CONGESTION AT AIRPORTS: IMPLEMENTING A TWO-PART LANDING FEE AT SAN FRANCISCO INTERNATIONAL AIRPORT

MIGUEL-ANGEL ALCOBENDAS*

Toulouse School of Economics
University of California-Irvine, Department of Economics
April 3, 2014

Abstract. We discuss the modification in 2008 of the U.S. policy that regulates airport rates and charges. Under the new regulatory framework, airports can charge a two-part landing fee to relieve congestion. Such a landing fee scheme consists of the standard aircraft weight-based charge plus an operation charge applied in peak hours. The question is relevant, since flight delay is a serious problem in the U.S. economy and, so far, no airport has put into practice this type of charging scheme. We develop and estimate a structural model to investigate the consequences of implementing such a two-part landing fee at San Francisco International Airport. Our simulations suggest the higher the operation charge, the lower are the number of flights arriving during peak hours and the bigger are the sizes of aircraft. As a result, the level of congestion and total demand at San Francisco International Airport decrease.

Our model captures important characteristics of the airline industry that most of the previous literature has neglected: endogeneity of airport charges with respect to decisions of travelers and carriers, correlation across markets, and two decision variables of airlines (ticket price and flight frequency).

1. Introduction

Airport congestion has been object of study since air-traffic growth has led a number of airports to operate at maximum capacity. There exists an important debate about how to manage and reduce the resulting flight delays. The U.S. Department of Transportation (DoT)

*E-mail: maalcobendas@gmail.com.

I wish to thank Professors 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 Jiawei Chen, Achim Czerny, Pierre Dubois, Ricardo Flores, Christian Hellwig, Laura Lasio, Dan Luo, Pierre Picard, Mar Reguant, Kevin Roth, Paul Scott, Kenneth Small, Damian Tago, and Anming Zhang for their insightful comments. I also thank participants at the TSE applied micro, UCI Transportation, and the University of Barcelona workshops.

is aware of this problem. For that reason, in 2008 the DoT introduced three amendments to the 1996 "Policy Regarding the Establishment of Airport Rates and Charges", a statement that sets the standards applicable to fees imposed for aeronautical use of airports. These amendments provide airports with a new set of tools to reduce flight delays without being in conflict with the regulatory framework. With these modifications, the DoT explicitly authorizes airports to impose a two-part landing fee scheme, consisting of both a congestion-based flight operation charge and an aircraft weight-based charge, in lieu of the standard weight-based charge. The current landing fee scheme is based on the weight of aircraft and bears no relationship to the level of airport congestion. Hence, airport operators cannot use it to give airlines incentives to reduce the number of scheduled flights during peak periods.[1] The new policy lets airports charge different prices depending on their level of congestion. According to the DoT (FR 73, No 135, July 14, 2008), this tool would let airports divert traffic to less congested hours, while increasing the size of the aircraft.

The objective of this paper is to analyze the consequences of implementing such a two-part landing fee at the San Francisco International Airport (SFO). Using data from the third quarter of 2006, we study the equilibrium behavior of travelers and carriers when we introduce a peak-hour charge to supplement the current weight-based landing fee scheme. In particular, we investigate how the congestion charge modifies the number of scheduled landings and the sizes of aircraft, leading to changes in flight delays at SFO. To conduct the analysis, we develop and estimate a structural model of air-travel demand and carrier supply.

So far, no U.S. airport has put into practice this type of charging scheme. From comments previous to the implementation of the amendments (FR-73, No12, January 17, 2008), it is clear that airport managers and airlines disagree about the desirability of this new regulatory framework. While airport operators welcome these modifications, airlines argue that the amendments allow airports to charge unreasonable and discriminatory fees, and they are

[1]The new regulation also permits airports to include in the peak-hours charge a portion of the costs of an airfield project under construction. Previously, only the costs of fully operational facilities could be taken into account. The last amendment lets peak-hour landing fees include airfield costs of other underutilized airports owned by the same proprietor, with the objective of diverting operations from congested to underutilized airports. For instance, Los Angeles International (LAX) and Ontario International (ONT) Airports are owned by Los Angeles World Airports. With the new regulation, LAX landing fees could include a portion of the costs of operating ONT to relieve LAX congestion and promote the use of ONT.

not compatible with the federal law.[2] These differences in preferences may explain why a new pricing methodology has not yet been implemented. Moreover the policy statement is just a guideline and specific rules are not imposed. Thus, each airport directly negotiates with carriers over the charging scheme for using its infrastructure. Airports and airlines usually set contracts called "Airport-Carrier Lease and Use Agreements" that carriers may be reluctant to change.[3] In addition, the new guidelines arrived at the same time as the economic crisis reduced air-travel demand, allowing airports to postpone the decision of whether to impose this new mechanism. However, demand is reaching pre-crisis levels, leading policymakers, carriers and airport operators to discuss again the necessity of introducing measures to control congestion.

There exist other alternatives to reduce flight delays at airports. For example, airports can relieve congestion by improving their infrastructure to accommodate more flight operations (e.g. constructing new runways, new terminals, or air-traffic control technology upgrades). However, this solution is not always feasible due to, for instance, budget limitations, space constraints, noise and environmental regulations, or opposition of cities surrounding the airport.[4] Moreover, this type of measure needs time to be implemented and it would not solve current problems. Another solution is to implement mechanisms to rationalize the use of airports by imposing administrative rules to constrain the number of operations per hour (e.g. slot constraints as at John F. Kennedy International or Ronald Reagan National airports). As Borenstein (1988) points out, such measures face the problem of how to allocate the slots and how to avoid creating barriers to entry. Finally and closely related with the DoT's amendments, policies based on the concept of "optimal congestion pricing" can be used to determine the price that an aircraft should pay for operating at a congested airport. With this price mechanism, the landing fees paid by airlines vary with the level of airport congestion. How to optimally compute and implement this type of charge is still under discussion by scholars and policymakers.

[2]On behalf of the U.S. airlines, the Air-Transport Association of America (ATA) appealed the DoT amendments and claimed that they are not legal under the federal law. However, the United States Court of Appeals for the District of Columbia Circuit denied the appeal.

[3]The Airport Cooperative Research Program (ACRP) defines Airport-Airline Use and Lease Agreement as an *"agreement that specifies the financial obligations, terms of use, and other responsibilities that each party assumes with respect to the use of the airport's facilities. The Agreement sets the commencement and termination dates for the use of airport facilities, identifies the facilities to be used and the degree of use, the rate-making methodology, and defines the approved uses of the facility"*

[4]The cost of constructing a new runway or terminal is very high. For example, the third runway at Seattle-Tacoma International Airport cost \$1.1 billion, and the London-Heathrow Terminal 5 \$6.42 billion.

Several reasons explain the convenience of choosing the SFO airport. First, SFO meets the requirements imposed by the Federal Aviation Administration (FAA) to be considered as a congested airport, thus being eligible to implement a two-part tariff charge scheme.[5] Second, SFO has one of the lowest-performing arrival rates of the national hubs due to a combination of foggy weather conditions and heavy airline traffic during peak hours. If we look at Figure 1, the dash-dotted line represents the average delay of arriving flights during the day.[6] There are two peaks, one in the morning and another in the evening. If we compare this line with the distribution of arrivals during the day (continuous line in Figure 1), we observe a positive correlation between the number of arrivals and delays.[7] The existence of peaks and valleys through the day may justify the use of a congestion charge, since a higher landing fee in peak hours may affect the distribution of flight frequency during the day. Carriers may decide to eliminate or reschedule the less profitable flights to less congested hours. Last, our empirical application uses several data-sets that are available for different U.S. airports. Nevertheless, we also use a unique survey done in 2006 by the Metropolitan Transportation Commission of the San Francisco Bay (MTC), which provides important demographic information on travelers using SFO that is very useful for the estimation of model parameters.

As we will see later, there are substantial differences between our study and previous work addressing the airport congestion problem. Zhang and Czerny (2012) present an excellent review of recent research about this topic. While there exists extensive theoretical work, empirical papers are scarce, mainly because congestion charges are not currently levied. Most of the existing papers use data to simulate the consequences of imposing congestion fees. They usually rely on previous literature developed to study road congestion. Daniel

[5]According to the "Policy Regarding the Establishment of Airport Rates and Charges", the U.S. Department of Transportation (DoT) *"considers a currently congested airport to be:*

(1) An airport at which the number of operating delays is one per cent or more of the total operating delays at the 55 airports with the highest number of operating delays; or

(2) An airport identified as congested by the Federal Aviation Administration listed in table 1 of the FAA's Airport Capacity Benchmark Report 2004, or the most recent version of the Airport Capacity Benchmark Report."

[6]We define delay as the difference between the shortest observed travel time on a given nonstop route and the actual travel time of a particular flight (Mayer and Sinai (2003)).

[7]We approximate the average delays and distribution of arrivals during the day using a Nadaraya-Watson Kernel with the Silverman Rule-of-Thumb bandwidth (h_n), $h_n = 0.9(min\{\hat{\sigma}, IQR/1.34\})N^{-1/5}$, where IQR is the interquartile range (the difference between the 75th and 25th percentile), N is the sample size, and $\hat{\sigma}$ is the standard deviation of the sample.

(1995, 2001) presents a stochastic-bottleneck model to simulate the consequences of congestion pricing at the Minneapolis-Saint Paul airport. Johnson and Savage (2006) and Ashley and Savage (2008) also apply a bottleneck model to simulate the effects on Chicago O'Hare International airport. Those models impose a time-varying congestion fee equivalent to the congestion cost that each aircraft imposes on all others.

Other related empirical work focuses on testing if a carrier with market power at an airport internalizes the congestion that each flight imposes on the other flights it operates. The question is relevant because it has big implications for the design of the optimal congestion fee. In spite of that, the answer is not clear. While the aforementioned papers by Daniel (1995, 2001) and Daniel and Harback (2008) claim that airlines do not internalize delays, Brueckner (2002) and Mayer and Sinai (2003) find evidence that they do. Morrison and Winston (2007) compare both approaches and evaluate the welfare loss from ignoring internalization in computing congestion charges, finding that the loss is not large.

Our study is different from the aforementioned work. It is the first one that quantifies the impact of establishing a two-part landing fee. Furthermore, we use a structural model where air travel demand and carrier behavior are specified. Earlier work based on bottleneck models does not explicitly model the preferences of travelers, and this omission is important since the decisions of carriers (fares, frequency of flights, and size of aircraft) are endogenously determined by travelers' demand. Another important contribution is the use of game theoretical tools to estimate the flight costs, while previous work relies on reports or other papers. Our model is easier to implement, since the two-part landing fee is constructed by using the current scheme applied by SFO and adding a congestion charge during peak hours. Since the congestion charge is fixed, airlines can anticipate the consequences of their decisions. This is not the case for bottleneck models, because fees depend on delays that an aircraft imposes on all others, and this quantity depends on several real time factors, including the time of the day or weather conditions. This type of endogeneity makes it difficult for airlines to decide on the frequency and the size of their aircraft, since these choices must be made well in advance. Finally, while existing related work treats non-congestion charges as exogenous, we let them be endogenously determined by traveler and carrier behavior.

Some other contributions are made from a methodological point of view: first, carriers face two decision variables (ticket prices and frequency of flights). Most of the previous literature in applied industrial organization focuses on a single decision variable (price). Second, our rich model specification also captures correlation across markets. As we will see in detail later on, passengers from different markets share aircraft reaching SFO, affecting the decision of carriers with respect to fares and frequency of their flights. Congestion at

SFO is another source of dependence across markets. Since changes in the frequency of one product influence flight delays at SFO, all products offered during the same period will be affected even if they do not belong to the same market. Last, we use spatially-based consumer characteristics to capture heterogeneity of travelers.

In our model there are three types of agents: travelers, airlines, and the SFO airport operator. Travelers are heterogeneous individuals with different locations (origins or destinations) in the San Francisco Bay who have different tastes. They choose the product that gives them the highest utility. Carriers choose fares, the frequency and the schedule of their flights taking into account the level of congestion at SFO during the day. Their decision will be the solution of a profit maximization problem, which is solved sequentially: carriers first decide on the flight frequency for peak and off-peak hours, and afterwards the price of tickets. On the other hand, the SFO airport is compensated by carriers for using its airfield (landing fee) and terminals (rental rate). These charges and the mechanism to determine them are set by local airport authorities with the objective to achieve financial self-sufficiency of the airport. They are computed according to a methodology that depends on traffic, revenues, and costs generated at the airport. While the current landing fees are established according to the weight of aircraft, the rental rate for using the terminals depends on the surface leased by each airline. As we will see, both charges are endogenous since they depend on travel demand, flight frequency, and ticket prices. Airlines take into account such endogeneity in their decision problem.

The model is solved in two stages: first, we estimate the model to characterize the preferences of travelers and carriers' behavior. Second, we modify the landing fee scheme to accommodate a congestion charge during peak hours. Then, we simulate changes in the equilibrium behavior of players and analyze variations in air-travel demand, fares, frequency of flights, size of aircraft, and delays at SFO.

As expected, our estimates suggest that travelers prefer to arrive at SFO in the morning or in the evening. They also prefer direct flights and more frequency. On the other hand, they dislike delays and their utility decreases with distance from their location (origin or final trip destination in the Bay Area) to SFO. If we look at the consequences of adding a congestion charge to the current weight-based landing fee scheme, the higher the charge, the lower is the number of flights landing at SFO during peak hours. As a consequence, the level of airport congestion decreases. In addition, the higher the charge, the higher are the incentives of airlines to increase the size of their planes. The reduction in the number of flights is also accompanied by an increase in fares, leading to a diminution in air-travel demand during peak-hours. Implementing a congestion charge does not only affect flights in the peak but also in off-peak hours. Part of the lost air travel demand in peak hours is diverted to off-peak hours, increasing the frequency of flights operating during periods of low congestion. However, total demand for SFO decreases. As a result, the weight-based component of the landing fee goes up, increasing the cost of operating a flight. For instance, if SFO imposes a congestion charge of 2,000 dollars per arrival during peak hours, the total number of flights reaching the airport during congested hours decreases by 3.86%, the average delay falls by 12.16%, the weight of aircraft increases by 9.10%, fares grow 0.30%, and demand for peak hours products decreases by 4.14%. A 2,000 dollars congestion charge also increases demand during off-peak periods around 1%, but total demand for SFO decreases by 2.66%, increasing the weight-based landing fee 2.43%.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 derives the optimality conditions for carriers. Section 4 explains the application and the data. Section 5 outlines the estimation methodology. Section 6 discusses the estimation results. Section 7 analyzes the consequences of implementing a peak-hour congestion charge. Section 8 concludes.

2. Model

The model determines the purchasing decisions of travelers as a function of their attributes and the characteristics of the products offered by carriers. At the same time, carriers' pricing and frequency decisions are affected by the landing fees and rental charges levied by SFO.

