in another quarter. As I only use itineraries with one connection, some of these multi-connection routes are dropped.

[22]This radius is chosen to match the sample of passengers from the SFO Customer Survey. I plan to conduct a robustness check that includes all airports within 250 miles of San Francisco.

Table 2: Airports in the San Francisco Bay region

Airport	Code	Nonstop destinations*	Originating passengers†	Distance from SFO‡
San Francisco International	SFO	42	9,816,060	–
Oakland International	OAK	36	3,896,110	11.02
San Jose International	SJC	29	3,563,650	30.24
Sacrameto International	SMF	33	3,348,980	85.68
Charles M. Schulz Sonoma County	STS	6	117,830	65.90
Monterey Regional	MRY	4	112,220	77.00
Stockton Metropolitan	SCK	3	79,140	65.10

*Among the sample of 75 destination cities in 2017Q2
†Total originating domestic passengers in 2017
‡Great circle distance in miles

(OAK), San Jose International (SJC), Sacramento International (SMF), Charles M. Schulz Sonoma County (STS), Monterey Regional (MRY), and Stockton Metropolitan (SCK). Table 2 shows the number of nonstop destinations and number of domestic passengers originating at each airport in 2017 as well as each airport's distance from San Francisco International. SFO is the dominant airport in the region and is the fifth-busiest airport in the country as measured by the number of originating domestic passengers. The airports in Oakland, San Jose, and Sacrameto are about one third as busy as SFO; each accounts for between three and four million originating passengers each year. Finally, the region contains three small airports that together only account for about 300,000 passengers a year. While excluding these three airports is unlikely to affect the estimation results,[23] I include them because they may be attractive options for nearby business travelers who have a high opportunity cost for ground travel.

To further clean the data, I restrict attention to round-trip itineraries. In addition, I keep only the outgoing leg of each itinerary, as this more closely matches the assumptions of the home address model. The remaining cleaning process follows the previous literature, especially Ciliberto and Williams (2014). In particular, I drop itineraries with more than one connection on the outgoing leg and obtain the fare for the outgoing leg of the itinerary by dividing the total fare by two. I then drop itineraries with directional fares that are less than $25 or more than $2,500, as these are often the result of data entry errors. Fares are converted to constant 2017 dollars using the consumer price index. I then aggregate itineraries to form products, which consist of a unique combination of an origin airport, destination airport, airline, and nonstop or connecting service type. As is the common practice, this product definition does not consider itineraries with different intermediate connections as distinct

[23]Together, these three airports account for fewer than 20 products per quarter. I plan to run a robustness check in which these three small airports are excluded from the estimation sample.

Table 3: Summary statistics for products (DB1B)

	Mean	(Std. Dev.)	5%	Median	95%
Fare (dollars)	257.12	(72.80)	131.93	262.74	366.19
Direct flight distance (miles)	1,778.38	(742.59)	397.43	1,943.37	2,608.99
Nonstop dummy	0.34	(0.47)	0	0	1
Vacation dummy	0.13	(0.34)	0	0	1
Hub dummy	0.17	(0.38)	0	0	1
Number of products per market					
Home address model	12.20	(10.95)	1	9	32
Airport pair model	2.77	(1.98)	1	2	7
Number of airlines per market					
Home address model	3.89	(2.38)	1	4	8
Airport pair model	2.28	(1.55)	1	2	5
Outside option share					
Home address model	0.9975	(0.0056)	0.9875	0.9995	1.0000
Airport pair model	0.9982	(0.0033)	0.9917	0.9994	0.9999
Number of products			6,757		

products.[24] The fare for a given product is computed as the passenger-weighted average fare across all itineraries. Finally, I drop any products purchased by fewer than 100 passengers in a quarter (i.e., those with fewer than 10 passengers in the DB1B). This process results in a sample of 6,757 products.

Market shares are calculated by dividing the total number of passengers who purchase the product in a given quarter by a measure of market size. The choice of market size must capture the number of consumers who could potentially choose to travel to the destination in a given quarter. In the airport pair model, I follow the previous literature and define market size as the geometric mean of the population of the endpoint MSAs.[25] However, this definition does not make sense in the home address model, as products departing from different MSAs are considered to be in the same market. I instead use the total population living within 100 miles of San Francisco as the market size in the home address model.[26]

Table 3 shows summary statistics for the products in my sample. The average fare is about \$257; there is wide variation in fares across products, ranging from near \$100 to \$400. The average nonstop flight distance is about 1,700 miles and ranges between 300 and 2,700

[24] Thus flying from San Francisco to Des Moines with a connection in Denver is indistinguishable from flying from San Francisco to Des Moines with a connection in Dallas.

[25] Admittedly, this definition is somewhat arbitrary. Some papers suggest alternatives to the geometric mean definition, but it is not clear whether these are any less arbitrary. For example, Li et al. (2018) regress the logarithm of the total number of passengers flying on a given route on route characteristics and then "multiply the predicted traveler number by 3.5 so that, on average, the combined market shares of carriers is just under 30%" (Li et al. (2018), Appendix A, p. 5).

[26] Again, the 100 mile radius was chosen to match the sample of passengers from the SFO Customer Survey. I plan to experiment with alternative market size definitions for the home address model. For example, one could define the market size as the geometric mean of the population within 100 miles of San Francisco and the population of the destination MSA.

miles. Nonstop flights account for about 34 percent of the products but transport about 88 percent of the passengers in the region. Additional product characteristics include a "vacation dummy" that equals one if the flight's destination is in Florida or Las Vegas and a "hub dummy" that equals one if either the origin or destination airport is a hub for the airline offering the product. Around 13 percent of the products go to a vacation destination, and about 17 percent of flights start or end at a hub.

The different market definitions in the home address and airport pair models yield important differences in the number of competitors in each market. The average number of products per market is around twelve in the home address model but less than three in the airport pair model. Similarly, there are roughly four airlines offering service in the median market of the home address model but only two in the airport pair model. Within each model, there is significant variation in both the number of products and number of airlines across markets.[27]

Finally, it is important to note that the share of the outside option (not flying) is typically over 99 percent in both models. Although this is intuitive (very few people in the population as a whole actually fly to a given city in a given quarter), it also means that the model's nesting structure is particularly important for cross-price elasticities. When the share of the outside option is large in a non-nested multinomial logit model, consumers' response to a price change is largely to switch to the outside option rather than to one of the other products. Since the nested logit model allows for tastes for products in the same nest to be correlated, substitution will mostly be towards other products within the nest so long as within-nest correlation in preferences is sufficiently strong.

4.3 Fraction of business travelers by ZIP code

The model also requires information on the masses of business and leisure travelers in each ZIP code, μ_l^r. Unfortunately, these quantities are not directly observed and must be estimated.[28] I estimate these quantities by combining information from the SFO Customer Survey with information on the distribution of household income in each ZIP code from the Internal Revenue Service (IRS).[29]

The SFO Customer Survey asks passengers for the primary purpose of their trip, their

[27] This is an important source of variation that will be useful to identify the nesting parameter σ and common price coefficient α. See Section 5.3 for further discussion.

