AIRLINE-AIRPORT AGREEMENTS IN THE SAN FRANCISCO BAY AREA: EFFECTS ON AIRLINE BEHAVIOR AND CONGESTION AT AIRPORTS

MIGUEL-ANGEL ALCOBENDAS*

Toulouse School of Economics
University of California, Irvine Department of Economics

February 25, 2014

Abstract. This paper provides a methodological framework to analyze the decisions of airlines and travelers taking into account the contractual agreement between airports and airlines. This contract sets the fees that carriers pay for landing, the rental rate for the terminal space that they occupy, as well as the methodology to determine these charges. Using data from San Francisco International Airport (SFO) and Metropolitan Oakland International Airport (OAK), we quantify the effects of changes in the agreement on the behavior of airlines and congestion at airports. In particular, we look at modifications in the design of charges and variations in the operating costs at airports. Counterfactuals suggest that different methodologies to compute charges and changes in airport costs may induce airlines to behave differently, affecting delays at airports.

Our structural model captures important characteristics of the airline industry: endogeneity of airport charges with respect to decisions of travelers and carriers, correlation across markets, and two decision variables of airlines (fares and frequency of flights).

1. Introduction

Interactions between airlines, travelers and airports in the U.S. have been the object of several studies since airline deregulation at the end of the 70's. The rise of air traffic and limited capacity of airports have led researchers to study the efficiency of carrier operations at airports. However, most of the empirical work does not take into account the relationship between airports and airlines. This contractual relationship sets the fees that airlines pay for landing and the rental rate for the terminal space that they occupy. In this paper,

*This paper is part of my PhD dissertation at Toulouse School of Economics. Email:maalcobendas@gmail.com.

I wish to thank my advisor Marc Ivaldi and Jan Brueckner for guidance and support. I am grateful to Miguel Urdanoz and the Metropolitan Transportation Commission of the San Francisco Bay for sharing their data. Thanks to Laura Lasio, Senay Sukullu, and the two anonymous referees for their insightful comments.

we analyze how these charges are determined and how they affect the strategic behavior of carriers and the level of congestion at airports. Our empirical application is based on the competition between the two main airports located in the San Francisco Bay Area: San Francisco International Airport (SFO) and Metropolitan Oakland International Airport (OAK).

In the U.S., landing fees and rental of terminals are designed to let airports achieve financial self-sufficiency. The methodology to determine these charges is airport specific and follows guidelines proposed by the Department of Transportation (DoT). Charges are the result of well defined pricing schemes that depend on measurable variables (for instance, parking revenues, maintenance costs, retail-shop revenues, weight of aircraft, and number of landings). If one of the components of the schemes changes, the airport operator may be obliged to modify the charges even if such a modification is unpopular among airlines and the press. For instance, Los Angeles International Airport (LAX) recently increased its landing fees for the 2014 fiscal year from \$4.46 to \$4.60 per 1,000 pounds of the maximum gross landing weight (MGLW) of passenger aircraft. This rise was motivated by an increase in the cost of operating the airport. Charges can also change due to shocks in the demand for airport services. For example, retail-shop revenues clearly depend on the number of travelers, and the number of flights reaching airports is another variable affecting airport fees. Using data from OAK and SFO, we characterize the equilibrium behavior of travelers and airlines, and quantify the response of carriers and congestion at airports when airport costs affecting the pricing schemes change. Since OAK and SFO apply different methodologies, the behavior of carriers is also expected to be different at these airports.

Charges may be designed to be low in order to attract carrier operations, but at the same time they can also cause congestion. This seems to be the case at SFO, since it uses revenues generated at the parking lots to reduce the amount that airlines pay for the use of its infrastructure. OAK, on the other hand, does not take these revenues into account when it computes landing fees and rental charges. While the SFO methodology is appropriate in periods of airport overcapacity, this is not the case when the airport operates at the maximum of its possibilities, as SFO will do in the near future. Without investments in new infrastructure (e.g. new terminals, runways, or air-traffic control technology upgrades), SFO needs to consider alternatives to manage congestion. One solution is to revise the methodology used to determine its charges. To explore the effects of such a change, the last part of the paper analyzes the consequences of SFO adopting the contract scheme used by OAK.

In order to characterize the interaction between travelers, carriers and airports, we use a structural model where demand and supply functions are specified. The demand is formed by heterogeneous travelers with different locations (origin or destination) in the San Francisco Bay and different tastes. The airline profit function depends on the pricing scheme (landing fees and rental rates) charged by airports, with carriers deciding on fares and the frequency of flights. At the same time, charges are endogenously determined by the behavior of travelers and airlines. We use recent advances in estimation of two stage games (Villas-Boas (2007) and Fan (2012)) to estimate the model.

There are few papers addressing empirically the role of airport charges. Van Dender (2007), Bilotkach et al. (2012a), and Bel and Fageda (2009) analyze how market factors affect the level of these charges. However, they do not explicitly model the relationship between airlines and airports. In contrast, Ivaldi et al. (2011) present a structural model introducing landing fees and travelers' charges. They treat airports as platforms and show the existence of two-sidedness effects. While they consider airports as profit maximizing monopolists, we model the determination of fees and rates using rules based on cost recovery, which is more consistent with the methodology applied by the two main airports in the San Francisco Bay. Moreover, none of the aforementioned works considers the fact that charges are endogenously determined by the behavior of travelers and carriers, and neither considers the role of charges as a tool to manage flight delays.

Some other contributions are also made from a methodological point of view. First, carriers behave as profit maximizing firms with respect to ticket prices and frequency of flights reaching the Bay Area. Most of the previous literature only focuses on prices. Second, our model captures two sources of correlation across markets: one comes from the possibility that travelers purchasing different products use the same aircraft to reach the San Francisco Bay. The model also captures the fact that planes contribute to congestion at airports, affecting other aircraft even if they operate in different markets.

For our application, we use U.S. domestic flight data from the third quarter of 2006. The Airline Origin and Destination Survey (DB1B), the T-100, and the Airline On-Time Performance data sets from the U.S. Bureau of Transportation Statistics let us include the supply side, analyze elasticities, and perform counterfactuals. In particular, these data sets give us detailed information about product characteristics and the choices of travelers. We will combine them with travelers' demographic information using the American Community Survey (ACS), financial airport information from the Federal Aviation Administration (FAA), and technical aircraft characteristics. Finally, in order to increase the precision of the estimates, we will add additional information obtained from the 2006 Airline Passenger Survey done by the Metropolitan Transportation Commission of the San Francisco Bay (MTC).

Consistent with previous literature, travelers, on average, prefer to use SFO rather than OAK. However, traveler heterogeneity is also an important factor to explain their purchasing pattern. For instance, their decision significantly depends on the distance from their location (origin or final trip destination in the Bay Area) to the airports.

If we look at the relationship between carriers and airports, we observe that changes in the cost of operating airports not only affect landing fees and rental charges, but also carrier decisions regarding the number of flights and size of aircraft, and congestion at the airports. For example, a rise in landing fees as a result of an increment to the operating costs of an airport is accompanied by a decrease in the daily frequency of flights, an increase in the average size of aircraft, and a reduction of airport congestion. These results hold for OAK and SFO, but they are much stronger at SFO. For instance, if the operating cost used to compute charges at SFO increases by 20%, the total number of daily flights reaching the airport decreases by 2.4%, the average weight of aircraft increases by 1.7%, and the average delay of flights at SFO decreases by 8.1%. Similarly, an increase in the cost of operating OAK by 20% reduces the number of daily flights by 0.7%, increases the average weight of aircraft by 1.6%, and reduces congestion at OAK by 2.1%. Our simulations also suggest that changes in the operating cost of one airport, barely change the behavior of carriers operating at the competing airport.

Finally, the design of charges may play an important role in the behavior of carriers and congestion at airports. When we analyze the effects of SFO adopting the contract used by OAK, we find a threshold that helps us to identify under which conditions the new contract is useful to reduce flight delays. As we will see in detail later on, each airport uses different cost components to determine charges. The results of SFO implementing the OAK charge scheme depend on the magnitude of the costs used in the new methodology relative to the costs currently used at SFO. For example, if the sum of the cost components in the new pricing scheme is 20% lower than the original one, the number of flights reaching SFO would decrease by almost 4%. Consequently, the level of congestion at SFO would be 12% lower.

The rest of the paper is structured as follows. Section 2 introduces general features of the San Francisco Bay and the airports operating in the area. Section 3 presents the model. Section 4 outlines the optimality conditions of carriers. Section 5 describes the application and data. Section 6 outlines the estimation methodology. Section 7 presents the estimation results. Section 8 analyzes the contractual relationship between airports and airlines. Finally, section 9 concludes.

2. The Nature of the Interaction between Airlines and Airports

To analyze the equilibrium behavior of airlines (fares and flight frequency) and traveler demand, it is necessary to understand the characteristics of each of the airports serving the San Francisco Bay Area, and the nature of the relationship between airports and carriers.

The Bay Area is a region located in Northern California that is home to 7.15 million people distributed around nine counties (Figure 1). It is served by 3 main airports: San Francisco International (SFO) located in San Mateo County, Metropolitan Oakland International (OAK) in Alameda County, and Mineta San Jose International (SJC) in Santa Clara County. OAK and SFO are located 11 miles apart, while SJC is around 30 miles from SFO and OAK. SFO is the busiest of the three airports and an important entrance to the U.S. from the Pacific.

We focus our attention on SFO and OAK. SJC will be included in the outside option of the model. Such a decision is driven by the lack of passenger data for SJC. The impact of this limitation in our analysis is expected to be low since, as several authors pointed out (Bilotkach et al. (2012b), Brueckner et al. (2013)), OAK and SFO are closer substitutes compared to SJC.

Most U.S. airports are operated as independent not-for-profit facilities overseen by a local governmental entity such as a county, city, or state government. OAK and SFO are not exceptions. The Port of Oakland owns and operates OAK and SFO is owned by the City and County of San Francisco.

Airlines operate at airports under a contract called "use and lease agreement", which details the fees and rental rates that an airline has to pay, and the method by which they are calculated. The charges that we consider are those related to landing operations (landing fees) and the rates that carriers must pay for using the terminals (rental rates). The magnitude of the two key elements of the contract are not negligible. If we look at the financial statements of the airports, in 2006 SFO reported \$74 million in landing fee revenues and \$145 million in revenues from rental of terminals. Similarly, OAK reported \$17 million and \$27 million respectively.

Both airports use a hybrid approach to determine charges.[1] Under such a methodology, operating costs and revenues are allocated to different cost centers. Three of these cost centers are used to compute landing fees and rental rates: the Terminal Cost Center, the Airfield Cost Center, and the Groundside Cost Center. The Terminal Cost Center includes all costs and revenues generated in the terminal buildings. For instance, maintenance and payments to the police in the terminals would be allocated to this cost center. Similarly, revenues generated from concessions (mainly food, beverage, and car rentals) are attributable to this cost center. The Airfield Cost Center includes, for example, the maintenance of the ramp and cost recovery of investments in capital.[2] Finally, the Groundside Cost Center is mainly related to costs and revenues from vehicle parking and ground transportation vehicle access (e.g. taxi cabs, charter buses, or limousines). The way that OAK and SFO compute rental rates and landing fees is different, and it depends on the weight that each of the aforementioned cost centers has in the charging rules.[3]

In the case of OAK, the rental charge is fully determined by the costs and revenues attributable to the Terminal Cost Center. In particular, this cost center must break even for the fiscal year. That is, the total amount that airlines reimburse the Port of Oakland for using its terminal is computed as the difference between the operating costs minus the revenues assigned to the Terminal Cost Center in the fiscal year. Then, this amount is distributed to each of the individual airlines depending on how much area each airline leases from the airport terminal (rental charge). Logically, the revenues and costs are not known during the fiscal year. That is why quantities are forecasted, with the amounts regularized the following year. If airlines pay in excess, the airport credits the corresponding amount, and otherwise, airlines pay the shortfall. Landing fees are computed in a similar way. In this case, it is the Airfield Cost Center which must break even. The difference between the costs

[1]Previous literature distinguishes three broad class of contracts: residual, compensatory, and hybrid. Under the residual contract, airlines pay the net cost of running the airport after taking into account aeronautical and non-aeronautical revenues. As a result, airlines are charged so that the airport breaks even (revenues=costs). By contrast, with the compensatory approach, airlines pay agreed charges based on recovery of costs allocated to the facilities and services they use. Finally, the hybrid method combines elements of the previous two types of contracts. Under such an approach, revenues and costs are assigned to different cost centers, and some of these centers are defined as residuals (break even) and others as compensatory (cost recovery) (Daniel (2001)).

[2]For instance, a percentage of the costs of constructing a new taxiway or ramp are yearly allocated to the airfield cost center until the total cost is recovered

[3]Further details about the method applied by each airport to compute fees and rates can be found in the Oakland International Master Plan (2006) and in the 2006 Annual Operating Budget document for San Francisco International Airport SFO.

and revenues assigned to this cost center in a fiscal year is divided by the total estimated landing weight of aircraft using the airport. This total weight is equal to the maximum gross landing weight (MGLW) of aircraft used by airlines times the number of landings at OAK each performs in the accounting period. This ratio yields landing fees. In our application, the landing fee rate in OAK is equal to \$1.460 per 1,000 pounds of aircraft MGLW.

In the case of SFO, the total amount that airlines face for using its terminals (rental charge) equals the amount needed to cover 3/2 of the net operating costs of the Terminal Cost Center plus 50% of the net operating surplus of the Groundside Cost Center. Note that the Groundside Cost Center includes the highly profitable car parking activity. Hence, including this term in the charge rules generally reduces the amount that airlines pay. Once again, these quantities are estimated for the current fiscal year and regularized afterwards. Once the amount is computed, it is allocated to airlines according to the surface they lease. Finally, required total landing fees equal the amount needed to cover the net operating costs of the Airfield Cost Center plus 50% of the sum of the Terminal Cost Center net costs and the net operating surplus of the Groundside Cost Center. The ratio of this amount and the forecasted total MGLW gives the landing fee rate. In 2006, SFO had a ratio equal to \$3.213 per 1,000 pounds of aircraft MGLW.

3. Model

Our model captures the strategic behavior of travelers and carriers. We use a discrete choice framework to model demand. On the supply side, carriers not only decide on fares, but also on the frequency of their flights reaching the Bay Area. Such decisions are affected by the landing fees and rental charges imposed by airports.

3.1. **Demand:** Demand for products offered by airlines is derived from the aggregation of individual choices of heterogeneous travelers. Preferences over products are represented as a function of individual characteristics and the attributes of products. Such an approach lets us incorporate individual tastes of travelers for product characteristics and add heterogeneity with respect to household income and distance between travelers' location and the airports in the Bay Area.

We define a market as a round trip directional city-pair. For instance, a market could be the directional pair San Francisco - Atlanta, where San Francisco is the origin and Atlanta is the destination. This market is different from Atlanta - San Francisco, where Atlanta is the origin and San Francisco is the destination. Within a market, travelers can choose among a set of differentiated products. We distinguish products according to the combination of their characteristics such as fare, frequency of flights, ticketing carrier, airport of origin and destination, and itinerary of the trip. According to our market definition, any product departing from either OAK or SFO with the same destination belongs to the same market. Similarly, products arriving at either OAK or SFO using the same origin airport also belong to the same market.

Suppose that we observe $t = 1, ..., T$ markets with $i = 1, ..., I_t$ consumers, and $j = 1, ..., J_t$ products. The utility that a potential traveler i obtains from purchasing a ticket j in market t is given by

$$
\begin{aligned}
(1) \quad u_{ijt} \quad &= \quad \alpha_{sfo}\hat{I}_{jt}^{sfo} + \underbrace{(\alpha_p + \alpha_y y_i + \sigma^p \nu_i^p)}_{\alpha_{ip}} p_{jt} + \underbrace{(\alpha_f + \sigma^f \nu_i^f)}_{\alpha_{if}} \hat{f}_{jt} + \underbrace{(\alpha_d + \sigma^d \nu_i^d)}_{\alpha_{id}} \hat{D}_{jt} + \xi_{jt} + \\
&+ \quad \lambda d(L_i) + x_{jt}\beta + \sigma^0 \nu_i^0 + \epsilon_{ijt}
\end{aligned}
$$

where

- $\hat{I}_{jt}^{sfo}$ is a dummy variable equal to one if the product is offered at SFO and zero if the product uses OAK.
- p_{jt} is the ticket price.
- y_i is household income with probability distribution P_Y.
- $\hat{f}_{jt}$ corresponds to the daily frequency of flights. We construct this variable as the mean of the frequencies for each of the flight segments of the product jt.
- $\hat{D}_{jt}$ is the average delay. This variable is equal to the mean of arrival delays for the connecting (if one exists) and destination airports used by product jt.
- x_{jt} is a vector of travel characteristics for product jt. Such characteristics are observed by the econometrician: ticketing carrier, flight distance, dummy for direct flight, and a dummy for destination airports with slot constraints.
- ξ_{jt} captures the unobserved-to-researcher characteristics of product j in market t. An increase in ξ_{jt} makes the product j in market t more attractive to all consumers.
- $d_{jt}(L_i)$ determines the distance of individual i to the airport (OAK or SFO) used in product j in market t. L_i denotes the location of individual i in the Bay Area, with probability distribution P_L. This variable is interpreted according to the nature of the traveler. If the individual i is originally departing from one of the airports of the Bay Area, $d_{jt}(L_i)$ may be considered as the distance from his residence or work place to the airport. On the other hand, if the traveler is arriving in the Bay Area, $d_{jt}(L_i)$ is interpreted as the distance from the airport to his final destination (hotel or office).

- ν_i^p, ν_i^f, ν_i^d and ν_i^0 account for the unobserved taste of travelers for fares, frequency, delays and a constant respectively. As we previously mentioned, we allow interaction between both price and other product characteristics and individual tastes to obtain richer patterns of substitution. We assume that each of these random variables is drawn from a normal distribution except the ones that interact with prices (ν_i^p). In this case the distribution is assumed to be lognormal.
- ϵ_{ijt} is a mean-zero error term, assumed to be i.i.d. across travelers and products and to follow a type-I extreme value distribution.