In our model, markets are defined as a round trip directional city-pair. For instance, a market could be the directional pair San Francisco International Airport - Miami International Airport where San Francisco is the origin and Miami is the destination. This market is different from Miami International Airport - San Francisco International Airport, where Miami 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: fare, distance, frequency of flights, ticketing carrier, frequency of flights, dummy for direct flights, dummy for slot constrained airports, delays, and scheduled arrival time at SFO (peak or off-peak hours).

2.1. **Demand:** We assume that the demand for a ticket follows a random coefficient logit representation. Such an approach lets us introduce heterogeneity in the demand for tickets. Travelers have different tastes with respect to product characteristics and they are also heterogeneous with respect to their incomes and their locations in the Bay Area.

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 product j in market t is given by

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

where

- $\hat{I}_{jt}^{peak}$ is a dummy equal to one if the product is operating during peak hours and zero otherwise. If the scheduled arrival of the product is in the morning (between 9:00AM and 12:30PM) or in the evening (between 6:30PM and 10:30PM), the product is assumed to be offered during peak hours (see Figure 1). In such a case, $I_{jt}^{peak} = 1$, and it equals zero otherwise.
- p_{jt} is the fare of the product.
- y_i is household income, with probability distribution P_Y.
- $\hat{f}_{jt}$ corresponds to the flight frequency of product jt. It depends on the daily frequency of flights for connecting and destination airports used by the product.[8]
- $\hat{D}_{jt}$ is the delay of product jt. It depends on the level of congestion at connecting and final airports used by the product.[9]
- x_{jt} is a vector of travel characteristics for product jt. Such characteristics are observed by the econometrician: a constant, ticketing carrier, flight distance, a dummy for direct flights, and a dummy for airports with slot constraints.
- ξ_{jt} is the unobserved-to-researcher characteristics of product j in market t. An increase in ξ_{jt} makes the product jt more attractive to all consumers.
- $d(L_i)$ determines the distance of individual i to SFO airport. 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 SFO, $d(L_i)$ 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, $d(L_i)$ is interpreted as the distance from the airport to his final destination (for instance, hotel or office).

[8]More details are provided in the data section.

[9]More details are provided in the data section.

- 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. We allow interactions between 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 arriving at SFO in peak hours (α_{peak}), for daily frequency (α_f, σ^f), for delays (α_d, σ^d), for other product characteristics β, for distance to SFO (λ), and the parameter σ^0 associated with the constant. 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. Those 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}. Thus, the use of appropriate instrumental variables will be necessary to avoid inconsistent estimates.

We distinguish the mean utility level of product j in market t (δ_{jt}) from the traveler-specific deviation ($\mu_{ijt} + \epsilon_{ijt}$):

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

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

Hence, the utility that individual i obtains from product j in market t is equal to

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

Let u_{i0t} denote the utility from the outside good of not flying from/to SFO. This term adds flexibility to the discrete choice model since travelers are not obliged to use SFO. They can decide to fly from alternate airports in the Bay area such as Oakland International or Mineta San Jose International airports, use another transportation mode as a car or bus, or not travel at all. This utility is random and is written as

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

Once the demand function is specified, the estimation of parameters depends on the capacity of our model to predict the product market shares. Following Nevo (2001), if we assume that the idiosyncratic unobservable component of utility ϵ_{ijt} is i.i.d. type I extreme value, the probability that a traveler i chooses alternative j in market t is

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

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

Assuming that the distributions of ν_i, ϵ, L_i, and y_i are independent, the model-predicted market share of product $j \in J_t$ is given by

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

where $P_\nu(\cdot)$ is the distribution of the unobservables, $P_L(\cdot)$ is the distribution of the locations of travelers in the Bay Area, and $P_Y(\cdot)$ is the distribution of household incomes.[10]

2.2. **Carriers and Airport Charges:** Airlines are assumed to be profit maximizing firms with respect to ticket prices, and the frequency and schedule of their flights during the day. They may operate at SFO during peak and/or off-peak hours. Moreover, they might offer differentiated products within a market. At the same time, profits depend on the landing fees and the terminal rental rate levied by the airport.

The equilibrium concept in the model is the subgame perfect Nash Equilibrium. The game has two stages: in the first stage, airlines simultaneously decide on the flight frequency (f) of the last trip segment arriving at SFO in each period (peak, off-peak). In the second stage, firms decide on fares (p). As we will see later, the size of aircraft will be the result of the interaction between the optimal decisions regarding fares and frequencies. It is also important to note that, in our model, airlines only decide the flight frequency of the trip

[10]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.

segment arriving at SFO. This spoke-route is directly affected by the charges levied by SFO. However, products may be composed of several segments, and we hence implicitly assume that the frequency decisions for trip segments are independent of each other. We use the optimality conditions with respect to fares and flight frequencies to estimate the parameters of the model and analyze the effects of imposing a congestion charge in peak hours.

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 SFO airport as an endpoint. Individual spokes are denoted by r, and l denotes the time period of the flight arrival ($l = \mathcal{H}$ for peak hours and $l = \mathcal{L}$ for off-peak hours). The optimal decision of airline c will be the solution of the following profit maximization problem:

$$(8)\ \max_f \max_p \Pi_c = \max_f \max_p \Bigg[\sum_{t\in T}\sum_{j\in\mathcal{J}_{ft}} ([p_{jt} - m_{jt}]s_{jt}(p, f, \delta(\theta); \theta) \times M_t) -$$

$$- \underbrace{\sum_{r\in\Omega_c}\sum_{l\in\{\mathcal{L},\mathcal{H}\}} \tilde{f}_{lrc}(FCost_{lrc} + \beta^d D_l(f) + \underbrace{fees(s, p, f) \times weight_{lrc}(s, p, f) + \rho_l}_{\text{Two-Part Landing Fee}})}_{\text{Total Operating Flight Cost}} -$$

$$- RC_c(s, p, f) - F_c\Bigg]$$

where Π_c corresponds to profits of airline c (in our application, third quarter of 2006). m_{jt} represents the product-specific cost for product j in market t. $s_{jt}(\cdot)$ is the previously defined market share. M_t is the total population that may be interested in traveling in market t, and equals the geometric mean population of the origin and destination metropolitan statistical areas.

The term "Total Operating Flight Cost" captures the total airline cost for operating flights landing at SFO. This term depends on the number of flights ($\tilde{f}_{lrc}$) that the airline operates on each spoke r during peak ($l = \mathcal{H}$) and off-peak ($l = \mathcal{L}$) hours, the weight-based landing fee ($fees$), the congestion charge (ρ_l), the weight of aircraft ($weight_{lrc}$), the average delay incurred at the airport for peak and off-peak hours (D_l), the monetary value of one minute delay (β^d), and the undelayed flight cost component ($FCost_{lrc}$). RC_c is the total rental cost for carrier c in using the terminals of SFO. Finally, F_c is the total fixed cost incurred by the airline operating at the airport. $fees$, $weight_{lrc}$, and RC_c are endogenously determined, depending on market shares, fares and flight frequencies. On the other hand, D_l only depends on the number of flights operating during period l.

Note that we make a distinction between product delay in the utility function ($\hat{D}_{jt}$) and airport delays in the profit function (D_l). While $\hat{D}_{jt}$ takes into account delays at each of connecting and destination airports used by product jt, D_l refers to the average delay at SFO in period l.[11] Similarly, we distinguish between the daily frequency of product jt ($\hat{f}_{jt}$), the daily frequency of flights of a carrier on one particular spoke operating in period l (f_{lrc}), and the total number of operations in the quarter for the carrier on the spoke r and period l ($\tilde{f}_{lrc}$). While $\hat{f}_{jt}$ takes into account the frequencies for each of the trip segments of airports used by product jt, f_{lrc} only considers the carriers' flight frequency on spokes arriving at SFO. Finally, we assume that f_{lrc} is the same for all days of the quarter. Consequently, $\tilde{f}_{lrc}$ is equal to f_{lrc} times the number of days of the quarter (92 days).

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 trip segment reaching SFO, even if they belong to different markets. Imagine, for instance, travelers flying from New York (JFK) to SFO via Boston (BOS) and arriving at SFO during peak hours. They may share the same aircraft in their last trip segment with travelers flying non-stop from BOS to SFO and also arriving during peak hours. Moreover, since markets are defined as round trip directional city-pairs, passengers may share the aircraft and belong to different markets even if they fly non-stop. That is, people traveling non-stop between any U.S. city and the Bay Area during the same congested hours 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 that markets are independent does not hold in our model. The optimality conditions will capture this dependence across markets.

The two major reasons for delays are weather conditions and the number of landings relative to airport capacity. In our case, we assume a deterministic relationship between the airport delay and the total number of daily flights arriving at the airport. Following Morrison and Winston (1989, 2007) the delay function is given by

$$D_l = \begin{cases} exp(\omega^d_{\mathcal{L}} \bar{f}_{\mathcal{L}}) \;\; if \;\; l = \mathcal{L} \\ \\ exp(\omega^d_{\mathcal{H}} \bar{f}_{\mathcal{H}}) \;\; if \;\; l = \mathcal{H} \end{cases} \tag{9}$$

[11] In our application, the average delay during peak hours is 28 minutes and 45 seconds ($D_{\mathcal{H}} = 28.75$) and 21 minutes and 40 seconds when the airport is operating during off-peak hours ($D_{\mathcal{L}} = 21.67$).

where $\omega_{\mathcal{L}}^d$ and $\omega_{\mathcal{H}}^d$ are the congestion parameters, and $\bar{f}_{\mathcal{L}}$ and $\bar{f}_{\mathcal{H}}$ are the total number of daily operations in off-peak and peak hours respectively. All three variables appearing in the expression are the same for all flights landing during the same period. While D_l, $\bar{f}_{\mathcal{L}}$ and $\bar{f}_{\mathcal{H}}$ are observed from the data, $\omega_{\mathcal{L}}^d$ and $\omega_{\mathcal{H}}^d$ 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. It also allows SFO to accommodate any number of flights in each period with exploding airport delays. By construction, average delay is a source of dependence across markets. Since changes in the frequency of one product affect the average congestion at the airport, all products offered during the same period (peak or off-peak hours) will be affected even if they do not belong to the same market. As we will see in detail when solving the maximization problem, this effect is considered by carriers when they decide the frequency of their own flights.

Landing fees have two components, the total weight-based fee ($fees \times weight_{lrc}$) and the congestion charge per operation (ρ_l for $l \in \{\mathcal{L}, \mathcal{H}\}$). The weight-based landing fee ($fees$) is the charge that airlines pay for each 1,000 lbs of maximum gross landing weight (MGLW) for each arriving aircraft ($weight_{lrc}$). Remember that each product is a round trip travel that may have several connections. However, for simplicity the only landing fees that we consider are the ones levied by SFO. The second component of the landing fee is the operation charge for peak and off-peak hours (ρ_l). Note that SFO does not currently charge any operation fee ($\rho_l = 0$ for $l \in \{\mathcal{L}, \mathcal{H}\}$). Thus, current landing fees bear no relationship to congestion. The objective of this paper is to simulate the effects of applying an operation charge in peak hours ($\rho_{\mathcal{H}} > \rho_{\mathcal{L}} = 0$).

The term $weight_{lrc}$ is related to the type of airplane that carriers use on the spoke-route r in congestion period l. We assume that all aircraft used by a carrier on a spoke and period have the same characteristics. Moreover, we also assume that airlines decide the type of plane according to the total daily demand for the segment and period (TDD_{lrc}), daily frequency (f_{lrc}), spoke distance ($dist_r$), airline identity ($carr_{lrc}$), and finally a dummy for peak hours (I_l^{peak}). Therefore,

$$weight_{lrc} = \tau_0 + \tau_1 TDD_{lrc} + \tau_2 f_{lrc} + \tau_3 dist_r + \tau_4 carr_c + \tau_5 I_l^{peak} + \epsilon_{lrc}^w \tag{10}$$

where ϵ_{lrc}^w is the disturbance term. Total demand (TDD_{lrc}) and frequency (f_{lrc}) 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_{lrc}) 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 considered in our model specification. Good examples are one-way or travelers connecting at SFO. Hence,

$$(11) \qquad TDD_{lrc} = \sum_{t \in T} \sum_{\substack{\{kt | r_{kt} = r, l_{kt} = l, \\ , k \in \mathcal{J}_{ct}\}}} \frac{s_{kt} * M_t}{92} + ResTDD_{lrc}$$

where the first term on the right hand side (RHS) captures the demand for products considered in our specification that use the spoke-route r, period l, and carrier c. r_{kt} denotes the last spoke used by product kt to reach SFO, and l_{kt} denotes the period when product kt is scheduled to arrive. As we previously noted, TDD_{lrc} is the daily demand for spoke r, period l, and carrier c. On the other hand, $s_{kt} * M_t$ is the demand for the whole quarter. If we assume that the 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. $ResTDD_{lrc}$ corresponds to travelers that do not use any of the products of the model but still use the same airline, spoke r, and arrive during period l. This term is assumed to be independent of the demand for products considered in our model.

The terms $m_{jt} \times s_{jt} \times M_{jt}$, *"Total Operating Flight Cost"*, and the rental charge of terminals (RC_c) in (8) are considered as 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 (8),

$$(12) \qquad mc_{jt} = m_{jt} + \sum_{r \in \Omega_c} \sum_{l \in \{\mathcal{L}, \mathcal{H}\}} \tilde{f}_{lrc} \left(\frac{\partial fees}{\partial q_{jt}} \times weight_{lrc} + fees \times \frac{\partial weight_{lrc}}{\partial q_{jt}} \right) + \frac{\partial RC_c}{\partial q_{jt}}$$

We note that the marginal cost (mc_{jt}) does not depend on the undelayed flight cost ($FCost_{lrc}$) or the congestion operation charge (ρ_l).

We assume that both the product marginal cost (mc_{jt}) and the undelayed flight cost ($FCost_{lrc}$) linearly depend on a vector of exogenous costs shifters (w^m_{jt}, w^f_{lrc}) via the respective parameters (γ_m, γ_f), and a random term that captures unobserved product characteristics (ω^m_{jt}, ω^f_{lrc}):[12]

[12] We could have considered the more standard log linear form for the product marginal cost. However, in our application we found that 0.8% of the estimated mc_{jt} are negative. This result prevents us from using the log form.

$$mc_{jt} = w^m_{jt}\gamma_m + \omega^m_{jt} \tag{13}$$

$$FCost_{lrc} = w^f_{lrc}\gamma_f + \omega^f_{lrc} \tag{14}$$

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

2.2.1. *Rental Building Rates and Landing Fees:* The way landing fees and rental rates are determined is airport specific and follows the guidelines proposed by the DoT. The design of these charges is important because it affects the decisions of carriers regarding fares and frequency of their flights. SFO uses a hybrid approach to determine charges.[13] Under such a methodology, operating costs and revenues during the fiscal year are allocated to different cost centers. Three of those 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 instance, the maintenance of the ramp and cost recovery of investments in capital.[14] 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).

Following the 2006 Annual Operating Budget document for the San Francisco International Airport, the total landing fee revenues equal the amount needed to cover the net operating costs of the Airfield Cost Center ($ARCost$), plus 50% of the operating deficit (or surplus) in the Terminal ($TCost$) and Groundside ($GCost$) Cost Centers. The ratio between

[13]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 those centers are defined as residuals (break even) and others as compensatory (cost recovery) (Daniel (2001)).