[28] Borenstein (2010) constructs an estimate of the fraction of business travelers departing from each MSA using the 1995 American Travel Survey, but this data is not available at the finer ZIP code level.

[29] Currently, this procedure is conducted separately from the estimation of the structural parameters. In the future, I plan to estimate both μ_l^r and the structural parameters jointly.

household income bracket,[30] and their home ZIP code. Using this data, I run a logit regression to determine the probability of business travel conditional on household income and home ZIP code population density.[31] Let $\widehat{\Pr(b|y,l)}$ be the predicted fraction of business travelers among people of income bracket y in ZIP code l implied by this logit regression, and let $\Pr(y|l)$ be the fraction of households in income bracket y in ZIP code l (which is directly observed in the IRS data). The probability of being a business traveler conditional on living in ZIP code l is therefore given by

$$\widehat{\Pr(b|l)} = \sum_y \widehat{\Pr(b|y,l)} \cdot \Pr(y|l).$$

The mass of business travelers living in l is then obtained by multiplying $\widehat{\Pr(b|l)}$ by the fraction of the population living in ZIP code l, so that $\sum_r \sum_l \mu_l^r = 1$.

One concern is that the business travel logit regression is estimated using a sample of business travelers who have chosen to fly from SFO; it might be the case that business travelers disproportionately choose SFO due to more frequent departures from the airport, for example. To obtain consistent estimates of the fraction of business travelers, I assume that consumers in a given income bracket and ZIP code who chose to fly from SFO are just as likely to be business travelers as are the people in their income bracket and ZIP code as a whole, so that $\Pr(b|y,l) = \Pr(b|y,l,SFO)$.[32]

The results of this procedure are displayed in Figure 3. As intuition would suggest, the areas with the highest proportions of business travelers are downtown San Francisco and West San Jose; however, it seems that the predicted fraction of business travelers in Sacramento is relatively low. The overall proportion of business travelers in the population is estimated to be 18 percent, much lower than the proportion in the SFO Customer Survey (32 percent). This is a result of two facts. First, when compared to leisure passengers, a larger proportion of business passengers come from high-income households (74 percent versus 46 percent). Second, relative to the population as a whole, a larger proportion of airline passengers come from high-income households; the fraction of surveyed passengers whose household income is above $100,000 is 55 percent, but the corresponding fraction in the population as a whole is only 19 percent. As a result, the fraction of business travelers in the population as a whole is much lower than in the sample of airline passengers.

[30] The household income brackets in the SFO survey are slightly more aggregated than in the IRS data. I synchronize both datasets into three income brackets: less than $50,000, $50,000–$100,000, and greater than $100,000.

[31] The results of this regression are displayed in Table 11 of Appendix A.

[32] This assumption can be relaxed by estimating μ_l^r jointly with the other structural parameters, as the model allows flying from SFO to be an endogenous outcome. I plan to use joint estimation in the future.

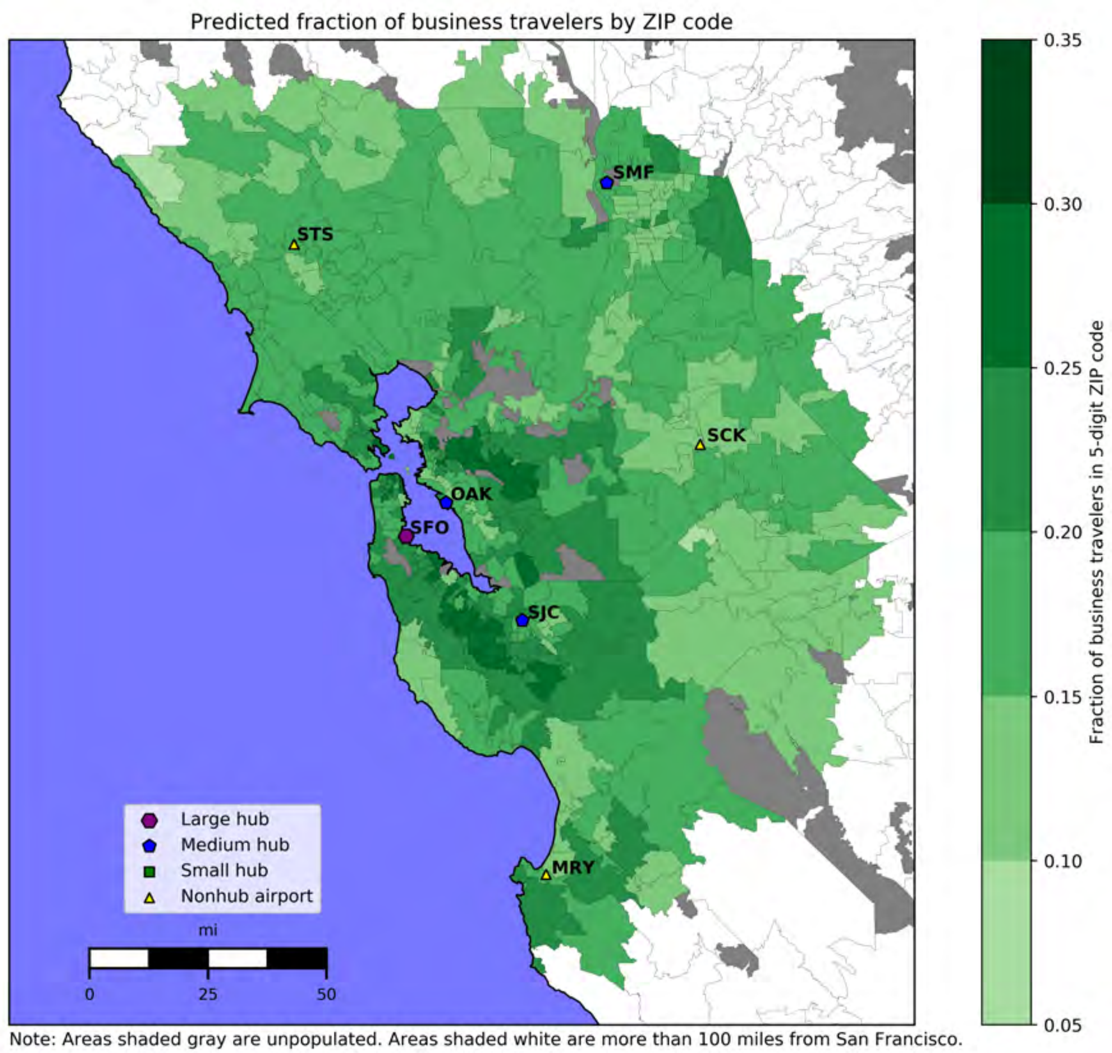

Figure 3: Predicted fraction of business travelers by ZIP code

5 Estimation and Identification

5.1 Estimation Methodology

One intuitive estimation approach would use the data from the SFO Customer Survey alone to estimate the parameters of the model using maximum likelihood. However, choice-based sampling introduces a problem in this framework. By definition, the surveyed passengers have chosen to take a flight from SFO. This means that passengers who choose not to fly at all are absent from the survey. As a result, the nesting parameter σ is not identified from the survey data alone; the degree of substitution between flying and not flying cannot be determined from a sample of passengers who choose to fly.

