Tracing the Woes: An Empirical Analysis of the Airline Industry*

Steven Berry[†] Panle Jia[‡]

This version: June, 2008

Abstract

The U.S. airline industry went through tremendous turmoil in the early 2000's. There were four major bankruptcies and two major mergers, with all legacy carriers reporting a large profit reduction. This paper presents a structural model of the airline industry, and estimates the impact of demand and supply changes on profitability. We find that, compared with the late 1990s, in 2006, a) air-travel demand was more price sensitive; b) passengers displayed a strong preference for direct flights, and the connection semi-elasticity was much higher; c) the changes of marginal cost significantly favored direct flights. These findings are present in all the specifications we estimated. Together with the expansion of low cost carriers, they explained more than 80% of the decrease in legacy carriers' variable profits, with changes in demand contributing to more than 50% of the reduction.

1 Introduction

The airline industry went through tremendous turmoil in the early 2000's with four major bankruptcies and two mergers. In August 2002, US Airways filed for bankruptcy. A year later, United Airlines followed suit. It stayed under Chapter 11 bankruptcy protection for more than three years, the largest and longest airline bankruptcy in history. In September 2005, Delta

*We have benefited from discussions with Glenn Ellison, Michael Greenstone, Jerry Hausman, and Marc Rysman. Special thanks go to Nancy Rose, who is extremely generous with her time and help. Michael Powell provided excellent research assistance. Comments are welcome.

†Department of Economics, Yale University. and NBER. Email: steven.berry@yale.edu.

‡Department of Economics, M.I.T. and NBER. Email: pjia@mit.edu.

Airlines and Northwest Airlines went bankrupt on the same day. By then, four of the six legacy carriers were under bankruptcy reorganization.[1] Only American and Continental managed to escape bankruptcy, but all legacy carriers reported a large reduction in profits. In the spring of 2008, four low cost carriers declared bankruptcy and discontinued their passenger service operations within a week.[2]

On the other hand, when measured by revenue passenger miles,[3] the industry's output had recovered from the sharp downturn after 9/11 by 2004 and has been trending up since (see Figure 1). The load factor,[4] another important measure of profitability, has increased steadily since 2001. According to Figure 2, the average load factors for U.S. airlines rose from 71.2% in 1999 to 79.7% in 2006, and posted a record high of 80.3% in the first half of 2007. If more passengers traveled and planes were fuller, what caused the financial stress of most airlines?

Several recent developments provide potential explanations. One category of explanations is related to changes in air-travel demand. Perhaps the bursting of the dot-com bubble, or improvements in electronic communications, decreased the willingness-to-pay of business travelers. As the economy cooled down, many companies imposed maximum reimbursement limits, and even business travelers started to shop around for cheaper flights.

Another potential change in demand stems from the tightened security regulations after 9/11. Passengers had to go through a strict security check, and many items were no longer allowed in carry-on luggage. The extra luggage handling, combined with stricter regulations, had lengthened the average connection time. At the same time, with most flights full, it became increasingly difficult for passengers to board a different plane in case of missed connections or flight cancellations. Consequently, carriers found it harder to charge high fares for connecting flights as passengers started to search for alternatives.[5]

The third important development is the option of purchasing airline tick-

[1] The legacy carriers are: American Airlines, Continental Airlines, Delta Airlines, Northwest Airlines, United Airlines, and U.S. Airways.

[2] These four low cost carriers are: Aloha Airlines, March 31st, 2008; ATA Airlines, April 2nd, 2008; Skyway Airlines, April 5th, 2008, and Skybus, April 6th, 2008. Frontier filed for reorganization under Chapter 11 on April 10th, 2008.

[3] Revenue passenger miles is the product of the number of revenue-paying passengers aboard and the distance traveled measured in miles.

[4] Load factor is the ratio of revenue passenger miles to available seat miles of a flight.

[5] For example, in announcing a 26% capacity cut at the Cincinnati hub, Delta said that connecting traffic was the least profitable for the carrier. (Source: Business Courier, September 7, 2005.)

ets on the internet. In 1997, all of Continental tickets were sold through the airline's reservation office or traditional travel agencies. By 2006, more than 40% of Continental's domestic tickets were sold through the internet.[6] The proliferation of various online sites that provided information previously limited to travel agents made consumers much more conscious of the fare availability and fare premiums across carriers. There were even websites like 'farecast.com' that predicted the fare trajectory for the near future. All of these changes were likely to affect both passengers' price sensitivity and their preference for flights with different attributes (direct vs. connecting, frequent vs. less frequent departures, etc.) For example, Brunger (2007) argued that the internet helped airlines to increase the average load factors by shifting consumers from crowded weekend flights to Tuesday and Wednesday flights that traditionally had fewer passengers.

On the supply side, a variety of changes affected the industry's market structure and profitability. The most cited transition was the expansion of the low cost carriers (LCC), whose market share of domestic origin-destination passengers increased steadily over the past decade, from 22.6% in 1999 to 32.9% in 2006.[7] As a result, the legacy carriers were forced to lower their fares and offer competing service. Many legacy carriers shifted their capacity to the more lucrative international markets, and reluctantly surrendered some of the domestic markets to the low cost carriers.

The recent aviation technology progress, in particular the advent of regional jets with different plane sizes, allowed carriers to better match the aircraft with the market size, and hence enabled carriers to offer direct flights to markets that used to rely on connecting services. In addition, with lower labor costs than the traditional jets, regional jets became a popular choice for carriers under financial pressure.[8] On the other hand, the cost of jet fuel, which accounts for 10-30% of the operation cost, more than doubled over the past decade.[9]

In this paper, we estimate a structural model of the airline industry, and disentangle how the various factors affect the profitability of the legacy carriers. We find that, compared with the late 1990s, in 2006, a) air-travel demand was more price sensitive; b) passengers displayed a strong preference for direct flights, and the connection semi-elasticity was much higher; c) the changes of marginal cost significantly favored direct flights. These findings

[6] See Bill Brunger (2007).

[7] Data source: http://www.darinlee.net/data/lccshare.html.

[8] See Mozdzanowska (2004).

[9] See Morrison & Winston (2005) for a discussion regarding the effect of fuel on carriers' profits.

are present in all specifications we estimated. These factors, together with the expansion of low cost carriers, explained more than 80% of the decrease in legacy carriers' variable profits, with changes in demand contributing to more than 50% of the reduction.

The remainder of the paper is structured as follows. Section 2 reviews the related literature. Section 3 presents the model. Section 4 describes the data sources. Section 5 proposes the empirical strategy. Section 6 discusses the results. Section 7 presents the conclusions.

2 Literature review

There have been many empirical papers that study the airline industry. Among the most recent ones, Borenstein (2005) reported that airline prices fell more than 20% adjusted for inflation from 1995 to 2004. He also found that premiums at hub airports declined, and that there was substantially less disparity between the cheaper and the more expensive airports than there had been a decade ago. Goolsbee and Syverson (2005) examined how incumbents responded to the threat of Southwest entry. Puller, Sengupta, and Wiggins (2007) tested theories of price dispersion and scarcity pricing in the airline industry. Ciliberto (2008) analyzed dynamic strategic deterrence in the airline industry. Dana and Orlov (2008) studied the impact of the internet penetration on airlines' capacity utilization. Forbes (2008) exploited a legislative change in takeoff and landing restrictions at LaGuardia Airport in 2000. She discovered that prices fell by $1.42 on average for each additional minute of flight delay.

There are only a few discrete choice applications in the airline literature. Peters (2006) simulated post-merger prices for five airline mergers in the late 1980s, and found evidence that supply-side effects, such as changes in marginal costs and deviations from the assumed model of firm conduct, were important factors in post-merger price increases. Berry, Carnall, and Spiller (hereafter BCS) (2007) focused on the evolution of the airline industry toward a hub-and-spoke system after the deregulation in 1970s. They found evidence of economies of density on longer routes. Armantier and Richard (2008) investigated the consumer welfare consequences of the code-share agreement between Continental Airlines and Northwest airlines. The results suggested that the code-share agreement increased the average surplus of connecting passengers, decreased the average surplus of nonstop passengers, and did not impact consumers significantly on average. We contribute to the literature by examining the developments in demand and supply from the late 1990s to the late 2000s that affected profitability in the airline industry.

3 Model

We consider a model of airline oligopoly "supply and demand" in the spirit of the recent literature on differentiated product markets following Berry, Levinsohn and Pakes (BLP) (1995). Our model is particularly close to BCS. The point of the present paper is not to provide any methodological innovation, but to make use of the existing models to understand the recent evolution of the industry.

For now, we think of U.S. airlines as offering a set of differentiated products in each of a large cross-section of "origin-and-destination" markets. Airline products are differentiated by price, by direct versus connect, by airline brand and so forth. Ticket restrictions (such as advanced-purchase and length-of-stay restrictions) are an important part of product differentiation that are not observed in our data. Neither are certain important flight-level details, such as the time of departure. Thus, there is a particularly important role for product-unobservable characteristics that are correlated with price.