[14]For example, 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.

the total landing fee revenues and the total annual scheduled landing weight of aircraft ($TWeight$) is the weight-based fee ($fees$) that airlines pay per 1000 pounds of the maximum gross landing weight (MGLW) of aircraft. That is,

$$fees = \frac{ARCost + \frac{1}{2}\left[TCost + GCost\right]}{TWeight} \tag{15}$$

Several remarks may be made with respect to the aforementioned charges: first, in 2006 this ratio was equal to \$3.213 per 1000 pounds of aircraft MGLW. Second, the Groundside Cost Center ($GCost$) tends to be profitable since it includes the lucrative car parking activity. Because of that, having this term in the charge rule generally reduces the amount that airlines must pay. Third, the above landing fee applies to airlines that sign the Airport-Airline Use and Lease Agreement (signatory airlines). Carriers operating in SFO without such a contract agreement (non signatory airlines) are usually charged more (the signatory landing fees plus a fixed amount). In our application, all airlines are assumed to pay landing fees according to the above methodology. Finally, as we already noted, the weight-based fee scheme does not take into account the level of congestion at the airport.

Now we turn to the analysis of each component appearing on the RHS of (15). The net operating costs of the Airfield Cost Center ($ARCost$) is assumed to be exogenous. On the other hand, the net operating costs of the Terminal Cost Center in SFO ($TCost$) are equal to the difference between operating expenditures (OE) and the operating revenues (OR):

$$TCost = OE - OR \tag{16}$$

While operating expenditures (OE) are assumed to be exogenous, operating revenues (OR) depend on the number of travelers using the airport ($TTravelers$). If we assume a linear relationship between both variables, then

$$OR = \psi_{terminal} \times TTravelers \tag{17}$$

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

Similarly, we define the net costs of the Groundside Cost Center ($GCost$) as the difference between costs (GC) and revenues ($GRev$) coming from groundside operations:

$$GCost = GC - GRev \tag{18}$$

Groundside costs (GC) are assumed to be exogenous. However, groundside revenues ($GRev$) depend on the total number of enplaned travelers ($TTravelers$). If we assume a linear relationship between both variables, then

$$GRev = \psi_{ground} \times TTravelers \tag{19}$$

where ψ_{ground} may be interpreted as the average revenue per enplaned passenger from groundside operations (for instance, revenues from parking the car at the airport).

Note that the total number of enplaned travelers using SFO is not necessarily equal to the total demand for products considered in our application. Total demand can be decomposed as follows,

$$TTravelers = \sum_{kt} s_{kt} \times M_t + ResTT \tag{20}$$

while the first term on the RHS accounts for demand considered in our model (domestic round trips with SFO as an origin or final destination), $ResTT$ captures the demand that is not included in the products of our model (connecting and international flights, as well as domestic flights from quarters of 2006 other than the 3rd). For simplicity, $ResTT$ is assumed to be independent of the demand for products considered in our model.

Finally, the denominator in (15) corresponds to the total scheduled landing weight of aircraft at SFO ($TWeight$) for the fiscal year, and it is equal to

$$TWeight = \sum_{l,r,c} weight_{lrc} \times \tilde{f}_{lrc} + ResTW \tag{21}$$

While the first term on the RHS accounts for the weight of aircraft used by products considered in our application, $ResTW$ 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 other than the 3rd). $ResTW$ is assumed to be independent of aircraft used by products considered in our application.

Apart from the landing fee, SFO is also compensated by carriers for the use of its terminals. Following again the 2006 Annual Operating Budget document for SFO, the total terminal rental charge paid by airlines equals the amount needed to cover 3/2 of the net operating costs of the Terminal Cost Center ($TCost$), plus 50% of the calculated net operating deficit (or surplus) of the Groundside Cost Center ($GCost$). Then the total rental charge that airline c must pay to SFO for using its terminals is equal to

$$RC_c = \left(\frac{3}{2}TCost + \frac{1}{2}GCost\right) \times Usage_c \tag{22}$$

where $Usage_c$ is the percentage of the total terminal surface leased by the airline c.

3. 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 for peak and off-peak hours and afterwards decide on the price of tickets. As usual, this game is solved backwards: first, we derive the optimality conditions for fares taking frequencies as given, and then we derive the first order conditions for frequencies taking into account the response of fares.

3.1. **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 given by

$$\begin{aligned} \frac{\partial \Pi_c}{\partial p_{j't'}} &= \sum_{t\in T}\sum_{j\in \mathcal{J}_{ct}} (p_{jt} - m_{jt})\frac{\partial s_{jt}}{\partial p_{j't'}} M_t + s_{j't'}M_{t'} - \\ &- \sum_{r\in\Omega_c}\sum_{l\in\{\mathcal{L},\mathcal{H}\}} \tilde{f}_{lrc}\left[\frac{\partial fees}{\partial p_{j't'}} weight_{lrc} + fees\frac{\partial weight_{lrc}}{\partial p_{j't'}}\right] - \frac{\partial RC_c}{\partial p_{j't'}} = 0 \end{aligned} \tag{23}$$

where $t' \in T$ and $j' \in \mathcal{J}_{ct'}$. Several remarks are in order: first, optimal fares follow from isolating the variable ticket price in (23). Second, the derivative of the profit function with respect to fares does not depend on the congestion charge (ρ_l) or the undelayed flight cost ($FCost_{lrc}$). That means that optimal fares are not directly affected by ρ_l, but indirectly through changes in frequencies. Finally, in our application, the product-specific cost (m_{jt}) is not observed, and we will use (23) to recover it. Then, we can compute the marginal costs (mc_{jt}) using (12) and estimate the parameters appearing in the marginal cost equation (13).

Now we turn to the computation of the derivatives appearing on the RHS of (23). As we will see, these derivatives 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 relationship in (10) are estimated. Following Nevo (2000), the derivative of the market share of product j in market t with respect to the price of product j' in market t' is

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

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 product j in market t (similar interpretation for $s_{ij't'}$). α_{ip} is the previously defined individual-specific coefficient associated with the ticket price.

We saw that the weight-based fee is a function of the revenues and costs assigned to different cost centers (equation 15). At the same time, these revenues depend on travel demand, flight frequency, and ticket prices. The gradient of the weight-based landing fee ($fees$) with respect to changes in fares, assuming that the net operating costs of the Airfield Cost Center ($ARCostC$) are exogenous,[15] is equal to

$$
(25) \qquad \frac{\partial fees}{\partial p_{j't'}} = \frac{1}{TWeight} \left[\frac{1}{2} \frac{\partial (TCost + GCost)}{\partial p_{j't'}} - \frac{\partial TWeight}{\partial p_{j't'}} fees \right]
$$

We remark that the expression is the same for products operating during peak or off-peak periods. We use (16) to compute the derivative of the net operating costs of the Terminal Cost Center ($TCost$). Similarly, using (18) we obtain the derivative of the net costs of the Groundside Cost Center ($GCost$). Finally, we use (21) to compute the derivative of the total scheduled landing weight ($TWeight$) with respect to fares.

If we look again at the RHS of (23), we use (10) to compute the derivative of the weight of aircraft ($weight_{lrc}$) with respect to ticket prices. The gradient of the terminal rental cost (RC_c) follows from computing the derivative of (22).

As we previously noted, 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. (23) and solve for the product-specific cost (m_{jt}). This result lets us obtain the marginal costs (mc_{jt}) and estimate the rest of parameters appearing in the first stage of the game, where frequencies are chosen.

[15] We may argue that $ARCostC$ depends on the number of landings at the airport. That is, the higher the number of operations, the higher are the costs of maintenance of the ramp. For simplicity, we do not consider this effect.

3.2. **First Stage: Frequencies.** Once we derive the optimality conditions for fares, 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 the spoke r' and congestion period l' is given by

$$
\begin{aligned}
(26)\ \frac{\partial \Pi_c}{\partial f_{l'r'c}} \quad &= \quad \sum_{t \in T} \sum_{j \in \mathcal{J}_{ct}} \left[(p_{jt} - m_{jt}) \frac{\partial s_{jt}}{\partial f_{l'r'c}} M_t + \frac{\partial p^*_{jt}}{\partial f_{l'r'c}} s_{jt} M_t \right] - \\
&- \quad 92 \left[fees \times weight_{l'r'c} + \beta^d D_{l'} + FCost_{l'r'c} + \rho_{l'} \right] - \\
&- \quad \sum_{r \in \Omega_c} \sum_{l \in \{\mathcal{L},\mathcal{H}\}} \tilde{f}_{lrc} \left[\frac{\partial fees}{\partial f_{l'r'c}} weight_{lrc} + fees \frac{\partial weight_{lrc}}{\partial f_{l'r'c}} + \beta^d \frac{\partial D_l}{\partial f_{l'r'c}} \right] - \frac{\partial RC_c}{\partial f_{l'r'c}} = 0
\end{aligned}
$$

where $r' \in \Omega_c$ and $l' \in \{\mathcal{L},\mathcal{H}\}$. $\frac{\partial p^*_{jt}}{\partial f_{l'r'c}}$ denotes the derivative of the optimal fare with respect to frequency. In our application, we will use (26) to estimate the monetary value of one minute of delay (β^d) and the undelayed flight cost component ($FCost_{lrc}$). Moreover, we also use this expression to analyze the impact of imposing a congestion charge in the landing fee rule.

The difficulty in (26) lies in computing the gradient of the optimal fare ($\frac{\partial p^*_{jt}}{\partial f_{l'r'c}}$) and the derivative of market shares with respect to frequencies ($\frac{\partial s_{jt}}{\partial f_{l'r'c}}$). Let us start with $\frac{\partial p^*_{jt}}{\partial f_{l'r'c}}$. 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 (23) with respect to fares ($dp_k, k = \{1, \cdots, J\}$) and daily flight frequency ($f_b, b = \{1, \cdots, |\Omega \times \{\mathcal{L},\mathcal{H}\}|\}$), where J is the total number of products ($J = \sum_{t \in T} J_t$), and $|\Omega \times \{\mathcal{L},\mathcal{H}\}|$ is the total number of spokes operated by airlines at SFO in each period. 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 denote the $J \times |\Omega \times \{\mathcal{L},\mathcal{H}\}|$ ownership matrix with the general element $\Psi^f_c(j't', b)$ equal to one if the product $j't'$ and the spoke-period pair b are operated by the same carrier c. Then, the total derivative of the fare F.O.C. (23) with respect to fares and frequencies for product $j't'$ and carrier c can be written as

$$
(27) \qquad \underbrace{\sum_k \Psi^p_c(j't', k) \frac{\partial^2 \Pi_c}{\partial p_{j't'} \partial p_k}}_{G^p_c(j't',k)} dp_k + \underbrace{\sum_b \Psi^f_c(j't', b) \frac{\partial^2 \Pi_c}{\partial p_{j't'} \partial f_b}}_{H^p_c(j't',b)} df_b = 0
$$

We can express (27) in a matrix form. Let G^p_c be a $J \times J$ dimensional matrix with component $G^p_c(j't', k)$. Similarly, let H^f_c be a $J \times |\Omega \times \{\mathcal{L},\mathcal{H}\}|$ dimensional matrix with component $H_c^f(j't',b)$. Then, condition (27) for the entire collection of products $j't'$ can be written as

$$G_c^p dp + H_c^f df = 0 \tag{28}$$

where dp and df are column-vectors of dp_k for $k = \{1, \cdots, J\}$ and df_b for $b = \{1, \cdots, |\Omega \times \{\mathcal{L}, \mathcal{H}\}|\}$ 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. Let $G_p = \sum_c G_c^p$ and $H_f = \sum_c H_c^f$, then the following equality also holds

$$G_p dp + H_f df = 0 \tag{29}$$

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

$$\frac{dp^*}{df} = -G_p^{-1} H_f \tag{30}$$

Once we derive $\frac{dp^*}{df}$, we can compute the derivative of the market share with respect to flight frequency ($\frac{\partial s_{jt}}{\partial f_{l'r'c}}$). Remember that travelers using different products may share the same aircraft. Moreover, the frequency of flights operating in one spoke during peak or off-peak hours affects the average delay of other flights operating during the same period. These two features are captured by the derivative of the market share, which is given by

$$\frac{\partial s_{jt}}{\partial f_{l'r'c}} = \int s_{ijt} \left(\kappa_{ijt}^{l'r'c} - \sum_{n=1}^{J_t} \kappa_{int}^{l'r'c} s_{int} \right) dP_\nu(\nu_i) dP_L(L_i) dP_Y(y_i) \tag{31}$$

Expression (31) follows from computing the derivative of the market share equation (7) with respect to the frequency of flights, where $\kappa_{i\Theta t}^{l'r'c}$, with $\Theta \in \{j, n\}$ in (31), is given by

$$\kappa_{i\Theta t}^{l'r'c} = \alpha_{ip} \frac{\partial p^*_{\Theta t}}{\partial f_{l'r'c}} + \frac{1}{e_{\Theta t}} \left(\alpha_{if} \mathbb{1}\{r_{\Theta t} = r' \,\cap\, \Theta \in \mathcal{J}_{ct} \,\cap\, l_{\Theta t} = l'\} + \alpha_{id} \frac{\partial D_l}{\partial f_{l'r'c}} \right) \tag{32}$$

Θ denotes a product of market t ($\Theta \in \{j, n\}$ in (31)), and $\frac{\partial p^*_{\Theta t}}{\partial f_{l'r'c}}$ is equal to the corresponding matrix element of the derivative of optimal fares with respect to flight frequencies (30). $\mathbb{1}\{\cdot\}$ is an indicator function equal to one if the condition inside brackets holds, and zero otherwise. As we previously noted, travelers care about the flight frequency and average delay at each of the connecting and destination airports used by products. That is why we introduce $e_{\odot t}$ as the number of connections of the product.[16] Also remember that we already defined α_{ip}, α_{if}, and α_{id} as individual-specific parameters associated with fares, frequencies, and delays respectively. The last term on the RHS of (32) corresponds to the derivative of the previously defined average airport delay in period l (equation 9), and it is given by

$$\frac{\partial D_l}{\partial f_{l'r'c}} = \begin{cases} exp(\omega_l^d \bar{f}_l)\omega_l^d \ \ if \ \ l = l' \\ \\ 0 \ \ if \ \ l \neq l' \end{cases} \tag{33}$$

The derivative of the market share of product jt in (31) not only depends on its own characteristics but also on characteristics of other products. This effect is captured by the summation in the expression. $\kappa_{ijt}^{l'r'c}$ depends on the relationship between the product jt, the operating carrier c, the spoke r', and the period l'. Looking at (32), if jt uses r' ($r_{jt} = r'$), during period l' ($l_{jt} = l'$), and carrier c, then $\kappa_{ijt}^{l'r'c}$ not only depends on the derivative of the optimal fare with respect to frequency ($\frac{\partial p_{jt}^*}{\partial f_{l'r'c'}}$), but also on the parameters linked to the demand for flight frequency (α_{if}) and also delay ($\alpha_{id}\frac{\partial D_{r_{jt}}}{\partial f_{l'r'c}}$). On the other hand, if the product jt does not use r' or carrier c but still is offered in the same period l', then $\kappa_{ijt}^{l'r'c}$ is no longer affected by α_{if} but still depends on the derivative of the airport delay in period l' and the derivative of the optimal fare with respect to frequency. Finally, if product jt is not scheduled to land in l', then $\kappa_{ijt}^{l'r'c}$ only depends on the derivative of the optimal fare with respect to frequency. Similar reasoning holds for $\kappa_{int}^{l'r'c}$.