As a result, I must rely on market-level variation in the choice set to identify the nesting parameter. However, the survey data is still useful for identification of travel costs and the type-specific parameters. I therefore estimate the model using a procedure similar to Berry, Levinsohn and Pakes (2004), which combines data on the purchase decisions of individual consumers with product-level data. In particular, I stack the first order conditions of the log-likelihood function along with standard product-level moment conditions in a nonlinear GMM estimator.

Let $r(i)$, $l(i)$, and $j(i)$ denote the traveler type, home location, and flight chosen by passenger i in the SFO survey. Passenger i's contribution to the log-likelihood function is

$$\mathcal{L}_i(\boldsymbol{\theta}) = \log\left(s^{r(i)}_{l(i),j(i),mt}(\boldsymbol{\theta})\right)$$

where $s^r_{ljmt}(\boldsymbol{\theta})$ is defined in equation (1) and $\boldsymbol{\theta} = (\beta^r, \alpha^r, \tau, \sigma)$. The first set of moment conditions is formulated as the score function of the log-likelihood:

$$\mathbb{E}\left[\frac{\partial \mathcal{L}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] = \mathbf{0}.$$

The second set of moment conditions is formed by interacting the unobserved product characteristic ξ_{jmt} with a vector of instruments $\boldsymbol{z}_{jmt}$ for each product:

$$\mathbb{E}\left[\xi_{jmt}(\boldsymbol{\theta}) \cdot \boldsymbol{z}_{jmt}\right] = \mathbf{0}.$$

After replacing the population moments with their sample analogs, the stacked moment conditions are

$$\boldsymbol{m}(\boldsymbol{\theta}) = \begin{bmatrix} \dfrac{1}{N}\displaystyle\sum_{i=1}^{N} \dfrac{\partial \mathcal{L}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ \\ \dfrac{1}{J}\displaystyle\sum_{j=1}^{J} \xi_j(\boldsymbol{\theta}) \cdot \boldsymbol{z}_j \end{bmatrix}$$

where N is the number of passengers from the SFO Customer Survey and J is the total number of products across all markets and time periods. The GMM estimator is

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \boldsymbol{m}(\boldsymbol{\theta})' \boldsymbol{W} \boldsymbol{m}(\boldsymbol{\theta})$$

where $\boldsymbol{W}$ is a block-diagonal weight matrix.[33]

I use a modification of the BLP contraction mapping so that the optimization algorithm

[33]The weight matrix is block-diagonal because the two sets of moment conditions are derived from two independent sampling processes. See Petrin (2002) and Houde (2012) for similar examples in the demand estimation context.

only needs to search for the values of the nonlinear parameters, namely β^r, α^r, τ, and σ. Berry (1994) showed that for every value of $\boldsymbol{\theta}$, there exists a unique vector of mean utilities $\boldsymbol{\delta}_{mt}$ that matches predicted market shares to observed market shares for every m and t. Grigolon and Verboven (2014) then showed that in the nested logit model with random coefficients, $\boldsymbol{\delta}_{mt}$ solves a slight modification of the original BLP contraction mapping, given by

$$T\left(\delta_{jmt}\right) = \delta_{jmt} + (1-\sigma)\left[\ln\left(s_{jmt}\right) - \ln\left(s_{jmt}(\boldsymbol{\theta})\right)\right].$$

The remaining parameters ($\boldsymbol{\beta}$ and α) can then be obtained from an instrumental variables regression of δ_{jmt} on product characteristics.

5.2 Empirical Specification

In addition to airfare and ground travel distance, several product characteristics are likely to be important to consumers and are included in $\boldsymbol{x}_{jmt}$. All else equal, passengers prefer to avoid layovers by taking nonstop flights, so I include a nonstop flight dummy to capture these preferences. Passengers may also prefer flights that depart from or arrive at a carrier's hub, as these airports may have additional amenities; I therefore include a dummy variable that indicates hub status. I also include a dummy variable for common vacation destinations in Las Vegas and Florida, which may have relatively high demand relative to market size or population. Air travel likely becomes more attractive as the distance between origin and destination increases (since alternatives such as driving become less attractive). However, this effect likely diminishes as the length of the flight becomes tedious. I therefore include both the nonstop flight distance and its square in the vector of product characteristics. In addition to preferring nonstop to connecting flights, passengers may dislike connecting flights with circuitous routes. I include an "extra distance" variable and its square, which measures how many additional miles consumers fly on the average connecting itinerary on a given airline relative to a nonstop flight with the same origin and destination. Finally, I include a set of origin airport, carrier, and year dummies to control for unobserved factors that may affect demand at a particular departure airport (e.g., parking fees), on a particular airline (e.g., baggage fees), or in a particular year (e.g., macroeconomic fluctuations).

5.3 Identification

This section outlines my identification strategy. Berry and Haile (2014) provide a non-parametric identification result for a general class of differentiated product models using only market-level data. In my context, the SFO Customer Survey data provides additional consumer-level variation that aids identification.

The moment conditions derived from the score of the log-likelihood function identify the type-specific coefficients β^r and α^r as well as τ, the disutility of ground travel. The type-specific constant β^r is identified by variation in the proportion of business passengers across different destinations. Similarly, the type-specific price coefficient α^r is identified by variation in fares across products chosen more frequently by business travelers.[34] Travel costs are identified by variation in ground travel distances for passengers who take a given flight.

The second set of moment conditions must identify the remaining parameters. The identification of the coefficients on nonprice product characteristics is straightforward, as they are assumed to be independent of unobserved product characteristics and act as their own instruments. This leaves two parameters: the nesting parameter σ and the common price coefficient α. As discussed at the beginning of this section, I rely on variation in the choice set to identify σ. Berry and Jia (2010) note that σ is identified by variation in the market share of the inside nest as the number of products varies. Thus, a natural instrument for σ is the number of products in each market. As discussed in Section 4.2, there is wide variation in this instrument across markets in both the home address model and the airport pair model. In order to keep the set of instruments identical across models, I use the number of products sharing the same origin and destination airports[35] as an instrument for σ.

Prices are endogenous; consumers observe ξ_{jmt} when making purchase decisions, and firms take this into account when setting prices. Failing to correct for endogeneity could yield biased estimates of the price coefficient α. I use two techniques to address this problem. First, the inclusion of origin airport, carrier, and year dummies eliminates the effect of any unobservables that are constant within these dimensions. Thus, the potential sources of endogeneity are limited to unobservables that vary across origin airports, carriers, and time.