The vector of demand parameters to be estimated is denoted by θ and includes: the taste for product price (α_p, α_y, σ^p), for using SFO (α_{sfo}), for daily frequency (α_f, σ^f), for delays (α_d, σ^d), for other product characteristics (β), for distance to airports (λ), and the parameter σ^0 associated with the constant. We assume that the marginal utility of income (α_y) is the same for all households independently of their income level and used airport. Similarly, we assume that the distance sensitivity to airports (λ) is the same for all travelers independently of their location and airport. Finally, α_{ip}, α_{if} and α_{id} are the individual-specific coefficients linked to fares, frequencies, and delays respectively.

We also use county dummy variables to capture county-specific tastes for products served by the two airports. These variables equal one if the traveler comes from (goes to) the specified county and zero otherwise.

Ticket prices (p_{jt}) and flight frequencies ($\hat{f}_{jt}$) are expected to be correlated with ξ_{jt}. Hence the use of appropriate instrumental variables will be necessary to avoid inconsistent estimates.

Following Berry, Levinsohn, and Pakes (1995), we distinguish the mean utility level of product j in market t (δ_{jt}) from the traveler-specific deviation ($\mu_{ijt} + \epsilon_{ijt}$):

$$\delta_{jt} = \alpha_{sfo}\hat{I}_{jt}^{sfo} + \alpha_p p_{jt} + \alpha_f \hat{f}_{jt} + \alpha_d \hat{D}_{jt} + x_{jt}\beta + \xi_{jt} \tag{2}$$

$$\mu_{ijt} = (\alpha_y y_i + \alpha^p \nu_i^p)p_{jt} + \sigma^f \nu_i^f f_{jt} + \sigma^d \nu_i^d D_{jt} + \lambda d_{jt}(L_i) + \sigma^0 \nu_i^0 \tag{3}$$

Hence, the utility function can be rewritten as

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt} \tag{4}$$

Let u_{i0t} denote the utility from the outside good of not flying from the considered airports. The utility is random and is written as

$$u_{i0t} = \epsilon_{i0t} \tag{5}$$

If we integrate over ϵ_{ijt}, the probability that a traveler i chooses product j in market t is

$$\begin{aligned} P(u_{ijt} \geq u_{ilt} \;\; l \neq j / \hat{I}^{sfo}, p, \hat{f}, \hat{D}, x, d, \delta, \nu_i, L_i, y_i, \theta) &= s_{ijt}(p, f, \delta(\theta); \theta) = \\ &= \frac{exp[\delta_{jt} + \mu_{ijt}]}{1 + \sum\limits_{m \in J_t} exp[\delta_{mt} + \mu_{imt}]}, \end{aligned} \tag{6}$$

where $\hat{I}^{sfo}$, p, $\hat{f}$, $\hat{D}$, x, d, and δ are vectors consisting of the corresponding variables.

Aggregate demand $s_{jt}(\cdot)$ follows from integration over i and equals

$$\begin{aligned} s_{jt}(p, f, \delta(\theta); \theta) &= \int \frac{exp[\delta_{jt} + \mu_{ijt}]}{1 + \sum\limits_{m \in J_t} exp[\delta_{mt} + \mu_{imt}]} dP_\nu(\nu_i) dP_L(L_i) dP_Y(y_i) = \\ &= \int s_{ijt}(p, f, \delta(\theta); \theta) dP_\nu(\nu_i) dP_L(L_i) dP_Y(y_i) \end{aligned} \tag{7}$$

For simplicity we assume that the distributions of ν_i, ϵ, L_i, and y_i are independent. $P_\nu(\cdot)$ is the distribution of the unobservables, $P_L(\cdot)$ is the distribution of the location of travelers in the Bay Area, and $P_Y(\cdot)$ is the distribution of household income.[4]

3.2. **Carriers and Airports:** Airlines are assumed to be profit maximizing firms with respect to ticket prices and frequencies. They may operate at one or several airports in the Bay Area, and offer differentiated products within a market. At the same time, their profits are a function of the landing fees and rental charges levied by the airports.

As we previously stated, the methodology used by each airport to compute these charges is different and depends on the behavior of travelers and airlines. How the design of these charges affects the strategy followed by carriers and travelers, and its effects on congestion at airports, are the main contributions of this paper. In this section, we first develop the profit function of airlines taking into account the fees levied by OAK and SFO, and then show how these charges are currently determined.

[4]Given data limitations, we assume that the distributions of airport distance and household income are independent. This is clearly not true since some correlation is expected between them.

3.2.1. *Profit function of airlines:* Airlines maximize their profits with respect to fares and frequency of flights reaching OAK or SFO. The equilibrium concept in the model is the subgame perfect Nash Equilibrium. The game has two stages: in the first stage, airlines simultaneously decide the flight frequency (f) of the last trip segment arriving in the Bay Area. In the second stage, firms decide fares (p).

In our model, airlines only decide on the flight frequency of the trip segment arriving in the Bay Area. This spoke route is directly affected by OAK or SFO airport charges. However, products may be composed of several segments, and we hence implicitly assume that the frequency decision for trip segments are independent of each other. As we will see later, we use the optimality conditions for fares and flight frequencies to estimate the parameters of the model and analyze the effects of changing the terms of the agreements between carriers and airports.

Let $\mathcal{J}_{ct}$ denote the set of products offered by carrier c in market t, and let Ω_c denote the set of spoke routes used by carrier c that have one of the Bay Area airports as an endpoint. Individual spokes are denoted by r. Carrier c decides on fares and frequencies according to the the following optimization problem:

$$
\begin{aligned}
(8)\quad \max_f \max_p \Pi_c \;=\; & \max_f \max_p \Bigg[\sum_{t\in T} \sum_{j\in\mathcal{J}_{ct}} ([p_{jt} - m_{jt}] s_{jt}(p, f, \delta(\theta); \theta) \times M_t) - \\
& - \underbrace{\sum_{r\in\Omega_c} \tilde{f}_{rc}(FCost_{rc} + \beta^d D_r(f) + fees_r(s,p,f) \times weight_{rc}(s,p,f))}_{\text{Total Operating Flight Cost}} - \\
& - RC_{c,sfo}(s,p,f) - RC_{c,oak}(s,p,f) - F_c \Bigg]
\end{aligned}
$$

where Π_c corresponds to profits of airline c. In the particular case of our application, profits obtained from operating U.S. domestic flights during the third quarter of 2006. The first line of the profit function corresponds to the usual form of oligopolistic models and captures revenues from products offered by airline c. m_{jt} represents the product-specific costs for product j in market t. M_t is the total population that may be interested in traveling in market t. In our application, M_t is computed as the geometric mean of the population of the origin and destination cities.

One of the novelties in the specification of this type of profit function is the inclusion of the term “Total Operating Flight Cost”, which captures the total airline cost for operating flights landing at each of the two airports. This term depends on the number of flights that the airline operates on each spoke r ($\tilde{f}_{rc}$), landing fees applied by each airport ($fees_r$), the weight of the aircraft ($weight_{rc}$), the level of congestion measured as the average arrival delay of the airport used by the spoke r (D_r), the monetary value of one minute of delay (β^d), and the undelayed flight cost component ($FCost_{rc}$). $RC_{c,oak}$ and $RC_{c,sfo}$ are the total rental costs for carrier c in using the terminals of OAK and SFO respectively. Finally, F_c is the total fixed cost incurred by the airline operating in the area.

Remember that landing fees ($fees_r$) are airport specific; the methodology to compute them is described below. Their values are endogenously determined, depending on fares, flight frequencies, and the vector of market shares (s). A similar endogeneity problem arises in the variables $weight_{rc}$, $RC_{c,oak}$ and $RC_{c,sfo}$. On the other hand, D_r only depends on flight frequencies.

Note that we make a distinction between product delays in the utility function ($\hat{D}_{jt}$) and airport delays in the profit function (D_r). While $\hat{D}_{jt}$ is the mean delay for each of the connecting and destination airports used by product jt, D_r refers to the average delay of the airport in the Bay Area used by the spoke r. In our application, $D_r = 25$ minutes if the spoke has SFO as an endpoint, and $D_r = 18$ minutes if the endpoint is OAK. Similarly, we distinguish between the daily frequency of product jt ($\hat{f}_{jt}$) used in the utility function (1), the daily frequency of flights of a carrier on one particular spoke (f_{rc}), and the total number of operations in the quarter for the carrier on spoke r ($\tilde{f}_{rc}$) appearing in the profit function (8). While $\hat{f}_{jt}$ corresponds to the mean of the frequencies for each of the segments of the product jt, f_{rc} only takes into account the carrier's flight frequency on spokes arriving at OAK or SFO. Finally, we assume that f_{rc} is the same for all days of the quarter. Hence, $\tilde{f}_{rc}$ is equal to f_{rc} times the number of days in the quarter (92 days).

Note that the first line of the profit maximization problem (8) refers to products. On the other hand, the term "Total Operating Flight Cost" is linked to aircraft operations. This distinction is important because it is possible that several products share the same aircraft in the last connection reaching the Bay Area, even if they belong to different markets. For instance, travelers flying from New York (JFK) to SFO via Boston (BOS) may share the same aircraft in their last trip segment with travelers flying non-stop from BOS to SFO. Since markets are defined as round trip directional city-pairs, passengers may belong to different markets even if they fly non-stop. Note that people traveling non-stop between any US city and the Bay Area may share the same aircraft but belong to different markets since they may be residing in the Bay Area or just visiting it. Hence, the common assumption used by the previous literature that markets are independent does not hold in our model. The optimality conditions will capture this dependence across markets.

Following Morrison and Winston (1989, 2007), we assume a deterministic relationship between the airport delay and the total number of daily flights arriving at the airport. Let $\mathcal{R}_{oak}$ and $\mathcal{R}_{sfo}$ denote the set of spokes reaching OAK and SFO respectively. Then, the delay function is given by

$$D_r = \begin{cases} exp(\omega^d_{oak}\bar{f}_{oak}) \;\; if \;\; r \in \mathcal{R}_{oak} \\ \\ exp(\omega^d_{sfo}\bar{f}_{sfo}) \;\; if \;\; r \in \mathcal{R}_{sfo} \end{cases} \tag{9}$$

where ω^d_{oak} and ω^d_{sfo} are the congestion parameters, and $\bar{f}_{oak}$ and $\bar{f}_{sfo}$ are the total number of daily operations at each of the airports. All three variables appearing in each line in (9) are the same for all flights landing at the same airport. While D_r, $\bar{f}_{oak}$ and $\bar{f}_{sfo}$ are observed from data, ω^d_{oak} and ω^d_{sfo} are computed to ensure that the equalities hold. As Morrison and Winston (1989, 2007) point out, this specification lets the marginal delay be an increasing function of the number of operations. Note that by construction, average delay is a source of dependence across markets. Since changes in the flight frequency of one carrier operating in a spoke route affect the average congestion at the airport, all products using the airport will be affected even if they do not use the same spoke.

Landing fees ($fees_r$) are the charge that airlines pay for each 1,000 lbs of maximum gross landing weight ($weight_r$) for each aircraft arrival.[5] Note that each product involves a round trip travel that may have several connections, where landing fees are levied. However, for simplicity and since our main interest is in airports located in the Bay Area, the only landing fees that we consider are those charged by OAK and SFO.

The weight of aircraft ($weight_{rc}$) is an indicator of its passenger capacity and depends on the type of airplane that carriers use on the spoke route r. For simplicity, all aircraft used by a carrier on a spoke are assumed to have the same characteristics. Moreover, we also assume that the weight of aircraft linearly depends on the total daily demand for the segment arriving in the Bay Area (TDD_{rc}), spoke route daily frequency (f_{rc}), the spoke distance ($dist_r$), airline identity ($carr_c$), and finally a dummy for SFO (I_{sfo_r}). Therefore,

$$weight_{rc}(s,p,f) \;\; = \;\; \tau_0 + \tau_1 TDD_{rc}(s,p,f) + \tau_2 f_{rc} + \tau_3 dist_r + \tau_4 carr_c + \tau_5 I_{sfo_r} + \epsilon^w_{rc} \tag{10}$$

[5] We assume the same fee per 1,000 pounds of aircraft MGLW applies for all flights landing at the same airport. In reality, some differences may apply depending on the relationship between carriers and airports and type of airplane.

where ϵ_{rc}^{w} is the disturbance term. Total daily demand (TDD_{rc}) and spoke route daily frequency (f_{rc}) are expected to be correlated with the error term. Consequently, the use of appropriate instruments is necessary to avoid inconsistent estimates.

Total daily demand (TDD_{rc}) does not necessarily equal the sum of the demand for products considered in our model. The reason is that we might find other travelers that use the same flight but do not belong to any of the products used in our model specification. Hence,

$$TDD_{rc}(s,p,f) = \sum_{t} \sum_{\substack{\{kt|r_{kt}=r \\ k\in\mathcal{J}_{ct}\}}} \frac{s_{kt}\times M_t}{92} + ResTDD_{rc} \tag{11}$$

where the first term on the right hand side (RHS) in (11) captures the demand for products considered in our specification that uses the spoke route r and carrier c. r_{kt} denotes the last spoke used by product kt to reach one of the airports in the Bay Area. While TDD_{rc} is the daily demand for spoke r and carrier c, $s_{kt}\times M_t$ is the product demand for the whole quarter. If we assume that demand is the same for any day of the quarter, we have to divide this term by the number of days of the quarter, set at 92. $ResTDD_{rc}$ corresponds to travelers that do not use any of the products of the model but still use the same airline and spoke r (for instance, one-way travelers or travelers connecting at SFO or OAK). This term is assumed to be independent of the demand for products accounted for in our model.

The terms $m_{jt}\times s_{jt}\times M_{jt}$, $RC_{c,oak}$, $RC_{c,sfo}$ and *"Total Operating Flight Cost"* are part of variable costs. Hence, their derivative with respect to the demand for a particular product will give us its marginal cost (mc_{jt}). Letting $q_{jt} = s_{jt}\times M_t$ and using the profit function (8),

$$\begin{aligned} mc_{jt}(s,p,f) &= m_{jt} + \sum_{r\in\Omega_c} \tilde{f}_{rc}\left(\frac{\partial fees_r}{\partial q_{jt}} \times weight_{rc} + fees_r \times \frac{\partial weight_{rc}}{\partial q_{jt}}\right) + \\ &+ \frac{\partial RC_{c,sfo}}{\partial q_{jt}} + \frac{\partial RC_{c,oak}}{\partial q_{jt}} \end{aligned} \tag{12}$$

Finally, we assume that both the product marginal cost (mc_{jt}) and the undelayed flight cost ($FCost_{rc}$) linearly depend on a vector of exogenous costs shifters (w_{jt}^{m}, w_{rc}^{f}) via the respective parameters (γ_m, γ_f) and a random term that captures unobserved product characteristics (ω_{jt}^{m}, ω_{rc}^{f}).[6] That is,

[6]We could have considered the more standard log linear form for the product marginal cost. However, in our application we found that 0.6% of the estimated mc_{jt} are negative. This result prevents us from using the log form.

$$mc_{jt}(s,p,f) = w_{jt}^{m}\gamma_m + \omega_{jt}^{m} \tag{13}$$

$$FCost_{rc} = w_{rc}^{f}\gamma_f + \omega_{rc}^{f} \tag{14}$$

The parameters w_{jt}^{m} will be estimated by equating (13) to (12), with the value of (12) generated as explained below.

3.2.2. *Landing Fees and Rental Building Rates:* An important question is how airports determine landing fees and terminal building rental rates, since these two variables will have an impact on the strategy of airlines. The methodology to compute these charges is airport specific and depends on the costs and revenues generated at airports. As we will see later, different methodologies lead to different responses of airlines and congestion at airports.

OAK airport:

As we previously stated, at OAK the Airfield Cost Center and the Terminal Cost Center must break even for the fiscal year. We compute total expenditures and deduct revenues. The residual is the amount that airlines must compensate the airport for using its infrastructure.

Landing fees at OAK are determined as the ratio between the difference of the costs and revenues attributed to the Airfield Cost Center ($ARCost_{oak}$) and the total scheduled maximum gross landing weight of carriers at OAK ($TWeight_{oak}$) for the fiscal year:

$$fees_r(s,p,f) = \frac{ARCost_{oak}}{TWeight_{oak}(s,p,f)} \quad for \;\; r \in \{\mathcal{R}_{oak}\} \tag{15}$$

where $\mathcal{R}_{oak}$ is the set of spoke routes at OAK.

Note that our application only includes product data from U.S. domestic flights during the 3rd quarter of 2006, while the methodology to determine landing fees and rental rates uses fiscal yearly data, as well as the weight of international, cargo, and general aviation flights. The lack of data, in some cases, and computer limitations, in others, constrain our analysis to U.S. domestic products for the 3rd quarter of 2006. Consequently, the variable $TWeight_{oak}$ is equal to

$$TWeight_{oak}(s,p,f) = \sum_{r \in \mathcal{R}_{oak}} \sum_{c} weight_{rc}(s,p,f) \times \tilde{f}_{rc} + ResTW_{oak} \tag{16}$$

The first term on the RHS in (16) corresponds to the total weight of all domestic flights operating at OAK that use one of the products of the model (U.S. domestic flights from the

3rd quarter of 2006). On the other hand, the term $ResTW_{oak}$ captures the total weight of flight operations that do not belong to products of our model. This term corresponds to the weight of international and cargo flights, as well as domestic flights from quarters of 2006 others than the 3rd. For simplicity, the latter term is assumed to be independent of the demand for products taken into account in our application.