3.1 Demand

The demand model is a simple random-coefficient discrete-choice model in the spirit of McFadden (1981) and BLP. Like BCS, we use a "discrete types" version of the random coefficient utility model. In the base specification, we assume that consumers are one of two types, tourists or business travelers.[10] For product j in market t, the utility of consumer i, who is of "type" r, is given by

$$u_{ijt} = x_{jt}\beta_r - \alpha_r p_{jt} + \xi_{jt} + \nu_{it}(\lambda) + \lambda\epsilon_{ijt}, \tag{1}$$

where

- x_{jt} is a vector of product characteristics,
- p_{jt} is the product price,
- β_r is the vector of "tastes for characteristics" for consumers of type r,
- α_r is the marginal disutility of a price increase for consumers of type r,

[10] In principle, if the data are rich enough, one could think of estimating a large number of types, R. See section 5.1 for more discussions.

- ξ_{jt} is the unobserved (to researchers) product characteristic of product j,
- ν_{it} is a "nested logit" random taste that is 1) constant across airline products; 2) differentiates "air travel" from the "outside" good,
- λ is the nested logit parameter, and
- ϵ_{ijt} is an i.i.d. (across products and consumers) "logit error."

The utility of the outside good is given by

$$u_{iot} = \epsilon_{i0t} \tag{2}$$

where ϵ_{i0t} is another logit error. The error structure

$$\nu_{it}(\lambda) + \lambda\epsilon_{ijt}$$

is assumed to follow the distributional assumption necessary to generate the classic "nested logit" purchase probability for consumers of type r, where the two "nests" consist of 1) all the airline products, and 2) the outside option of not flying. If $\lambda = 1$, then $\nu \equiv 0$, and the purchase probability of type r consumers takes the simple multinomial logit form. If $\lambda = 0$, then the i.i.d. ϵ's have no effect and all type r consumers buy the "best" product with probability one. When $\lambda \in (0, 1)$, the product shares have the traditional nested logit form.

Specifically, conditional on purchasing some airline product, the percentage of type r consumers who purchase product j in market t is given by

$$\frac{e^{(x_{jt}\beta_r - \alpha_r p_{jt} + \xi_{jt})/\lambda}}{D_{rt}}$$

where the denominator is

$$D_{rt} = \sum_{k=1}^{J} e^{(x_{kt}\beta_r - \alpha_r p_{kt} + \xi_{kt})/\lambda} \tag{3}$$

Given this specification, the "within market" share of type r consumers is then

$$s_t^r(x_t, p_t, \xi_t, \theta_d) \equiv \frac{D_{rt}^{\lambda}}{1 + D_{rt}^{\lambda}}. \tag{4}$$

Let γ_t denote the percentage of type r consumers in the population. The overall market share of product j in market t is

$$s_{jt}(x_t, p_t, \xi_t, \theta_d) \equiv \sum_r \gamma_r \frac{e^{(x_{jt}\beta_r - \alpha_r p_{jt} + \xi_{jt})/\lambda}}{D_{rt}} s_t^r(x_t, p_t, \xi_t, \theta_d). \tag{5}$$

Notice that the vector of demand parameters to be estimated, θ_d, includes the tastes for product characteristics, β_r, the disutility of price, α_r, the nested logit parameter, λ (which governs substitution to the outside good), and the consumer-type probabilities γ_r. With two types of consumers, there is only one γ parameter, since the probability of the other type is $(1-\gamma)$.

Following BLP, we form moments that are the expectations of the unobservable ξ interacted with exogenous instruments that are discussed below. Further details of the estimation method are found in BLP and the related literature, but we provide a brief review here.

We first invert the market share equations (5) to solve for the vector of demand unobservables ξ_t, as a function of the parameters and the observed data on market shares, prices and characteristics.

$$\xi_t = s^{-1}(x_t, p_t, s_t, \theta_d) \tag{6}$$

As in BCS, the multiple-type nested logit model requires us to slightly modify the "contraction mapping" method used in BLP. In particular, the "step" between each iteration is multiplied by λ, the nested logit parameter:

$$\xi_{jt}^M = \xi_{jt}^{M-1} + \lambda \left[\ln s_{jt} - \ln s_{jt}(x_t, p_t, \xi_t, \theta_d)\right]$$

where M denotes the Mth iteration, s_{jt} is the observed share, and $s_{jt}(x_t, p_t, \xi_t, \theta_d)$ is defined by equation (5).

The moment conditions used in estimation are based on restrictions of the form

$$E(\xi(x_t, p_t, s_t, \theta_d) \mid z_t) = 0, \tag{7}$$

where z_t is a vector of instruments. A classic GMM estimation routine notes that these moment conditions imply

$$E(h(z_t)\xi(x_t, p_t, s_t, \theta_d)) = 0, \tag{8}$$

for any vector of functions $h(\cdot)$. Intuitively, a method of moments estimation routine chooses θ_d to make sample analogs of the expectations in (8) as close to zero as possible.

The product-level unobservable ξ_{jt} accounts for a large number of product characteristics, such as ticket restrictions and departure time, that are absent from our data source. Therefore, it is especially important for us to allow for a correlation between ξ_{jt} and price, and instrument prices. We also allow for the possible endogeneity of "flight frequency" on a route. As we cannot allow for the full endogeneity of all product characteristics, we treat a number of product characteristics (such as "distance", and "direct vs. connect") as exogenous.

Obviously, the instrument set must include exogenous variables that help to predict endogenous characteristics (prices and flight frequencies). The instruments also have to identify the parameters that govern substitution patterns across products in a market. These parameters include the type specific taste parameters, λ, and the shares of each type γ_r. Intuitively, exogenous variation in choice sets across markets greatly helps to identify substitution patterns.[11] Our specific choice of demand instruments is considered in section 5.2, after we introduce the data in more detail.

3.2 Markups and Marginal Cost

We assume that prices are set according to a static Nash equilibrium with multi-product firms. Again following BLP, we compute equilibrium markups from knowledge of the demand data and parameters. Let $b_{jt}(s_t, x_t, p_t, \theta_d)$ denote these markups. Marginal cost of product j in market t is:[12]

$$mc_{jt} = p_{jt} - b_{jt}(s_t, x_t, p_t, \theta_d) \tag{9}$$

We posit a somewhat simpler version of marginal costs as compared to BCS. The marginal cost function is given by

$$mc_{jt} = \mathrm{w}_{jt}\psi + \omega_{jt} \tag{10}$$

where

[11] Berry and Haile (2008) consider this argument more formally.

[12] The markup equation in matrix form is:

$$MC = P + \left(\frac{\partial Q}{\partial P}\right)^{-1} Q$$

where $Q = (q_{1t}, ..., q_{J_f,t}) = (s_{1t}, ..., s_{J_f,t}) * M_t$, $\left(\frac{\partial Q}{\partial P}\right) = \begin{pmatrix} \frac{\partial q_{1t}}{\partial p_{1t}} & ... & \frac{\partial q_{J_f,t}}{\partial p_{1t}} \\ & ... & \\ \frac{\partial q_{1t}}{\partial p_{J_f,t}} & & \frac{\partial q_{J_f,t}}{\partial p_{J_f,t}} \end{pmatrix}$. J_f is the number of products by firm f in market t, M_t is the market size, and s_{jt} is defined by equation (5).

- w_{jt} is a vector of observed cost-shifters,
- ω_{jt} is an unobserved cost shock and
- ψ is a vector of cost parameters to be estimated.

Equation (9) and (10) imply that the cost-side unobservable is the difference between prices, the markups, and the deterministic part of marginal cost:

$$\omega_{jt} = p_{jt} - b_{jt}(s_t, x_t, p_t, \theta_d) - \mathrm{w}_{jt}\psi \tag{11}$$

As with demand, we form moments that are the expectations of the cost-side unobservable ω interacted with cost-side instruments:

$$E(h(z_t)\omega(x_t, p_t, s_t, \theta, \psi)) = 0, \tag{12}$$

where z_t is a vector of instruments. These instruments can include:

- exogenous elements of the marginal-cost shifters, w,
- exogenous demand-side instruments that help to predict the markup term, $b_{jt}(\cdot)$, that enters the pricing equation.

In addition to estimating the marginal cost parameter ψ, the supply side restrictions in (12) also help to estimate the demand parameters, because these parameters enter the markup term. Once again, we leave a detailed discussion of the choice of instruments until after a more detailed discussion of the data. Note, however, that nothing in the estimation method allows us to estimate fixed costs.

4 Data

There are three main data sources for this study. The Airline Origin and Destination Survey (DB1B), published by the U.S. Department of Transportation (DOT), provides detailed information on the fare, itinerary (origin, destination, and all connecting airports), the ticketing and operating carrier for each segment, and the number of passengers traveled on the itinerary at the given fare in a quarter.[13] The flight frequency is constructed using the scheduling data from Back Aviation Solutions, Inc. Flight delays are extracted from the Airline On-Time Performance Data, also published by DOT. In the following, we explain our market definition and sample selection. See the appendix for further details.