Second, I use a collection of instrumental variables to identify the fare coefficient. To satisfy the relevance and exogeneity conditions, these instruments must be correlated with prices but uncorrelated with the remaining component of unobserved product characteristics. In addition to the number of products with the same origin-destination pair discussed above, I include the number of airlines per origin-destination pair and the number of nonstop flights per origin-destination pair. These instruments reflect the competitiveness of a particular route and are thus strongly correlated with prices. The assumption required to satisfy the exogeneity condition is that airlines make their entry decisions before observing ξ_{jmt}. I

[34] The data does not contain any variation in fares for a given product since the surveyed passengers were not asked to provide information on the fare they paid. Intuitively, such variation would also aid in the identification of α^r, as business travelers likely pay higher fares than leisure travelers, *ceteris paribus.*

[35] This is exactly the same as the number of products per market in the airport pair model.

Table 4: Estimation results by model

	Home address model		Airport pair model	
	Estimate	(Std. Err.)	Estimate	(Std. Err.)
η^r_{iljmt}:				
Ground travel (100 miles) (τ)	-3.586^{***}	(0.012)		
Nesting parameter (σ)	0.134^{***}	(0.000)	0.204^{***}	(0.000)
Business dummy (β^b)	0.357^{***}	(0.006)	0.682^{***}	(0.006)
Business × Fare (\$100) (α^b)	0.123^{***}	(0.002)	0.009^{***}	(0.002)
δ_{jmt}:				
Fare (\$100) (α)	-2.469^{***}	(0.141)	-1.374^{***}	(0.114)
Nonstop dummy	0.804^{***}	(0.056)	0.824^{***}	(0.046)
Vacation dummy	-0.215^{***}	(0.047)	0.193^{***}	(0.041)
Hub dummy	1.009^{***}	(0.061)	0.461^{***}	(0.047)
Direct flight distance (1,000 miles)	1.216^{***}	(0.222)	0.647^{***}	(0.188)
Direct flight distance squared	-0.069	(0.049)	-0.158^{***}	(0.041)
Extra flight distance (1,000 miles)	-0.210	(0.313)	-0.153	(0.262)
Extra flight distance squared	-2.130^{***}	(0.566)	-1.593^{***}	(0.493)
Constant	-4.497^{***}	(0.221)	-6.779^{***}	(0.199)
Origin airport dummies	Yes		Yes	
Carrier dummies	Yes		Yes	
Year dummies	Yes		Yes	
Number of surveyed passengers	3,802		3,802	
Number of products	6,757		6,757	

Standard errors are robust to heteroskedasticity. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

also include several "Differentiation IVs" introduced by Gandhi and Houde (2017), which measure the deviation of a product's exogenous characteristics from the average characteristics of its rivals. These instruments are correlated with prices since they reflect the relative attractiveness of each product but are uncorrelated with unobserved product characteristics since they are nonlinear transformations of exogenous variables. Again, these instruments are constructed at the origin-destination pair level so that the set of instruments is identical across models.

6 Results

6.1 Parameter Estimates and Travel Costs

Table 4 shows parameter estimates and standard errors for both the home address model and the airport pair model. Many of the results are as expected. Business travelers find air travel more attractive than do leisure travelers in both models. Passengers also prefer nonstop flights and flights departing from or arriving at a carrier's hub airport. Air travel demand is increasing but concave in the distance between the origin and destination. In addition, consumers dislike connecting flights with long flight segments.

Table 5: Travel costs and elasticities by traveler type

	Home address model		Airport pair model
	Implied travel cost: $\tau/(\alpha+\alpha^r)$	Median price elasticity	Median price elasticity
Business	\$1.53/mile	-6.05	-4.15
Leisure	\$1.45/mile	-6.40	-4.17

However, there are key differences between the models. The most important of these is the estimated fare coefficient, which I find to be -2.47 in the home address model but -1.37 in the airport pair model. This reflects the fact that consumers have a much larger choice set in the home address model and are able to substitute away from high-priced products at one airport to cheaper products at another airport. Similarly, the nesting parameter σ, which governs the degree of substitution among air travel products in a given market, is somewhat smaller in the home address model than in the airport pair model. Since markets are restricted to products that share the same origin and destination airport, products in a given market are more similar in the airport pair model than in the home address model, which is in turn reflected in the nesting parameter. The differences between these coefficients are reflected in price elasticities, which I discuss below.

Table 5 shows the median price elasticity for each traveler type in each model as well as implied ground travel costs in dollars per mile from the home address model. I find that business travelers are less price sensitive than leisure travelers, though this difference is extremely small relative to what has previously been found in the literature. For example, in their baseline estimates with Bertrand-Nash competition, Ciliberto and Williams (2014) find that the median price elasticity for business and leisure travelers is -0.56 and -6.26 respectively. In my airport pair model, I find that these are nearly indistinguishable: -4.15 for business travelers and -4.17 for leisure travelers. This is despite the fact that the magnitude of the median own-price elasticity across products in our studies is nearly identical: the median own-price elasticity in the base case of Ciliberto and Williams (2014) is -4.32, and the median own-price elasticity in my airport pair model is -4.16.[36] Previous studies treated passenger types as a dimension of unobserved heterogeneity. In contrast, the SFO Customer Survey allows me to directly distinguish business from leisure passengers. Intuition would suggest that I should therefore see similar or even larger differences between these passengers, but this does not appear to be the case. One possible explanation is that due to the high cost of living in the area, business and leisure travelers may be more similar

[36]Though Ciliberto and Williams (2014) use an earlier sample period (2006Q1–2008Q4) and a larger sample of products (flights between all large U.S. cities rather than only those departing from the Bay Area), our product definitions and empirical specifications are nearly identical, meaning our elasticites are directly comparable.

in the Bay Area than they are in the rest of the country. I am continuing to explore what is driving these results.

Travel costs are calculated by dividing the marginal disutility of ground travel (τ) by the price sensitivity for a given type of traveler ($\alpha + \alpha^r$). Due to the relatively small difference in price sensitivity between business and leisure travelers, I find that ground travel costs are similar as well; the travel costs for business and leisure travelers are respectively \$1.53 and \$1.45 per mile. This implies that a leisure passenger is willing to travel up to 69 additional miles on the ground in order to save \$100 on airfare. This willingness to travel is reflected in the cross-price elasticities for flights at different airports, which I will discuss below. In addition, I find large differences in own-price elasticities between the home address and airport pair models; the median price elasticity for leisure travelers is -6.40 in the home address model but -4.17 in the airport pair model. These results have very different implications for substitution patterns and the degree of market power, which I discuss below.

6.2 Elasticities and Markups

Table 6 shows the cross-price elasticity matrix generated by the home address model for nonstop flights to Denver in the second quarter of 2017; while this is only one example, it is representative of the patterns in other markets.[37] The cross-price terms shaded in gray are equal to zero in the airport pair model because these flights depart from different airports. This matrix shows that the airport pair model is missing important substitution patterns; the estimated travel costs are sufficiently low to allow significant substitution between airports. Though flights departing from the same airport are closer substitutes than flights departing from a different airport, flights from some airports are nearly as close substitutes. For example, United's flight from San Francisco to Denver is nearly as close a substitute for Southwest's flight from San Jose to Denver as it is for Southwest's flight from San Francisco to Denver. On the other hand, flights from Sacramento tend to be less attractive substitutes for flights departing from San Francisco, Oakland, or San Jose because Sacramento is significantly further away.