Once $\frac{\partial p_{jt}^*}{\partial f_{l'r'c}}$ and $\frac{\partial s_{jt}}{\partial f_{l'r'c}}$ are computed, the rest of derivatives appearing in the frequency F.O.C. (26) are straightforward. In particular, the derivative of the weight-based landing fee with respect to frequency is equal to

$$\frac{\partial fees}{\partial f_{l'r'c}} = \frac{1}{TWeight}\left[\frac{1}{2}\frac{\partial\left(TCost + GCost\right)}{\partial f_{l'r'c}} - \frac{\partial TWeight}{\partial f_{l'r'c}} fees\right] \tag{34}$$

where the derivative of the net operating costs of the Terminal Cost Center ($TCost$) follows from computing the derivative of (16). Similarly, we use (18) for the derivative of the net costs of the Groundside Cost Center ($GCost$). In addition, we use (21) to compute the derivative of the total scheduled landing weight ($TWeight$).

If we look again at the RHS of (26), we use (10) to compute the derivative of the weight of the aircraft ($weight_{lrc}$), and equation (9) for the derivative of the average airport delay

[16] More details are provided in the data section.

in each period (D_l). Finally, the derivative of the terminal rental charge of carrier c (RC_c) follows from computing the derivative of (22).

4. Data and Statistics

4.1. **Data Description:** We perform the estimation of the model for the third quarter of 2006. Note that the new policy was issued in 2008, when the economic downturn reduced air-travel demand. We use pre-crisis data (2006) to better capture the effects of the new guidelines. In order to conduct the analysis, several data sources are used. First, we use the Airline Origin and Destination Survey (DB1B), the T-100, the Airline On-Time Performance, and the Schedule B-43 Aircraft Inventory data-sets from the U.S. Bureau of Transportation Statistics. These sources, jointly with the aircraft manufacturers' websites, are used to obtain information about the product choices of travelers and their characteristics. Second, the Federal Aviation Administration website and the Annual Operating Budget for SFO give detailed financial information about the airport and the methodology used to determine the weight-based landing fee and rental charges. Finally, we use demographic data from the American Community Survey (ACS) and the 2006 MTC Airline Passenger Survey.

The Appendix contains further details about the data sources.

4.2. **Construction of the data-set:** We restrict our attention to data from the third quarter of 2006. Following Urdanoz and Sampaio (2011) and using the DB1B data-set, 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 segment; (8) that do not have an airport coded as NYC, since we cannot identify which of the 4 airports in the New York Metropolitan area was used; and finally, (9) with fares between \$50 and \$3000.

A product is considered to be offered in peak hours if it is scheduled to arrive at SFO when the airport congestion is high (from 9:00AM to 12:30PM, and from 6:30PM to 10:30PM), and off-peak otherwise. We use the Airline On-time Performance and DB1B data-sets to define the characteristics of products in each period: market shares, daily frequency of flights, and delays for peak and off-peak hours.

It is important to note that the DB1B only reports quarterly data about product characteristics and number of passengers. Therefore, it does not specify if passengers arrived at

SFO during peak or off-peak hours. The way we distribute travelers between the two periods is as follows: first, we use the Airline On-time Performance and Schedule B-43 Aircraft Inventory data-sets to compute the daily capacity of carriers offering flights at SFO by spoke and period. As we will see in detail later, daily capacity is equal to the maximum gross landing weight of aircraft times the number of daily operations for the spoke-route, period and carrier. As a result, for each spoke-route and carrier we know the percentage of daily capacity assigned to each period. We use those percentages to determine which passengers (from the DB1B) arrive at SFO during peak and which ones during off-peak hours. This assumption seems reasonable since the total daily capacity at each period is highly correlated with the number of passengers.

A similar problem arises with fares. We use the average fare paid by travelers for the same product independently if they decide to flight during peak or off-peak hours. We did not apply any discount in off-peak hours because we are not aware of any empirical work analyzing differences in fares between periods. In any case, we repeated the exercise with a discount of 15% and found minor changes in the main results of the paper.

We also use the Airline On-time Performance data-set to construct the frequency variable for products. We know the number and the arrival time of flights operated by the main U.S. carriers. Hence, we know which flights land during peak and which ones during off-peak hours. Since products are round trip and may have several connecting flights, some other assumptions are needed. If travelers land at SFO during peak hours, we assume that they also leave SFO during congested periods, and off-peak otherwise. For connecting flights, we alternate peak and off-peak depending on the arrival time. For example, imagine that a traveler arrived at SFO during peak hours from John F. Kennedy International via Chicago O'Hare Airport. According to our assumption, the traveler landed at Chicago in off-peak and departed from John F. Kennedy International in peak hours. For simplicity, we assume that each U.S. airport has similar congested periods. That is, flights arriving between 9:00AM and 12:30PM and between 6:30PM and 10:30PM are considered operating in peak hours, and off-peak otherwise. Given those assumptions, the flight frequency of a product is constructed as the mean of the frequency for each trip segment of the product and corresponding period.

We follow the approach proposed by Mayer and Sinai (2003) to construct the delay variable. Rather than defining flight delay as the percentage of flights arriving more than 15 minutes after the scheduled arrival time, we use the difference between actual and minimum feasible flight durations. Such an approach avoids the use of scheduled arrival times that may be subject to airline manipulation (padding) to increase on-time performance. We construct the flight duration of a trip segment as the difference between the actual arrival time and the wheels off time at the airport of departure. We do not take into account the aircraft taxi time at the airport of origin, since differences in this variable may be explained by reasons other than congestion at the destination airport. Following the same approach as for flight frequency, we first compute the average delay at connecting and final airports in each period, and then we derive the mean average delay of airports and periods used by the product.

As we previously noted, landing fees depend on the maximum gross landing weight (MGLW) of aircraft (variable $weight_{lrc}$ in our model). We use the Airline On-Time Performance and the Schedule B-43 Aircraft Inventory data-sets to construct this variable. The Airline On-time Performance does not show the model of aircraft used by carriers, but it reports the aircraft tail number (civil registration serial number). We match this number with the one appearing in the Schedule B-43 Aircraft Inventory to obtain the model of aircraft. Given the type of airplane, we check its technical specifications in the manufacturer's website to obtain its MGLW. Carriers may use several types of aircraft for the same spoke-route and period. We compute the average MGLW of the different aircraft operating in the same spoke-route and period weighted by their number of operations during 2006. The daily capacity of carriers at each spoke-period pair is derived as the average MGLW of aircraft times the corresponding number of operations.

We use the American Community Survey (ACS) and the 2006 MTC Airline Passenger Survey to obtain demographic information about income and travelers' locations. Using the 2006 MTC data-set, we construct the empirical distributions of household income and distance conditional on the departure period (see extra moments conditions (40) and (41)). On the other hand, we create the variable "traveler distance" ($d(L_i)$) by taking random draws from the Bay Area population distribution provided by the ACS. We also use the ACS to create the variable "household income" for individual i (y_i). We take random draws of the income distribution for each of the metropolitan statistical areas where the airports of origin are located (see utility function (1)). We do not have the exact address of respondents in the ACS or 2006 MTC surveys, but only the area where they come from. Two different location measures are used: the 2006 MTC survey classifies the place of origin according to Regional Travel Analysis Zones (TAZ), dividing the San Francisco Bay area into 1454 zones. Instead, the ACS partitions the Bay Area using census tracts (1099 zones). Fortunately there exists an equivalence between the two measures. Once the measures are homogenized, we compute the distance from the traveler's location to SFO as the Euclidean distance between the airport and the centroid of the census tract where the traveler is located.[17] Our model uses county fixed effects, so we group the census tracts according to their respective county. Similarly, we need to transform the household income information in order to compare both surveys.

As noted above, we use products from the 3rd quarter of 2006 to estimate the model. However, the weight-based landing fee and rental charge are computed using data from the fiscal year.[18] For that reason, to construct some variables we use yearly data rather than just one quarter. The variable $TWeight$, used as denominator in the weight-based landing fee expression (15), will be the sum of the MGLW of each aircraft times the number of operations for the year 2006. This variable is constructed using the T-100 data-set. We use the total number of enplaned passengers (using also the T-100), the Groundside Cost Center revenues ($GRev$), and the Terminal Cost Center revenues (OR) for all quarters of 2006 to determine the average revenue per passenger from terminals ($\psi_{terminal}$) and groundside operations (ψ_{ground}).

We do not have information about the terminal areas leased by airlines. Instead, we use the number of operations performed by each airline at SFO as a proxy for the usage of terminals ($Usage_c$).

4.3. **Summary Statistics:**

4.3.1. *Choice and Flight Characteristics Statistics:* Table 1 reports statistics for product characteristics. We differentiate products offered during peak hours (columns 4 and 5) from off-peak products (columns 2 and 3). Columns 6 and 7 report the mean and the standard deviation for all products. As explained in the data section, we did not assume any fare discount in off-peak products, which explains why the mean fare in both periods are similar. On the other hand, we do observe differences in the number of purchased tickets. Travelers purchased, on average, 149.24 tickets of each product offered during peak and only 95.97 during off-peak hours. The percentage of direct flights, frequency, and delays are also higher in peak hours. There are not significant differences in the distribution of tickets by carrier between periods. In both periods, United Airlines (UA) is the carrier with the highest

[17]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 SFO. Another option would be to compute the time needed to reach the airport.

[18]The fiscal year for SFO starts in July 1 and finishes June 30. For simplicity we assume that the fiscal year and the natural year are the same.

presence, followed by American Airlines (AA) and Delta Airlines (DL). Finally, 30% of products use one of the airports with slot constraints.[19]

Table 2 reports the number of products and the number of operating carriers within markets. On average, there are more products and operating carriers in peak hours. On top of that, in almost all markets there are products from both periods. That is, in 435 out of 437 markets, there exist products with scheduled flights landing during peak and off-peak hours.

Table 3 summarizes the characteristics of spokes reaching SFO by carrier and period. It provides information about the daily frequency of flights, the weight of aircraft (MGLW), and the total daily capacity of the spoke-carrier pairs for peak and off-peak hours (measured as the product of MGLW times the daily frequency of flights by carrier and period). Although the average frequency of flights is higher in off-peak hours, the total offered capacity is lower. This is explained by the use of bigger planes during congested hours. Note also that the number of spokes served in peak hours is higher (56 spokes in peak vs 49 in off-peak).

Tables 4 and 5 show statistics for spoke-route traffic during peak and off-peak hours respectively and broken down by carrier. They report the mean and standard deviation of the number of daily flight arrivals, the average weight of aircraft (MGLW), daily capacity of the spoke-carrier pairs, and the number of spokes operated by carriers in each period. SFO is one of the hubs for United Airlines (UA), which explains why UA operates a high number of spokes. Moreover, in both periods DL operates the biggest planes and US the smallest ones.

Table 6 reports financial details of SFO for the year 2006. This information is incorporated in the estimation of the supply side of the model. It includes the landing fee per 1,000 pounds of MGLW, the total number of enplaned passengers, the operating revenues of the Terminal Cost Center (OR), and the revenues of the Groundside Cost Center ($GRev$). The average operating revenue per traveler in the terminals ($\psi_{terminal}$) is computed as the ratio between the operating revenues (OR) and total enplaned passengers. The average revenue per enplaned passenger from groundside operations (ψ_{ground}) is equal to the ratio between the groundside revenues ($GRev$) and the total enplaned passengers. Table 6 also reports the total weight of landed aircraft in 2006 ($TWeight$).

4.3.2. *2006 Airline Passenger Survey Statistics (2006 MTC):.* This survey helps us to identify the parameters of the demand equation (1). Table 7 and Table 8 summarize the distribution of household income and location in the Bay Area of the survey respondents. Table 7

[19]Slot constrained markets are the ones with one of the following airports: JFK, LGA, DCA, or ORD.

shows the joint distribution of respondents departing during peak hours. Similarly, Table 8 analyzes the distribution in off-peak hours. 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). Most of respondents come from the county in which the airport is located (San Mateo County) or the county next to it (San Francisco County). If we look at household incomes, the distributions are skewed to household groups with higher income, with most of respondents belonging to the \$100,000 and \$150,000 group. In both cases, the distributions for peak and off-peak hours are very similar.

5. 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 (26) to estimate the valuation of one minute of delay (β^d) and the parameters 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.

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

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

5.1.1. *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}:

$$s_{jt}(\delta(\theta), \cdot; \theta) = s_{jt} \;\; for \;\; j = \{0, \cdots, J_t\} \;\; and \;\; t = \{1, \cdots, T\} \tag{35}$$

Berry (1994) shows 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

$$\delta_{jt}^{h+1} = \delta_{jt}^{h} + ln(s_{jt}) - ln(s_{jt}(\delta(\theta), \cdot; \theta)) \tag{36}$$

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

$$s_{jt}(\delta(\theta), \cdot; \theta) = \frac{1}{g} \sum_{i=1}^{g} \frac{exp\left[\delta_{jt} + \mu_{ijt}\right]}{1 + \sum\limits_{m \in J_t} exp\left[\delta_{mt} + \mu_{imt}\right]} \tag{37}$$

For the observable individual characteristics (distance and income) we use g random draws from the empirical distribution.[20] 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 (2) and given δ_{jt}, θ and product characteristics, we can derive the moment condition related to the unobserved-to-researcher characteristics of products j in market t (ξ_{jt}). That is,

$$\xi_{jt} = \delta_{jt} - \alpha_{peak} \hat{I}_{jt}^{peak} - \alpha_p p_{jt} - \alpha_f \hat{f}_{jt} - \alpha_d \hat{D}_{jt} - x_{jt}\beta \tag{38}$$

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

$$E[z_{jt}^{d} \xi_{jt}] = 0 \tag{39}$$

5.1.2. *Additional Demand Information 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 departing during peak or off-peak hours. In particular, we use information about their distance to the airport and their household income.