[13] The URL is (as of April, 2008): http://www.transtats.bts.gov/DataIndex.asp.

4.1 Sample selection

The DB1B data is a 10% random sample of airline tickets from U.S. reporting carriers. Following Brueckner and Spiller (1994), and BCS, we kept round-trip itineraries within U.S. continent with at most four segments. We eliminated tickets cheaper than $25, with multiple ticketing carriers, or containing the ground traffic as part of the itinerary.

A market is defined as a directional pair of an origin and a destination airport. For example, Atlanta - Las Vegas is a different market from Las Vegas - Atlanta. This allows for the characteristics of the origin city to affect demand. As in BCS, market size is the geometric mean of the Metropolitan Statistical Area population of the end-point cities.[14]

We focused on airports located in medium to large metropolitan areas with at least 850,000 people in 2006. There were 3,998 such markets in 1999 and 4,300 markets in 2006. These markets accounted for around 80% of total passengers, and roughly overlapped with the top 4000 most traveled markets, which is the scope of focus in many empirical studies.[15]

There are two main reasons for excluding small markets. The first one is computational: the estimation time increases substantially with the number of markets and products. The small airports accounted for only one-fifth of the passengers, but they constituted around three-quarters of the markets and a third of products. The main reason for excluding small markets, however, is the drastic difference between large and small markets. Even within our selected sample, the number of passengers and revenues in the largest markets are hundreds of times larger than the smallest markets. As the demand pattern and the operation cost are likely to be different among markets with diverse sizes, it is difficult for our stylized model to explain all these differences.

Six groups of airports are geographically close.[16] Carriers in nearby airports might compete against each other as consumers can choose which airport to fly from. In one of our specifications, we group these nearby airports, and define a market based on the grouped airports.

In 2006, our sample contains 700,000 unique records, or 163 records per

[14] Data source (as of April 2008) for the MSA population: http://www.census.gov/population/www/estimates/CBSA-est2006-annual.html.

[15] For example, the Government Accounting Office (GAO) focuses on the top 5,000 most traveled markets in their annual report of the airline industry.

[16] The six groups of airports are: Dallas-Ft Worth Intl and Love Field in Dallas TX, Dulles and National in D.C., Midway and O'Hare in Chicago IL, Kennedy, La Guardia, and Newark in New York NY, Los Angeles, Burbank, and Long Beach in Los Angeles CA, San Francisco, Oakland, and San Jose in San Francisco CA.

market. In comparison, BCS reported 9 records per market using the 1985's DB1B data. Given that the product shares need to be inverted in every market at each iteration, both the memory requirement and the estimation time increase substantially with a large number of products. In addition, conditioning on observed characteristics, many observations have very similar fares (for example, a \$325 ticket and a \$328 ticket), and are not likely to be viewed by consumers as distinctive products. Therefore, we aggregate the records using a set of progressive bins conditioning on the itinerary and the ticketing carrier.[17] In summary, our product is a unique combination of origin, connection, destination, the ticketing carrier, and the binned fare. We have 226,532 products in 2006 and 214,809 products in 1999.

Back Aviation Solutions' schedule data report the departure time and arrival time for all domestic flights. To generate the number of departures for direct flights, we aggregate over all carriers that operate for a ticketing carrier in a given market. The number of departures for connecting flights are route specific. We restrict the connecting time to 45 minutes and 4 hours. When there are multiple feasible connections, we only include the connection with the shortest layover time. Using other departure measures, including all feasible connections between 45 minutes to 4 hours, and the minimum number of departures for the two connecting segments, does not make much difference.

To evaluate changes in demand and supply between the late 1990s and the late 2000s, we conducted the empirical analysis using two cross-section data: the second quarter in 1999 and the second quarter in 2006. We chose 2006 to avoid the few years right after 9/11 when carriers were adjusting for the changing security regulations.

4.2 Data summary

Table 1 reports the summary statistics of our sample. The top panel displays the mean and standard deviation for all regressors used in the estimation. There were several noticeable changes between 1999 and 2006. The average fare, in 2006 dollars, decreased from \$493 to \$451, a reduction of 8.5%. In 1999, 7.6% of the products were priced above \$1,000; the fraction was reduced to 4% in 2006. The average fare for connecting flights dropped by 12%, while the average fare for direct flights fell by only 4%. Figure 3 and Figure 4 plot the fare density for direct and connecting flights, respectively.

[17] We use the following set of bins: \$20 for all tickets under \$700 (so tickets between \$300 and \$320 with the same itinerary and ticketing carrier are aggregated as one product), \$50 for tickets between \$700 and \$1,000, and \$100 for tickets above \$1,000.

Compared to 1999, fares of connecting flights were lower at each quantile of the distribution in 2006. For direct flights, the fraction of both high fare products ($\geq$\$1,000) and low fare products ($\leq$\$200) shrank, while that of medium fare ones increased.

The second pronounced development was the increasing number of direct passengers. Figure 5 displays the percentage of all U.S. domestic passengers who flew with direct flights from 1995 to 2006. It varied between 63% to 64.5% from 1995 to 2001, and steadily trended up since then. By 2006, more than 67.3% of passengers traveled on direct flights. In our sample markets, the number of direct passengers per market increased by 13% from 1999 to 2006, while that of connecting passengers diminished by 23%.

The trend away from connecting flights was universal – all legacy carriers flew fewer connecting passengers in 2006. American and Delta experienced the largest reduction, with the total number of connecting passengers decreased by 29% and 40% from 1999 to 2006, respectively. The number of connecting passengers was lower in relative terms as well. According to Figure 6, the percentage of connecting passengers among all passengers in 2006 was lower than 1999 for all carriers except for Continental. After the 1999 code-share agreement with Northwest, Continental started to issue connecting tickets with part of the route operated by Northwest, which led to a slightly higher faction of connecting passengers in 2006.

The declining number of connecting passengers during the sample period appeared to be closely related to the recent 'dehubbing' phenomenon in the airline industry. For example, Delta closed its hub in Dallas Ft. Worth International airport in January 2005, and cut 26% of flights at the Cincinnati hub in September 2005. US Airways downgraded Pittsburgh from a hub to a focus city in 2004. By October 2007, it had reduced the daily departures out of Pittsburgh from over 500 in 2000 to fewer than 70, and canceled service to more than 90 destination cities. With the exception of a few airports, most hubs serviced fewer connecting passengers in the recent years.

As a result of the increasing number of direct flights, the number of destination cities, which is the number of cities to which a carrier flies nonstop flights from the origin airport, increased from 17 to 19. The average number of daily departures dropped from 5.3 to 4.8, due to the legacy carriers' recent capacity reduction. The average plane size reduced from 135 seats to 123 seats, which reflected the increasing penetration of regional jets. All together, the six legacy carriers offered 77-78% of the products, accounted for 66% of passengers in 1999 and 61% of the passengers in 2006.

The bottom panel of Table 1 documents the market average summary statistics. Both the number of products and the number of carriers per market declined slightly after the two major mergers.[18] During the sample period, 39% of the markets experienced LCC entry.

5 Empirical model

5.1 Model specification

As illustrated in the previous sections, two of the most salient changes in the airline industry during the past decade were the decrease in fares (in real dollars) and the increase in the ratio of direct over connecting passengers. Therefore, we allowed three type-specific parameters: a constant, the fare coefficient, and the coefficient of the number of connections. The type-specific constant turned out to be important in improving the fit of the model, as it allowed the model to fit the aggregate shares for both expensive and inexpensive tickets.[19]

We spent a considerable amount of time experimenting with three or more types of passengers, without much success. The demand parameters common to all types were fairly robust, but the type-specific parameters, the λ and the γ parameters appeared to be sensitive to small changes in the model's specifications or the choice of instruments. Sometimes multiple parameter vectors delivered a similar fit for the data. Our conclusion is that the limited variation in the instruments prevents precise estimates for an overly flexible model.

We also tried to model carriers' choices of flight frequencies together with the pricing decisions, but faced three major challenges. First, some carriers mixed different aircraft on the same route. For example, large jets were typically reserved for dense traffic during peak time, while smaller regional jets or turbo planes were often the choice for off-peak flights. Second, it was difficult to measure flight frequencies for indirect flights, which affected our ability to estimate marginal revenues created by an additional departure. Lastly, to model how carriers trade off larger planes with fewer flights versus smaller planes with more frequent flights, we would need information on the age and type of the aircraft used, the flight schedule, and the number of passengers on each flight. In lack of such detailed data, we treated flight frequencies as endogenous without explicitly modeling how departures were

[18] American merged with Trans World in 2001, and American West merged with U.S. Airways in 2005.

[19] We also estimated the model with type-specific parameters for flight frequencies and the tour dummy. The parameters were similar across types, and there was not much improvement in the model's fit.

determined. The exercise of modeling departures directly is left for future research.