Figure 4 shows the distribution of own-price elasticities generated by each model. There is a clear distinction between the home address model and the airport pair model; the median own-price elasticity is -4.16 in the airport pair model but -7.23 in the home address model. Thus own-price elasticities are about 74 percent larger in the home address model than in the airport pair model. Under the assumption of Bertrand-Nash pricing behavior, marginal costs

[37] The relatively small magnitudes of the cross-price elasticities are a result of the large outside option market share; though σ is non-zero, it is does not induce enough correlation in preferences to overcome the strong pull of the outside option.

Table 6: Cross-price elasticity matrix for nonstop flights to Denver in 2017Q2

	OAK-WN	SFO-AS	SFO-F9	SFO-UA	SFO-VX	SFO-WN	SJC-UA	SJC-WN	SMF-UA	SMF-WN
OAK-WN	−5.01	0.13	0.13	0.13	0.13	0.13	0.11	0.11	0.06	0.06
SFO-AS	0.01	−3.10	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00
SFO-F9	0.02	0.02	−2.52	0.02	0.02	0.02	0.02	0.02	0.01	0.01
SFO-UA	0.22	0.22	0.22	−4.94	0.22	0.22	0.19	0.19	0.08	0.08
SFO-VX	0.06	0.06	0.06	0.06	−3.87	0.06	0.05	0.05	0.02	0.02
SFO-WN	0.06	0.06	0.06	0.06	0.06	−4.17	0.05	0.05	0.02	0.02
SJC-UA	0.03	0.04	0.04	0.04	0.04	0.04	−6.00	0.06	0.01	0.01
SJC-WN	0.07	0.08	0.08	0.08	0.08	0.08	0.13	−5.28	0.03	0.03
SMF-UA	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	−5.90	0.11
SMF-WN	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.39	−5.79

Note: Each element shows the percent change in the market share of the row product resulting from a 1% increase in the price of the column product. Elements shaded in gray are equal to zero by definition in the airport pair model.
Airports: Oakland (OAK), San Francisco (SFO), San Jose (SJC), Sacramento (SMF)
Airlines: Southwest (WN), Alaska (AS), Frontier (F9), United (UA), Virgin (VX)

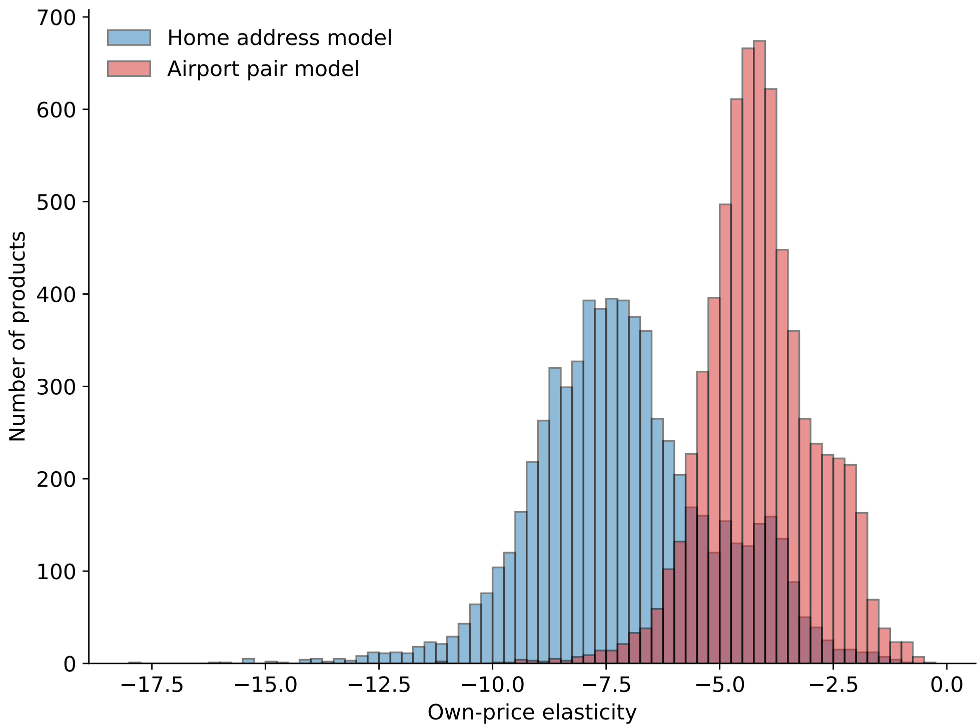

Figure 4: Distribution of own-price elasticities by model

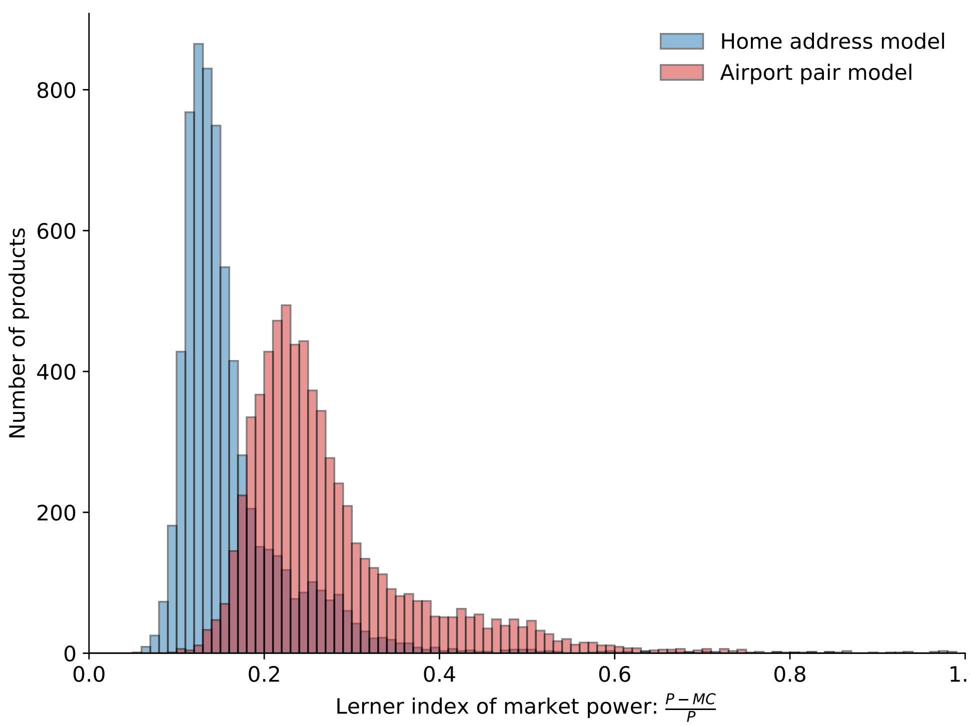

Figure 5: Distribution of Lerner indices by model

Table 7: Comparison of elasticities, marginal costs, and market power by model

	Home address model	Airport pair model
Median own-price elasticity	-7.23	-4.16
Median marginal cost (dollars)	225.55	198.08
Median markup (dollars)	36.86	62.86
Median Lerner index $\frac{P-MC}{P}$	0.14	0.25