Basically we will try to construct moments that match the predicted average consumer demographics obtained from the BLP moments with the average consumer demographic

[20]In our application g=1200.

characteristics from the MTC survey. The extra moment conditions will match the probability that a traveler i departing from SFO in period l (where $l \in \{\mathcal{L}, \mathcal{H}\}$), comes from/goes to a specific county ($\mathcal{C}$) and belongs to a income group ($\mathcal{Y}$). That is,

$$\eta_c(\mathcal{C}, l) = E\left[L_i \in \mathcal{C} | \{i \ \ departs \ \ in \ \ l\}\right] \tag{40}$$

$$\eta_y(\mathcal{Y}, l) = E\left[y_i \in \mathcal{Y} | \{i \ \ departs \ \ in \ \ l\}\right] \tag{41}$$

where

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

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

$$l \in \{\mathcal{L}, \mathcal{H}\}$$

where L_i and y_i are the location and household income group of individual i. $\eta_c(\mathcal{C}, l)$ and $\eta_y(\mathcal{Y}, l)$ are the probabilities from the 2006 MTC survey (Tables (7) and (8)). The RHS expressions in (40) and (41) are the expected values predicted by our model and computed using the simulated market shares (37). These extra conditions apply for all income groups, counties and levels of congestion.

Since the probabilities $\eta_c(\mathcal{C}, l)$ and $\eta_y(\mathcal{Y}, l)$ conditional on the level of congestion 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 departing during peak or off-peak hours, we need to apply the definition of conditional probability to match the predicted probabilities with the MTC survey probabilities. Thus, the sample analog of the additional information moments can be written as

$$\eta_c(\mathcal{C}, l) - \frac{\sum_{i=1}^{g} \sum_{\{jt | l_{jt} = l\}} s_{ijt}(\delta(\theta), \cdot; \theta) M_t \mathbb{1}\{L_i \in \mathcal{C}\}}{\sum_{i=1}^{g} \sum_{\{jt | l_{jt} = l\}} s_{ijt}(\delta(\theta), \cdot; \theta) M_t} \tag{42}$$

$$\eta_y(\mathcal{Y}, l) - \frac{\sum_{i=1}^{g} \sum_{\{jt|l_{jt}=l\}} s_{ijt}(\delta(\theta), \cdot; \theta) M_t \mathbb{1}\{y_i \in \mathcal{Y}\}}{\sum_{i=1}^{g} \sum_{\{jt|l_{jt}=l\}} s_{ijt}(\delta(\theta), \cdot; \theta) M_t} \tag{43}$$

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 $\mathcal{Y}$ lands at SFO in period l. As we previously noted, l_{jt} denotes the period when product jt is scheduled to arrive. Finally, $\mathbb{1}\{\cdot\}$ is an indicator function equal to one if the condition inside brackets holds, and zero otherwise.

5.1.3. *Marginal Cost Moment:* In this case, the corresponding moment condition is

$$E[z_{jt}^m \omega_{jt}^m] = 0 \tag{44}$$

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

5.1.4. *GMM Estimation:* Let $\vartheta = \begin{bmatrix} \theta \\ \gamma_m \end{bmatrix}$ denote the set of parameters to be estimated using GMM. Our optimal 2-step GMM estimators will be

$$\hat{\vartheta} = arg\min_{\vartheta} \hat{G}(\vartheta)'\Phi^{-1}\hat{G}(\vartheta) \tag{45}$$

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 $\delta(\dot{\theta})$), afterwards the GMM optimization problem (given $\delta(\dot{\theta})$, solve for $\ddot{\theta}$ and $\ddot{\gamma}_m$), and iterate again until convergence is reached.[21]

As we previously pointed out, the product-specific cost (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

[21] 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/CG (barrier) algorithm to solve the GMM minimization problem.

use the fare F.O.C. (23) 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 parameters of the weight equation, the only unknown in the optimality condition for fares is $\dot{m}_{jt}$. Knowing the product-specific cost ($\dot{m}_{jt}$), 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

$$J^{1/2}(\hat{\vartheta} - \vartheta^0) \sim \mathcal{N}(0, (\Gamma'\Phi^{-1}\Gamma)^{-1}) \tag{46}$$

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

5.2. **Estimation of β^d and the undelayed flight cost ($FCost_{lrc}$):** As in the case of the product-specific cost (m_{jt}), the undelayed flight cost ($FCost_{lrc}$) is not observed by the econometrician. Neither do we observe the monetary value of one minute of delay (β^d). Under the assumption that $FCost_{lrc}$ depends linearly on a vector of cost shifters (w^f_{lrc}) and a random term (ω^f_{lrc}) (equation 14), we use the frequency F.O.C. (26) to identify β^d and the parameters linked to the undelayed flight cost component (γ_f).

Combining equations (26) and (14), the expression we use for the estimation of β^d and w^f_{rc} is given by

$$\begin{aligned}
&\frac{1}{92}\sum_{t\in T}\sum_{j\in\mathcal{J}_{ct}}\left[(p_{jt}-m_{jt})\frac{\partial s_{jt}}{\partial f_{l'r'c}}M_t + \frac{\partial p^*_{jt}}{\partial f_{l'r'c}}s_{jt}M_t\right] - fees \times weight_{l'r'c} - \rho_{l'} - \\
&-\frac{1}{92}\left[\sum_{r\in\Omega_c}\sum_{l\in\{L,H\}}\tilde{f}_{lrc}\left[\frac{\partial fees}{\partial f_{l'r'c}}weight_{lrc} + fees\frac{\partial weight_{lrc}}{\partial f_{l'r'c}}\right] - \frac{\partial RC_c}{\partial f_{l'r'c}}\right] = \\
&= w^f_{lrc}\gamma_f + \beta^d\left[D_l + \frac{1}{92}\sum_{r\in\Omega_c}\sum_{l\in\{\mathcal{L},\mathcal{H}\}}\tilde{f}_{lrc}\frac{\partial D_l}{\partial f_{l'r'c}}\right] + \omega^f_{lrc}
\end{aligned} \tag{47}$$

All elements on the LHS of (47) are known. Note also that in our application $\rho_{l'} = 0$, since SFO does not impose any operation charge except the weight-based landing fee ($fees$). Once the value of the LHS expression is computed, the estimation of the parameters follows from applying OLS.

Finally, the fitted value for $FCost_{lrc}$ equals the product of the estimates for γ_f times the cost shifters (w^f_{lrc}) (see (14)). Once all parameters of the model are estimated, we will recompute the optimal response of carriers with respect to the frequency of their flights and fares for different values of the congestion charge ($\rho_{l'}$).

5.3. **Instruments and Identification:** We allow for the possible endogeneity of fares and frequencies with the unobserved-to-researcher variable ξ_{it} in the demand equation (see (1)) and the random term in the marginal cost equation (13). Similarly, total demand and frequency are also likely correlated with the disturbance term in the weight equation (10). We correct for endogeneity using similar instruments to those proposed by Berry and Jia (2010) and Nevo (2001):

- Demand characteristics considered exogenous: distance, dummy for peak hours, operating carrier, dummy for direct flight, and a dummy for airports 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.

To construct instruments for flight frequency, we first regress the number of departures on distance, market size (measured by the geometric mean population of the origin and destination metropolitan statistical areas), 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 peak hours. Once we obtain the estimates, we compute the residuals, actual minus fitted frequencies, and include them as instruments. The rationale of this approach is as follows: we assume that what is left after controlling for several factors (residuals) is correlated with the marginal cost but uncorrelated with the demand unobservable (ξ_{jt}).

The approach used for the instruments in the weight equation (10) is similar. The only difference is that we do not use connecting flights, and 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 (40) and (41)), 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 ($FCost_{lrc}$) and the cost of one minute delay (β^d) (equation (47)). In order to identify these two unobservable variables, we assume that

$FCost_{lrc}$ linearly depends on some factors: distance, operating carrier, dummy for peak hours, and a disturbance term (see (14)).

6. Estimation Results

6.1. **Demand Parameters:** Table 9 presents the demand estimates (1). The standard errors are reported in parenthesis. The second column corresponds to a logit model without instruments (*OLS* column). The third column reports estimates using instruments (*IV* column). In these basic specifications we assume that demand depends on fares, distance, flight frequency, a dummy for direct flights, a dummy for peak hours, carrier dummies, slot constrained destinations, and delays. Most of the estimates have the expected sign. As noted above, fares and flight frequencies are likely to be correlated with the unobserved-to-researcher characteristic (ξ). ξ can be interpreted as a quality index and it is positively correlated with fares and frequency of flights. That explains why the fare (α_p) and frequency (α_f) estimates in the OLS specification are biased upwards and downwards respectively. These two models do not capture important aspects of travelers' heterogeneity.

Column 4 reports estimates for the full model. The last ten variables control for the heterogeneity of travelers: the absolute values of σ^0, σ^f and σ^d are equal to the standard deviation of normal distributions and they are linked to a constant, frequencies and delays respectively. 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. Having α_y positive is consistent with the idea that the higher the traveler's income, the less is the sensitivity to changes in fares. The estimates associated with flight distance are also in line with the previous literature. As Berry and Jia (2010) point out, demand is increasing in flight distance because travel alternatives (e.g. car, bus, or train) become less attractive, but at the same time the trip is less and less pleasant. The estimate associated with distance from travelers' location to the airport (λ) is negative, indicating that the farther the airport from the traveler's final/origin location in the Bay Area, the higher is the disutility of flying from SFO. Travelers prefer higher frequency because they would find it easier to choose a flight that better matches their preferred departure time. As predicted, the direct flight coefficient is also positive. The delay estimate (α_d) also has the expected sign. Increasing delays reduce the willingness to pay for a ticket. Lastly, travelers prefer to fly in the morning and late in the afternoon, periods of time when the level of congestion at the airport is high.

Table 10 reports the mean and standard deviation of the product own-price elasticities. Columns 2 and 3 display the elasticities by period and carrier. The last column analyzes the case where we do not distinguish if the product is scheduled to arrive at SFO during peak or off-peak hours. We do not observe large differences across airlines or periods. If we compare them with the elasticities provided by Berry and Jia (2010), they are quite similar. Although our estimates are somewhat lower, they are of the same order of magnitude. In particular, our average price elasticity is equal to -3.44, higher (in absolute value) than the -2.10 value that Berry and Jia (2010) report. These similarities in results are remarkable if we take into account that Berry and Jia (2010) study the whole U.S. domestic market and that their model differs from ours.

Table 11 displays the frequency semi-elasticities of demand. Increasing by one the number of daily flights during peak and off-peak hours, the demand increases, on average, 7.319%.

6.2. **Supply estimates:** As we pointed out in the model section, the weight of aircraft is an important determinant of the total landing fees that carriers must pay. Weight can be seen as a proxy for the passenger capacity of the plane. Thus, it depends on air-travel demand and also on the strategy of carriers. Table 12 reports estimates for the aircraft weight equation (10). Note that we use spoke-route and period data rather than product data. That explains why we only have 165 observations instead of 11,316. The regression includes a constant, total number of passengers, daily frequency, distance, carriers and a dummy variable indicating if aircraft land in peak or off-peak hours. We are interested in the demand coefficient (τ_1) and the coefficient of daily flight frequency (τ_2). Both estimates are directly related to the decision of carriers with respect to fares and flight frequency. As expected, the τ_1 estimate is positive and significant, which means that the higher the demand, the larger are the size of the planes. Regarding τ_2, the higher the frequency of flights, the lower are the size of aircraft. Consistent with previous literature (Borenstein and Rosen (2008)), the longer the trip, the bigger are the planes. We do not observe significant differences between the size of the planes operating during peak and off-peak hours.

Table 13 reports the average marginal cost per passenger-mile (second column) and the mean Lerner index, defined as the ratio between the markups and fares (third column). Looking at the second column, only small differences arise between airlines. The average marginal cost per passenger-mile is equal to 6 cents, the same as reported in Berry and Jia

(2010). If we look at the third column, the average Lerner index equals 31%. Furthermore, direct flight products have a higher margin than those using connecting flights. If we compare our Lerner index estimates with those of Berry and Jia (2010), ours are much lower. They report an average of 63% for all flights, a much higher value than the one we found. Differences may be explained because Berry and Jia (2010) analyze the whole U.S. market and we only focus on the markets that use SFO as an origin or final destination.

Table 14 reports estimates for the linear specification of marginal costs (13), which includes a constant, carrier dummies, flight distance, square of flight distance, a dummy indicating if the connecting or destination airports are a hub of the operating carrier, a dummy for peak hours, and number of connections. These estimates are obtained in the GMM step. The results are again consistent with Berry and Jia (2010). Marginal costs increase non-linearly with distance. This finding makes sense if we think that an important fraction of aircraft fuel is consumed at take-off. A similar argument can be used to explain the positive estimate for the number of connections. The hub estimate is positive but not significant. This result is unexpected since we may think that airlines use their hubs to take advantage of economies of density. Hence, we predicted a negative sign rather than a positive one (Brueckner and Spiller (1994)).

In order to estimate the monetary value of one minute delay (β^d), we assume that the undelayed cost component ($FCost_{rc}$) linearly depends on distance, squared distance, and carrier dummies. Table 15 shows the parameter estimates for equation (47). We tried several specifications, and in all of them we obtained the unexpected result of β^d being negative. On the other hand, the distance and squared distance coefficients have the expected signs. That is, operating flight costs are non-linear increasing functions with respect to distance. There are not significant differences in undelayed flight costs between American Airlines (AA) and other carriers except for United Airlines (UA) and US Airways (US).

Table 16 summarizes the fitted values of the undelayed cost of operating a flight ($F\hat{C}ost_{rc}$) (the RHS of equation (14) without the error term). This variable may be interpreted as the fixed cost of operating a flight in a spoke-route. We did not include the weight of the aircraft as an explanatory variable because its endogeneity (it depends on fares and flight frequency) would increase the complexity of the model. The average fixed cost of operating a new flight is around $10,550.

7. Implementing an operation charge in peak hours

This section analyzes the consequences of implementing a congestion charge for flights arriving during peak hours ($\rho_{\mathcal{H}} > 0$ and $\rho_{\mathcal{L}} = 0$). Figures 2 to 12 show the effects on the main variables of interest: daily frequency, fares, air-travel demand, size of aircraft, delays, fees, and landing fee revenues. We simultaneously solve for flight frequency, ticket prices, and demand as a result of changes in $\rho_{\mathcal{H}}$. Fares and market shares are recomputed using their respective elasticities with respect to frequencies, and the new equilibrium in frequencies is obtained using (47).[22] We repeat the process for different values of $\rho_{\mathcal{H}}$ and interpolate the results to construct the figures.

While Figures 2 to 6 show the direct effect of implementing a congestion charge during peak hours, Figures 7 to 10 display the consequences in off-peak hours.

Figure 2 shows changes in the number of daily flights offered at SFO during peak hours. The x-axis represents the imposed congestion charge ($\rho_{\mathcal{H}}$) in dollars. Displayed on the y-axis is the change in the total number of daily flights arriving during peak hours. The variation in flight frequency has the expected sign: the higher the congestion charge, the lower is the number of flights during peak hours. For instance, a $\rho_{\mathcal{H}} = 2,000$ leads to a reduction in the number of flights of 3.86% (from 217 daily flights during peak hours to 208). The result is consistent with the idea that the higher the $\rho_{\mathcal{H}}$, the more expensive is operating a flight. Therefore, carriers have incentives to reduce the number of flights.