Similarly, airlines also compensate OAK for the use of its terminals. The total terminal rental charges paid by airlines are equal to the difference between the operating expenditures and revenues of the Terminal Cost Center ($TCost_{oak}$):

$$TCost_{oak}(s,p,f) = (OE_{oak} - OR_{oak}(s,p,f)) \tag{17}$$

where operating revenues (OR_{oak}) are generated by concessions located in the terminals (mainly retail shops, restaurants, and car rentals). Operating expenditures (OE_{oak}) are costs associated with operating and maintaining the buildings and the cost recovery of capital investments (for instance, construction of a new terminal). Such airport net costs are allocated among airlines according to the percentage of the total terminal surface leased by each airline ($Usage_{c,oak}$). Then, the total rental charge that airline c must pay OAK is equal to

$$\begin{aligned} RC_{c,oak}(s,p,f) &= TCost_{oak}(s,p,f) \times Usage_{c,oak} = \\ &= (OE_{oak} - OR_{oak}(s,p,f)) \times Usage_{c,oak} \end{aligned} \tag{18}$$

We assume that expenditures (OE_{oak}) are exogenous. On the other hand, operating revenues (OR_{oak}) depend on the number of travelers. As we previously stated, OR_{oak} basically comes from shops, restaurants, and car rentals located at terminals, and such revenues depend on the number of travelers using the airport ($TTravelers_{oak}$). If we assume a linear relationship between concession revenues and the number of travelers, then

$$OR_{oak}(s,p,f) = \psi_{terminal,oak} \times TTravelers_{oak}(s,p,f) \tag{19}$$

where $\psi_{terminal,oak}$ is considered as the average operating revenue per traveler. Other specifications could be considered. For instance, we could assume that the concession revenue per traveler is decreasing in the total quantity of passengers. That would be more consistent with Van Dender's (2007) empirical results.

Once again, the total number of travelers using OAK may not be equal to the sum of the demand for products considered in our model. That is,

$$(20) \qquad TTravelers_{oak}(s,p,f) = \sum_{\{kt|r_{kt}\in\mathcal{R}_{oak}\}} s_{kt}(p,f,\delta(\theta);\theta) \times M_t + ResTT_{oak}$$

where the first term on the RHS in (20) corresponds to the total demand for products considered in our model, and $ResTT_{oak}$ is the demand not accounted in the products of our specification. For simplicity, the latter term is assumed to be independent of the demand for products took into account in our application.

SFO airport:

The way that SFO determines charges is different than the one used by OAK. In this case, the total landing fee revenues equal the amount needed to cover the net operating costs of the Airfield Cost Center ($ARCost_{sfo}$), plus 50% of the operating deficit (or surplus) in the Terminal ($TCost_{sfo}$) and Groundside ($GCost_{sfo}$) Cost Centers. The ratio between the total landing fee revenues and the total scheduled maximum gross landing weight of carriers ($TWeight_{sfo}$) is the fee ($fees_r$) that airlines pay per 1,000 pounds of MGLW of aircraft. That is,

$$(21) \qquad fees_r(s,p,f) = \frac{ARCost_{sfo} + \frac{1}{2}\left[TCost_{sfo}(s,p,f) + GCost_{sfo}(s,p,f)\right]}{TWeight_{sfo}(s,p,f)}$$

for $r \in \{\mathcal{R}_{sfo}\}$. $\mathcal{R}_{sfo}$ corresponds to the set of spoke routes at SFO. Note that while SFO includes the highly profitable vehicle parking and ground transportation vehicle access activities ($GCost_{sfo}$) in computing landing fees, OAK only considers the costs assigned to its Airfield Cost Center ($ARCost_{oak}$). The net operating cost of the Terminal Cost Center in SFO ($TCost_{sfo}$) is equal to the difference between operating expenditures (OE_{sfo}) and the operating revenues (OR_{sfo}):

$$(22) \qquad TCost_{sfo}(s,p,f) = (OE_{sfo} - OR_{sfo}(s,p,f))$$

As at OAK, while operating expenditures (OE_{sfo}) are assumed to be exogenous, operating revenues (OR_{sfo}) linearly depend on the number of travelers. Thus,

$$(23) \qquad OR_{sfo}(s,p,f) = \psi_{terminal,sfo} \times TTravelers_{sfo}(s,p,f)$$

where $\psi_{terminal,sfo}$ is the average operating revenue per traveler.

We define the net costs of the Groundside Cost Center ($GCost_{sfo}$) as the difference between costs (GC_{sfo}) and revenues ($GRev_{sfo}$) coming from groundside operations:

$$GCost_{sfo}(s,p,f) = GC_{sfo} - GRev_{sfo}(s,p,f) \tag{24}$$

where the term GC_{sfo} is assumed to be exogenous. We assume a linear relationship between $GRev_{sfo}$ and the total number of enplaned travelers:

$$GRev_{sfo}(s,p,f) = \psi_{ground,sfo} \times TTravelers_{sfo}(s,p,f) \tag{25}$$

where $\psi_{ground,sfo}$ is interpreted as the average revenue per enplaned passenger coming from the groundside operations (for instance, revenues from parking the car at the airport). The total number of enplaned travelers using SFO is not necessarily equal to the total demand for products considered in our application:

$$TTravelers_{sfo}(s,p,f) = \sum_{\{kt|r_{kt}\in\mathcal{R}_{sfo}\}} s_{kt}(p,f,\delta(\theta);\theta) \times M_t + ResTT_{sfo} \tag{26}$$

where $ResTT_{sfo}$ corresponds to demand that is not included in products of our model.

If we look again at the RHS in the fee rule (equation (21)), the total scheduled maximum gross landing weight of carriers at SFO ($TWeight_{sfo}$) is equal to,

$$TWeight_{sfo}(s,p,f) = \sum_{r\in\mathcal{R}_{sfo}} \sum_{c} weight_{rc}(s,p,f) \times \tilde{f}_{rc} + ResTW_{sfo} \tag{27}$$

where the first term of the RHS accounts for the weight of aircraft using products considered in our application, and $ResTW_{sfo}$ captures the total weight of flight operations that does not belong to products of our model (mainly international and cargo flights, as well as domestic flights from quarters of 2006 others than the 3rd).

Finally, the total terminal rental charges paid by airlines at SFO equal the amount needed to cover 3/2 of the net operating costs of the Terminal Cost Center ($TCost_{sfo}$), plus 50% of the calculated net operating surplus of the Groundside Cost Center ($GCost_{sfo}$). Then the total rental charge that airline c pays SFO for using its terminal is equal to

$$RC_{c,sfo}(s,p,f) = \left(\frac{3}{2}TCost_{sfo}(s,p,f) + \frac{1}{2}GCost_{sfo}(s,p,f)\right) \times Usage_{c,sfo} \tag{28}$$

where $Usage_{c,sfo}$ is the percentage of the total terminal surface leased by the airline. Remember that OAK only uses revenues and costs assigned to the Terminal Cost Center ($TCost_{oak}$) to compute its rental rate ($RC_{c,oak}$).

4. Solving the carriers' decision problem

In this section we describe the optimality conditions for airlines. As we previously noted, our model is a two stage game where carriers first decide on flight frequencies and afterwards decide on the price of tickets. As usual, this game is solved backwards: first, we derive the optimality conditions for fares given frequencies, and then we derive the first order conditions for frequencies taking into account the response of fares.

4.1. **Second Stage: Fares.** Solving the second stage, the first order condition for maximizing the profit function of airline c with respect to the fare $p_{j't'}$ is equal to

$$
\begin{aligned}
(29)\quad \frac{\partial \Pi_c}{\partial p_{j't'}} &= \sum_{t\in T}\sum_{j\in \mathcal{J}_{ct}} (p_{jt}-m_{jt})\frac{\partial s_{jt}}{\partial p_{j't'}}M_t + s_{j't'}M_{t'} - \\
&- \sum_{r\in\Omega_c}\tilde{f}_{rc}\left[\frac{\partial fees_r}{\partial p_{j't'}}weight_{rc} + fees_r\frac{\partial weight_{rc}}{\partial p_{j't'}}\right] - \frac{\partial RC_{c,oak}}{\partial p_{j't'}} - \frac{\partial RC_{c,sfo}}{\partial p_{j't'}} = 0
\end{aligned}
$$

where $t' \in T$ and $j' \in \mathcal{J}_{ct'}$. In our application, the product-specific costs (m_{jt}) are not observed. We will use the fare first order conditions to recover them. Then, we can compute the marginal costs (mc_{jt}) using (12) and estimate the parameters appearing in the first order conditions for flight frequency, which are presented below.

Now we turn to computation of the various derivatives appearing on the RHS of (29). As we will see, derivatives in the F.O.C. end up being functions of the derivatives of the market shares with respect to fares. Hence, those derivatives are easily computed once the demand (1) and the aircraft weight expressions (10) are estimated. Following Nevo (2000a), the derivative of the market share of product j in market t with respect to the price of product j' in market t' is

$$
(30)\qquad \frac{\partial s_{jt}}{\partial p_{j't'}} = \begin{cases} \int \alpha_i s_{ijt}(1-s_{ijt})dP_\nu(\nu_i)dP_L(L_i)dP_Y(y_i) & if \quad j=j' \;\&\; t=t' \\ -\int \alpha_i s_{ijt}s_{ij't'}dP_\nu(\nu_i)dP_L(L_i)dP_Y(y_i) & if \quad j\neq j' \;\&\; t=t' \\ 0 & if \qquad t\neq t' \end{cases}
$$

where $s_{ijt} = exp(\delta_{jt} + \mu_{ijt}) / \left[1 + \sum_{m \in J_t} exp(\delta_{mt} + \mu_{imt})\right]$ is the probability of individual i purchasing the product j in market t (similar interpretation for $s_{ij't'}$). α_i is the previously defined individual-specific coefficient associated with the ticket price.

If we look at the gradient of landing fees ($fees$) with respect to fares in (29), and assuming that the net operating costs of the Airfield Cost Center ($ARCostC$)[7] are exogenous, the derivative of the landing fees with respect to fares at OAK is equal to

$$\frac{\partial fees_r}{\partial p_{j't'}} = -\frac{fees_r}{TWeight_{oak}} \frac{\partial TWeight_{oak}}{\partial p_{j't'}} \quad for \ \ all \ \ r \in \{\mathcal{R}_{oak}\} \tag{31}$$

In the above expression, we use (16) to compute the derivative of the total scheduled landing weight ($TWeight_{oak}$) with respect to fares.

In the case of SFO,

$$\frac{\partial fees_r}{\partial p_{j't'}} = \frac{1}{TWeight_{sfo}} \left[\frac{1}{2} \frac{\partial (TCost_{sfo} + GCost_{sfo})}{\partial p_{j't'}} - \frac{\partial TWeight_{sfo}}{\partial p_{j't'}} fees_r\right] \tag{32}$$

for all $r \in \mathcal{R}_{sfo}$. Differences in the derivatives arise because SFO landing fees depend on the Terminal and Groundside Cost Centers while in OAK they do not. We use (22) and (24) to compute the derivatives of the net costs of the Terminal Cost Center ($TCost_{sfo}$) and the Groundside Cost Center ($GCost_{sfo}$) respectively. Similarly, using (27) we obtain the derivative of the total scheduled landing weight ($TWeight_{sfo}$) with respect to fares.

If we look again at the RHS in (29), we use (10) to compute the derivative of the weight of aircraft ($weight_{rc}$) with respect to ticket prices. Finally, the derivatives of the cost for carrier c associated with the rental of terminals are given by

$$\frac{\partial RC_{c,oak}}{\partial p_{j't'}} = -\frac{\partial OR_{oak}}{\partial p_{j't'}} \times Usage_{c,oak} \tag{33}$$

$$\frac{\partial RC_{c,sfo}}{\partial p_{j't'}} = -\left[\frac{3}{2} \frac{\partial OR_{sfo}}{\partial p_{j't'}} + \frac{1}{2} \frac{\partial GRev_{sfo}}{\partial p_{j't'}}\right] \times Usage_{c,sfo} \tag{34}$$

for OAK and SFO respectively. Once again, differences in both derivatives arise because OAK and SFO rental charges are determined using different mechanisms.

As noted above, the previous derivatives are computed using estimates from the demand (1) and aircraft weight equations (10). Then we plug their values in the fare F.O.C. (29)

[7]We may argue that $ARCostC$ depends on the number of landings at the airports. That is, the higher the number of operations, the higher are the costs of maintenance of the ramp. That would affect the costs attributed to the Airfield Cost Center. For simplicity, we do not consider this effect.

and solve for the product-specific cost (m_{jt}). This result lets us obtain the marginal costs (mc_{jt}), estimate the marginal cost equation (13), and the rest of parameters appearing in the first stage of the game (optimal decision of carriers with respect to frequencies).

4.2. **First Stage: Frequencies.** Once we derive the optimality conditions for prices, we solve the first stage of the game. The first order condition of the profit function of carrier c with respect to the daily frequency of its flights operating on a particular spoke r' is given by

$$
\begin{aligned}
(35)\quad \frac{\partial \Pi_c}{\partial f_{r'c}} &= \sum_{t \in T} \sum_{j \in \mathcal{J}_{ct}} \left[(p_{jt} - m_{jt}) \frac{\partial s_{jt}}{\partial f_{r'c}} M_t + \frac{\partial p^*_{jt}}{\partial f_{r'c}} s_{jt} M_t \right] - \\
&- 92 \left[fees_{r'} \times weight_{r'c} + \beta^d D_{r'} + FCost_{r'c} \right] - \\
&- \sum_{r \in \Omega_c} \tilde{f}_{rc} \left[\frac{\partial fees_r}{\partial f_{r'c}} weight_{rc} + fees_r \frac{\partial weight_{rc}}{\partial f_{r'c}} + \beta^d \frac{\partial D_r}{\partial f_{r'c}} \right] - \frac{\partial RC_{c,oak}}{\partial f_{r'c}} - \frac{\partial RC_{c,sfo}}{\partial f_{r'c}} = 0
\end{aligned}
$$

where $r' \in \Omega_c$. $\frac{\partial p^*_{jt}}{\partial f_{r'c}}$ denotes the derivative of the optimal fare with respect to frequency. In our application, we will use this condition to estimate the monetary value of one minute delay (β^d) and the undelayed flight cost component ($FCost_{rc}$). Moreover, we also use this expression to analyze the impact of changing the structure of the landing fee and rental charge rules.

The difficulty in (35) lies in computing the gradient of the optimal fare with respect to frequencies ($\frac{\partial p^*_{jt}}{\partial f_{r'c}}$), and the derivative of market shares with respect to frequencies ($\frac{\partial s_{jt}}{\partial f_{r'c}}$). Let us start with $\frac{\partial p^*_{jt}}{\partial f_{r'c}}$. We assume that the equilibrium pricing function is smooth with respect to flight frequency and take an approach similar to Villas-Boas (2007) and Fan (2012). We compute the total derivative of the price optimality condition (29) with respect to fares ($dp_k, k = \{1, \cdots, J\}$) and daily flight frequency ($f_b, b = \{1, \cdots, |\Omega|\}$), where J is the total number of offered products ($J = \sum_{t \in T} J_t$), and $|\Omega|$ is the total number of spokes operated by airlines at both airports in the Bay Area. Let Ψ^p_c denote the $J \times J$ ownership matrix with the general element $\Psi^p_c(j't', k)$ equal to one when both products $j't'$ and k are offered by carrier c and zero otherwise. Similarly, let Ψ^f_c denotes the $J \times |\Omega|$ ownership matrix with the general element $\Psi^f_c(j't', b)$ equal to one if the product $j't'$ and the spoke b are operated by carrier c. Then the total derivative of the first order condition (29) with respect to fares for product $j't'$ and carrier c is given by

$$(36) \qquad \underbrace{\sum_k \Psi_c^p(j't',k)\frac{\partial^2 \Pi_c}{\partial p_{j't'}\partial p_k}}_{G_c^p(j't',k)} dp_k + \underbrace{\sum_b \Psi_c^f(j't',b)\frac{\partial^2 \Pi_c}{\partial p_{j't'}\partial f_b}}_{H_c^p(j't',b)} df_b = 0$$

We can rewrite the previous expression in a matrix form. Let G_c^p be a $J \times J$ dimensional matrix with component $G_c^p(j't',k)$. Similarly, let H_c^f be a $J \times |\Omega|$ dimensional matrix with component $H_c^f(j't',b)$. Then, condition (36) can be written as

$$(37) \qquad G_c^p dp + H_c^f df = 0$$

where dp and df are column-vectors of dp_k for $k = \{1, \cdots, J\}$ and df_b for $b = \{1, \cdots, |\Omega|\}$ respectively. Note that the components of the matrices G_c^p and H_c^f are different from zero only if the pairs $(p_{j't'}, p_k)$ and $(p_{j't'}, f_b)$ belong to carrier c.

We can express the previous equality in a more general form where all carriers are included. Letting $G_p = \sum_c G_c^p$ and $H_f = \sum_c H_c^f$, then the following equality also holds:

$$(38) \qquad G_p dp + H_f df = 0$$

If G_p is a full rank matrix, the derivative of optimal fares with respect to flight frequencies is given by

$$(39) \qquad \frac{dp^*}{df} = -G_p^{-1} H_f$$

Once we derive $\frac{dp^*}{df}$, we compute the derivative of market shares with respect to flight frequencies ($\frac{\partial s_{jt}}{\partial f_{r'c}}$). The derivative of the market share of product j in market t with respect to the flight frequency of carrier c operating in the spoke r' ($f_{r'c}$) is given by

$$(40) \qquad \frac{\partial s_{jt}}{\partial f_{r'c}} = \int s_{ijt}\left(\kappa_{ijt}^{r'c} - \sum_{n=1}^{J_t} \kappa_{int}^{r'c} s_{int}\right) dP_\nu(\nu_i) dP_L(L_i) dP_Y(y_i)$$

where

$$(41) \qquad \kappa_{i\Theta t}^{r'c} = \alpha_{ip}\frac{\partial p_{\Theta t}^*}{\partial f_{r'c}} + \frac{1}{e_{\Theta t}}\left(\alpha_{if}\mathbb{1}\{r_{\Theta t} = r' \cap \Theta \in \mathcal{J}_{ct}\} + \alpha_{id}\frac{\partial D_{r_{\Theta t}}}{\partial f_{r'c}}\right)$$

In (41), Θ denotes a product of market t ($\Theta \in \{j,n\}$ in equation (40)), and $\mathbb{1}\{\cdot\}$ is an indicator function equal to one if the condition inside brackets holds and zero otherwise. As we previously noted, travelers care about the mean of the flight frequency and average delay of each of the connecting and destination airports used by products, but the derivatives (40) refer to flight frequencies of the last spoke reaching OAK or SFO. That is why we introduce $e_{\Theta t}$ as the number of connections of the product. Also remember that we already defined α_{ip}, α_{if}, and α_{id} as parameters associated with fares, frequencies, and delays respectively.