Table 8: Operating profit margins for large airlines in 2017Q2

	Accounting data*	Home address model	Airport pair model
Alaska	0.28	0.26	0.45
American	0.18	0.15	0.26
Delta	0.19	0.15	0.25
JetBlue	0.17	0.21	0.39
Southwest	0.22	0.24	0.42
United	0.17	0.13	0.23
Virgin	0.17	0.18	0.30

*Calculated from Form 41 Financial Statements, Schedule P-1.2, domestic region

and markups can be backed out from demand estimates, prices, and market shares.[38] Figure 5 shows the distribution of Lerner indices (i.e., percentage markups) implied by each model; the Lerner index measures the deviation from the perfectly competitive outcome where markups are zero. Again, the differences between models are clear; the median Lerner index is 0.14 in the home address model but 0.25 in the airport pair model. Table 7 summarizes these results. In dollar terms, the median markup is \$62.86 in the airport pair model but \$36.86 in the home address model, or about 41 percent lower.

Finally, to examine the fit of each model, Table 8 compares the operating profit margins (i.e., variable profits divided by revenues) implied by each model to those reported by the airlines in accounting statements submitted to the Department of Transportation. It is important to note that the accounting data covers all routes in the United States, while my sample only contains products originating in the San Francisco area; such routes may not be representative of the other routes flown by the major airlines. However, the home address model seems to match these profit margins more closely than the airport pair model, which tends to overstate profitability.

[38]Figure 6 in Appendix A shows the distribution of marginal costs implied by each model. Though the distributions look similar, for any given product there is a \$25–\$30 difference in marginal costs between the home address model and the airport pair model.

7 Merger Simulations

To further explore the effects of spatial competition in the airline industry, I use both models to simulate the price effects of two recent mergers: the 2015 merger between American Airlines and US Airways and the 2018 merger between Alaska Airlines and Virgin America. Readers should note that these results are relatively preliminary compared to those in the rest of the paper.[39]

Both of these mergers were challenged by the Department of Justice, though for different reasons. In its complaint against American and US Airways, the Department of Justice noted that the merger would create the world's largest airline and substantially lessen competition on over 1,000 routes, including many in the Bay Area. Though Alaska Airlines and Virgin America were relatively small airlines, their merger is particularly interesting for my study. Both Alaska and Virgin primarily operated on routes up and down the West Coast, and SFO was one of only two airports that were hubs for both airlines.

To simulate the price effects of each merger, I use marginal costs and demand unobservables from the period prior to the merger in order to solve for the new Bertrand-Nash equilibrium prices in each market and in each model.[40] For products that were offered by both firms before the merger (e.g., nonstop service from SFO to JFK), I assume that the merged airline offers a product with the highest unobserved quality and lowest marginal cost of the premerger products. While this is essentially the best-case scenario for the profitability of the merged firm, it is a standard assumption in the literature (Ciliberto, Murry and Tamer, 2016; Li et al., 2018).

Tables 9 and 10 show the predicted postmerger price changes for nonstop routes in which the merging firms competed directly. For the American-US merger, the airport pair model predicts that prices would increase by 2.47 percent on average for these routes, while the home address model predicts an increase of 1.08 percent on average. Similarly, for the Alaska-Virgin merger, the airport pair model predicts a price increase of 3.46 percent on average while the home address model predicts a 1.16 percent average increase. Again, failing to account for spatial competition can lead to misleading conclusions about the exercise of market power.

As stated above, these results are preliminary. In the near future, I plan to use a differences-in-differences technique to compare the predicted price changes implied by each

[39]For example, the magnitudes of the predicted postmerger price changes for the American-US merger are very small compared to those found by Li et al. (2018). I suspect this is due in part to the larger market share of the outside option in my study. In addition, it may be helpful to estimate marginal costs as a function of observed cost shifters in order to eliminate the effects of any transitory shocks.

[40]I use 2015Q2 and 2017Q2 as the premerger periods for the American-US and Alaska-Virgin simulations respectively.

Table 9: American Airlines–US Airways merger: predicted change in prices by model (nonstop routes in which the merging parties competed directly)

		Percentage change in price postmerger	
Origin	Destination	Home address model	Airport pair model
SFO	ATL	+0.12%	+0.04%
SFO	DFW	+1.18%	+4.47%
SFO	JFK	+6.12%	+4.29%
SFO	LAX	+0.09%	+0.62%
SFO	MIA	+0.77%	+3.20%
SFO	ORD	+0.02%	+0.01%
SFO	TPA	+0.38%	+0.64%
SJC	DFW	+0.88%	+3.45%
SJC	ORD	+0.33%	+3.13%
SMF	DFW	+0.89%	+4.81%
Mean		+1.08%	+2.47%
Median		+0.58%	+3.16%

Table 10: Alaska Airlines–Virgin America merger: predicted change in prices by model (nonstop routes in which the merging parties competed directly)

		Percentage change in price postmerger	
Origin	Destination	Home address model	Airport pair model
SFO	AUS	+1.29%	+3.31%
SFO	BOS	+1.50%	+2.53%
SFO	DAL	+1.82%	+9.17%
SFO	DCA	+1.06%	+3.10%
SFO	DEN	+0.92%	+2.71%
SFO	EWR	+1.85%	+2.36%
SFO	FLL	+1.20%	+3.47%
SFO	IAD	+1.21%	+2.09%
SFO	JFK	+1.97%	+3.61%
SFO	LAS	+1.70%	+6.36%
SFO	LAX	+1.22%	+6.79%
SFO	MCO	+0.34%	+0.57%
SFO	ORD	+1.14%	+2.36%
SFO	PDX	+0.11%	+1.37%
SFO	SAN	+1.01%	+4.67%
SFO	SEA	+0.21%	+0.86%
Mean		+1.16%	+3.46%
Median		+1.21%	+2.90%

model to the actual price changes observed in the data.[41] Admittedly, the magnitudes of the predicted price increases seem low. On the other hand, recent evidence suggests that these mergers may not have been anti-competitive. Using a differences-in-differences approach, Carlton et al. (2018) find that prices on nonstop routes directly affected by the American–US merger actually *declined* by 12.3 percent on average.

8 Conclusion

In this paper, I relax the assumption that airline markets consist of origin-destination city pairs by estimating a structural model that allows consumers to choose among flights departing from airports in different cities. This model accounts for the cost of ground travel between a passenger's home and the departure airport for her flight, which generates a trade-off between ground travel and lower fares. I compare the results from my model to those generated by a conventional airport pair model of airline markets.

I find that leisure passengers are willing to travel up to 69 miles to save $100 on airfare. These costs are sufficiently low to allow significant substitution between flights departing from different airports. As a result, demand is 74 percent more elastic and markups are 41 percent lower when spatial competition is accounted for. These results show that airlines face substantial competition from flights departing from nearby airports and that the narrow airport pair definition of airline markets overstates the degree of market power.