To gain an idea about the order of magnitude of these results, we compare them with those presented by Ashley and Savage (2010) and Daniel (2001). The comparison is not straightforward, since the application and methodology are different. Remember that these papers use a bottleneck model where aircraft are charged according to the congestion externality imposed on other flights. In the case of Ashley and Savage (2010), they use data from 2004 to compute the congestion fees for flights arriving at Chicago O'Hare Airport. In one of their simulations, they analyze the change in the optimal congestion fees as a result of an exogenous reduction of 2% in the number of flights arriving during peak hours. They claim that such a change would reduce the average congestion fee paid by aircraft from $8,100 to $5,800. Our results suggest that if the SFO operator decides that the number of flights arriving at the airport during peak hours should be 2% lower, then it should levy a congestion charge of around $1,000 (Figure 2). This amount is much lower than the one obtained by Ashley and Savage (2010). While in their paper changes in frequencies are treated as exogenous, we take them into account in the strategic behavior of carriers.

If we look at Daniel (2001), he uses data from 1990 to compute the optimal congestion fees and the effects of their implementation at the airport of Minneapolis-St. Paul (MSP). MSP in 1990 and SFO in 2006 are difficult to compare. In particular, MSP is not considered

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

as congested as SFO. For purposes of comparison, we look at the scenario where MSP is heavily congested, the demand elasticity equals 2, and only considering large commercial airlines (see Daniel (2001), Table 1(b), columns 2 and 6). In his simulation, implementing an optimal congestion fee reduces the number of arriving flights by 5%, and the average congestion fee during the day is $375 ($578 in 2006 dollars). This average charge is far lower than the $2,500 that our model predicts would be needed for a similar reduction in flights (Figure 2). However, while Daniel's result corresponds to the average charge during the day, our congestion charge is only levied in peak hours.

Due to the deterministic relationship between flight frequency and delays (equation 9), the average delay at SFO decreases with imposition of a congestion fee (Figure 3). If $\rho_{\mathcal{H}} = 2,000$, the average delay during peak hours falls by 12.16% (from an average of 28 minutes and 45 seconds to 25 minutes and 15 seconds). The non-linear relationship between flight frequency and delays explains the larger change in the level of congestion.

If we look at the estimates of equation (10), the size of aircraft is positively correlated with demand and negatively correlated with flight frequency. As we will see later, peak demand decreases with the congestion fee. However, since the decrease in the frequency of flights is relatively larger than the decrease in demand during peak hours, the net effect is an increase in the size of aircraft (Figure 4). A congestion charge of $2,000 increases the average size of aircraft operating during peak hours by 9.10% (from 137,940 pounds to 150,492 pounds).

Increasing the congestion charge also leads to higher fares (Figure 5). The impact is very small, but the curve is increasing. For example, $\rho_{\mathcal{H}} = 2,000$ leads to a modest fare growth of 0.30% (from $427.47 to $428.75). The increase is so small because the optimal fares (using (23)) are not directly affected by ρ, but are only affected through changes in the frequency of flights. Our results are in line with those presented by Brueckner (2010). His theoretical model also shows that optimal fares indirectly depend on flight costs through frequencies. Implementing a congestion charge leads to an increase in the costs of operating a flight, but it also increases the number of passengers per aircraft (size increase). Consequently, having more passengers in the plane reduces the impact of the congestion charge on fares.

Note that the demand estimates (Table 9) show that travelers prefer higher frequency of flights and lower fares. Since the congestion charge decreases frequency and increases ticket prices, the average demand for products in peak hours decreases (Figure 6). If $\rho_{\mathcal{H}}$ equals $2,000, then the average demand for peak hours products decreases by 4.14% (from an average of 149.24 purchased tickets per product to 143.06).

Implementing a congestion charge ($\rho_{\mathcal{H}}$) in peak periods also has an impact on off-peak hours. The loss of demand in congested hours is partly diverted to off-peak hours (Figure 7 and Figure 8). While Figure 7 shows an increase in the average demand for flights offered during low congested hours, Figure 8 displays a reduction in the average demand for products offered at SFO independently of the scheduled arrival time. Some travelers may prefer to use alternate airports or other modes of transportation rather than using SFO at less preferred times. Such an increase in off-peak demand leads to a slight growth in the number of flights (Figure 9), an increase in the size of aircraft (Figure 10), and an almost negligible increase in fares.[23]

If we look at equation (15), the weight-based fee depends on the size and the number of flights landing at SFO, and the revenues from travelers using the groundside and terminal services. The reductions in total demand and the number of flights during peak hours push up the weight-based fee, increasing the costs of operating a flight independently of the arrival time (Figure 11).

To conclude this section, if we look at the revenues from the congestion charge (the sum of the number of daily operations during peak hours times the congestion charge), they are increasing with $\rho_{\mathcal{H}}$ (Figure 12). The amount is not negligible: a $\rho_{\mathcal{H}}$ of \$2,000 leads to a quarterly revenue of \$38.39 million. This fact raises a question about what to do with those revenues. According to the DoT, the total revenues from the two-part landing fee cannot exceed the allowable costs of the airfield. If the total landing fee revenues systematically violates this DoT requirement, the landing fee scheme should be modified to include, for instance, a rebate program that preserves the incentive effects from the introduction of the congestion charge. Alternative rebate programs are discussed by Daniel (2001). Among them, the lump-sum rebates per aircraft, rebate per aircraft miles flown, and rebates from reduced fuel taxes also work in our framework. It is important to mention that those programs should be applied at national level (U.S.) rather than local (SFO). If not, the rebates would attract additional aircraft at SFO, reducing the power of the congestion charge. The lump-sum rebate per aircraft would distribute the revenues among all active aircraft operating in the U.S.. The problem with this program is that it would induce airlines to operate some aircraft just for the rebate. A rebate based on the miles flown by aircraft would eliminate this problem. However, this program would introduce some distortions in the aircraft cost per mile. Finally, a rebate from reduced fuel taxes is a similar program, but applied per gallon of fuel rather than per flown mile. We could also consider using the revenues from the

[23]Fares changes in off-peak hours are not reported here but are available upon request.

congestion charge to finance the Next Generation Air Transportation System (NextGen). This is a program that has the objective of updating the U.S. traffic control system from the current ground-based system to a more efficient satellite based system. This technology would reduce traffic delays, increase the capacity of airports, and reduce fuel consumption of aircraft. While the benefits of this technology are clear, its full implementation is being delayed because of its high costs.

8. Conclusions

Using data from the San Francisco International Airport (SFO), our study is the first one that measures the effects of implementing a two-part landing fee on congestion at airports. This new pricing scheme consists of a fee that depends on the weight of aircraft and a congestion charge levied in peak periods. The U.S. Department of Transportation claims that this price mechanism would reduce the level of airport congestion by decreasing the number of landings, but at the same time, the size of aircraft would increase in order to fulfill travel demand. We do not have empirical confirmation of DoT's expectations since no airport has put into practice such a scheme. Instead, airports charge landing fees according to the weight of aircraft without taking into account the level of congestion. Our simulation not only is in line with DoT's expectations, but it also measures the quantitative impact of implementing this scheme in the particular case of SFO. As expected, the higher the congestion charge, the lower are the number of flights arriving during peak hours and the bigger are the size of aircraft. As a result, flight delay decreases. On the other hand, a congestion charge also leads to a reduction in the total demand for products at SFO, and an increase in the weight-based component of the landing fee.

We also make some contributions from a methodological point of view that let us capture important characteristics of the airline industry: first, while in our analysis carriers face two decision variables (ticket prices and frequency of flights), most of the previous literature in applied industrial organization focuses on a single decision variable (price). Second, 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 SFO. The second source of dependence is congestion at SFO. Since changes in the frequency of one product influence flight delays at SFO, all products offered during the same period will be affected even if they do not belong to the same market. Last, we use spatially-based consumer characteristics to capture heterogeneity of travelers.

We acknowledge some limitations of our study. Our application only considers round trip tickets with U.S. domestic passengers landing at SFO. However, landing fees are also paid by international and cargo flights. The lack of data prevents us from including those flights in our estimation procedure. Furthermore, we did not use information about passengers connecting at SFO, but including those travelers in our model is challenging because their preferences are very different from those with SFO as a final destination. Similarly, the strategy of carriers with regard to connecting passengers should also consider the entire route network rather than just the spoke-routes arriving at SFO. We believe that including information about connecting and international flights could weaken our results. Airlines may be reluctant to change the schedule of their flights to avoid increasing the layover time of connecting passengers. The outcome would depend on how valuable the domestic travelers with SFO as a final destination are compared to international or connecting passengers. The lack of data also prevents us from including general aviation operations.[24] Since those flights usually use small aircraft, they are more sensitive to congestion fees than big commercial airplanes. Hence, including information from general aviation will strengthen the impact of implementing a congestion charge. In addition, we only include data from the third quarter of 2006, while the methodology to compute landing fees and rental charges uses yearly information. However, including data from the other quarters will only reinforce our results. For example, the loss of demand as a consequence of the congestion charge $\rho_{\mathcal{H}}$ would be greater if we used four quarters, raising still more the weight-based landing fee. Because of that outcome, the reduction in the frequency of flights arriving at SFO during peak hours would be greater than just considering the 3rd quarter of 2006.

Future research can study a more general-equilibrium type of model, where all U.S. congested airports introduce a two-part landing fee. This type of project would let us assess the impact of the new policy for the whole U.S. economy.

Finally, our model could also be used to investigate the impact of the third amendment of the 1996 "Policy Regarding the Establishment of Airport Rates and Charges". That is, how a congestion charge diverts operations from a congested airport to an underutilized airport owned and operated by the same proprietor. Good candidates could be Los Angeles International and Ontario International airports. Both are owned by Los Angeles World Airports and meet the requirements by the DoT for implementing a congestion fee.

[24] According to the International Civil Aviation Organization (ICAO), *"general aviation is defined as all civil aviation operations other than scheduled air services and non-scheduled air transport operations for remuneration or hire. The general aviation activities are classified into instructional flying, business flying, pleasure flying, aerial work and other flying."*

References

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

[2] Ashley, K. , and Savage, I., (2010): "Pricing Congestion for Arriving Flights at Chicago O'Hare Airport", Journal of Air Transportation Management, 16:36-41.

[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] Blundell, R., and Powell, J. (2004): "Endogeneity in Semiparametric Binary Response Models", Review of Economic Studies, 71:655-679.

[7] 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.

[8] Borenstein, S. (1988): "On the Efficiency of Competitive Markets for Operating Licenses", Quaterly Journal of Economics, 103:357-385.

[9] Brueckner, J. (2010): "'Schedule Competition Revisited", Journal of Transport Economics and Policy, 44(3):261-285.

[10] Brueckner, J. (2002): "Airport Congestion When Carriers have Market Power", The American Economic Review, 92(5):1357-1375.

[11] Brueckner, J., and Spiller, P. (1994): "Economies of Traffic Density in the Deregulated Airline Industry", Journal of Law and Economics, 37(2):379-415.

[12] Daniel, J. and Harback, K. T. (2008): "(When) Do hub airlines internalize their self-imposed congestion delays?", Journal of Urban Economics, 63:583-612.

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

[14] Daniel, J. (1995): "Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues", Econometrica, 63(2):327-370.

[15] 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.

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

[17] Federal Aviation Administration (2008): "Policy Regarding Airport Rates and Charges", FR 73, No 135, July 14, 2008.

[18] Federal Aviation Administration (2008): "Policy Regarding Airport Rates and Charges", FR 73, No 12, January 17, 2008.

[19] Federal Aviation Administration (1996): "Policy Regarding Airport Rates and Charges", FR 61, No 121, June 21, 1996.

[20] Johnson, T. , and Savage, I., (2006): “Departure Delays, the Pricing of Congestion, and Expansion Proposals at Chicago O’Hare Airport”, Journal of Air Transportation Management, 12:182-190.

[21] 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.

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

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

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

[25] 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.

[26] 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.

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

[28] San Francisco International Airport (2010): “Five-Year Capital Plan. Fiscal Year 2010-2011 - Fiscal Year 2014/15”.

[29] Train K. (2009): “Discrete Choice Methods with Simulation”. Cambridge University Press.

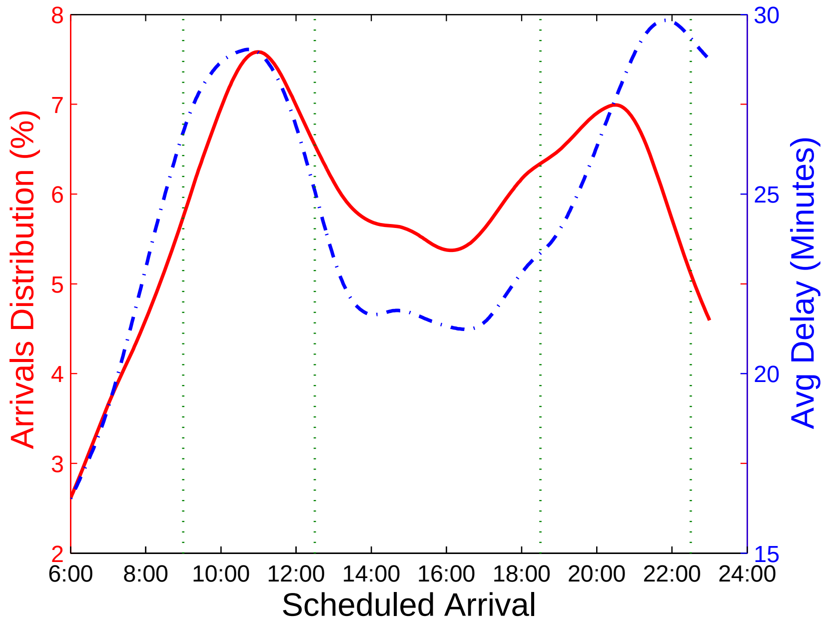
[30] Urdanoz M., and Sampaio R. (2011): “Airlines’ Cooperation in the U.S. Domestic Market: Measuring the Evolution of Price Gaps”, mimeo.

[31] United States Court of Appeals for the District of Columbia Circuit (2010): “On Petition for Review of an Order of the Department of Transportation and the Federal Aviation Administration”, No. 08-1293.

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

[33] Zhang, A. , and Czerny A., (2012): “Airports and Airlines Economics and Policy: An Interpretive Review of Recent Research”, Economics of Transportation, 1:15-34.

FIGURE 1. Arrival Frequency and Delay



- Dash-dotted line: average delay of arriving flights.
- Continuous line: distribution of arrivals during the day.