The last term on the RHS in (41) corresponds to the derivative of the previously defined average flight delay at each airport (9), and it is given by

$$\frac{\partial D_{r_{\Theta t}}}{\partial f_{r'c}} = \begin{cases} exp(\omega_a^d \bar{f}_a)\omega_a^d \;\; if \;\; r_{\Theta t} \in \mathcal{R}_a \\ \\ 0 \;\; if \;\; r_{\Theta t} \notin \mathcal{R}_a \end{cases} \tag{42}$$

where $a \in \{oak, sfo\}$.

The derivative of the market share with respect to flight frequency in (40) not only depends on own characteristics but also on characteristics of other products. This effect is captured by the summation in the expression. $\kappa_{ijt}^{r'c}$ depends on the relationship between the product jt and the spoke-carrier pair $r'c$. Looking at (41), if jt uses r' ($r_{jt} = r'$) and carrier c, then $\kappa_{ijt}^{r'c}$ not only depends on the derivative of the optimal fare with respect to frequency ($\frac{\partial p^*_{\Theta t}}{\partial f_{r'c}}$), but also on the parameters linked to the demand for flight frequency (α_{if}) and also delay ($\alpha_{id}\frac{\partial D_{r_{jt}}}{\partial f_{r'c}}$). On the other hand, if the product jt does not use r' and carrier c but still uses the same airport, then $\kappa_{ijt}^{r'c}$ is no longer affected by α_{if} but still depends on the airport delay and the optimal fare derivative. Finally, if jt does not use airport a then $\kappa_{ijt}^{r'c}$ only depends on the derivative of the optimal fare with respect to frequency. Similar reasoning holds for $\kappa_{int}^{r'c}$.

Once we compute $\frac{\partial p^*_{jt}}{\partial f_{r'c}}$ and $\frac{\partial s_{jt}}{\partial f_{r'c}}$, the rest of the derivatives appearing in the frequency F.O.C. (35) are straightforward. In particular, the derivative of the landing fee with respect to frequency for OAK is equal to

$$\frac{\partial fees_r}{\partial f_{r'c}} = -\frac{fees_r}{TWeight_{oak}} \frac{\partial TWeight_{oak}}{\partial f_{r'c}} \;\; for \;\; r' \in \mathcal{R}_{oak} \tag{43}$$

where we use (16) for the derivative of the total scheduled landing weight.

In the case of SFO the derivative is equal to

$$\frac{\partial fees_r}{\partial f_{r'c}} = \frac{1}{TWeight_{sfo}} \left[\frac{1}{2} \frac{\partial \left(TCost_{sfo} + GCost_{sfo} \right)}{\partial f_{r'c}} - \frac{\partial TWeight_{sfo}}{\partial f_{r'c}} fees_r \right] \tag{44}$$

for $r' \in \mathcal{R}_{sfo}$, where the derivative of the net operating costs of the Terminal Cost Center ($TCost_{sfo}$) follows from computing the derivative of (22). Similarly, we use (24) for the derivative of the net costs of the Groundside Cost Center ($GCost_{sfo}$). In addition, we use (27) to compute the derivative of the total scheduled landing weight ($TWeight_{sfo}$).

If we look again at the RHS in (35), we use (10) to compute the derivative of the weight of aircraft ($weight_{rc}$) with respect to ticket prices. Finally, the derivatives of airline c's terminal rental charges with respect to frequency for OAK and SFO are respectively given by

$$\frac{\partial RC_{c,oak}}{\partial f_{r'c}} = -\frac{\partial OR_{oak}}{\partial f_{r'c}} \times Usage_{c,oak} \tag{45}$$

$$\frac{\partial RC_{c,sfo}}{\partial f_{r'c}} = -\left[\frac{3}{2}\frac{\partial OR_{sfo}}{\partial f_{r'c}} + \frac{1}{2}\frac{\partial GRev_{sfo}}{\partial f_{r'c}}\right] \times Usage_{c,sfo} \tag{46}$$

5. Data and Statistics

5.1. **Data Description:** We perform the estimation of the model for the third quarter of 2006. In order to conduct the analysis, several data sources are used. We can classify them according to the information they provide. First, we use the Airline Origin and Destination Survey (DB1B), the T-100, the Airline On-Time Performance data sets from the U.S. Bureau of Transportation Statistics, and aircraft manufacturers' websites to obtain information about the product choices of travelers and their characteristics. Second, the Federal Aviation Administration website, competition plans, and airports' board meeting proceedings give detailed financial information about airports and the methodology they use to determine landing fees and rental charges. Finally, we use demographic data from the American Community Survey (ACS), and add extremely useful demographic information provided by the 2006 Airline Passenger Survey conducted by the Metropolitan Transportation Commission of the San Francisco Bay (2006 MTC Survey).

The Appendix contains further details about data sources and construction of the data sets.

5.2. **Summary Statistics:**

5.2.1. *Choice and Flight Characteristics Statistics:* Summary statistics for the main flight characteristics variables are provided in Table 1. It shows the mean and standard deviation of product characteristics for each of the airports and also for both airports (column denoted as "*Both*"). As expected, the busiest airport (SFO) is the one with the highest mean fares and daily frequency. On average, products using SFO carry fewer passengers but the number of these products is higher than at OAK. In the 3rd quarter of 2006, travelers purchased, on average, 171.84 tickets of each product served at SFO, and 305.88 at OAK. If we look at distance, measured as the sum of the distance of each trip segment, products from SFO offer slightly longer trips than OAK. The table also shows the percentage of products offered by carriers. We remark that United Airlines (UA) is the carrier with the highest presence at SFO, with Southwest (WN) having the highest presence at OAK. It is also interesting to note that several carriers are established in both airports. This is not the case with WN and Northwest Airlines (NW), which only operate in OAK and SFO respectively in the sample year. In our application, the percentage of products using airports with slot constraints is higher at SFO than at OAK.[8] Finally, we observe that products using SFO have, on average, more delays. This is partly explained by the higher average delay at SFO (25 minutes) than at OAK (18 minutes).

Table 2 summarizes the number of products and number of operating carriers within markets. Both airports offer similar destinations, with 402 of the 446 existing markets having products using OAK and SFO. We also observe that SFO offers on average more products within a market. Similarly, SFO has, on average, higher number of operating carriers per market.

Table 3 reports statistics for spoke-carrier pairs using OAK and SFO. In particular, it shows the mean and standard deviation of MGLW, the daily flight frequency, and the total daily capacity (measured as the product of MGLW times frequency) of the spokes operated by each carrier. As we previously mentioned, the spoke-carrier pair characteristics are not the same as the product characteristics since several products may use the same spoke and carrier. On average, the daily frequency of flights arriving at SFO is higher than at OAK. We also observe differences in the average MGLW of aircraft landing at the two airports. That explains why SFO offers 30% more capacity than OAK. Finally, the total number of spokes operated by airlines is also higher at SFO (87 vs 72).

Tables 4 and 5 show statistics for spoke routes arriving at each airport broken down by carrier. They report the mean and standard deviation of daily flight arrivals, MGLW, and daily capacity for the spoke-carrier pairs at each of the two airports. In OAK (Table 4), WN is not only the airline that operates the most spokes (45), but also one of the carriers offering, on average, the highest number of flights and highest daily capacity on each of the spokes. The table also shows the degree of standardization of aircraft. In the particular case

[8] Slot constrained markets are the ones that have as destinations the airports JFK, LGA, DCA , and ORD.

of WN, the standard aircraft is the Boeing 737. That explains its low weight variance. If we look at Table 5, SFO is one of the hubs for United Airlines (UA). That may be the reason why UA operates a high number of spokes and offers the highest daily capacity. Finally, except for WN (in 2006 only operating in OAK), the presence of airlines (measured as the number of spokes they operate) is higher at SFO than at OAK.

Table 6 shows financial details of the airports for the year 2006. Such information is useful when we incorporate landing fees and rental rates in our model. While OAK charges \$1.460 per landed 1,000 pounds, SFO charges \$3.213. We also observe that the number of enplaned passengers at SFO is more than double that of OAK. The terminal operating revenues (OR_a for $a \in \{oak, sfo\}$) are the amount that airports receive from concessions: food, beverage, retail stores, and rental cars. The bigger shopping and restaurant area in SFO terminals explains the large difference between the operating revenues of both airports (17 million dollars at OAK vs 126 million dollars at SFO). Assuming a linear relationship between operating revenues and enplaned passengers, we can obtain the average concession revenue per enplaned passenger ($\psi_{terminal,a}$). We also observe that the groundside revenues at SFO ($GRev_{sfo}$) are double those at OAK. If we divide this term by the total number of enplaned passengers in 2006, we obtain the groundside revenue per enplaned passenger ($\psi_{ground,a}$). Interestingly, this value is higher at OAK than SFO, which may be explained because of the higher OAK car parking rates. Finally, as expected, the total weight (aircraft weights times the number of operations in 2006) at SFO is more than double that of OAK.

5.2.2. *2006 Airline Passenger Survey Statistics:* In our estimation, this survey helps us to identify demographic variables for passengers (household income and distance to airports). Table 7 and Table 8 summarize the results for the 2006 MTC survey. Table 7 shows the joint distribution of the survey respondents using SFO according to their county of origin/destination and household income. Similarly, Table 8 analyzes the distribution of respondents for OAK. We group counties according to their geographical location and similarities in income distributions. This decision has been made because of the low number of observations in some cases (second column of both tables). We observe that it is more likely that a traveler chooses the closest airport for his trip. In the case of SFO, 66.8% of the respondents are located in the same county as the airport or in the county next to it (San Mateo and San Francisco counties). Similar comments may be made for OAK, where travelers are more likely to come from Alameda (where OAK is located) and Contra Costa counties (56.8% of the respondents). If we look at household incomes, differences between the two airports arise for the populations with very low and very high income. If we compare the last row of both tables, we observe that the probability that a traveler belongs to the household income group below $25,000 is 3.8% for SFO and 6.8% for OAK. For the highest income group, the probability is higher for travelers using SFO rather than OAK (16.7% vs 11.9%). For the rest of the income groups the probabilities are quite similar. As we will explain more carefully below, we will use the probability distributions appearing in both tables to increase the efficiency of the estimates of our model (columns 3 and 4, and row 7 of each table).

6. Estimation

The model is estimated as follows: first, we estimate the parameters of the weight equation (10) by two-stage least squares (TSLS); second, we estimate the demand (1) and marginal cost equations (13) by the general method of moments (GMM); finally, we use the first order conditions with respect to frequencies in (35) to estimate the valuation of one minute delay (β^d) and the parameters (γ_f) of the undelayed cost equation (14). While the estimation of (10) is straightforward, some remarks are necessary for the GMM procedure, and the estimation of β^d and γ_f. In this section, we sketch the estimation procedure. Further details are provided in the appendix.

The GMM estimation procedure follows the nested fixed point approach suggested by Petrin (2002). Petrin extended the algorithm proposed by Berry, Levinsohn and Pakes (1995) (BLP) by combining data from different sources. The model is estimated using a nonlinear GMM method. Three sets of moment conditions are used: one derived from the difference between the observed market shares and predicted market shares, the marginal cost moments, and other moment conditions that add extra demand information using the 2006 MTC Survey.

From the mean utility equation (1) and given δ_{jt}, θ and product characteristics, we can derive the moment condition related to the unobserved-to-researcher characteristics of product j in market t (ξ_{jt}). Using appropriate instruments (z^d) to control for price and frequency endogeneity, our moment condition can be written as

$$E[z^d_{jt}\xi_{jt}] = 0 \tag{47}$$

The second set of moments corresponds to the marginal cost condition and equals

$$E[z^m_{jt}\omega^m_{jt}] = 0 \tag{48}$$

where ω_{jt}^m is the residual of the marginal cost equation (13) and z_{jt}^m are cost instruments.

The last set of moments of the GMM procedure is constructed using the 2006 MTC Airline Survey Data. Adding other data sources is an extremely useful tool for identification (Petrin (2002)). In our application, the 2006 MTC Airline Survey Data give us interesting demographic information about travelers conditional on the use of one of the two airports in the Bay Area. In particular, we use information about travelers' distances to the airports and their household incomes. Basically we construct moments that match the predicted average consumer demographics obtained from the moments (47) with the average consumer demographic characteristics from the 2006 MTC survey. In particular, the extra moment conditions will match the probability that a traveler i using one of the airports (a) comes from/goes to a specific county ($\mathcal{C}$) and belongs to a income group ($\mathcal{Y}$). That is,

$$\eta_c(\mathcal{C}, a) = E\left[L_i \in \mathcal{C} | \{i \;\; uses \;\; airport \;\; a\}\right] \tag{49}$$

$$\eta_y(\mathcal{Y}, a) = E\left[y_i \in \mathcal{Y} | \{i \;\; uses \;\; airport \;\; a\}\right] \tag{50}$$

where

$$\mathcal{C} \in \{\text{S.Francisco-S.Mateo, Sta Clara, Alameda-C.Costa, Solano-Napa, Sonoma-Marin}\}$$

$$\mathcal{Y} \in \{\text{<\$25k, \$25k-50k, \$50k-75k, \$75k-100k, \$100k-150k, \$150k-200k,>\$200k}\}$$

$$a \in \{OAK, SFO\}$$

and where L_i and y_i in (49) and (50) are the county of origin/destination and household income group of individual i respectively. $\eta_c(\mathcal{C}, a)$ and $\eta_y(\mathcal{Y}, a)$ are the probabilities from the 2006 MTC survey (Tables (7) and (8)). The RHS expressions in (49) and (50) are the expected values predicted by our model and computed using the simulated market shares in (7). These extra conditions apply for all income groups, counties and airports.

In a final step, we estimate the undelayed flight cost component ($FCost_{rc}$) and the monetary value of one minute delay (β^d). Under the assumption that $FCost_{rc}$ linearly depends on a vector of cost shifters (w_{rc}^f) and a random term (ω_{rc}^f) (equation 14), we use the F.O.C. with respect to frequencies in (35) to identify β^d and the parameters linked to the undelayed flight cost component (γ_f). We combine equations (35) and (14), substituting the variable $FCost_{rc}$ in (35) using the RHS of the equation (14). Then, the expression we use for the estimation of β^d and γ_f is given by

$$\text{(51)} \qquad \frac{1}{92}\sum_{t\in T}\sum_{j\in\mathcal{J}_{ct}}\left[(p_{jt}-m_{jt})\frac{\partial s_{jt}}{\partial f_{r'c}}M_t+\frac{\partial p^*_{jt}}{\partial f_{r'c}}s_{jt}M_t\right]-fees_{r'}\times weight_{r'c}-$$
$$-\frac{1}{92}\left[\sum_{r\in\Omega_c}\tilde{f}_{rc}\left[\frac{\partial fees_r}{\partial f_{r'c}}weight_{rc}+fees_r\frac{\partial weight_{rc}}{\partial f_{r'c}}\right]-\frac{\partial RC_{c,oak}}{\partial f_{r'c}}-\frac{\partial RC_{c,sfo}}{\partial f_{r'c}}\right]=$$
$$=w^f_{rc}\gamma_f+\beta^d\left[D_r+\frac{1}{92}\sum_{r\in\Omega_c}\tilde{f}_{rc}\frac{\partial D_r}{\partial f_{r'c}}\right]+\omega^f_{rc}$$

Once the value of the LHS expression is computed, the estimation of the parameters follows from applying OLS. We did not include (51) in the GMM algorithm because the estimation procedure would be infeasible from a computation-time point of view.

Finally, the fitted value for $FCost_{rc}$ equals the product of the estimates for γ_f times the cost shifters (w^f_{rc}).

6.1. **Instruments and Identification:** As we previously stated, fares and frequencies are likely correlated with the unobserved-to-researcher variable ξ_{it} (see (1)) and the residual of the marginal cost equation (13). Similarly, total demand and frequency are also correlated with the disturbance term in the weight equation (10). Consequently, we need instruments to avoid inconsistent estimates. We use similar instruments to those proposed by Berry and Jia (2010) and Nevo (2001).

We use:

- Demand characteristics considered exogenous: distance, airport, operating carrier, direct flight, and airport with slot constraints.
- Number of flight connections.
- Dummy indicating if the connecting/destination airport is a hub for the carrier operating the flight.
- Dummy for trips longer than 1,500 miles.
- The mean of the distance of all products offered by competing carriers in the market.
- The mean of the distance of all products offered by the own carrier in the market.

The power of using other product characteristics depends on the proximity in characteristics space between products; products with closer substitutes likely have lower prices and/or higher frequency of flights.

In equation (1), we also control for the flight frequency endogeneity as follows: we first regress the number of daily departures on distance, market size (measured by geometric mean population of the origin and destination cities), number of competitors, carrier dummies, dummies for trips longer than 1,500 miles, a dummy indicating if the connecting/destination airport is a hub for the carrier operating the flight, and a dummy for SFO. Then we include the residuals, actual minus fitted frequencies, as instruments. The idea is that what is left after controlling for several factors (residuals) is correlated with the marginal cost but uncorrelated with the demand unobservable (ξ_{jt}).

A similar approach is used for the instruments in the weight equation (10). The only difference is that there are not connecting flights. Therefore, we cannot use the number of connections and average distance of competing products as instruments. Finally, we also add potential demand as an instrument (geometric mean population of the origin and destination metropolitan statistical areas).

The identification strategy is similar to the previous literature: 1) Reliance on substantial variation of product and demographic characteristics across markets, 2) Use of micro data (2006 MTC Survey), which lets us add extra moment conditions that match the predictions of our model with the survey (equations (49) and (50)), 3) Imposing a Bertrand-Nash equilibrium in prices, 4) Using the profit first order conditions with respect to flight frequency to estimate the cost of operating a flight ($FCost_{rc}$) and the cost of one minute delay (β^d) (equation (51)). In order to identify these two unobservable variables, we assume that $FCost_{rc}$ linearly depends on some factors: distance, operating carrier, airport, and a disturbance term (see (14)).

7. Estimation Results

7.1. **Demand Parameters:** Table 9 reports the demand estimates (1), which are consistent with the previous literature (see Berry and Jia (2010)). The standard errors are reported in parenthesis. While the second column corresponds to a logit model without instruments (OLS column), the third column reports estimates using instruments (IV column). In these model specifications, demand linearly depends on fares, distance, flight frequency, a dummy for direct flights, a dummy for SFO airport, carrier dummies, a dummy for slot constrained destinations, and delays. Most of the estimates have the expected sign. As we previously noted, fares and flight frequencies are likely correlated with the unobserved-to-researcher characteristics (ξ_{jt}). Without correcting for endogeneity, the price and frequency coefficients are biased upwards and downwards respectively (OLS column). One of the limitations of these two model specifications is that they do not capture important aspects of travelers' heterogeneity.