5.2 Instruments

As is typical of most demand studies with endogenous prices, we need instruments to identify the fare coefficients. One common strategy is to exploit the rival product attributes and the competitiveness of the market environment. All else being equal, products with closer substitutes have lower prices. A standard instrument is the number of products. In our data, the number of products in a market varies from 3 to 223, with an average of 53. However, we were concerned about the endogeneity of this variable because of the way it is constructed. A product is a group of tickets whose fares fall in a fixed bin. By construction, a market with a wider price dispersion has a larger number of products. Similar concerns extend to using rival product attributes as instruments.[20] We used the route level characteristics instead. Our instruments along this line include the percentage of rival routes that offer direct flights, the average distance of rival routes, the number of rival routes, the number of all carriers, etc.

A second identification strategy searches for variables that affect costs but not demand. One candidate is whether the destination is a hub. It affects the marginal cost of a flight, because larger and more fuel efficient planes can be used on routes with denser traffic, but is excluded from demand. The number of cities to which a carrier flies nonstop flights from the destination airport serves a similar role. We also included a dummy for transferring at the hub, using similar arguments that costs were lower if a flight connected at a hub.

The third group of instruments included the 25th and the 75th quantile of fitted fares.[21] As documented by Borenstein and Rose (1994, 2007), there was a wide fare dispersion across passengers on the same route. The 25th and the 75th fitted fare quantiles are nonlinear functions of the first-stage regressors, and allow us to better capture the exogenous price dispersion.

To construct instruments for flight frequencies, we first regressed segment departures on characteristics of the end cities (including whether they

[20] For example, with a wide price dispersion and a large number of products, the sum of rival product attributes will be high as well.

[21] In the first stage fare regression, the regressors included carrier dummies, segment and route level characteristics (distance, difference in January temperatures between the end cities, whether in tourist places, etc), market size (measured by population), number of competitors, and the carrier's shares of cities connected via nonstop flights at both the origin and the destination airport.

are located in tourist places, the average January temperature difference between these airports, population, distance, etc.), and then included the fitted segment departures as instrument.

The last group of instruments were the exogenous variables that directly entered the share equation (5) and the marginal cost equation (9). Finally, we included the interaction terms of the above variables provided they were not highly colinear.

5.3 Identification

The identification of most parameters is straightforward. Here we focus on λ and the type-specific parameters. λ is identified from changes in the aggregate market share when the number of products varies. In the extreme case of $\lambda = 0$, all products are perfect substitutes. The aggregate share remains fixed as the number of products changes, as long as the 'best product' does not change. On the other hand, if $\lambda = 1$, the nested logit demand is reduced to a simple logit, and the aggregate market share is close to $\frac{K}{K+1}$ if there are K products with similar product attributes. Identification of the type-specific parameters follows from the random coefficient literature, as our model is a special case where the random coefficients take two values. Briefly, these type specific parameters are identified from the substitution patterns among similar products when the mix of products varies across markets.

5.4 Model limitations

One implicit assumption of our empirical model is that the hub structure and the carriers' entry decisions in each market are exogenously given. Ideally, we would like to model a three-stage game: a) first, carriers form their hubs; b) given the hub structure, each carrier chooses the set of markets to serve; and c) given these entry decisions, carriers compete in prices and flight departures. However, solving this game with a dozen of carriers and thousands of markets is beyond our capability.

Instead, we focused on the last stage of the game and modeled a) consumers' choices between different products, and b) carriers' price decisions. We were concerned about the endogeneity of prices and departures induced by temporal demand shocks, and assumed that the instruments we used (the hub structure and the number of carriers, etc.) were pre-determined and uncorrelated with these temporal demand shocks. This is admittedly a strong assumption, but is analogous to the standard assumption in the discrete-choice demand literature, that variation in the set of available products and the number of firms across markets is exogenous.

As we did not observe the day-to-day variation in fares and flight availability, we did not allow consumers to choose strategically the date of purchase. We also ruled out the dynamic considerations in firms' pricing decisions. Modeling the dynamic aspect is a difficult but promising topic. See Ciliberto (2008) for an interesting study on the strategic deterrence in the airline industry.

Lastly, we did not observe the fixed costs of operating a flight, which limited our ability to estimate the changes in the net profit. Our profit estimates were based on the variable profits.

6 Result

The parameters from the base case specification was presented first, followed by results from eight other specifications. The profit estimates were discussed next. Finally, we reported results from the counter-factual exercises designed to isolate the effects of changes in demand, supply and competition on legacy carriers' profits.

6.1 Parameters

6.1.1 Demand parameters

Demand is affected by the following product attributes: fares, the number of total connections round trip, the number of destinations,[22] the average daily departures, the total distance (in thousand miles) round trip, distance squared, a tour dummy for airports in Florida and Las Vegas, the number of slot-controlled airports that the flight passes through,[23] and carrier dummies.[24]

[22] A product's number of destinations is the total number of cities to which its ticketing carrier serves direct flights from the origin airport.

[23] Four airports were under the slot control during the sample period: the LaGuardia airport and the Kenney airport in New York, the National airport in D.C., and the O'Hare airport in Chicago.

[24] In 1999, we included carrier dummies for American (the default group), American West, Continental, Delta, Northwest, Trans World, United, U.S. Airways, Southwest, and a dummy for all other carriers. In 2006, we added a dummy for Jetblue (which started operation in 2000), and excluded dummies for American West (merged with U.S. Airways in 2005) and Trans World (merged with American in 2001).

We expect consumers' utility to decrease with connections. The number of destination cities captures the value of the frequent flier programs. The larger the number of cities for which consumers can redeem frequent miles, the higher the value of these loyalty programs. In addition, a carrier that flies to a large number of destination cities is likely to have more convenient gate access and offer better service.

The air-travel demand is usually U-shaped in distance. Short-haul flights compete with cars and trains, which become a worse substitute as distance increases, so demand initially grows with distance. As distance increases further, travel becomes less pleasant, and demand starts to decrease. We include both distance and distance squared to capture the curvature of demand.

The tour dummy helps to fit the relatively high product shares in Florida and Las Vegas that cannot be explained by the observed product attributes. The slot variable captures the potential negative effect of congestion in these slot-controlled airports on air-travel demand.

The first two columns in Table 2 present the parameters for the base case specification. Most parameters were precisely estimated. Consistent with the story of the dot-com bubble burst and the introduction of online ticketing sites, demand was more price sensitive in 2006. The tourists' price coefficient increased (in absolute value) from 0.78 to 1.05, and the business travelers' price coefficient rose from 0.07 to 0.10. In both cases, the differences are statistically significant. The price elasticity was 31% larger for the tourist travellers and 43% larger for the business travellers. In the meantime, the estimated percentage of business travellers rose from 41% to 49%, which moderated the increase in demand's overall price sensitivity. With both groups, the average price elasticity at the product level increased from 1.96 to 2.10. The aggregate price elasticity, which is the percent change in total demand when all products' prices increase by 1%, was 1.55 in 1999, and rose to 1.67 in 2006. Gillen *et al.* (2003) conducted a survey that collected 85 demand elasticity estimates from cross-sectional studies.[25] The elasticities ranged from 0.181 to 2.01, with a median of 1.33. Our estimates seemed quite reasonable.

Both the tourists and the business passengers exhibited a stronger preference for direct flights in 2006. The connection semi-elasticity, or the percentage change in demand when a direct flight becomes a connecting flight,

[25] Out of these 85 estimates, 80 were taken from Oum et al. (1986) and represented U.S. city -pair routes. All 85 studies were conducted between 1981 and 1986 and are slightly dated.

increased from 0.75 to 0.80 for tourists and from 0.55 to 0.75 for the business travellers. Combining both groups, the average connection semi-elasticity increased by 17%, up from 66% to 77%. In other words, the number of passengers on a direct flight would reduce by almost four-fifths when a layover is added to the route.

These two results – a higher price sensitivity and a higher aversion toward connecting flights – were the most pronounced findings of demand changes, and were present in all specifications that we estimated. Both findings are supported by the trends (fare reductions and reducing number of connecting passengers) documented in the data section. While a fare reduction could also be rationalized by increasing competition or decreasing costs, the fact that fares dropped in markets with and without LCC entry, and that fares reduced more for connecting flights that became more costly to operate, provided ample evidence for a demand change during our sample period.

As we did not model carriers' choice of hub airports or their entry decisions, we could not examine how changes in demand affected the hub structure. However, it seems quite possible that reduced demand for connecting flights is directly related to several recent hub downsizings. Explaining changes in the airlines' network structure is a promising topic for future research.

Consumers also displayed a stronger preference for frequent flights. The willingness to pay for an additional daily flight almost doubled: in 1999, the tourist type was willing to pay $5 for an additional flight; the business type, $61. In 2006, their willingness to pay increased to $10 and $105, respectively.