TABLE 1. Summary Statistics for 3rd Quarter 2006

	Off-Peak		Peak		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Fare (p) ($100)	4.29	2.09	4.27	2.08	4.28	2.09
Nb Passengers	95.97	662.03	149.24	950.81	124.04	827.26
Direct Flight	0.02	0.15	0.03	0.16	0.02	0.15
Daily Flight Frequency ($\hat{f}$)	1.74	0.72	1.98	0.70	1.86	0.72
Distance (1000 miles)	4.51	1.27	4.55	1.23	4.53	1.25
AA	0.17	0.37	0.17	0.37	0.17	0.37
CO	0.05	0.22	0.06	0.24	0.06	0.23
DL	0.17	0.37	0.16	0.36	0.16	0.37
NW	0.08	0.28	0.09	0.28	0.09	0.28
UA	0.38	0.48	0.37	0.48	0.37	0.48
US	0.13	0.34	0.12	0.33	0.13	0.33
Others	0.02	0.15	0.04	0.19	0.03	0.17
Slots	0.30	0.46	0.30	0.46	0.30	0.46
Delay ($\hat{D}$) (minutes)	16.11	1.78	16.99	1.73	16.57	1.81
Nb Products	5,353	-	5,963	-	11,316	-

TABLE 2. Market Average Statistics for 3rd Quarter 2006

	Off-Peak		Peak		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Nb Products	12.25	19.11	13.65	21.73	25.89	40.77
Nb Carriers	3.34	2.10	3.83	2.20	3.31	2.11
Nb Markets	435	-	437	-	437	-

TABLE 3. Supply Statistics for 3rd Quarter 2006 (On Time Performance)

	Off-Peak		Peak		Both	
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**
Daily Frequency (f)	2.64	1.58	2.38	1.42	2.50	1.49
Aircraft MGLW (*weight*) (10^3 pounds)	124.86	79.66	137.94	81.81	132.07	80.87
Total Daily Capacity (10^3 pounds)	290.38	250.18	336.78	311.34	315.97	285.64
Nb Spokes	49	-	56	-	58	-

TABLE 4. Supply Statistics for 3rd Quarter 2006 in Off-Peak Hours (On Time Performance)

	Daily Freq		MGLW (10^3 pounds)		Daily Capacity (10^3 pounds)		Spokes
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	**Nb Spokes**
Off-Peak	2.64	1.58	124.86	79.66	290.38	250.18	49
AA	1.98	1.17	172.81	43.84	323.00	178.95	7
CO	1.43	0.61	146.30	0.60	210.09	90.82	2
DL	1.98	1.75	222.21	88.16	381.24	231.19	4
NW	2.19	1.41	180.44	25.48	369.10	173.05	4
UA	2.94	1.73	110.65	85.52	300.01	300.06	42
US	2.44	1.36	89.53	61.84	167.43	104.31	9
Others	2.70	1.25	112.34	40.25	283.13	155.83	6

Table 5. Supply Statistics for 3rd Quarter 2006 in Peak Hours (On Time Performance)

	Daily Freq		**MGLW (10^3 pounds)**		**Daily Capacity (10^3 pounds)**		**Spokes**
	Mean	**Sd**	**Mean**	**Sd**	**Mean**	**Sd**	**Nb Spokes**
Peak	2.38	1.42	137.94	81.81	336.78	311.34	56
AA	2.28	1.51	186.47	79.12	415.18	288.52	11
CO	2.42	1.59	120.32	37.38	324.53	257.70	4
DL	3.30	1.25	191.82	60.82	638.96	329.14	4
NW	1.89	1.01	144.45	48.62	297.50	205.81	6
UA	2.48	1.57	128.99	88.68	341.43	367.47	46
US	2.35	1.01	110.81	81.00	241.83	173.32	12
Others	1.96	1.32	140.31	70.69	229.22	121.96	8

Table 6. Financial Information for year 2006

	SFO
Landing Fee (\$ per 10^3 pounds of MGLW)	3.213
Enplaned Passengers in 2006 (10^3)	16,574
Operating Revenues (OR) **(\$$10^3$)**	125,656
Groundside Revenues ($GRev$) **(\$$10^3$)**	57,686
$\psi_{terminal}$ **(\$ per passenger)**	7.580
ψ_{ground} **(\$ per passenger)**	3.476
Total Weight in 2006 ($TWeight$) **(10^6 pounds)**	20,095

TABLE 7. Peak Hours 2006 MTC Survey: County - Income (Joint Distribution)

	Nb Obs	Peak	Off-Peak	less $25k	$25k-$50k	$50k-$75k	$75k-$100k	$100k-$150k	$150k-$200k	more $200k
S.Francisco-S.Mateo	1217	0.659	0.677	0.014	0.055	0.097	0.124	0.147	0.095	0.126
Santa Clara	167	0.097	0.086	0.002	0.008	0.008	0.019	0.019	0.020	0.022
Alameda-C.Costa	223	0.127	0.118	0.007	0.009	0.015	0.027	0.022	0.029	0.019
Solano-Napa	59	0.029	0.036	0.000	0.007	0.008	0.001	0.007	0.002	0.004
Sonoma-Marin	156	0.088	0.083	0.003	0.012	0.017	0.019	0.011	0.013	0.013
Total	1,822	1	1	0.027	0.091	0.145	0.189	0.206	0.159	0.185

TABLE 8. Off-Peak Hours 2006 MTC Survey: County - Income (Joint Distribution)

	Nb Obs	Peak	Off-Peak	less $25k	$25k-$50k	$50k-$75k	$75k-$100k	$100k-$150k	$150k-$200k	more $200k
S.Francisco-S.Mateo	1217	0.659	0.677	0.022	0.059	0.100	0.132	0.138	0.101	0.124
Santa Clara	167	0.097	0.086	0.003	0.003	0.013	0.016	0.020	0.013	0.017
Alameda-C.Costa	223	0.127	0.118	0.007	0.008	0.013	0.016	0.035	0.021	0.019
Solano-Napa	59	0.029	0.036	0.000	0.001	0.008	0.004	0.005	0.010	0.008
Sonoma-Marin	156	0.088	0.083	0.004	0.007	0.016	0.014	0.024	0.011	0.007
Total	1,822	1	1	0.036	0.077	0.150	0.183	0.222	0.156	0.174

Table 9. Demand Estimates

	OLS	**IV**	**RCM**
Intercept	-10.785**	-10.536**	-6.924**
	(0.170)	(0.226)	(0.272)
Fare (α_p)	-0.078**	-0.606**	-1.236**
	(0.006)	(0.057)	(0.079)
Distance (1000 miles)	0.864**	1.423**	1.692**
	(0.050)	(0.081)	(0.098)
Distance Squared	-0.125**	-0.169**	-0.191**
	(0.006)	(0.009)	(0.011)
Frequency (α_f)	0.061**	0.340**	0.384**
	(0.021)	(0.028)	(0.032)
Direct	4.305**	4.463**	4.413**
	(0.091)	(0.117)	(0.108)
Peak Hours (α_{peak})	0.490**	0.421**	0.430**
	(0.026)	(0.034)	(0.038)
CO	-0.173**	0.367**	0.418**
	(0.062)	(0.093)	(0.101)
DL	-0.070	-0.101	-0.255**
	(0.046)	(0.062)	(0.074)
NW	-0.248**	-0.212**	-0.259**
	(0.055)	(0.073)	(0.078)
UA	-0.342**	0.355**	0.374**
	(0.037)	(0.070)	(0.081)
US	-0.418**	-0.238**	-0.225**
	(0.050)	(0.064)	(0.066)
Others	0.666**	0.738**	0.714**
	(0.080)	(0.103)	(0.112)
Slots	-0.315**	-0.233**	-0.279**
	(0.032)	(0.043)	(0.048)
Delay (α_d)	-0.108**	-0.126**	-0.409**
	(0.011)	(0.014)	(0.017)
Random Constant (σ^0)	-	-	-0.874**
	(-)	(-)	(0.145)
Random Price (σ^p)	-	-	-0.066**
	(-)	(-)	(0.004)
Distance to Airports (λ)	-	-	-0.283**
	(-)	(-)	(0.081)
Random Price-Income (α_y)	-	-	0.054**
	(-)	(-)	(0.001)
Random Frequency (σ^f)	-	-	0.122
	(-)	(-)	(0.131)
Random Delay (σ^d)	-	-	-0.155**
	(-)	(-)	(0.001)
Sfo-Mateo	-	-	1.129**
	(-)	(-)	(0.249)
Sta Clara	-	-	-1.564**
	(-)	(-)	(0.128)
Alameda-Costa	-	-	0.142
	(-)	(-)	(0.418)
Sonoma-Marin	-	-	-0.448
	(-)	(-)	(0.329)
Nb Observations	11,316	11,316	11,316
J-Statistic	-	-	44.54
Sargan Statistic	-	170.36	-
R^2	0.3376	-	-
Degrees of Freedom	-	4	33
χ^2 Critical Value (5%)	-	9.49	47.40
χ^2 Critical Value (1%)	-	13.28	54.78

** Significant at the 5 percent level.

Table 10. Price Elasticities Estimates

	Off-Peak	**Peak**	**Both**
All Carriers	-3.43	-3.44	-3.44
	(1.03)	(1.04)	(1.04)
AA	-3.38	-3.39	-3.38
	(0.91)	(0.93)	(0.92)
CO	-3.76	-3.77	-3.76
	(1.03)	(1.06)	(1.04)
DL	-3.07	-3.07	-3.07
	(1.11)	(1.12)	(1.12)
NW	-3.25	-3.24	-3.24
	(0.81)	(0.81)	(0.81)
UA	-3.66	-3.70	-3.68
	(1.03)	(1.06)	(1.04)
US	-3.35	-3.33	-3.34
	(0.85)	(0.86)	(0.85)
Others	-2.96	-3.04	-3.01
	(0.92)	(0.89)	(0.90)

Table 11. Frequency Semi-Elasticities of Demand

	Off-Peak	**Peak**	**Both Airports**
Mean	7.468	7.183	7.319
Median	7.206	6.985	7.088
Std	3.181	3.186	3.187

Table 12. Weight Estimates (TSLS)

Weight (10^3 pounds)	**Estimate**
Intercept (τ_0)	95.185** (17.826)
Daily Pax (τ_1)	0.439** (0.054)
Daily Frequency (τ_2)	-12.974** (3.748)
Distance (τ_3) (1000 miles)	17.295** (2.899)
CO	-79.079** (23.640)
DL	21.191 (21.349)
NW	-7.819 (19.750)
UA	-5.914 (13.510)
US	-13.880 (16.671)
Other	-5.772 (18.223)
I_l^{peak} (τ_5)	-7.498 (8.059)
Nb Observations	165
Sargan Statistic	4.648
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 mile	**Lerner Index**
All Flights	0.06	0.31
Off-Peak	0.06	0.31
Peak	0.06	0.31
Direct Flights	0.09	0.35
Connecting Flights	0.06	0.31
AA	0.06	0.31
CO	0.06	0.28
DL	0.05	0.34
NW	0.05	0.33
UA	0.08	0.29
US	0.06	0.31
Others	0.05	0.35

Table 14. Marginal Cost Estimates

mc ($100)	Estimate
Intercept	0.246
	(0.237)
Peak	-0.007
	(0.060)
Distance (1000 miles)	0.818**
	(0.116)
Distance2	-0.069**
	(0.015)
Hub	0.142
	(0.085)
Nb Connections	0.098**
	(0.048)
CO	0.673**
	(0.162)
DL	-0.379**
	(0.104)
NW	-0.149
	(0.146)
UA	0.679**
	(0.102)
US	-0.029
	(0.113)
Others	-0.235
	(0.188)
Nb Observations	11,316

** Significant at the 5 percent level.

TABLE 15. Total Cost Frequency Estimates

$100	**Estimate**
Intercept	74.78**
	(9.21)
β^d	-3.25**
	(0.21)
Distance (1000 miles)	7.88**
	(3.05)
Distance2	-1.06
	(0.73)
CO	15.22
	(8.86)
DL	10.21
	(7.97)
NW	6.21
	(7.48)
UA	31.58**
	(9.40)
US	22.02**
	(6.12)
Others	9.47
	(6.76)
Nb Observations	165
R^2	0.909

** Significant at the 5 percent level.

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

$F\hat{C}ost$ **($)**	**Off-Peak**	**Peak**	**Both**
Mean	10,606	10,547	10,573
Median	10,928	10,874	10,921
Std	1,136	1,156	1,145

FIGURE 2. Δ Total Daily Flights - Peak (%)

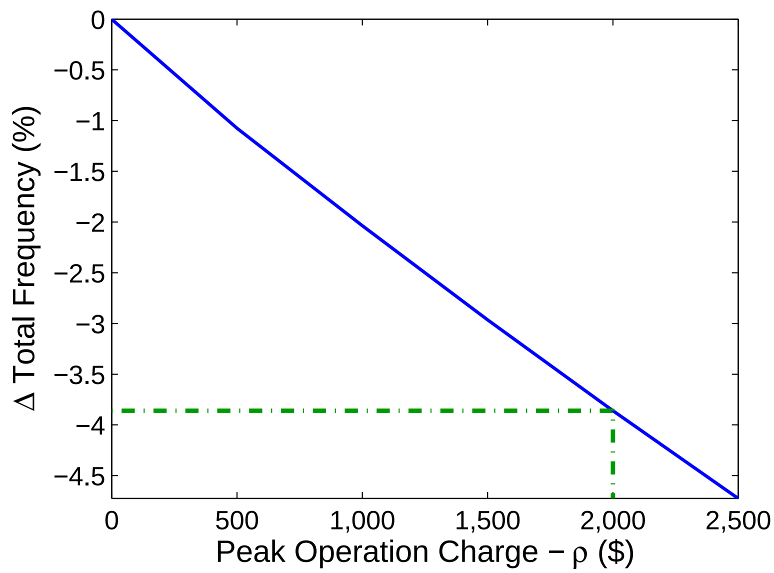


FIGURE 3. Δ Average Delay - Peak (%)

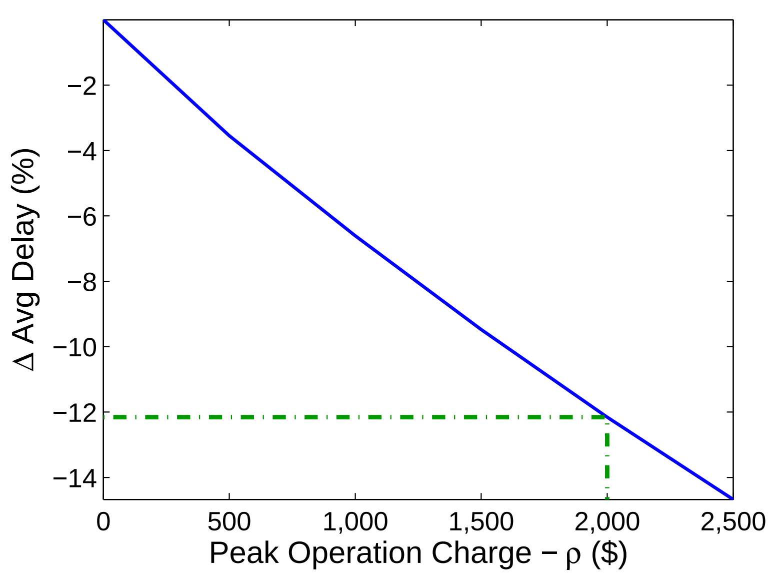


Figure 4. Δ Aircraft Average Weight - Peak (%)