Column 4 reports estimates for the full model introducing heterogeneity of travelers. If we look at the last ten variables, σ^0, σ^f, and σ^d capture the individual taste for a constant, frequency of flights and delays, σ^p is the parameter associated with individual fare taste, λ is the airport-distance sensitivity of travelers, and α_y is the marginal utility of income. The last four parameters correspond to county fixed effects. Remember that the price coefficient has three components: the parameter common to all travelers (α_p), the marginal utility of income (α_y) times the household income (y_i), and the component that captures the heterogeneity that is not related to income ($\sigma^p \nu_i^p$). α_p and α_y have the expected sign. α_y is positive, indicating that the higher the traveler's income, the lower is the sensitivity to changes in fares. The estimates associated with flight distance are also consistent with the previous literature. The estimate linked to the distance from travelers' location to the airport (λ) is negative and significant. Hence, the farther are the airports, the more attractive is the outside option (e.g. use another mode of transportation or not traveling). Travelers prefer higher flight frequency because it is easier for them to find a flight that better matches their preferred departure time. Moreover, they dislike delays and prefer direct flights. Once we control for income and distance of travelers to the airports, SFO is preferred to OAK. This result does not hold in the OLS and IV cases because these two models do not account for travelers' heterogeneity.

Table 10 reports the mean and standard deviation of the product own-price elasticities. Columns 2 and 3 display the elasticities by airport and carrier. The column denoted by "*Both*" analyzes elasticities without distinguishing the airport. We do not observe large differences across airlines or airports. The mean own-elasticities in SFO are slightly higher than in OAK. This result may be partly explained by a higher number of competitors operating in markets using SFO. However, given the standard deviations of the estimates, differences are not statistically significant. If we compare the elasticities with the ones provided by Berry and Jia (2010), they are quite similar. Although our estimates are lower, they are of the same order of magnitude. Note that while we focus on products using airports located in the Bay Area, Berry and Jia (2010) study the whole U.S. domestic market.

Table 11 displays the frequency semi-elasticities of demand. For instance, if each carrier increases by one the number of daily flights operating on each spoke at both airports, the demand would increase on average 2.796%.

7.2. **Supply estimates:** As we pointed out in the model section, the weight of aircraft is an indicator of passenger capacity and an important determinant of the total landing fees that carriers pay. Table 12 reports estimates for the aircraft weight equation (10). Note that we use spoke route data (Table 2) rather than product data (Table 1). That explains why we use 315 observations instead of 12,790. We assume that the weight linearly depends on a constant, total number of passengers, daily frequency, distance, carriers, and a dummy variable indicating if the aircraft lands at SFO or OAK. The most important estimates are those linked to the demand (τ_1) and the daily flight frequency (τ_2). As expected, the τ_1 estimate is positive and significant, which means that the higher is the demand, the larger is the passenger capacity of the planes. Regarding τ_2, the larger is the frequency of flights operated by carriers on a spoke route, the smaller is the plane size. As Borenstein and Rosen (2008) remark, the higher the flight distance, the bigger are the planes. Finally, we do not observe significant differences between the size of aircraft operating at SFO and OAK.

Table 13 reports the average marginal cost per passenger-mile. The third column reports the mean Lerner index, defined as the ratio between the markups and fares. There are not significant differences across airlines.[9] We observe that the average marginal cost per passenger-mile for all products equals 6 cents, while the average Lerner index equals 36%. Berry and Jia (2010) reported the same marginal cost per passenger-mile, but a much higher Lerner Index (63%). Differences may be explained because Berry and Jia (2010) analyze the whole U.S. market and we only focus on markets using the main airports in the Bay Area.

Table 14 presents estimates for the linear specification of marginal costs (13), which includes a constant, carrier dummies, a SFO dummy, flight distance, the square of flight distance, a dummy indicating if the connecting or destination airports are hubs of the operating carrier, and number of connections. The results are again consistent with Berry and Jia (2010). Marginal costs increase nonlinearly with distance since an important fraction of the fuel is consumed during landings and take-offs. A similar argument can be used to explain the positive estimate for the number of connections. Ceteris paribus, marginal costs are higher at SFO. Finally, the hub estimate is unexpected. As Brueckner and Spiller (1994) point out, airlines use their hubs to take advantage of economies of traffic density. However, our estimate for the hub coefficient is positive rather than having the expected negative sign.

We assume that the undelayed cost component ($FCost_{rc}$) linearly depends on distance, squared distance, carriers and airport dummies. Table 15 shows the parameter estimates for equation (51). We tried several specifications, and in all of them β^d is negative but not significant. The distance and squared distance coefficients have the expected signs and they are significant. We do not observe significant differences among carriers or airports of arrival.

To conclude this section, Table 16 presents the fitted values for the undelayed fixed cost of operating a flight ($F\hat{C}ost_{rc}$), that equals equation (14) without the error term. The average fixed cost of operating a new flight is almost \$10,562. Finally, we do observe that the cost of operating flights at SFO are slightly higher than at OAK, but not significant.

[9]Standard errors are not reported, but they are relatively big. As a result, we cannot reject the hypothesis that the difference between any two means is zero.

8. Airline-Airport Agreements: Effects on Airline Behavior and Congestion at Airports

This section analyzes the effects of the contractual relationship between carriers and airports. Remember that charges are designed to let airports achieve financial self-sufficiency. Hence, changes in the operating costs of airports affect landing fees and rental rates of terminals. We first study how variations in expenditures at airports affect the behavior of carriers and flight delays. In a second counterfactual exercise, we show that different methodologies may induce airlines to behave differently. In particular, we look at the consequences of SFO adopting the contract used by OAK in order to manage congestion.

8.1. **Airport Operating Costs:** We study how changes in the cost of airports affect charges and their consequences on the strategy followed by airlines. Since delays depend on the daily frequency of flights, the response of airlines will also affect airport congestion. We assume that the new costs are attributable to the cost centers used to compute landing fees and rental charges: the costs assigned to the Airfield Cost Center, the operating expenditures (OE_a) assigned to the Terminal Cost Center, and the costs (GC_a) of the Groundside Cost Center for $a \in \{oak, sfo\}$. As we previously pointed out, we assume that these terms are exogenous.

The nature of such changes may be diverse, from minor repairs of the runway to more complex projects such as a new drainage system or the construction of a new taxiway or runway. However, our analysis assumes that changes in costs do not affect the capacity of airports. For example, in the case of constructing a new runway, our model would be valid until the new facility is operational.[10]

As we pointed out, OAK and SFO have different mechanisms to determine landing fees and rental rates. Let EXC_{oak} and EXC_{sfo} denote the expenditure element used in the computation of landing fees for OAK and SFO respectively (see (15) and (21)), then

$$EXC_{oak} = ARCost_{oak} \tag{52}$$

$$EXC_{sfo} = ARCost_{sfo} + \frac{1}{2}\left[OE_{sfo} + GC_{sfo}\right] \tag{53}$$

Since $ARCost_{oak}$ and $ARCost_{sfo}$ are assumed to be exogenous, using these terms or just the expenditures linked to these cost centers does not affect our analysis.

[10]Under certain conditions, the DoT let airports add a portion of the costs of airfield projects under construction to the landing-fee rule. See DoT FR-73, No 135, July 14, 2008.

Note that the optimal decisions of airlines are not affected by the cost component included in the rental rate methodology. If we look at the fare and frequency first order conditions (see (29) and (35)), they depend on the derivative of the rental rates and this derivative does not include costs.

Figures 2 to 9 show how changes in EXC_a for $a \in \{oak, sfo\}$ affect the most relevant variables of our model: landing fees, flight frequency, weight of aircraft and airport congestion. We simultaneously solve for flight frequency using (51) and obtain the new equilibrium in frequencies as a result of changes in the level of expenditure EXC_a.[11] We repeat the process for different values of EXC_a and interpolate the results. We decided to keep demand and fares fixed to reduce the computational burden. The effects of recomputing fares are expected to be low, since the percentage of the marginal cost of a product attributable to landing fees and rental of terminals is very small. Similarly, changes in total demand at OAK and SFO are also expected to be modest, since the cross frequency elasticities of demand between products using OAK and SFO are small. That means when the flight frequency of a product changes, travelers still prefer using products at the airport where the most preferred option is located rather than changing of airport or using the outside option.

Figures 2 and 3 display changes in landing fees as a function of SFO and OAK costs respectively. The x-axis represents the variation in the expenditures (denoted as ΔEXC_a for $a \in \{oak, sfo\}$). For instance, a ΔEXC_{sfo} equal to 5% indicates that the new cost at SFO is 5% higher than the original one. Displayed on the y-axis is the landing fee as a result of changes in expenditure. First, we observe that the landing fee curves have the expected slope: the higher are ΔEXC_{oak} and ΔEXC_{sfo} the higher are the fees that airlines must pay. We can also see that landing fees are slightly more sensitive to cost changes at SFO than at OAK. For instance, if $\Delta EXC_{sfo} = 20\%$, the SFO landing fee would increase by 44% (from \$3.213 to \$4.631). On the other hand, if the cost of operating OAK increases by 20%, the OAK landing fee would increase by 19% (\$1.460 to \$1.741).

If we look at the other variables of interest, an increase in landing fees is accompanied by a decrease in the total number of flights arriving at the airports (Figures 4 and 5), an increase in the size of aircraft (Figures 6 and 7), and a reduction in the level of congestion (Figures 8 and 9). These effects are nonlinear and much stronger at SFO. For example, if ΔEXC_{sfo} increases by 20%, the total number of daily flights reaching SFO decreases by 2.4% (from 620 daily flights to 605), the average weight of aircraft increases by 1.7% (from 147,312 to 149,816 pounds) and, as a consequence, the average delay of flights reaching SFO

[11] We use the Knitro package for Matlab to solve the system of non-linear simultaneous equations.

decreases by 8.1% (from 25 minutes and 22 seconds to 23 minutes and 20 seconds). Similarly, an increase in the costs of operating OAK by 20% ($\Delta EXC_{oak} = 20\%$), reduces the number of daily flights by 0.7% (from 326 daily flights to 303), increases the average weight by 1.6% (from 139,714 pounds to 141,949), and reduces congestion at OAK by 2.1% (from an average flight delay of 18 minutes and 40 seconds to 18 minutes and 16 seconds).

Finally, increasing the operating cost in one airport barely affects the behavior of airlines in the competing airport. However, the direction of the effects is the expected one. Imagine, for example, that the costs of operating SFO increase. Under this scenario, some carriers decide to increase the frequency of their flights at OAK in response to the decrease in the number of operations at SFO. Similar effect is observed when OAK increases its operating costs. However, the magnitude of these responses are so small that results are not reported.

8.2. **Contract Design:** One interesting exercise is to analyze the consequences of SFO adopting the airport-airline contract applied by OAK. This decision could be driven by the interest of the SFO managers in keeping the surplus from the Groundside Cost Center rather than sharing them with airlines. It could also be explained as a measure to reduce congestion, since the current methodology includes the net revenues from the Groundside Cost Center that likely reduce the amount that airlines pay for the use of SFO infrastructure.[12] The OAK methodology, on the other hand, does not take these revenues into account when it computes landing fees and rental charges. As a result, SFO would then not use the net revenues of the Groundside Cost Center ($GCost_{sfo}$) and a percentage of the net costs of the Terminal Cost Center ($TCost_{sfo}$) in the landing fee and rental charge equations (see (21) and (28)). That is, the proposed contract is given by

$$fees^*_{sfo} = \frac{ARCost_{sfo}}{TWeight^*_{sfo}} \tag{54}$$

and

$$RC^*_{c,sfo} = TCost^*_{sfo} \times Usage_{c,sfo} = (OE_{sfo} - OR^*_{sfo}) \times Usage_{c,sfo} \tag{55}$$

where $fees^*_{sfo}$ and $RC^*_{c,sfo}$ are the new fees and rental charges as a result of the new rule. Similarly, $TWeight^*_{sfo}$ and OR^*_{sfo} are equal to the new total weight and terminal operating revenues. In this scenario, $ARCost_{sfo}$ is the only expenditure taken into account to compute the landing fee.

[12] Remember that the Groundside Cost Center includes the usually highly profitable vehicle parking and ground transportation vehicle activities.

We use the first order conditions with respect to frequencies (51), substituting the original landing fees and rental rates with the new charge scheme ((54) and (55)), to recompute the new equilibrium in flight frequencies resulting from the new contract. If we look at the optimality conditions, we do not need $RC^{*}_{c,sfo}$ but do need its derivative. This derivative is known because it only depends on the average operating revenue per traveler ($\psi_{terminal,sfo}$) and the derivative of market shares with respect to frequencies. The problem is that the cost of operating the SFO airfield ($ARCost_{sfo}$) is unknown. We have data about the landing fee and the total weight of aircraft arriving at SFO, as well as the terminal and groundside revenues generated at the airport. With this information we can compute the total expenditures (EXC_{sfo}) used to determine landing fees. If we look at (21), the landing fee rule can be rewritten as follows:

$$(56) \quad fees_{sfo} \times TWeight_{sfo} + \frac{1}{2}\left[OR_{sfo} + GRev_{sfo}\right] = \underbrace{ARCost_{sfo} + \frac{1}{2}\left[OE_{sfo} + GC_{sfo}\right]}_{EXC_{sfo}}$$

As we previously noted, the LHS in (56) is observed. Consequently, the total expenditure (EXC_{sfo}) is also known. However, we do not have information about the expenditure components: net costs of the Airfield Cost Center ($ARCost_{sfo}$), terminal operating expenditure (OE_{sfo}), and groundside operation costs (GC_{sfo}).[13] The only thing we know about them is the total value for EXC_{sfo}. Hence, the term $ARCost_{sfo}$ is not identified.

In this section, we simulate for different values of $ARCost_{sfo}$ the new equilibrium in landing fees, frequencies, aircraft weight, and congestion using the new charge scheme. The results of the counterfactual exercise are displayed in Figures 10, 11, 12, and 13. The x-axis in the figures represents the percentage in the expenditure term attributable to $ARCost_{sfo}$. For instance, a value equal to 60% indicates that sixty percent of the cost component (EXC_{sfo}) corresponds to $ARCost_{sfo}$. Obviously, the figures are bounded above for a value of $ARCost_{sfo}$ equal to 100% of EXC_{sfo}, since the terms OE_{sfo} and GC_{sfo} are strictly positive.

Figure 10 shows the different values of landing fees for hypothetical $ARCost_{sfo}$ percentages. As expected, the higher is the exogenous component which remains in the landing fee rule, the higher is the fee. For instance, if SFO implements the contract used by OAK and the $ARCost_{sfo}$ is equal to 80% of the original expenditures, then the landing fee would be

[13]While the "Annual Operating Budget for SFO" and financial details that the airport reports to the FAA clearly state the revenues attributable to each cost center, the cost centers expenditures are not clear. For instance, the cost concept "salaries" appearing in the SFO balance accounts for personnel working at the terminal, airfield, and groundside. However, we do not observe which percentage belongs to each cost center.

around \$6, much higher than the original landing fee (\$3.213). $ARCost_{sfo}$ should be equal to 39% of the original costs to have the same landing fee as under the old contract.

Similarly, if changing the methodology reduces the level of expenditure used to compute the landing fee by 20% ($ARCost_{sfo}$ corresponds to 80% of EXC_{sfo}), total frequency would decrease by almost 4% (from 620 daily flights to 595) (see Figure 11), and the average size of airplanes would increase by 1.7% (from 140,289 pounds to 142,674) (see Figure 12). That would reduce congestion by 12% (from an average flight delay of 25 minutes and 22 seconds to 22 minutes and 19 seconds) (see Figure 13). Again, both contracts would lead to the same level of congestion, aircraft size, and flight frequency if $ARCost_{sfo}$ is equal to 39% of the original expenditures.

Given the fee scheme applied by SFO and the proposed contract, we derive the following inequality to analyze the effectiveness of the new agreement as a tool to reduce congestion:

$$ARCost_{sfo} > 0.39 \times \underbrace{\left[ARCost_{sfo} + \frac{1}{2}\left[OE_{sfo} + GC_{sfo}\right]\right]}_{EXC_{sfo}} \tag{57}$$

where the LHS in (57) corresponds to the proposed expenditure term used in the new landing fee scheme (54), and the RHS displays the original costs (EXC_{sfo}) multiplied by the previously discussed factor 0.39.

We can conclude that if the inequality in (57) holds (see Figure 13), the new rule would be an effective tool to reduce congestion. Otherwise, the landing fee would be lower than the original one, leading to an increase in the number of operations and higher congestion.

9. Conclusions

Using data from the Metropolitan Oakland International Airport (OAK) and San Francisco International Airport (SFO), we analyze the decisions of carriers and travelers taking into account the contractual agreement between airports and airlines. This contract sets the fees that carriers pay for landing, the rental rate for the terminal space that they occupy, as well as the methodology to determine these charges. Airlines consider these charges and how they are determined when they decide ticket prices and the frequency of their flights. Since OAK and SFO airports have different charge rules, the effects of changes in the terms of the contracts are also different. At both airports, an increase in landing fees is accompanied by a decrease in the total number of flights arriving at the airports, an increase in the size of aircraft, and a reduction in the level of congestion.

While aircraft weight-based methodologies for determining landing fees, as used in OAK and SFO, are not appropriate to reduce congestion at airports with big differences between the number of operations during peak and off-peak hours, they can be useful at airports continuously operating at the maximum capacity. The reason is that weight-based landing fees only depend on the weight of aircraft without taking into account at what time of the day the planes land. In our application, we show that under certain conditions, implementing at SFO the contract used by OAK can reduce congestion. This counterfactual exercise shows that different methodologies may induce airlines to behave differently. Further research could be addressed to study the optimal design of contracts given the preferences of travelers, strategy of carriers, and airport regulations.