Many previous studies pointed out the existence of a hub premium: carriers were able to charge higher fares for hub-originating flights, either because they offered more convenient gate access, or the frequent flier program was more valuable at hub airports. Borenstein (2005) and Boreistein and Rose (2007) pointed out that the hub premium declined over the past several years.[26] Our parameter estimates were consistent with their findings. The coefficient of the number of destinations – which we used to capture a carrier's presence at the airport – dropped from 0.38 to 0.27. The result was very similar with the hub dummy. Either the loyalty programs became less valuable, or the difference in service between hub airports (or airports with a large carrier presence) and non-hub airports (or airports with a small carrier presence) had narrowed.

All other demand parameters had the expected signs. For example,

[26] Boreistein (2005) found that fare premium at the 10 most expensive U.S. airports (all but one served as a hub) fell from 33% in 1995 to 24% in 2004.

demand increased in distance up to 1,600 miles (one-way) and then decreased in distance. Tourist places attracted more consumers, and flights through slot controlled airports had fewer passengers.

The business type accounted for 41% and 49% of total passengers in 1999 and 2006, respectively (see the third panel in Table 6). According to the 2001-2002 National Household Travel Survey, roughly 39% to 47% of air travel was taken for business purposes, depending on whether personal business trips were treated as business trips.[27] Our model's predictions match closely with the survey.

Interestingly, λ decreased from 0.77 in 1999 to 0.72 in 2006, which suggests that products became closer substitutes. It is probably because most carriers had cut down their services, which reduced the differentiation among products offered by different carriers.

Overall, the carrier dummies were broadly consistent with the news reports. In 1999, American (the omitted carrier) and United had the highest parameter values. They were also the most popular and successful carriers in the late 1990s. During the sharp downturn following 9/11, the legacy carriers, especially American and Delta, began to shift capacities to the more lucrative international markets. These structural changes were reflected in their negative carrier dummies in 2006. JetBlue had a large positive coefficient, which is consistent with its popularity due to high on-time performance, new planes, free TV programs, etc. In fact, by 2006, it had been voted the number one U.S. domestic airline by Conde Nast Traveler five years in a row.[28]

6.1.2 Marginal cost parameters

Column 3 and 4 in Table 2 report the marginal cost parameters, which includes a constant, the distance in thousand miles, and the number of connections. Two offsetting factors affect the marginal cost of connecting flights. On one hand, by channeling passengers from different origins and to different destinations through the connecting airport, carriers can generate denser traffic, increase the load factor, and defray costs with more passengers. On the other hand, a large fraction of the fuel is consumed at

[27] The National Household Travel Survey was conducted on 26,000 households. According to the survey, 56% of the trips longer than 50 miles were taken for pleasure, 16% for business, 13% each for commuting and for personal business(trips taken for family, personal, religious or medical reasons), and 3% for other reasons. Air travel accounted for 7% of pleasure trips, 18% of business trips, 5% of personal business trips, and none of the commuting trips.

[28] Data source: http://en.wikipedia.org/wiki/JetBlue_Airways#Awards.

the landings and takeoffs. With two extra landings and takeoffs, the fuel component of a connecting flight's marginal cost is much higher than that of a direct flight. The connection coefficient reflects the net effect of these two countervailing factors.

As different aircraft were used for short-medium haul routes and long haul routes, we allowed two sets of cost parameters: one for markets shorter than 1500 miles, and the other for markets longer than 1500 miles. We also included a hub dummy (equal to 1 if the flight departs from, transfers at, or lands at a hub airport), a slot dummy (equal to 1 if the flight passes through a slot-controlled airport), and carrier dummies. The same scale economy argument for connecting flights also applies to flights at the hub airports that tend to have denser traffic. Costs are higher at slot controlled airports due to the higher landing fees, etc.

The most noticeable difference between 1999 and 2006 was the connection coefficient, which changed signs during the sample period. In 1999, there was evidence of scale economies for connecting flights. Conditioning on distance, congestion, and the hub status, the marginal cost of serving a connecting passenger on a long route was \$18 less than that of a direct passenger, or roughly 12% of the average marginal cost. Unlike BCS that reported the existence of scale economies only on longer routes, our estimated marginal cost of connecting flights was lower on both long and short-medium routes in 1999.

The cost advantage of connecting flights disappeared in 2006. Conditioning on other cost shifters, the marginal cost of a connecting flight was \$12 more expensive than that of a direct flight. This is likely driven by the increasing fuel cost in the sample period. Since the fraction of fuel consumed at the takeoffs and landings could be as high as 40%, rising fuel cost offset the benefit of denser traffic created by connecting flights.[29]

All other parameters (except for the carrier dummies) were similar between the two periods, with the expected signs. Marginal cost increased with distance, and was higher for routes that passed through slot-controlled airports. Flights through hubs had a lower marginal cost.

The distance coefficient was smaller in 2006, which seemed somewhat puzzling given the higher fuel cost. The change probably reflected a combination of several factors, including reduced services, improved fuel efficiencies, etc.

[29] One might argue that the penetration of regional jets in the later period had an impact on marginal cost. However, adding the fraction of departures operated by regional jets to the regressors had no effect on other cost parameters, including the connection's coefficient.

As expected, the low cost carriers had lower marginal costs. Both JetBlue and SouthWest had lower marginal costs than the legacy carriers. Interestingly, American West also had a smaller marginal cost than the legacy carriers. According to U.S. DOT Form 41, its total operating cost per available seat mile (CASM) was comparable to SouthWest.[30] We do not fully understand Continental's coefficient in 2006, which suggested that Continental's marginal cost was comparable to that of SouthWest. It could be that these dummy variables reflected various carrier specific factors that were not captured by the model.

As in most empirical studies, marginal cost is not directly observed. The parameters are identified from a 'residual' regression where we 'regress' the difference between the price and markup on cost instruments. To examine the sensitivity of the marginal cost parameters to the over-identifying restrictions, we regressed the predicted marginal cost (the difference between prices and the estimated markup) on the variables that affected marginal cost directly. The coefficients from this OLS regression were very similar to the structural estimates, which suggested the robustness of the marginal cost instruments.

Finally, we compared our cost estimates with the carriers' reported operating costs per available seat mile. The average was 11.4 cents (in \$2006) in 1999 and increased by 10% to 12.5 cents in 2006. Our estimated marginal cost per mile was around 6 cents, about half of the average reported operating costs, which seemed plausible.

6.1.3 Other specifications

In the base specification, we estimated demand parameters (especially the price sensitivity) using both the share equation (5) and the pricing equation (9). As we were concerned about specification errors associated with our stylized pricing equation, we estimated the model again using only the share equation. The estimates are presented in the second column of Table 3A and Table 4A. As in BLP, we found that business travelers' 2006 price coefficient could not be reliably estimated without the markup equation.[31] However, most other demand parameters were similar to the base case. The aggregate price elasticity in 1999 was 1.69, similar to other specifications. It is particularly reassuring that the pattern of a stronger preference for direct flights remained: the connection semi-elasticity was 0.68 in 1999 and 0.76 in 2006.

[30] Data source: MIT Airline Data Project.

[31] The 2006 price coefficient was pushed to the pre-imposed boundary of zero.

Flight delays could potentially explain the changes in the preference for connecting flights, since the possibility of missing a connection is directly affected by delays. In the third specification, we added the delay variable, which is the percentage of flights arriving more than 30 minutes later than the scheduled arrival time, including canceled or diverted flights. The delay had the wrong sign in 1999 – demand was higher for flights with more delays. This is probably because delays are endogenous: crowded airports and popular flights were more likely to experience delays. The other coefficients, especially the connection's coefficients, barely changed. We experimented with various other measures of delays and flight time, including the percentage of delays longer than 15 minutes, the total taxi-in and taxi-out time, the total flight time, etc. None of these measures explained the increased disutility of connecting flights.

According to the official statistics, the percentage of flights with more than 30-minute delays was 14% in 1999 and 13% in 2006. However, the longest delays – those resulting from missed connections and canceled flights did not officially get counted. Bratu and Barnhart (2005) studied the August 2000 passenger itineraries on Continental Airlines, and discovered that when missed connections and flight cancellations were factored in, the average passenger delay was two-thirds longer than the official statistics. The problem of delays was much worse in 2006 as it was harder to find seats on later flights given the higher load factors. We expect that a better measured delay variable that reflects the actual connecting time would help to explain the increased disutilities of connections.

Six groups of airports are geographically close.[32] Combining these airports affected 38% of the markets, and raised the average number of products per market from 53 to 66. Perhaps not surprisingly, the λ coefficient was smaller, since consumers faced more similar choices in the grouped markets. Most other parameters were similar to the base case.

Our products were generated by a given set of fare bins. To examine the robustness of the parameter estimates to changes in the bin size, column five and six in Table 3 and 4 report results with a finer set of bins and a rougher set of bins, respectively.[33] Similar to column four with grouped airports, product level demand was more elastic with a larger number of products, while other parameters remained very similar to the base case.

Some of the airports might have higher fares or demand, either because

[32] See footnote [16] for a list of these airports.