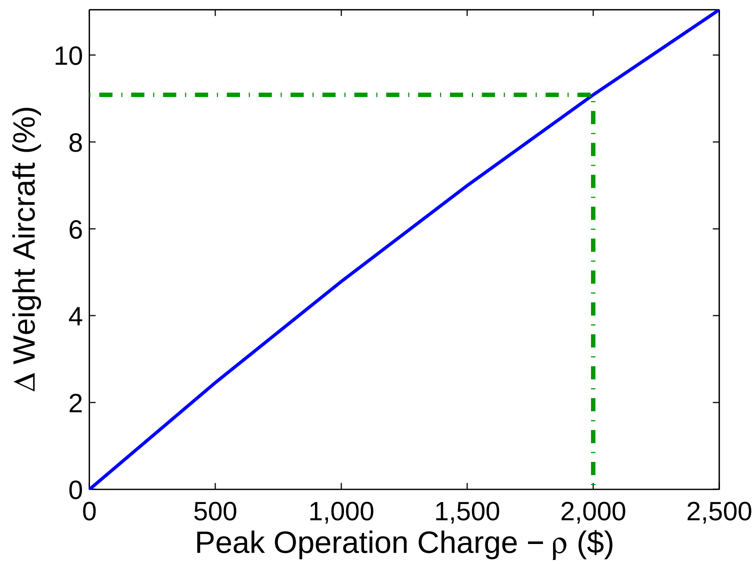


Figure 5. Δ Average Fare - Peak (%)

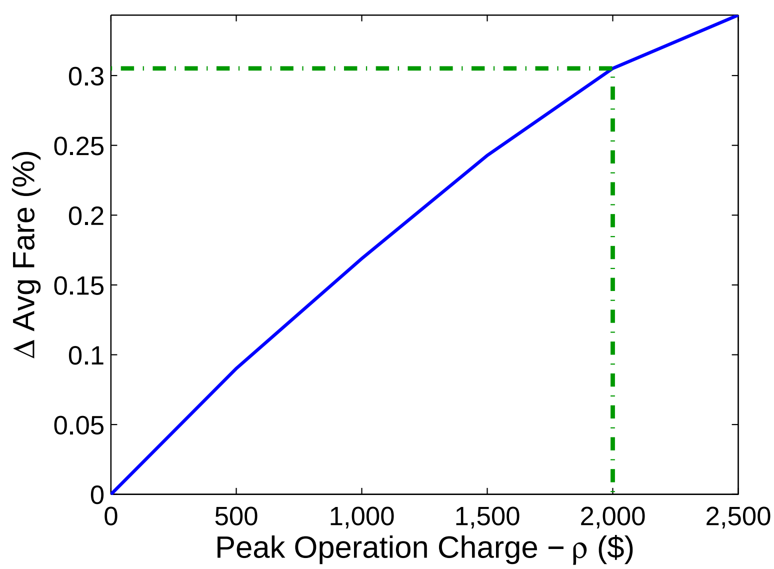


Figure 6. Δ Average Demand - Peak (%)

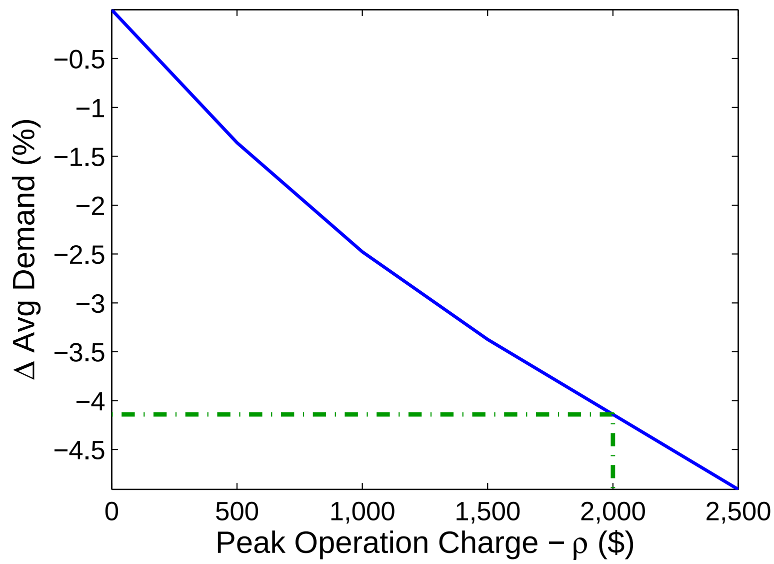


Figure 7. Δ Average Demand - Off-Peak (%)

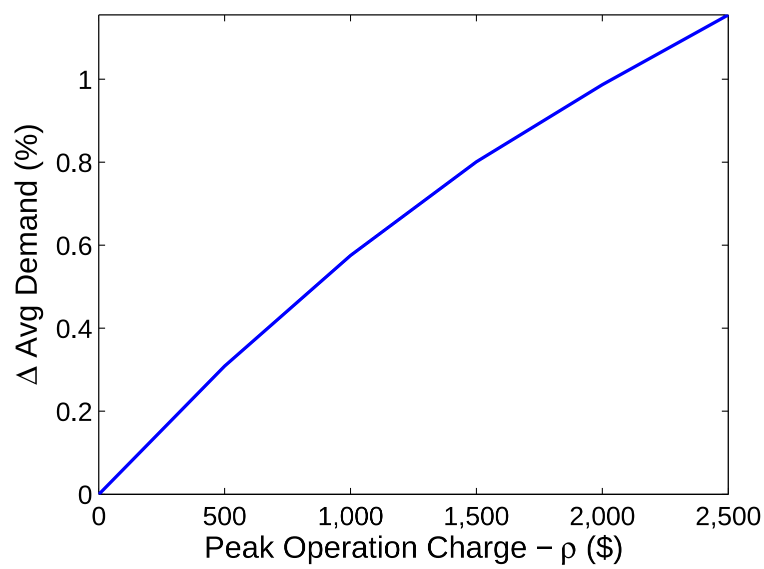


FIGURE 8. Δ Average Demand - SFO (%)

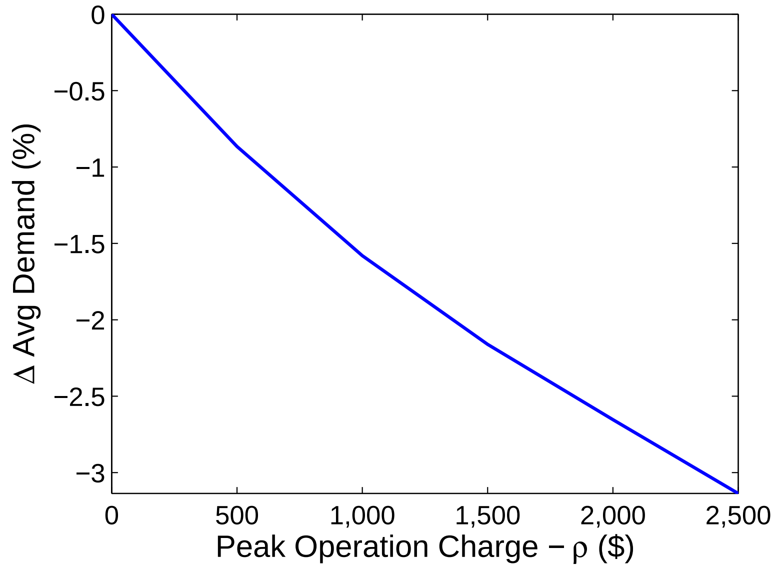


FIGURE 9. Δ Total Daily Flights - Off-Peak (%)

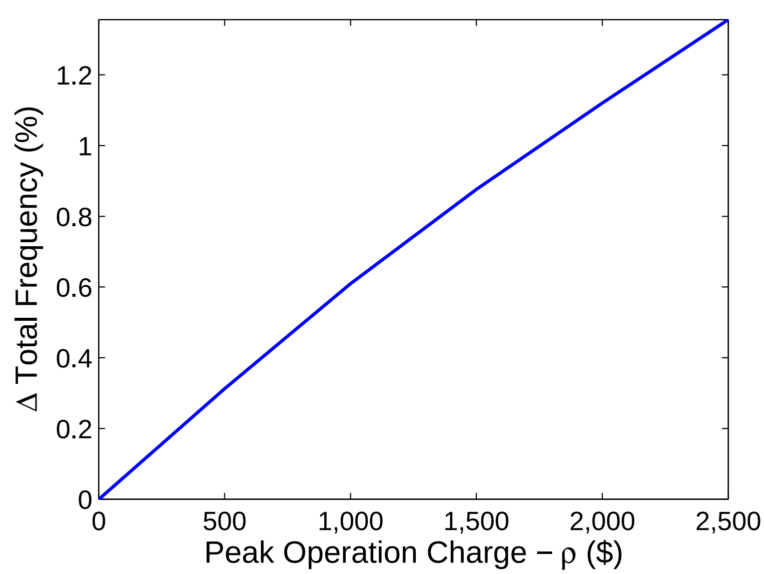


Figure 10. Δ Aircraft Average Weight - Off-Peak (%)

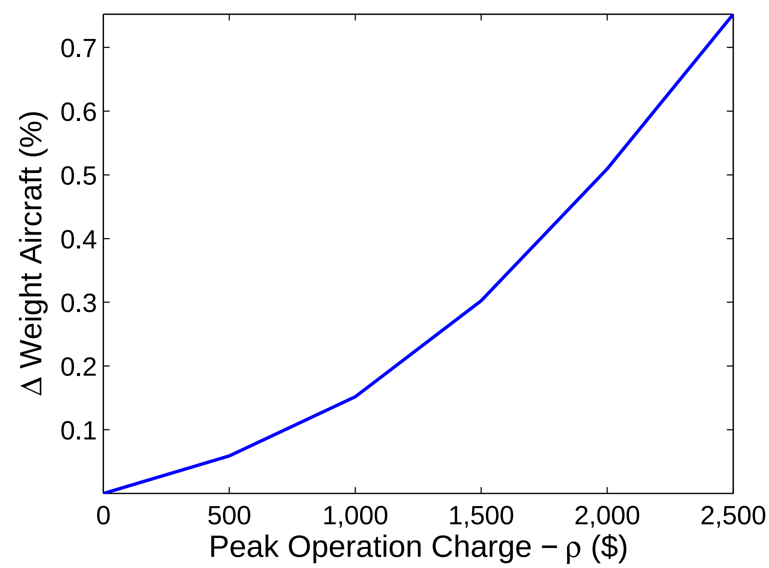


Figure 11. Weight-Based Fees ($)

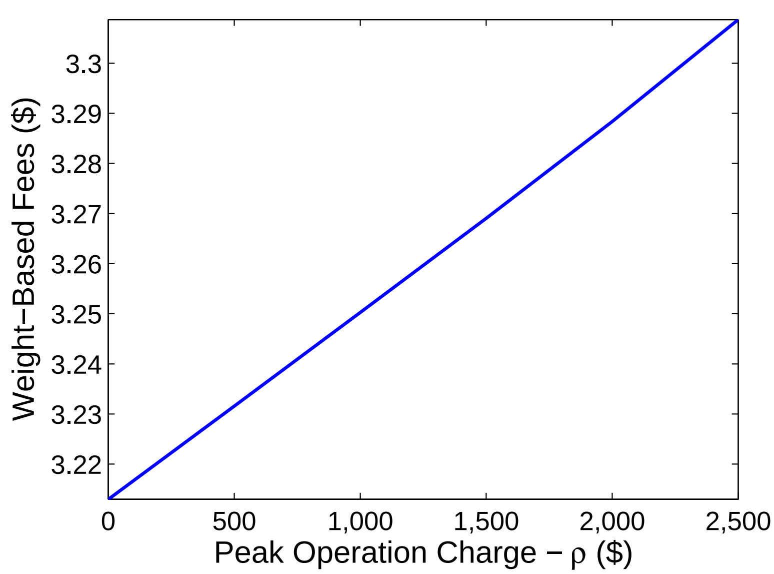
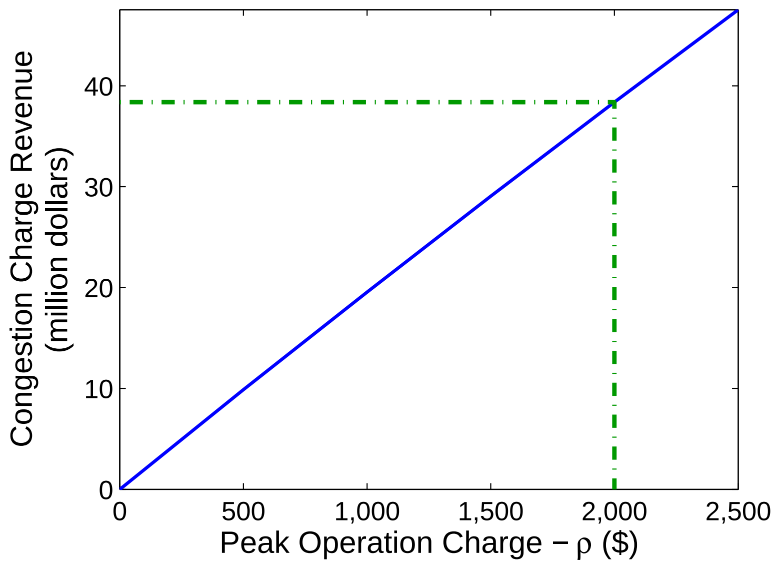


FIGURE 12. Congestion Charge Revenues (million dollars)



9. Appendix

9.1. **Description of Data Sources:**

9.1.1. *Choice and Flight Characteristics Sources:*

- Airline Origin and Destination Survey (DB1B): a 10 % sample of all passengers traveling within the U.S., with detailed ticket information such as the operating and ticketing carrier of each coupon, origin and destination airports, the airports in which the passenger made a connection, if any, and the fare.
- Airline On-Time Performance Data: provides information about on-time, delay, and daily frequency of flights for non-stop domestic flights by major air carriers.
- Schedule B-43 Aircraft Inventory: contains information about aircraft operating in the U.S.. In particular, it details characteristics for each plane such as the owner, model, tail number (civil registration number), and acquisition date.
- T-100 database: has information on frequency of flights, total number of passengers and type of aircraft for all segments in the U.S..

9.1.2. *Airport Financial Reports:* We use financial details that airports report to the FAA (summary report form 127) to obtain information about the airfields and terminals' operating revenues. Such information is useful when we introduce landing fees and terminal rental charges in our model. The amounts and the methodology to determine the landing fee and the rental charge are obtained from the 2006 Annual Operating Budget document for the San Francisco International Airport.

9.1.3. *American Community Survey (ACS):.* The ACS is a household survey developed by the 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. 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 household income and distance to the airport.

9.1.4. *2006 Airline Passenger Survey (2006 MTC):.* This survey gives detailed information about travelers using the San Francisco International Airport (SFO) and is provided by the Metropolitan Transportation Commission of the San Francisco Bay. The survey contains: household income, location of the traveler in the Bay Area, airport of origin/destination, carrier, transportation access to the airport, and departure time. 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 the household income and travelers distance to SFO conditional on arriving during peak or off-peak hours.