The paper also introduces methodological innovations to capture characteristics of the airline industry: first, we let charges being endogenously determined by the behavior of travelers and carriers. Second, while in our analysis carriers face two decision variables (fares and frequency of flights), most of the previous literature focuses on a single decision variable (fares). Third, our rich model specification captures two sources of correlation across markets. The first source is the possibility that passengers from different markets share aircraft reaching the Bay Area. The second source of dependence is congestion at airports. Changes in the frequency of one product impact the level of congestion. Thus, products using the same airport will be affected even if they belong to different markets.

We acknowledge some limitations in our analysis. Our application only includes data from the third quarter of 2006, while landing fees and rental rates methodologies use yearly data. Including the other quarters will lessen the impact of changes in the cost of operating airports since they will be distributed among more flights. The lack of data also prevents us from including cargo flights and general aviation operations. The direction of the effects of adding this information is similar to the case of using yearly data. Finally, we did not use data from international and connecting passengers at OAK and SFO. Including information about these types of passengers would complicate the model since their preferences are very different from passengers with the Bay Area as final destination. In that case, the reaction of airlines to changes in the "use and lease agreement" would depend on how valuable the domestic travelers with SFO or OAK as final destinations are compared to international or connecting passengers at these airports.

References

[1] Airport Commission of San Francisco International Airport (2006): "Annual Operating Budget FY 2006".

[2] Bel, G., and Fageda X., (2009): “Privatization, regulation and airport pricing: an empirical analysis for Europe”.Journal of Regulatory Economics, 37(2).

[3] Berry, S. , and Jia P., (2010): “Tracing the Woes: An Empirical Analysis of the Airline Industry.” The American Economic Journal:Microeconomics, 2(3): 1–43.

[4] Berry, S., Levinsohn, J., and Pakes, A.(1995): “Automobile Prices in Market Equilibrium”, Econometrica, 63(4):841-890.

[5] Berry, S. (1994): “Estimating Discrete-Choice Models of Product Differentiation”, The RAND Journal of Economics, 25(2):242-262.

[6] Bilotkach, V., Clougherty, J.A., Mueller, J., and Zhang, A. (2012): “Regulation, Privatization, and Aeronautical Charges: Panel Data Evidence from European Airports”, Journal of Regulatory Economics, 42:73-94.

[7] Bilotkach, V., Gorodnichenko, Y., and Talavera O. (2012): “Sensitivity of Prices to Demand Shocks: A Natural Experiment in the San Francisco Bay Area”, Transportation Research Part A: Policy and Practice, 46(7):1137-1151.

[8] Borenstein, S., and Rose N. (2008): “How Airline Markets Work...Or Do They? Regulatory Reform in the Airline Industry”, National Bureau of Economic Research working paper # 13452, forthcoming in N. Rose ed., Economic Regulation and Its Reform: What Have We Learned?, University of Chicago Press.

[9] Brueckner J.K., D. Lee, and E. Singer (2013): “City-Pairs vs. Airport-Pairs: A Market-Definition Methodology for the Airline Industry”, Review of Industrial Organization.

[10] Daniel, J. (2001): “Distributional Consequences of Airport Congestion Pricing”, Journal of Urban Economics, 50:230-258.

[11] Dube, J-P., Fox, J., and Su C-L. (2012): “Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation”, Econometrica, 80(5):2231-2267.

[12] Fan, Y. (2012): “Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market”, American Economic Review, forthcoming.

[13] Ivaldi, M., S. Sokullu and T. Toru (2011): “Are Airports Two-sided Platforms?: A Methodological Approach”, Advances in Airline Economics, Vol 3. J. Peoples Ed.

[14] Mayer, C., and Sinai, T. (2003): “Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil”, American Economic Review, 93(4):1194-1215.

[15] Morrison, S. A., and Winston, C. (2007): “Another Look at Airport Congestion Pricing”, American Economic Review, 97(5):1970-1977.

[16] Morrison, S. A., and Winston, C. (1989): “Enhancing the Performance of the Deregulated Air Transportation System”, Brookings Papers on Economics Activity: Microeconomics, 61-123.

[17] Nevo A. (2001): “Measuring Market Power in the Ready-to-Eat Cereal Industry”, Econometrica, 69(2):307-342.

[18] Nevo A. (2000): “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry”, The Rand Journal of Economics, 31(3):395-421.

[19] Nevo A. (2000): “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand”, Journal of Economics & Management Strategy, 9(4):513-548.

[20] Oakland International Airport: Master Plan (2006).

[21] Petrin (2002): "Quantifying the Benefits of New Products: The Case of the Minivan", Journal of Political Economy, 110(4):705-729.

[22] Port of Oakland: Board of Port Commissioners 2007. Proceedings.

[23] Train K. (2009): "Discrete Choice Methods with Simulation". Cambridge University Press.

[24] Urdanoz M., and Sampaio R. (2011): "Airlines' Cooperation in the US Domestic Market: Measuring the Evolution of Price Gaps", mimeo.

[25] Van Dender K. (2007): "Determinants of Fares and Operating Revenues at US Airports", Journal of Urban Economics, 62:317-336.

[26] Villas-Boas S. (2007): "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data", The Review of Economic Studies, 74(2):625-652.

Figure 1. San Francisco Bay

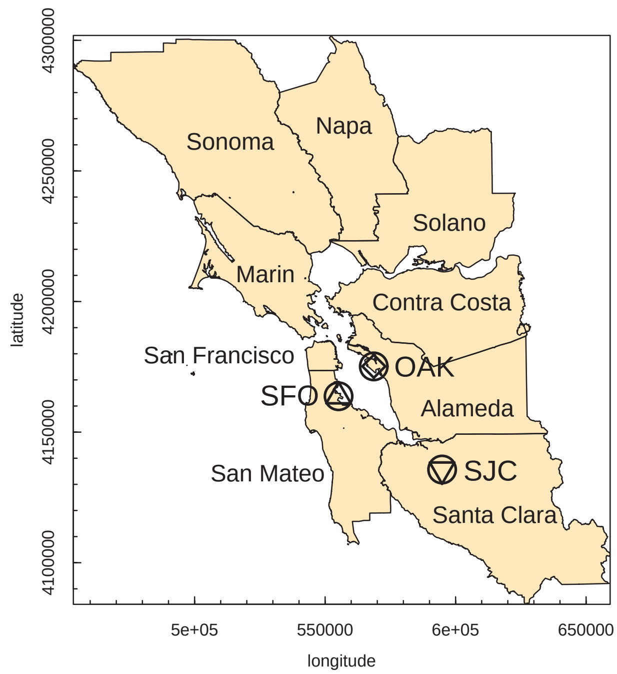


Table 1. Summary Statistics of Products (3Q. 2006)

	OAK		SFO		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Fare (p) ($100)	3.81	1.48	4.26	2.14	4.11	1.96
Nb Passengers	305.88	2,634.33	171.84	1,234.18	214.37	1,801.45
Direct Flight	0.03	0.16	0.02	0.15	0.02	0.15
Daily Frequency ($\hat{f}$)	4.56	2.39	4.89	2.11	4.79	2.21
Distance (1,000 miles)	4.23	1.33	4.64	1.15	4.51	1.22
AA	0.07	0.26	0.16	0.36	0.13	0.34
CO	0.04	0.20	0.05	0.22	0.05	0.22
DL	0.12	0.32	0.13	0.34	0.13	0.34
NW	-	-	0.11	0.31	0.07	0.26
UA	0.20	0.40	0.35	0.48	0.31	0.46
US	0.12	0.32	0.15	0.36	0.14	0.35
WN	0.37	0.48	-	-	0.12	0.32
Others	0.07	0.25	0.05	0.21	0.05	0.22
Slot Constraints	0.18	0.38	0.28	0.45	0.25	0.43
Delay ($\hat{D}$) (minutes)	22.48	2.56	25.11	2.14	24.28	2.59
Nb Products	4,058	-	8,732	-	12,790	-

Table 2. Market Average Statistics (3Q.2006)

	OAK		SFO		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Nb Products	9.10	12.89	19.58	30.15	28.68	40.88
Nb Carriers	2.56	1.77	3.58	2.23	4.24	2.30
Nb Markets	405	-	443	-	446	-

Table 3. Supply Statistics: Spoke-Carrier in the Bay Area (3Q.2006)

	OAK		SFO		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Daily Frequency	2.77	3.36	3.12	2.57	2.99	2.89
Aircraft MGLW (10^3 pounds)	139.71	24.84	147.31	66.29	144.48	54.73
Total Daily Capacity (10^3 pounds)	367.08	403.99	437.07	452.70	411.02	435.87
Nb Spokes	72.00	-	87.00	-	103.00	-

Table 4. Supply Statistics: Spoke-Routes Arriving at OAK (3Q.2006)

	Daily Freq		**MGLW (10^3 pounds)**		**Daily Capacity (10^3 pounds)**		**Spokes**
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	
OAK	2.77	3.36	139.71	24.84	367.08	403.99	72
AA	3.85	2.27	137.47	25.23	525.26	317.95	6
CO	2.64	0.62	115.42	60.44	289.07	138.63	3
DL	1.54	0.54	166.97	30.34	257.11	105.67	12
NW	-	-	-	-	-	-	-
UA	1.78	1.31	159.43	23.90	280.19	191.51	14
US	2.00	1.09	147.77	24.02	281.44	131.94	13
WN	3.63	4.96	127.58	9.54	446.03	587.68	45
Others	2.42	1.96	133.64	17.15	319.50	253.51	23

Table 5. Supply Statistics: Spoke-Routes Arriving at SFO (3Q.2006)

	Daily Freq		**MGLW (10^3 pounds)**		**Daily Capacity (10^3 pounds)**		**Spokes**
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	
SFO	3.12	2.57	147.31	66.29	437.07	452.70	87
AA	2.77	2.25	151.38	58.05	435.97	409.83	24
CO	2.80	2.11	142.71	35.11	420.40	329.11	9
DL	1.92	1.28	192.04	67.38	359.26	316.56	28
NW	2.53	1.89	155.93	46.23	380.20	265.10	17
UA	3.79	3.29	130.37	76.75	485.11	613.80	61
US	3.63	2.40	128.80	58.01	449.83	411.07	35
WN	-	-	-	-	-	-	-
Others	2.93	2.22	156.33	53.49	434.91	370.45	25

Table 6. Financial Information for year 2006

	OAK	**SFO**
Landing Fees (\$) per 10^3 pounds of MGLW	1.460	3.213
Year Enplaned Pax (10^3)	14,639	33,148
Operating Revenues (OR) (10^3\$)	17,323	125,656
Groundside Revenues ($GRev$) (10^3\$)	37,769	57,686
$\psi_{terminal,a}$	0.0237	0.0758
$\psi_{ground,a}$	0.0542	0.0473
Total Weight ($TWeight$) (10^6 pounds)	8,866	20,095

Table 7. SFO 2006 MTC Survey: County - Income (Joint Distribution)

	Nb Obs	**SFO**	**OAK**	**less $25k**	**$25k-$50k**	**$50k-$75k**	**$75k-$100k**	**$100k-$150k**	**$150k-$200k**	**more $200k**
S.Francisco-S.Mateo	2060	0.668	0.266	0.022	0.059	0.099	0.133	0.144	0.096	0.116
Santa Clara	259	0.093	0.025	0.003	0.006	0.011	0.018	0.019	0.016	0.020
Alameda-C.Costa	1555	0.126	0.568	0.008	0.011	0.014	0.021	0.031	0.024	0.016
Solano-Napa	159	0.032	0.040	0.000	0.003	0.007	0.005	0.005	0.005	0.006
Sonoma-Marin	406	0.081	0.101	0.005	0.009	0.016	0.016	0.016	0.011	0.010
Total	4439	1	1	0.038	0.089	0.147	0.193	0.214	0.152	0.167

Table 8. OAK 2006 MTC Survey: County - Income (Joint Distribution)

	Nb Obs	**SFO**	**OAK**	**less $25k**	**$25k-$50k**	**$50k-$75k**	**$75k-$100k**	**$100k-$150k**	**$150k-$200k**	**more $200k**
S.Francisco-S.Mateo	2060	0.668	0.266	0.017	0.029	0.048	0.052	0.052	0.032	0.035
Santa Clara	259	0.093	0.025	0.001	0.002	0.004	0.004	0.007	0.002	0.004
Alameda-C.Costa	1555	0.126	0.568	0.038	0.062	0.087	0.112	0.119	0.085	0.065
Solano-Napa	159	0.032	0.040	0.002	0.006	0.007	0.009	0.005	0.006	0.005
Sonoma-Marin	406	0.081	0.101	0.010	0.013	0.019	0.019	0.017	0.012	0.011
Total	4439	1	1	0.068	0.113	0.166	0.196	0.200	0.137	0.119

Table 9. Demand Estimates

	OLS	IV	RCM
Intercept	-10.557**	-10.010**	-5.861**
	(0.160)	(0.225)	(0.363)
Fare (α_p)	-0.056**	-0.581**	-1.301**
	(0.006)	(0.059)	(0.087)
Distance (1,000 miles)	0.384**	0.876**	0.997**
	(0.045)	(0.073)	(0.147)
Distance Squared	-0.078**	-0.120**	-0.128**
	(0.006)	(0.008)	(0.019)
Frequency (α_f)	0.021**	0.113**	0.063**
	(0.006)	(0.008)	(0.009)
Direct	4.071**	4.084**	4.137**
	(0.075)	(0.098)	(0.110)
SFO	-0.089**	-0.058	0.199**
	(0.029)	(0.039)	(0.043)
CO	-0.091	0.375**	0.396**
	(0.058)	(0.087)	(0.100)
DL	-0.042	0.004	-0.145
	(0.044)	(0.057)	(0.067)
NW	-0.248**	-0.238**	-0.317**
	(0.051)	(0.067)	(0.071)
UA	-0.258**	0.170**	0.326**
	(0.036)	(0.065)	(0.080)
US	-0.371**	-0.090	-0.038
	(0.043)	(0.059)	(0.070)
WN	-0.461**	-0.626**	-0.731**
	(0.052)	(0.068)	(0.073)
Others	0.368**	0.486**	0.510**
	(0.056)	(0.074)	(0.084)
Slots	-0.380**	-0.327**	-0.360**
	(0.029)	(0.038)	(0.043)
Delay (α_d)	-0.002	-0.017	-0.144
	(0.007)	(0.009)	(0.012)
Random Constant (σ^0)	-	-	-0.137
	(-)	(-)	(0.105)
Random Price (σ^p)	-	-	-0.059**
	(-)	(-)	(0.005)
Distance to Airports (λ)	-	-	-1.042**
	(-)	(-)	(0.002)
Random Price-Income (α_y)	-	-	0.087**
	(-)	(-)	(0.001)
Random Frequency (σ^f)	-	-	0.125**
	(-)	(-)	(0.002)
Random Delay (σ^d)	-	-	-0.081**
	(-)	(-)	(0.001)
Sfo-Mateo	-	-	1.461**
	(-)	(-)	(0.009)
Sta Clara	-	-	0.413**
	(-)	(-)	(0.005)
Alameda-Costa	-	-	2.213**
	(-)	(-)	(0.068)
Sonoma-Marin	-	-	3.125**
	(-)	(-)	(0.008)
Nb Observations	12,790	12,790	12,790
J-Statistic	-	-	27.9
Sargan Statistic	-	28.4	-
R^2	0.311	-	-
Degrees of Freedom	-	4	31
χ^2 Critical Value (5%)	-	9.49	44.99

** Significant at the 5 percent level.

Table 10. Price Elasticities Estimates

	OAK	**SFO**	**Both**
All Carriers	-3.00	-3.12	-3.08
	(0.79)	(0.94)	(0.90)
AA	-2.95	-3.06	-3.04
	(0.76)	(0.86)	(0.84)
CO	-3.21	-3.36	-3.32
	(0.79)	(0.88)	(0.86)
DL	-2.81	-2.79	-2.79
	(0.95)	(0.99)	(0.98)
NW	-	-2.90	-2.90
	-	(0.76)	(0.76)
UA	-3.29	-3.34	-3.33
	(0.87)	(0.99)	(0.97)
US	-3.09	-3.14	-3.12
	(0.73)	(0.86)	(0.83)
WN	-2.83	-	-2.83
	(0.60)	-	(0.60)
Others	-3.19	-2.82	-2.97
	(0.89)	(0.85)	(0.89)

Table 11. Frequency Semi-Elasticities of Demand

	OAK	**SFO**	**Both**
Mean	2.805	2.834	2.796
Median	2.665	2.687	2.647
Std	1.072	1.198	1.306

Table 12. Weight Estimates (TSLS)

Weight (10^3 pounds)	**Estimate**
Intercept	123.990**
	(11.527)
Daily Pax (TDD) (τ_1)	0.138**
	(0.018)
Daily Frequency (τ_2)	-14.445**
	(1.924)
Distance ($dist$) (1,000 miles)	28.223**
	(3.212)
CO	-16.969
	(14.939)
DL	34.3325**
	(11.960)
NW	27.925
	(14.782)
UA	7.332
	(11.090)
US	-10.379
	(14.749)
WN	-23.169
	(12.686)
Other	-24.089**
	(10.726)
SFO (I_{sfo})	-5.413
	(5.584)
Nb Observations	315
Sargan Statistic	4.187
Degrees of Freedom	3
χ^2 Critical Value (5%)	7.814

** Significant at the 5 percent level.

Table 13. Average Marginal Costs and Lerner Index

	mc ($) per pax-mile	**Lerner Index**
All Flights	0.065	0.363
OAK	0.065	0.366
SFO	0.066	0.362
Direct Flights	0.096	0.450
Connecting Flights	0.065	0.361
AA	0.059	0.367
CO	0.065	0.327
DL	0.054	0.398
NW	0.052	0.375
UA	0.079	0.346
US	0.063	0.349
WN	0.063	0.381
Others	0.068	0.380

Table 14. Marginal Cost Estimates

mc ($100)	**Estimate**
Intercept	0.297**
	(0.108)
SFO	0.078**
	(0.033)
Distance (1,000 miles)	0.659**
	(0.050)
Distance2	-0.060**
	(0.006)
Hub	0.237**
	(0.028)
Nb Connections	0.143**
	(0.019)
CO	0.783**
	(0.070)
DL	-0.213**
	(0.049)
NW	-0.018
	(0.063)
UA	0.449**
	(0.045)
US	0.242**
	(0.048)
WN	-0.029
	(0.057)
Others	0.091
	(0.064)
Nb Observations	12,790