[33] Column five used \$10 for fares under \$300, \$20 for fares between \$300 and \$700, \$50 for fares between \$700 and \$1000, and \$100 for fares above \$1000. Column six used \$50 for fares under \$1000, and \$100 for fares above.

of historical reasons, or because of convenient geographic locations that were not captured by the model. In column seven, we added airport dummies to the largest 25 airports. Most parameters were similar to the base case, except that demand was less price sensitive, which lead to a lower estimate of the marginal cost in both years.

One might argue that the discovery of a stronger preference for direct flights was driven by changes in the supply side, rather than changes in taste. During our sample period, low cost carriers expanded steadily, and offered a much higher fraction of point-to-point service. The more negative connection coefficient in 2006 could be driven by the decreasing shares of the legacy carriers who happened to offer more connecting flights. To address this concern, we estimated the model using only markets that did not experience LCC entry between 1999 and 2006. The results were presented in the first two columns of Table 5. As we did not model carriers' entry decisions, this sample potentially suffered from an endogenous selection problem, so we did not want to interpret the coefficients literally. Again, the parameter estimates were very close to the base case, which showed that consumers had a stronger preference for direct flights even in markets that were not affected by LCC entry.[34]

As mentioned in the introduction section, the advent of new regional jets allowed carriers to tailor the aircraft size to the size of the market and provide point-to-point service to markets traditionally dependent on connecting service. Another competing explanation is that consumers' preference has not changed, but there are more direct flights available. To tease out the regional-jet effect, we restricted the sample to long-haul markets with a great circle distance longer than 1500 miles, since few regional jets were used for direct flights in these markets. We lost about 70% of the markets, and our instruments had much less variation compared to the full sample. Distance squared was colinear with the distance variable (the correlation coefficient was 0.996) and was omitted from the regressors.[35] Demand was much more elastic than the base case, but the pattern of a stronger preference for direct flights remained: the connection semi-elasticity was 0.63 in 1999 and 0.80 in 2006.

We estimated many other specifications that are not reported here. For example, we estimated a model restricting the cost parameters to be the same across all markets for all seven specifications.[36] We also experimented

[34] The carrier dummies were not reported here, as there are too many parameters. Results are available upon request.

[35] There is only one set of cost parameters, since all markets are longer than 1500 miles.

[36] Results are available upon request. The cost parameters were more robust when we

with type-specific tour and flight frequency parameters. Our major findings – more price sensitive demand, a much stronger preference for direct flights, and changes in marginal cost favoring direct flights – were extremely robust and appeared in almost every set of parameter estimates. We are convinced that these findings revealed inherent data patterns and were not fabricated results of our modeling assumptions. The intuition for these results is straight forward: a negative supply shock should induce a small quantity *and* a high price. In our data, fewer passengers flew connecting flights even though fares were lower uniformly in 2006 – at each quantile of the fare distribution and in markets with or without entry of LCCs.

6.1.4 Marginal effects

To better understand the magnitude of the parameters, Table 6 tabulates changes in demand when product attribute changes. The effect of carrier airport presence on demand appears to be mild. Doubling the number of destinations for all products raises the aggregate demand by 11% in 1999 and 9% in 2006. On the other hand, adding one daily departure to all flights drives up the aggregate demand by 6% in 1999, and 16% in 2006. Changes in distance barely affect demand; in contrast, both the tour dummy and the slot variable have a significant impact. Adding the tour dummy to all products boosts the number of passengers by 32% in 1999 and 39% in 2006, while congestion in slot controlled airports reduces demand by 22%. These marginal effects do not vary much across specifications.

6.1.5 Elasticities, marginal cost, and markups

In Table 7 and 8, we summarize the elasticities, the percentage of each type of passengers, marginal cost and markups for different specifications. The aggregate price elasticity ranged from -1.35 to -1.69 in 1999, and -1.58 to -2.01 in 2006. The increase in price elasticity over the sample period varied from 6.5% to 24%, with an average of 13%. The connection semi-elasticity was much more stable across specifications, with an average increase of 16%. Changes in the marginal cost significantly favored direct flights: the increase in connecting flights' marginal cost was much bigger than that of direct flights. In fact, in four out of six specifications, direct flights' marginal cost was lower in 2006. Rising costs, combined with lower fares, led to

restricted them to be the same across the long-haul and the short-medium haul markets, but we prefer our reported specifications as there are probably significant cost differences between these markets (for example, the type of aircraft).

a sizeable reduction of the markups of connecting flights, ranging from 5 percentage points to 12 percentage points. The markup for the top 10% most expensive products dropped even further, from 90% to less than 70%. The reduction in the profitability of these high-end products, together with the shrinking profit of connecting flights, was an important explanation of the legacy carriers' financial stress in the recent years.

6.2 Profit and revenue estimates

The average number of products offered by a carrier in a given market was slightly different between 1999 and 2006. To avoid the complication of the changing number of products (which might reflect the changing dispersion of prices rather than the changing number of distinct products), we analyze a carrier's average profit and revenue per market, instead of the average profit from a product. We also focus on the legacy carriers throughout this analysis. We first report the profit estimates and the counter factual results using the base case parameters, then describe the general patterns over all counter-factual exercises.

Table 9 displays the legacy carriers' average profit and revenue per market. We discuss connecting flights and direct flights separately, as these products exhibit distinct patterns. For connecting flights, 2006 witnessed fewer passengers, lower revenues and lower profits. The reduction was fairly uniform and happened to every part of the fare distribution. Compared with 1999, the average demand shrank by 14%, and the average fare was 12% lower. As a result, the average revenue was reduced by 25%, and profit fell even further, by 32%. Profit for the top 10% most expensive products decreased by 56%, which was driven by a bigger reduction in fares among these high-end products.

The picture for direct flights was much more complicated. The average number of direct passengers per carrier per market increased by 8%, but the average revenue was down by 6%, and the average profit was 16% lower. A closer look at the changes across different quantiles of the fare distribution revealed that all of the profit reduction occurred among the 10% most expensive products. In 1999, these 10% products generated $477k in profits per market, and accounted for 32% of total profits from all direct flights. By 2006, profits from the top 10% products declined to merely $150k, and constituted only 12% of total profits. As our parameter estimates suggested, demand in 2006 was more price sensitive. Even though consumers showed a stronger preference for direct flights, they had in general stayed away from the high-end products and switched to flights with low or medium fares.

Profits and revenues for the bottom 90% direct flights were about 8-10% higher in 2006 than in 1999. However, the higher profitability from the low- and medium-fare flights was overwhelmed by the drastic profit declines among the most expensive flights. Profits from all direct flights fell by 16%.

When we combine both direct and connecting flight, the legacy carriers transported 4% more passengers, but generated 9% less revenues and 19% less profits in 2006 than in 1999.

6.3 Counter-factual analysis

To examine how the legacy carriers' profits were affected by a) the change of consumer tastes; b) the change of marginal cost; and c) LCC's expansion, we calculated the counter-factual profits and revenues for the following five different scenarios:

- using 2006's observed product attributes and marginal cost parameters, but 1999's taste parameters;
- using 2006's observed product attributes and marginal cost parameters, but 1999's taste parameters and ξ_j that 'replicates' its distribution in 1999;
- using 2006's observed product attributes and taste parameters, but 1999's marginal cost parameters;
- using 2006's observed product attributes, taste and marginal cost parameters, but excluding LCCs from the markets they entered between 1999 and 2006;
- using 2006's observed product attributes, but 1999's taste and marginal cost parameters, ξ_j that replicates its distribution in 1999, and excluding LCCs from the markets they entered during the sample period.

In each exercise, we solve for a new vector of the optimal prices that satisfy the first order conditions incorporating changes specified above.[37] The first exercise quantifies the effect of the changes of consumer preference, including the increased price sensitivity and the higher aversion to connecting flights.

[37] In solving for the optimal prices, we restricted the first order condition to be smaller than 10^{-9}. The convergence was slow for the third and fourth counter-factual exercise, so we set the tolerance level to 10^{-8}. There was not much difference in the profit estimates using different tolerance level.

As discussed in section 3.1, ξ_j, the utility from the unobserved product attributes (like the refund restrictions, advance purchase requirements, etc.), plays an important role in determining demand. If these product attributes were similar between 1999 and 2006, then the difference between ξ_j^{1999} and ξ_j^{2006} reflected changes in consumers' taste for these attributes, and constituted an important component of the preference changes. On the other hand, if the unobserved product attributes altered, then the difference was a combination of changes in taste and changes in product characteristics. In the second exercise, we incorporated changes in ξ_j by replicating its 1999 distribution conditioning on fares separately for direct and connecting flights. For example, given all direct flights priced at \$350, we replaced the first quantile of ξ_j^{06} with the first quantile of ξ_j^{99}, etc. Then we solved for the counter-factual prices using 1999 taste parameters and the constructed vector of ξ_j.

The third exercise isolated the effect of changes in the marginal cost, the fourth one examined the competition from LCCs, while the last exercise combined all factors discussed above.