These results have important implications for antitrust policy. The airport pair market definition appears to fail the hypothetical monopolist test used by the antitrust agencies to define the relevant geographic market; price increases at one airport are defeated by consumer substitution towards other airports. Though the city pair definition may do slightly better, it is clear that spatial competition is a more important factor in the airline industry than previously realized.

The results reported here are somewhat preliminary, as there are shortcomings in the analysis that remain to be addressed. For example, travel costs are calculated using great circle distances between points. This is likely a particularly poor approximation in my study area due to the presence of San Francisco Bay. In reality, residents must either go around the Bay or use one five bridges in the area, which can be congested. Furthermore, ground travel costs can vary significantly depending on the mode of travel and time of day. In the very near future, I plan to incorporate average travel times rather than distances to measure travel costs. Second, it is unclear whether travel costs in the San Francisco area are similar to those in the rest of the country. Though I believe these results are applicable to other

[41] Postmerger data for the Alaska-Virgin merger was just released in early October 2018.

large cities such as Los Angeles and New York, travel through rural areas of the country is much faster. As a result, travel costs per mile should be lower, and residents in rural areas may therefore travel much further in search of lower fares. In turn, this means that even more distant airports may be competitors in relatively rural areas of the country.

References

Armantier, Olivier and Oliver Richard. 2008. “Domestic Airline Alliances and Consumer Welfare.” *RAND Journal of Economics* 39(3):875–904.

Berry, Steven, James Levinsohn and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica* 63(4):841–890.

Berry, Steven, James Levinsohn and Ariel Pakes. 2004. “Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market.” *Journal of Political Economy* 112(1):68–105.

Berry, Steven, Michaeel Carnall and Pablo T. Spiller. 2006. Airline Hubs: Costs, Markups and the Implications of Customer Heterogenity. Vol. 1 of *Advances in Airline Economics: Competition Policy and Antitrust* Amsterdam: Elsevier pp. 183–214.

Berry, Steven and Panle Jia. 2010. “Tracing the Woes: An Empirical Analysis of the Airline Industry.” *American Economic Journal: Microeconomics* 2(3):1–43.

Berry, Steven T. 1994. “Estimating Discrete-Choice Models of Product Differentiation.” *RAND Journal of Economics* 25(2):242–262.

Berry, Steven T. and Philip A. Haile. 2014. “Identification in Differentiated Products Markets Using Market Level Data.” *Econometrica* 82(5):1749–1797.

Borenstein, Severin. 2010. “An Index of Inter-City Business Travel for Use in Domestic Airline Competition Analysis.” NBER Working Paper.

Brueckner, Jan K., Darin Lee and Ethan Singer. 2014. “City-Pairs Versus Airport-Pairs: A Market Definition Methodology for the Airline Industry.” *Review of Industrial Organization* 44(1):1–25.

Carlton, Dennis, Mark Israel, Ian MacSwain and Eugene Orlov. 2018. “Are Legacy Airline Mergers Pro- or Anti-Competitive? Evidence from Recent U.S. Airline Mergers.” Forthcoming*, International Journal of Industrial Organization* .

Ciliberto, Federico, Charles Murry and Elie Tamer. 2016. “Market Structure and Competition in Airline Markets.” Working Paper.

Ciliberto, Federico, Eddie Watkins and Jonathan W. Williams. 2018. “Collusive Pricing Patterns in the U.S. Airline Industry.” Forthcoming*, International Journal of Industrial Organization* .

Ciliberto, Federico and Jonathan W. Williams. 2014. "Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry." *RAND Journal of Economics* 45(4):764–791.

Davis, Peter. 2006. "Competition in Retail Markets: Movie Theaters." *RAND Journal of Economics* 37(4):964–982.

Evans, William N. and Ioannis N. Kessides. 1994. "Living by the "Golden Rule": Multimarket Contact in the U.S. Airline Industry." *Quarterly Journal of Economics* 109(2):341–366.

Gandhi, Amit and Jean-François Houde. 2017. "Measuring Substitution Patterns in Differentiated Products Industries." Working Paper.

Grigolon, Laura and Frank Verboven. 2014. "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation." *Review of Economics and Statistics* 96(5):916–935.

Ho, Katherine and Joy Ishii. 2011. "Location and Competition in Retail Banking." *International Journal of Industrial Organization* 29(5):537–546.

Hotelling, Harold. 1929. "Stability in Competition." *Economic Journal* 39(153):41–57.

Houde, Jean-François. 2012. "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline." *American Economic Review* 102(5):2147–2182.

Ishii, Jun, Sunyoung Jun and Kurt Van Dender. 2009. "Air Travel Choices in Multi-Airport Markets." *Journal of Urban Economics* 65(2):216–227.

Lazarev, John. 2013. "The Welfare Effects of Intertemporal Price Discrimination: An Empirical Analysis of Airline Pricing in U.S. Monopoly Markets." Working Paper.

Li, Sophia Ying, Joe Mazur, Yongjoon Park, James W. Roberts, Andrew Sweeting and Jun Zhang. 2018. "Endogenous and Selective Service Choices after Airline Mergers." NBER Working Paper.

Manuszak, Mark D. 2001. "The Impact of Upstream Mergers on Retail Gasoline Markets." Working Paper.

Manuszak, Mark D. and Charles C. Moul. 2009. "How Far for a Buck? Tax Differences and the Location of Retail Gasoline Activity in Southeast Chicagoland." *Review of Economics and Statistics* 91(4):744–765.

McManus, Brian. 2007. “Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee.” *RAND Journal of Economics* 38(2):512–532.

Peters, Craig. 2006. “Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry.” *Journal of Law and Economics* 49(2):627–649.

Petrin, Amil. 2002. “Quantifying the Benefits of New Products: The Case of the Minivan.” *Journal of Political Economy* 110(4):705–729.

Seim, Katja and Joel Waldfogel. 2013. “Public Monopoly and Economic Efficiency: Evidence from the Pennsylvania Liquor Control Board’s Entry Decisions.” *American Economic Review* 103(2):831–862.

Smith, Howard. 2004. “Supermarket Choice and Supermarket Competition in Market Equilibrium.” *Review of Economic Studies* 71(1):235–263.

Sweeting, Andrew, James W. Roberts and Chris Gedge. 2018. “A Model of Dynamic Limit Pricing with an Application to the Airline Industry.” Working Paper.

Thomadsen, Raphael. 2005. “The Effect of Ownership Structure on Prices in Geographically Differentiated Industries.” *RAND Journal of Economics* 36(4):908–929.

United States Department of Justice and Federal Trade Commission. 2010. “Horizontal Merger Guidelines.”.

United States et al. v. US Airways Group, Inc. and AMR Corporation. 2013. Civil complaint filed in United States District Court for the District of Columbia.

Williams, Kevin R. 2018. “Dynamic Airline Pricing and Seat Availability.” Working Paper.