** Significant at the 5 percent level.

Table 15. Total Cost Frequency Estimates

\$100	**Estimate**
Intercept	19.28** (4.02)
β^d	-0.37 (0.24)
Distance (1,000 miles)	4.55** (1.63)
Distance2	-0.42 (0.27)
CO	1.91 (3.30)
DL	-3.35 (2.30)
NW	-1.78 (2.82)
UA	-2.28 (3.17)
US	-0.61 (2.37)
WN	-1.02 (4.52)
Others	-0.36 (2.33)
SFO	2.59 (2.62)
Nb Observations	315
R^2	0.216

** Significant at the 5 percent level.

Table 16. Undelayed Cost Frequency Estimates ($F\hat{C}ost$)

$F\hat{C}ost$ **(\$)**	**OAK**	**SFO**	**Both**
Mean	8,837	11,585	10,562
Median	7,070	10,026	9,570
Std	3,585	3,970	4,050

Figure 2. Landing Fees SFO ($)

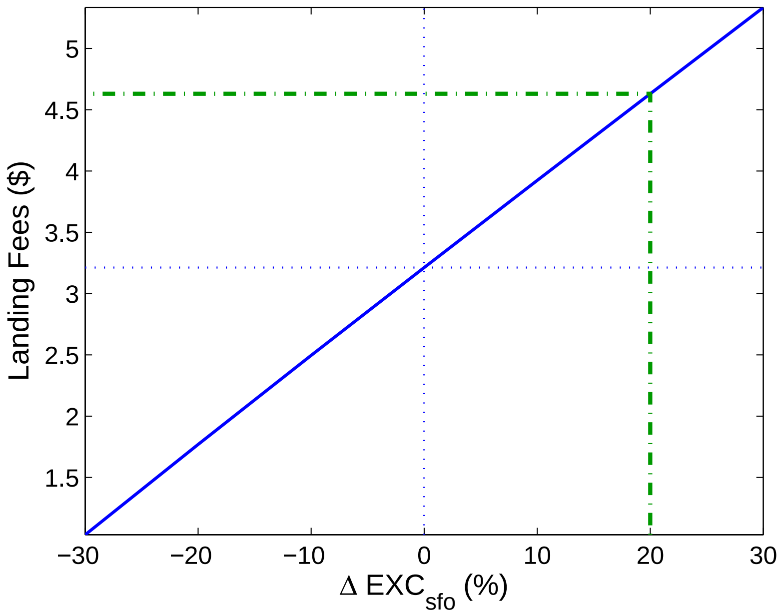


Figure 3. Landing Fees OAK ($)

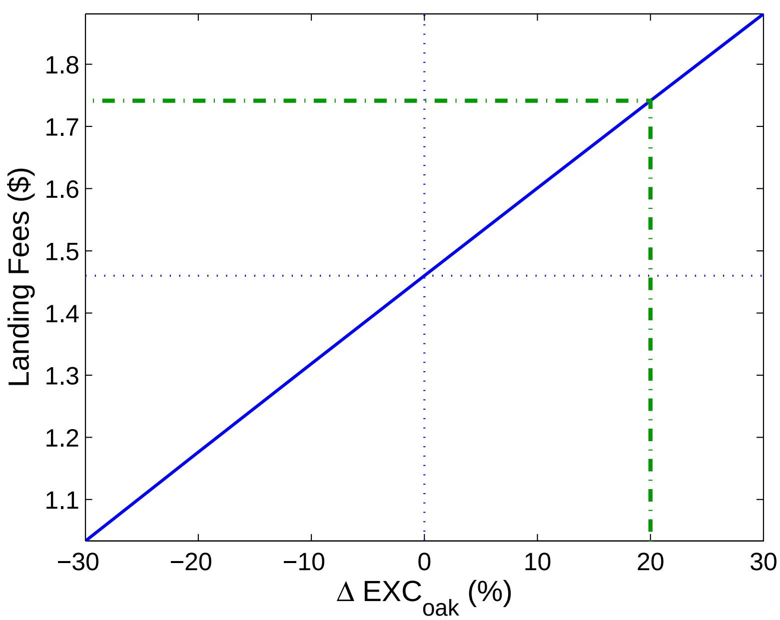


FIGURE 4. Δ Total Daily Flights SFO (%)

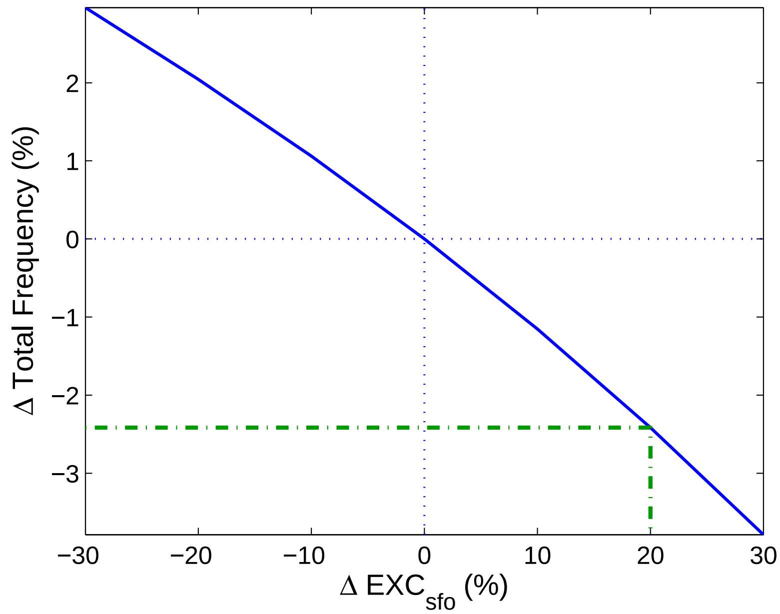


FIGURE 5. Δ Total Daily Flights OAK (%)

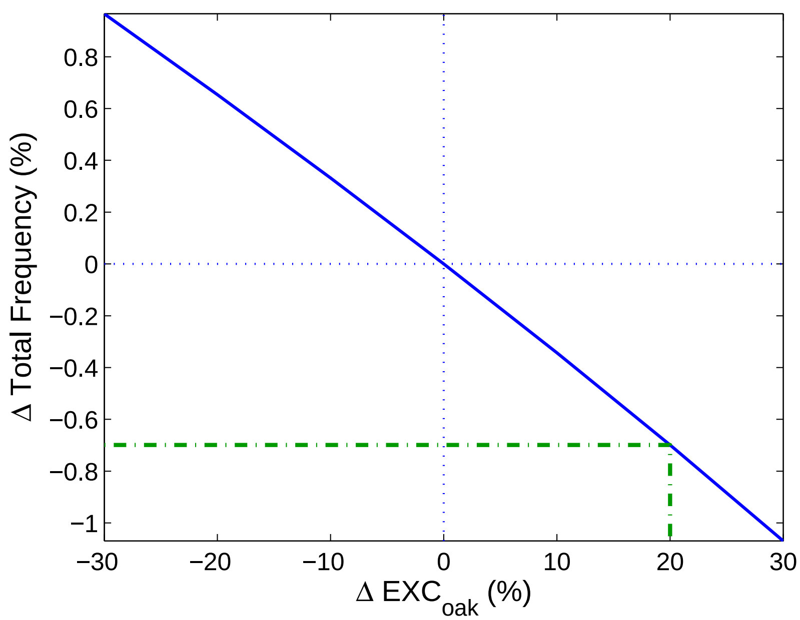


Figure 6. Δ Aircraft Average Weight SFO (%)

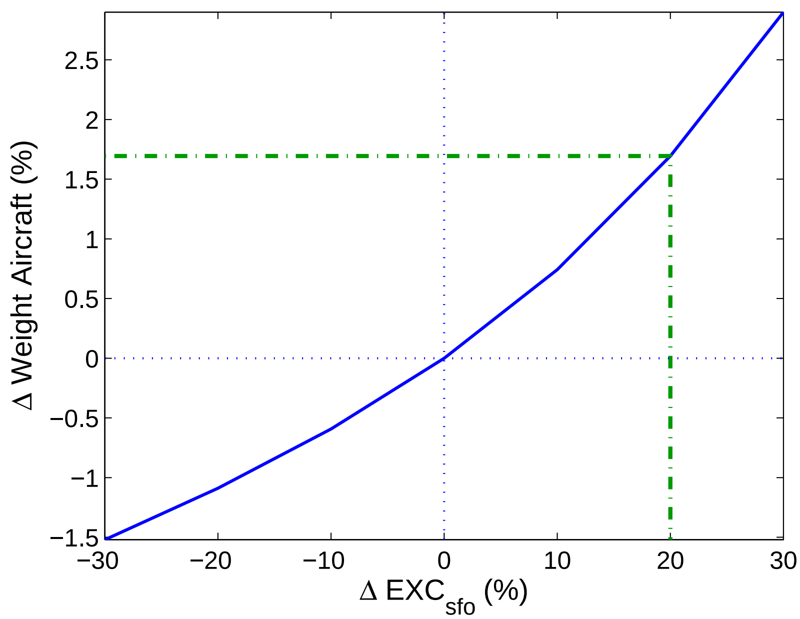


Figure 7. Δ Aircraft Average Weight OAK (%)

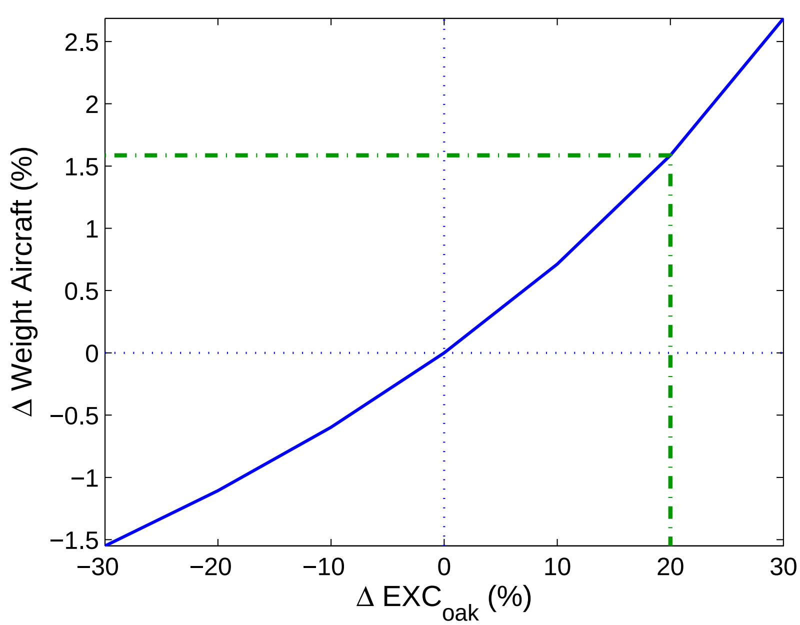


FIGURE 8. Δ Average Delay SFO (%)

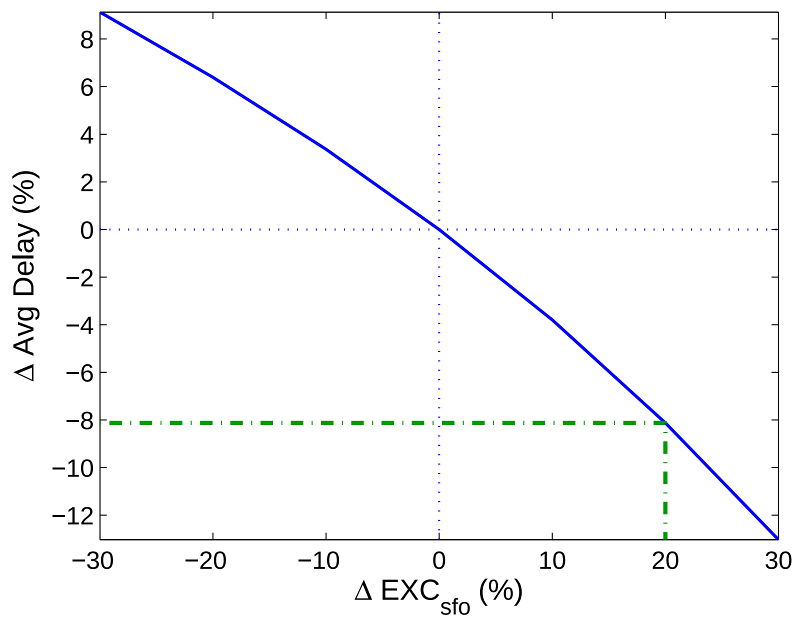


FIGURE 9. Δ Average Delay OAK (%)

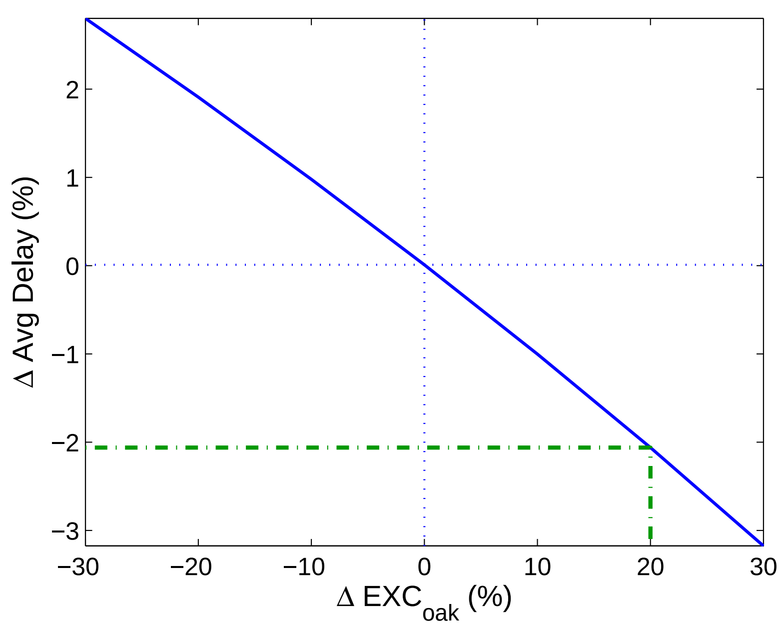


Figure 10. Landing Fees SFO ($) (OAK Methodology)

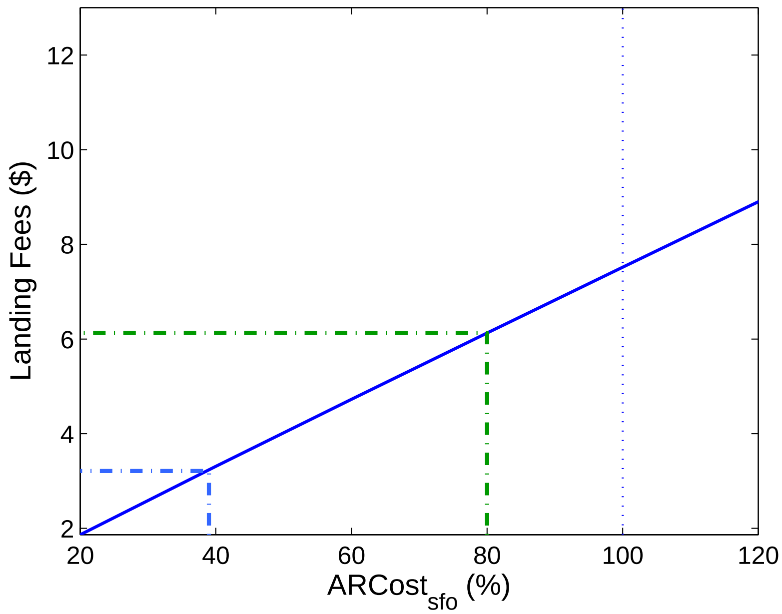


Figure 11. Δ Total Daily Flights SFO (%) (OAK Methodology)

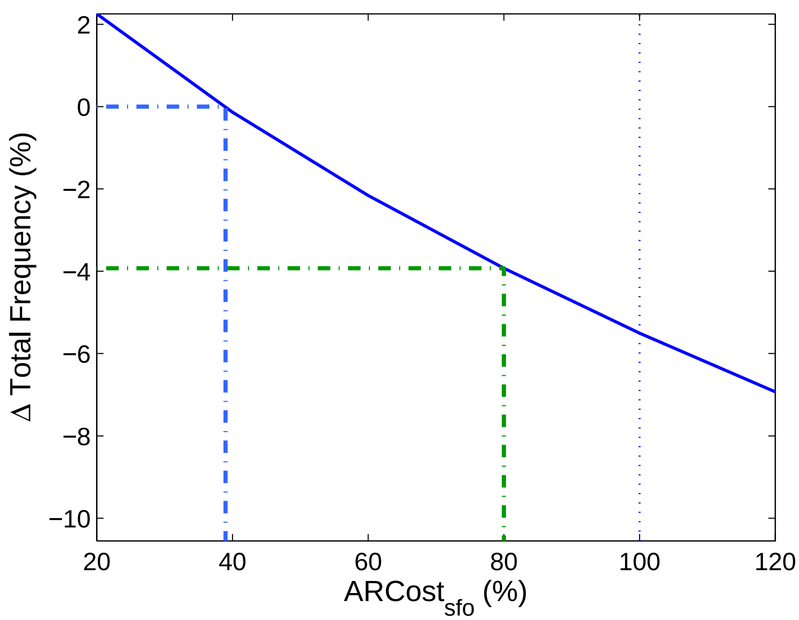


Figure 12. Δ Aircraft Average Weight SFO (%) (OAK Methodology)