Table 10 summarized the counter factual results for connecting flights. Overall, the model did a decent job explaining the profit change for connecting flights. Replacing the 2006's demand parameters with the 1999's explained 58% and 61% of the profit and revenue reduction, respectively. Results were similar when we incorporated ξ_j's 1999 distribution.

Using the 2006's taste parameters but the 1999's cost parameters accounted for about 9% of the profit and revenue decrease between 1999 and 2006. The marginal cost was higher in 2006, which led to higher fares, a lower demand, and a lower profit.

Around 40% of the markets experienced LCC entry during the sample period.[38] Compared with the change of tastes, competition from LCCs had a modest impact on connecting flights' profit. Removing LCCs explained 15% of a legacy carrier's profit drop in markets affected, or 8% when averaged over all markets. There are a couple of explanations for this finding. First, many new products introduced by the low cost carriers were direct flights. As discussed below, LCC entry explained a much larger fraction of the direct flights' profit reduction. Second, the legacy carriers had gradually developed strategies (for example, lowering fares, adding departures) to compete with low cost carriers.

[38] Some of these markets already had low cost carriers in 1999 (like Air Tran, Frontier, or Southwest). A market experienced LCC entry if a new low cost carrier established a service in that market between 1999 and 2006.

When we incorporated all factors, the model was able to replicate 72% of the profit decrease during the sample period. The model performed well even when we look at high-fare and low-fare products separately. It explained 81% of the profit change for the bottom 90% products, and 60% of the profit change for the top 10% most expensive products.

In comparison, the model roughly replicated direct flights' average profit, but could not fit very well the profit increase for low- and medium-fare products and the profit decrease for the high-fare products (see Table 11). Using the 1999 demand parameters, the predicted profit was comparable to the observed 1999 profit for the bottom 90% products, but was only 28% of the observed 1999 profit for the top 10% products. When we combined the 1999 demand parameters with the ξ_j's 1999 distribution, the predicted profit from all direct flights was close to the observed level in 1999.[39] However, even though we were able to match the average, our prediction was higher than the observed profit for the bottom 90%, and lower for the top 10% products. It became clear to us that the model did not have the ability to fit the various parts of the data's distribution, and could only explain changes in the mean.

As the marginal cost was higher in 1999, using 2006's demand parameters and 1999's cost parameters reduced profits by 4%. Low cost carriers' expansion had a bigger impact on direct flights than on connecting flights. Removing LCCs explained 25% of the legacy carriers' profit reduction in markets that experienced LCC entry, and 12% when averaged over all markets. Combining all factors explained 94% of the observed change in direct flights' profits.

We repeated the above counter-factual exercise for all other specifications and summarized the results in Table 12. For connecting flights, changes in demand accounted for around 50% of the profit reduction, changes in cost, 10-30%, and entry of LCCs, 8%. For direct flights, demand was by far the most important factor. LCC's expansion contributed to 8-18% of the profit drop. The change of marginal cost had mixed signs: it led to higher profits in four specifications, and lower profits in two specifications.

[39] It turns out that ξ_j was an important factor in determining demand for the high-end products. The ξ_j's dispersion among the high-end direct flights was much wider in 1999 than in 2006. Replicating ξ_j's distribution in 1999 helped us to generate demand for the high-fare flights.

7 Conclusions

We found that compared to the late 1990s, in 2006, air-travel demand was more price sensitive. Passengers displayed a much stronger preference for direct flights. In addition, the change of marginal cost significantly favored direct flights. These three factors, together with the expansion of LCCs, explained more than 80% of the observed reduction in legacy carriers' profits. Despite the press' emphasis on the increasing fuel cost and competition from LCCs, the change in demand was the most important explanation for the legacy carriers' profit losses.

We conclude with two caveats. First, our estimates were changes in variable profits, not changes in net profits, as we did not observe fixed costs. Second, we found that the impact of LCC entry in the 2000s was modest compared to the changes in consumers' preference. If the expansion of LCCs contributed to the change of the taste parameters via affecting consumers' search behavior, then their general equilibrium effect could be much larger. Lastly, reduced demand for connecting flights probably contributes to the recent 'dehubbing' phenomenon in the airline industry. Modeling the effect of LCC entry and demand changes on the airline industry's network structure is an interesting question for future research.

8 Appendix: constructing departures and flight delays

In this section, we explain how we constructed flight frequencies and the delay variable. The scheduling data from Back Aviation Solutions reported the scheduled departure time and arrival time for all flights operated by U.S. carriers that file with Official Airline Guides. We cross-examined this data with T100, and found that the departures matched closely (the correlation between Back departures and T100 departures exceeded 99.9%).[40] In addition, the Back data contained the scheduling information for some regional carriers that did not directly report to T100.[41]

Obtaining the number of departures for direct flights was straight forward: we counted the total number of direct flights by all carriers that

[40] The difference between these two data sets is consistent with the average completion ratio, since the Back data report scheduled departures, while T100 reports actual performed departures.

[41] For example, Comair filed form 298c prior to 2000 and Piedmont filed 298c until 2001. Both carriers showed up in the Back scheduling data.

operated for a ticketing carrier in a given market. The DB1B data reported the ticketing carrier and the operating carriers for all itineraries, so we indirectly observed the list of operating carriers that provided service for a ticketing carrier in a given airport pair. Constructing the number of departures for connecting flights was slightly more involved. We restricted the connecting time to 45 minutes and 4 hours. When there were multiple feasible connections, we only included the connection with the shortest layover time.[42] The complication arises when a ticketing carrier issued tickets operated by more than one operating carrier. We constructed the number of connecting departures using flights by all carriers that operated for the ticketing carrier in a given market.

DOT publishes the flight-level on-time arrival data for non-stop domestic flights.[43] We first obtained the on-time performance for each operating carrier for each airport pair, and then aggregated the delay variable (weighted by departures) by ticketing carriers and airport pairs. For connecting flights, the delay variable was the average over the two segments.

[42] Suppose some carrier offers the following flight schedule among airports A, B, C: 1) flight 1001 departs from A at 8am, arrives at B at 2pm; 2) flight 1002 departs from A at 10am and arrives at B at 4pm; 3) flight 1003 departs from B at 5:30pm and arrives at C at 7:30pm. Even though both flight 1001 and 1002 can be connected with flight 1003, we only count the connection with the shortest layover time. In this example, the carrier operates one connecting departure in market A-C.

[43] In 1999, only the major carriers – American, Continental, Delta, Northwest, Trans World, United, and US Air, plus Alaska, American West, and Southwest reported the delay statistics. In 2006, both major carriers and the largest regional carriers reported the delay statistics to DOT.

References

[1] Armentier, Olivier and Oliver Richard (2008), Domestic Airline Alliances and Consumer Welfare, Forthcoming Rand Journal of Economics.

[2] Berry, Steven, James Levinsohn, and Ariel Pakes (1995), Automobile Prices in Market Equilibrium, Econometrica, vol. 60-4, pp889-917.

[3] Borenstein, Severin (2005), U.S. Domestic Airline Pricing, 1995-2004. UC Berkeley Competition Policy Center working paper, No. CPC05-48.

[4] Borenstein, Severin and Nancy Rose (1994), Competition and Price Dispersion in the US Airline Industry, Journal of Political Economy, 102, 653-683.

[5] Borenstein, Severin and Nancy Rose (2007), How Airline Markets Work ... Or Do They? Regulatory Reform in the Airline Industry, NBER working paper No 13452.

[6] Bratu, Stephane, and Cynthia Barnhart (2005), Air Traffic Control Quarterly, Vol 13(1), pp. 1-27.

[7] Brunger, William (2007), The Impact of Internet on Airline Pricing. Case Western Reserve University working paper.

[8] Ciliberto, Federico (2008), Strategic Deterrence in a Dynamic Game: Evidence from the Airline Industry, University of Virginia working paper.

[9] Dana, James D. and Eugene Orlov (2008), Internet Penetration and Capacity Utilization in the US Airline Industry, Northeastern University working paper.

[10] Facchinei, Francisco and Jong-Shi Pang (2003), Finite-dimensional variational inequalities and complementarity problems. New York Springer Publisher.

[11] Forbes, Silke J. (2008), The Effect of Air Traffic Delays on Airline Prices, forthcoming in International Journal of Industrial Organization.

[12] Gayle, Philip (2004), Does Price Matter? Price and Non-Price Competition in the Airline Industry, Kansas State University Working Paper.

[13] Gillen, David W., William G. Morrison, and Christopher Stewart MBA (2003), Air Travel Demand Elasticities: Concepts, Issues and Measurement, Department of Finance Canada report.

[14] Goolsbee and Syverson (2008), How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines. University of Chicago working paper.

[15] Ito, Harumi, and Darin Lee (2003), Incumbent Responses to Lower Cost Entry: Evidence from the U.S. Airline Industry, Brown University Working Paper.