A Additional Tables and Figures

Table 11: Business travel logit regression

Outcome: probability of business travel	Estimate	(Std. Err.)
ZIP code population density (1,000 people/sq. mile)	0.011***	(0.003)
HH income above \$100,000	2.001***	(0.147)
HH income \$50,000–\$100,000	1.069***	(0.157)
Constant	−2.404***	(0.145)
Number of surveyed passengers	3,802	
Pseudo-R^2	0.070	

Standard errors are robust to heteroskedasticity. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

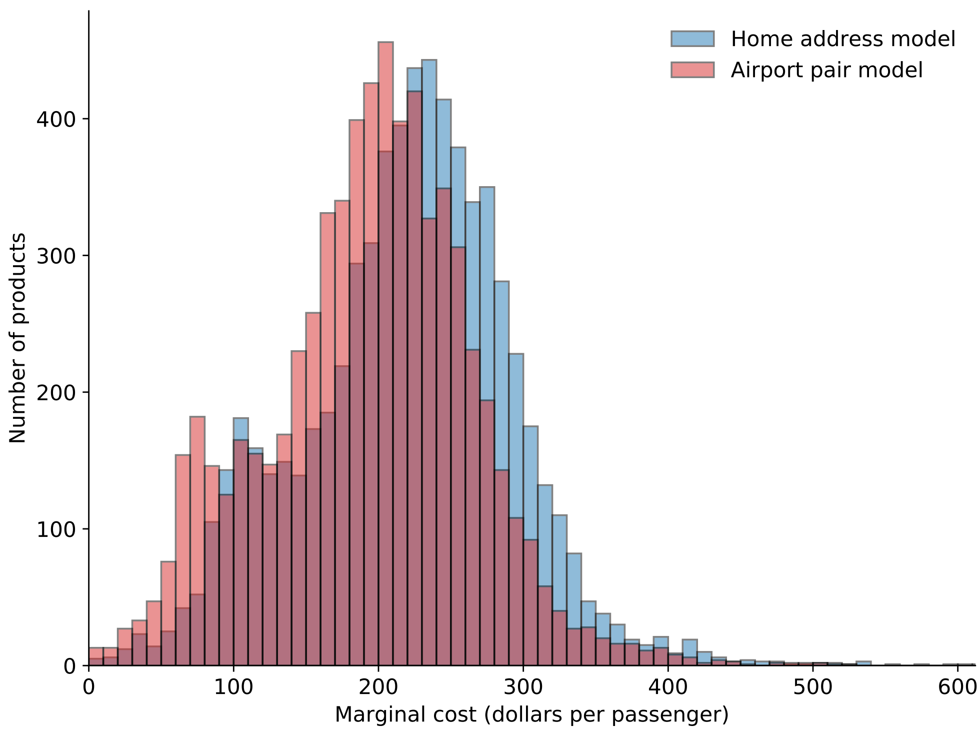

Figure 6: Distribution of marginal costs by model

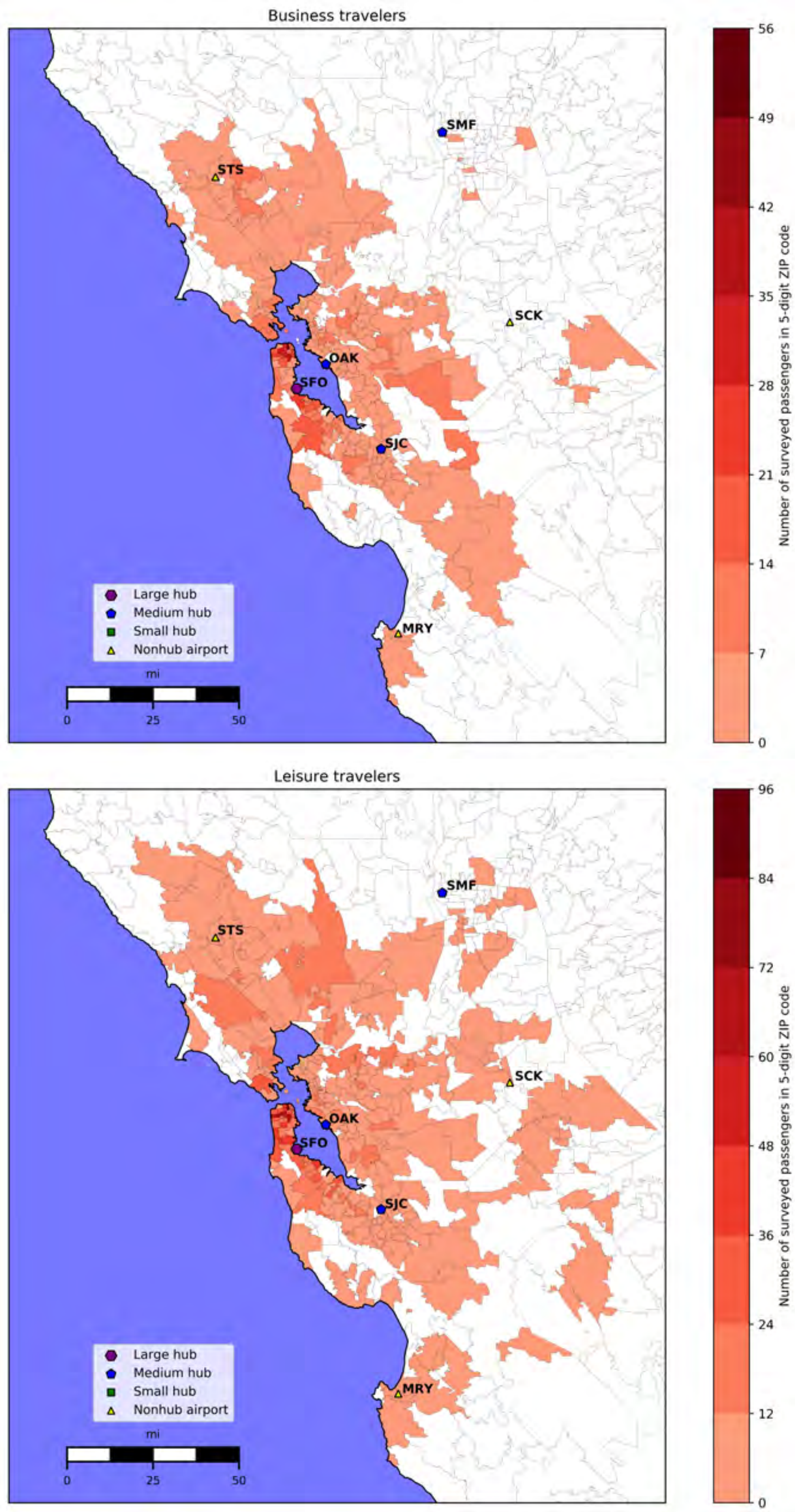

Figure 7: Spatial distribution of home locations for business travelers (top) and leisure travelers (bottom) departing from San Francisco International Airport (SFO), 2010–2017