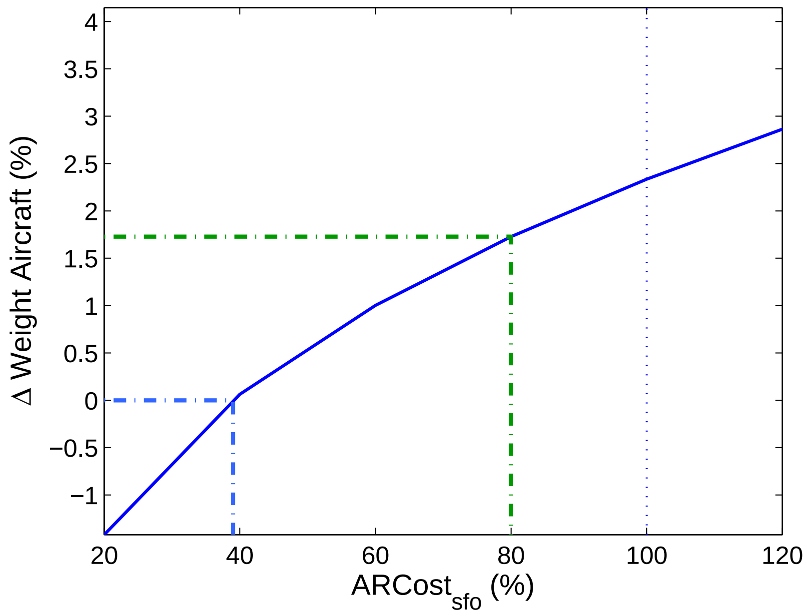
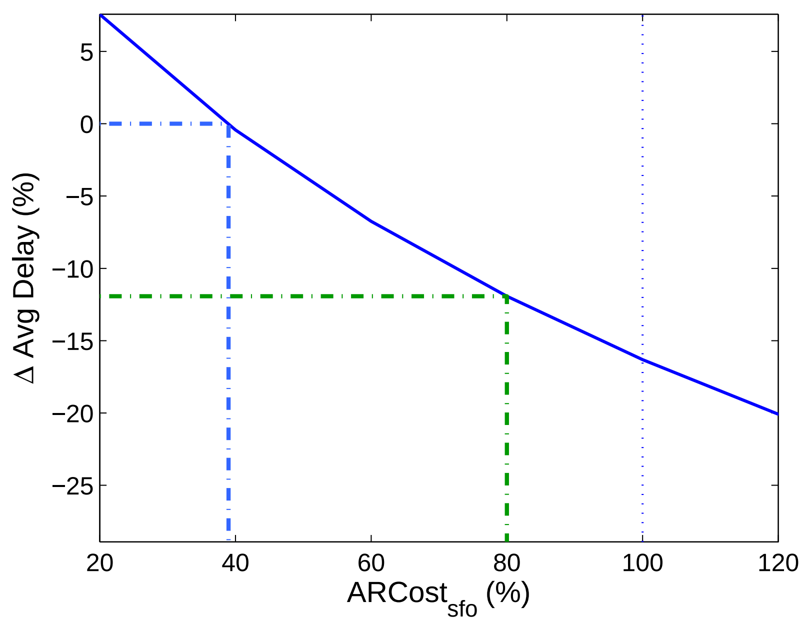


Figure 13. Δ Average Delay SFO (%) (OAK Methodology)



10. Appendix

10.1. **GMM discussion:** The GMM estimation procedure follows the nested fixed point approach suggested by Petrin (2002). Petrin extended the algorithm proposed by Berry, Levinsohn and Pakes (1995) (BLP) by combining data from different sources.

The model is estimated using a nonlinear GMM method. Three sets of moment conditions are used: one derived from the difference between the observed market shares and predicted market shares and the marginal cost moments (we call them BLP moments), and other moment conditions that add extra demand information using the 2006 MTC Survey.

10.1.1. *BLP Set of Moments:* First, as in BLP we want to match the predicted market shares $s_{jt}(\delta(\theta), \cdot; \theta)$ with those observed in the data s_{jt}:

$$s_{jt}(\delta(\theta), \cdot; \theta) = s_{jt} \ \ for \ \ j = \{0, \cdots, J_t\} \ \ and \ \ t = \{1, \cdots, T\} \tag{58}$$

Berry (1994) showed that under certain conditions, the previous equality holds for a unique value of the mean utility level (δ_{jt}). This property is useful because it will allow us to solve numerically for δ_{jt} by using a contraction mapping procedure. This is equivalent to computing the series

$$\delta_{jt}^{h+1} = \delta_{jt}^{h} + ln(s_{jt}) - ln(s_{jt}(\delta(\theta), \cdot; \theta)) \tag{59}$$

for $j = \{1, \cdots, J_t\}$, $t = \{1, \cdots, T\}$, and $h = \{0, \cdots, H\}$. Our approximation for δ_{jt} will be δ_{jt}^{H} such that $||\delta_{jt}^{H} - \delta_{jt}^{H-1}||$ is smaller than some tolerance (in our application 10^{-14}).

As usual in this type of model, we are not able to calculate analytically the integral associated with the market shares $s_{jt}(\delta(\theta), \cdot; \theta)$. So we simulate the market shares by taking g draws from the approximated distributions of distance to the airport (P_L), household income (P_Y) and the distribution of unobservables (P_ν). Hence, the simulated market shares are given by

$$s_{jt}(\delta(\theta), \cdot; \theta) = \frac{1}{g} \sum_{i=1}^{g} \frac{exp\left[\delta_{jt} + \mu_{ijt}\right]}{1 + \sum_{m \in J_t} exp\left[\delta_{mt} + \mu_{imt}\right]} \tag{60}$$

For the observable individual characteristics (distance and income) we use g random draws from the empirical distribution.[14] For the unobserved taste of travelers (ν's) we use Halton sequences rather than Monte Carlo simulations. This approach allows us to obtain a better approximation to the normal and lognormal distributions (Train (2009)).

From the mean utility equation (1) and given δ_{jt}, θ and product characteristics, we can derive the moment condition related to the unobserved-to-researcher characteristics of product j in market t (ξ_{jt}). That is

$$\xi_{jt} = \delta_{jt} - \alpha_0 p_{jt} - \alpha_f \hat{f}_{jt} - \alpha_d \hat{D}_{jt} - x_{jt}\beta \tag{61}$$

Using appropriate instruments (z^d) to control for price and frequency endogeneity, our moment condition can be written as

[14] In our application g=1200.

$$E[z_{jt}^{d}\xi_{jt}] = 0 \tag{62}$$

10.1.2. *Marginal Cost Moment:* The marginal cost moments are derived from (13) and equal

$$E[z_{jt}^{m}\omega_{jt}^{m}] = 0 \tag{63}$$

where ω_{jt}^{m} is the residual of the marginal cost equation and z_{jt}^{m} are cost instruments.

10.1.3. *Additional Demand Moments:* Following Petrin (2002), we extend the BLP model by adding moment conditions constructed using the 2006 MTC Airline Survey Data. Such survey data give us interesting demographic information about travelers conditional on the use of one of the two airports in the Bay Area. In particular, we use information about their distance to the airports and their household income.

Basically we construct moments that match the predicted average consumer demographics obtained from the BLP moments with the average consumer demographic characteristics from the MTC survey. In particular, the extra moment conditions will match the probability that a traveler i using one of the airports (a) comes from/goes to a specific county ($\mathcal{C}$) and belongs to a income group ($\mathcal{Y}$). That is

$$\eta_c(\mathcal{C}, a) = E\left[L_i \in \mathcal{C} | \{i \ \ uses \ \ airport \ \ a\}\right] \tag{64}$$

$$\eta_y(\mathcal{Y}, a) = E\left[y_i \in \mathcal{Y} | \{i \ \ uses \ \ airport \ \ a\}\right] \tag{65}$$

where

$$\mathcal{C} \in \{\text{S.Francisco-S.Mateo, Sta Clara, Alameda-C.Costa, Solano-Napa, Sonoma-Marin}\}$$

$$\mathcal{Y} \in \{<\$25\text{k}, \$25\text{k-}50\text{k}, \$50\text{k-}75\text{k}, \$75\text{k-}100\text{k}, \$100\text{k-}150\text{k}, \$150\text{k-}200\text{k}, >\$200\text{k}\}$$

$$a \in \{OAK, SFO\}$$

where L_i and y_i are the county of origin/destination and household income group of individual i. $\eta_c(\mathcal{C}, a)$ and $\eta_y(\mathcal{Y}, a)$ are the probabilities from the 2006 MTC survey (Tables (7) and (8)). The RHS expressions in (64) and (65) are the expected value predicted by our model and computed using the simulated market shares in (60). This extra condition applies for all income groups, counties and airports.

Since the probabilities $\eta_c(\mathcal{C}, a)$ and $\eta_y(\mathcal{Y}, a)$ conditional on each airport must sum to one, we do not include one of the options in the moment conditions. In particular, we do not include the county couple Solano-Napa nor household group with income less than \$25,000.

As we will see later, to minimize the GMM objective function, it is necessary to use the sample analogs of the previous moments. Since the MTC survey gives information conditional on using one of the airports, we will need to apply the definition of conditional probability to match the predicted probabilities with the MTC survey probabilities. Hence the sample analog of the additional information moments can be written as

$$\eta_c(\mathcal{C}, a) - \frac{\sum_{i=1}^{g} \sum_{r_{jt} \in \mathcal{R}_a} s_{ijt}(\delta(\theta), \cdot; \theta) M_t \mathbb{I}\{L_i \in \mathcal{C}\}}{\sum_{i=1}^{g} \sum_{r_{jt} \in \mathcal{R}_a} s_{ijt}(\delta(\theta), \cdot; \theta) M_t} \tag{66}$$

$$\eta_y(\mathcal{Y}, a) - \frac{\sum_{i=1}^{g} \sum_{r_{jt} \in \mathcal{R}_a} s_{ijt}(\delta(\theta), \cdot; \theta) M_t \mathbb{I}\{y_i \in \mathcal{Y}\}}{\sum_{i=1}^{g} \sum_{r_{jt} \in \mathcal{R}_a} s_{ijt}(\delta(\theta), \cdot; \theta) M_t} \tag{67}$$

where the second term of the expressions corresponds to the model predicted probabilities that a traveler i coming from county $\mathcal{C}$ and belonging to the household income group $\mathcal{Y}$ uses airport a respectively. $\mathbb{I}\{\cdot\}$ is a indicator function equal to one if the condition inside brackets holds and zero otherwise.

10.1.4. *GMM Estimation:* Let $\vartheta = \begin{bmatrix} \theta \\ \gamma_m \end{bmatrix}$ denote the set of parameters to be estimated using GMM (see equations (1) and (13)). Our optimal 2-step GMM estimators will be

$$\hat{\vartheta} = arg \min_{\vartheta} \hat{G}(\vartheta)' \Phi^{-1} \hat{G}(\vartheta) \tag{68}$$

where $\hat{G}(\vartheta)$ is the vector of sample analogs of the moment conditions noted above, and Φ is a consistent estimate of the variance-covariance matrix of the moments using the parameter estimates of the first step.

Detailed practical information about how to estimate this type of model can be found in Nevo (2000b). The algorithm is an iterative procedure characterized by first solving the contraction mapping (given initial values for $\dot{\theta}$ and $\dot{\gamma}_m$, solve for $\delta(\dot{\theta})$) and afterwards the GMM optimization problem (given $\delta(\dot{\theta})$, solve for $\ddot{\theta}$ and $\ddot{\gamma}_m$) and iterate again until convergence is reached. As suggested by Dube, Fox and Su (2012) we use tight tolerances: $1e^{-14}$ for the contraction and $1e^{-7}$ for the GMM function. We use the Knitro optimization package for Matlab and its interior/direct algorithm to solve the GMM minimization problem. In order to improve the performance of the minimization algorithm, we provide analytical gradients of the moment conditions. We repeat the algorithm for 50 different random starting points and choose the solution with the lowest objective function value.

As we previously pointed out, the product-specific costs (m_{jt}) and marginal costs (mc_{jt}) are unobserved by the econometrician. The GMM procedure also gives us estimates for both variables. They are part of the iteration process. Given an initial value for $\dot{\theta}$, we can use the F.O.C. with respect to fares (29) and estimates from the weight equation (10) to obtain $\dot{m}_{jt}$. The procedure is relatively simple, since once we use $\dot{\theta}$ and estimates for the weight of aircraft equation, the only unknown in the optimality condition for fares is $\dot{m}_{jt}$. Knowing the product-specific costs, marginal costs ($\dot{mc}_{jt}$) follow from computing the derivative of the variable costs with respect to the demand for product jt (12). New values of estimates for m_{jt} and mc_{jt} are obtained in each iteration until convergence is reached.

Finally, if we rely on the asymptotic properties of the estimates, then

$$J^{1/2}(\hat{\vartheta} - \vartheta^0) \sim \mathcal{N}(0, (\Gamma'\Phi^{-1}\Gamma)^{-1}) \tag{69}$$

where $\Gamma = E\left[\frac{\partial \hat{G}(\vartheta^0)}{\partial \vartheta}\right]$. We report standard errors using consistent estimates of Γ and Φ.

10.2. **Description of Data Sources:**

10.2.1. *Choice and Flight Characteristics Sources:*

- Airline Origin and Destination Survey (DB1B): a 10 % sample of all passengers travelling within the US, with detailed ticket information such as the operating and ticketing carrier of each coupon, the airports in which the passenger made a connection, if any, and the fare.
- DOT 100 database: has information on frequency of flights, total number of passengers and type of aircraft for all segments in the US. Given the aircraft type, we check the aircraft technical specifications in the manufacturers' website to obtain the maximum gross landing weight (MGLW) of the plane.
- Airline On-Time Performance Data: provides information about on-time and delay information for non-stop domestic flights by major air carriers.

10.2.2. *Airport Financial Reports:* We use financial details that airports report to the Federal Aviation Administration (FAA) to obtain information about the airfields and terminals' operating revenues. Such information is useful when we introduce landing fees and terminal rental rules in our model. The values of the landing fees and the methodology to determine landing fees and rental charges are obtained from the airports' board meeting proceedings.

10.2.3. *American Community Survey (ACS):.* The ACS is a household survey developed by the US Census Bureau to replace the long form of the decennial census program. The ACS is a large demographic survey collected throughout the year using mailed questionnaires, telephone interviews, and visits from Census Bureau field representatives to about 3 million household addresses annually. Starting in 2005, the ACS produced social, housing, and economic characteristics for demographic groups in areas with populations of 65,000 or more. It also produced estimates for smaller geographic areas, including census tracts and block groups. We use this database to construct the distribution of the household income and distance to the airports.

10.2.4. *2006 Airline Passenger Survey:* This survey gives detailed information about travelers using San Francisco International Airport (SFO) and Metropolitan Oakland International Airport (OAK). The primary purpose of the survey is not to analyze behavior of travelers with respect to the choice of carrier products but to permit the analysis of alternative policies with regard to airport access. The survey contains: household income, location of the traveler in the Bay Area, destination, airport of origin/destination, carrier, transportation access to the airport. The survey does not provide information about prices, nor the chosen itinerary. However, its rich demographic information can complement the DB1B and ACS databases. We will use this survey to construct the distributions of household income and distance conditional on the airport travelers use.

10.3. **Data Construction:** We restrict our attention to data from the third quarter of 2006. Following Urdanoz and Sampaio (2011), we only consider products with the following characteristics: (1) round trip itineraries starting and ending at the same airport, thus excluding one-way trips and "open jaws"; (2) products with up to three coupons per direction; (3) with at most one ticketing company; (4) with at most two operating carriers; (5) that are not operated by a foreign carrier; (6) that do not involve a coupon operated by an unknown carrier; (7) that do not involve a ground trip; (8) that do not have an airport coded as NYC since we cannot identify which of the 4 airports in New York Metropolitan area was used and finally (9) with fares between \$50 and \$3000.

We follow the approach proposed by Mayer and Sinai (2003) to construct the delay variable. Rather than defining flight delays as the percentage of flights arriving more than 15 minutes of scheduled arrival time, we use the difference between actual time and the minimum feasible travel time. Then we define the average delay at each airport as the average of the aforementioned difference for all flights arriving at the airport. Such an approach avoids the use of scheduled arrival times that may be subject to airline manipulation (padding) to increase their on-time performance. To obtain the average delay of each product, we compute the mean delay of each connecting airport.[15]

As we previously noted, total landing fees paid by each carrier depend on the weight of aircraft used on the trip segment arriving in the Bay Area. However, carriers may use several types of airplanes for the same segment. To determine the value of the maximum gross landing weight (MGLW)(the variable $weight_{rc}$ of our model), we compute the average MGLW of the different aircraft weighted by the total number of operations they performed in the quarter.

Regarding the 2006 MTC data set, we restrict our attention to observations where income and travelers' location information is provided. Then we construct the empirical distribution of those two variables conditional on each of the airports in the Bay Area.

We use the ACS survey to obtain demographic information about income and traveler locations. We do not have the exact address of respondents in the ACS survey, but only the area where they come from. The ACS classifies areas according to census tracts. In the particular case of the Bay Area, the ACS partitions the area into 1099 zones. We compute the distance from the traveler's location to airports as the Euclidean distance between the airports and the centroid of the census tract where the traveler comes from.[16] Finally, our model uses county fixed effects, so we group census tracts according to their respective county. We create the variable "traveler income (y_i)" by taking random draws of the income distributions provided by the ACS for each of the metropolitan statistical areas where the airports of origin are located.

As noted above, we use products from the 3rd quarter of 2006 to estimate the model. However, landing fees and rental charges are computed using the fiscal year.[17] Hence, to

[15]Other alternatives may be considered, such as the difference between the scheduled and the actual flight durations.

[16]Alternative measures can be used. For instance rather than using the Euclidean distance we can compute the topographic distance taking into account roads and access to airports. Another option would be to compute the time necessary to reach the airport

[17]The fiscal years for OAK and SFO start in July 1 and finish June 30. For simplicity we assume that the fiscal year and the natural year are the same.

construct some variables, we use yearly data rather than just one quarter. First, using the T-100 data set and manufactures's website we construct the total weight of aircraft landing at airports during 2006 ($TWeight_a$ with $a \in \{oak, sfo\}$), which is used as denominator in the landing fees formula (see (15) and (21)). Similarly, we use the total number of enplaned passengers, the Groundside Cost Center Revenues ($GRev_{sfo}$), and the Terminal Cost Center revenues (OR_a for $a \in \{oak, sfo\}$) for all of 2006 to determine the parameter values associated with the terminal ($\psi_{terminal,a}$ for $a \in \{oak, sfo\}$) and the groundside revenues ($\psi_{ground,sfo}$).

We do not have information about the terminal areas leased by airlines. Instead, we use the number of operations performed at each airport during 2006 as a proxy for the usage of terminals ($Usage_{c,a}$).