[16] McFadden, Daniel L. (1981), Econometric Models of Probabilistic Choice, in Charles F. Manski and Daniel McFadden, eds., Structural Analysis of Discrete Data with Econometric Applications, Cambridge MA: MIT Press.

[17] Morrison, Steven A. and Clifford Winston (2005), What's Wrong with the Airline Industry? Diagnosis and Possible Cures. Hearing before the subcommittee on Aviation Committee on Transportation and Infrastructure United States House of Representatives.

[18] Mozdzanowska, Aleksandra (2004), Evaluation of Regional Jet Operating Patterns in the Continental United States, MIT thesis.

[19] Peters, Craig (2006), Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry, Jounal of Law and Economics, vol. XLIX., pp627-649.

[20] Puller, Steven L., Anirban Sengupta and Steven Wiggins (2007), Testing Theories of Price Dispersion and Scarcity Pricing in the Airline Industry, University of Texas A&M working paper.

Table 1: Summary Statistics for the Data Set

	1999		2006	
Variable	Mean	Std.	Mean	Std.
Fare (2006 $100)	4.93	3.17	4.51	2.59
Product Share	1.42E-04	6.37E-04	1.42E-04	5.26E-04
Direct Flight	0.37	0.48	0.43	0.49
No. Daily Departures	5.25	3.41	4.83	2.85
No. Destinations (100 cities)	0.17	0.28	0.19	0.31
Hub	0.16	0.37	0.16	0.36
HubMC	0.85	0.36	0.72	0.45
Distance (1000 miles)	2.73	1.40	2.78	1.42
Distance^2 (1000 miles)	9.42	8.44	9.72	8.66
Tourist Place (FL/LAS)	0.13	0.33	0.13	0.34
Slot-Control	0.36	0.76	0.36	0.75
SlotMC	0.21	0.41	0.21	0.40
Plane Size (100)	1.35	0.33	1.23	0.34
Delay>=30 Minutes	0.14	0.07	0.13	0.07
American	0.16	0.37	0.18	0.39
Continental	0.10	0.29	0.08	0.28
Delta	0.19	0.39	0.15	0.36
American West	0.05	0.22		
NorthWest	0.09	0.28	0.08	0.28
Trans World	0.09	0.28		
United	0.13	0.34	0.14	0.34
US Air	0.10	0.30	0.15	0.36
JetBlue			0.01	0.12
SouthWest	0.04	0.20	0.09	0.29
Other Carrier	0.05	0.22	0.11	0.31
No. Observations	214809		226532	
Market Average				
No. Products	53.73	38.52	52.68	36.67
No. Carriers	3.51	2.00	3.30	1.88
No. Direct Passengers (1000)	20.13	40.45	22.75	43.66
No. Connecting Passengers (1000)	3.52	4.10	2.71	3.13
No. Markets w/ LCC Entry			1569	
No. Observations	3998		4300	

Note: Hub=1 if the origin airport is a hub; HubMC=1 if either the origin, the connecting airport, or the destination is a hub. Tourist Place=1 if the origin airport is in Las Vegas or Florida. Slot-Control is the number of slot-controlled airports the route of product j passes through. SlotMC=1 if Slot-Control>0. Delay is the percentage of flights arriving more than 30 minutes later than the scheduled arrival time.

Table 2: Base Case Parameter Estimates -- 1999 & 2006

Demand Variables	1999	2006	Cost Variables	1999	2006
Fare 1	-0.78*	-1.05*	Constant 1	1.07*	1.16*
	(0.01)	(0.01)		(0.03)	(0.03)
Connection 1	-0.53*	-0.59*	Distance 1	0.26*	0.19*
	(0.01)	(0.02)		(0.00)	(0.00)
Constant 1	-5.79*	-5.68*	Connection 1	-0.06*	0.07*
	(0.13)	(0.12)		(0.02)	(0.02)
Fare 2	-0.07*	-0.10*	Constant 2	1.61*	1.59*
	(0.00)	(0.00)		(0.04)	(0.04)
Connection 2	-0.31*	-0.51*	Distance 2	0.09*	0.04*
	(0.01)	(0.01)		(0.01)	(0.01)
Constant 2	-8.56*	-8.60*	Connection 2	-0.09*	0.06*
	(0.27)	(0.19)		(0.02)	(0.03)
No. Destination	0.38*	0.27*	HubMC	-0.02†	-0.05*
	(0.01)	(0.01)		(0.01)	(0.01)
No. Departures	0.04*	0.11*	SlotMC	0.08*	0.03*
	(0.00)	(0.00)		(0.01)	(0.01)
Distance	0.30*	0.53*			
	(0.02)	(0.01)			
Distance^2	-0.05*	-0.08*			
	(0.00)	(0.00)			
Tour	0.30*	0.36*			
	(0.01)	(0.01)			
Slot-Control	-0.19*	-0.18*			
	(0.00)	(0.00)			
lambda	0.77*	0.72*			
	(0.00)	(0.00)			
gamma	0.69*	0.63*			
	(0.08)	(0.07)			
Demand Carrier Dummy			Cost Carrier Dummy		
Other Carriers	-0.18*	0.06*	Other Carriers	-0.03†	-0.22*
	(0.01)	(0.01)		(0.01)	(0.01)
American West	-0.19*		American West	-0.22*	
	(0.01)			(0.01)	
Continental	-0.22*	0.07*	Continental	-0.03*	-0.19*
	(0.01)	(0.01)		(0.01)	(0.01)
Delta	-0.13*	-0.21*	Delta	-0.10*	-0.15*
	(0.01)	(0.01)		(0.01)	(0.01)
NorthWest	-0.15*	0.07*	NorthWest	-0.02†	-0.04*
	(0.01)	(0.01)		(0.01)	(0.01)
Trans World	-0.17*		Trans World	0.02	
	(0.01)			(0.01)	
United	0.16*	0.08*	United	-0.05*	-0.06*
	(0.01)	(0.01)		(0.01)	(0.01)
US Air	-0.19*	0.06*	US Air	-0.08*	-0.11*
	(0.01)	(0.01)		(0.01)	(0.01)
JetBlue		0.39*	JetBlue		-0.32*
		(0.03)			(0.03)
SouthWest	-0.05*	0.08*	SouthWest	-0.12*	-0.19*
	(0.02)	(0.01)		(0.01)	(0.02)
Function Value	49.37	58.07			
Observations	214.8k	226.5k			

Note: See Table 1 for the variable definitions. * and † denote significance at the 5% and 10% confidence level, respectively. Standard errors are in parentheses.

Table 3A: Demand Parameter Estimates from Different Specifications -- 1999

Demand Variables	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Fare 1	-0.78*	-1.14*	-0.78*	-0.80*	-0.80*	-0.69*	-0.74*
	(0.01)	(0.16)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Connection 1	-0.53*	-0.47*	-0.53*	-0.53*	-0.45*	-0.63*	-0.55*
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant 1	-5.79*	-4.66*	-5.84*	-5.47*	-6.05*	-5.77*	-6.33*
	(0.13)	(1.01)	(0.13)	(0.11)	(0.12)	(0.13)	(0.10)
Fare 2	-0.07*	-0.09*	-0.07*	-0.06*	-0.07*	-0.07*	-0.07*
	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Connection 2	-0.31*	-0.37*	-0.31*	-0.31*	-0.28*	-0.40*	-0.36*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant 2	-8.56*	-8.35*	-8.59*	-8.48*	-8.64*	-8.07*	-8.64*
	(0.27)	(0.94)	(0.27)	(0.26)	(0.25)	(0.29)	(0.20)
No. Destination	0.38*	0.32*	0.34*	0.34*	0.36*	0.40*	0.48*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
No. Departures	0.04*	0.05*	0.04*	0.06*	0.03*	0.05*	0.05*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance	0.30*	0.35*	0.27*	0.33*	0.35*	0.26*	0.29*
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Distance^2	-0.05*	-0.05*	-0.05*	-0.05*	-0.05*	-0.05*	-0.05*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tour	0.30*	0.32*	0.30*	0.34*	0.27*	0.31*	0.29*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Slot-Control	-0.19*	-0.18*	-0.21*	-0.11*	-0.19*	-0.20*	-0.13*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Delay			0.76*				
			(0.05)				
lambda	0.77*	0.72*	0.77*	0.69*	0.76*	0.79*	0.83*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
gamma	0.69*	0.52	0.70*	0.72*	0.70*	0.70*	0.68*
	(0.08)	(0.44)	(0.08)	(0.07)	(0.08)	(0.09)	(0.06)
Demand Carrier Dummy							
Other Carriers	-0.18*	-0.14*	-0.08*	-0.02†	-0.19*	-0.18*	-0.10*
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
American West	-0.19*	-0.19*	-0.17*	-0.11*	-0.22*	-0.14*	-0.13*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Continental	-0.22*	-0.20*	-0.20*	-0.17*	-0.23*	-0.21*	-0.14*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Delta	-0.13*	-0.13*	-0.10*	-0.10*	-0.10*	-0.20*	-0.11*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NorthWest	-0.15*	-0.13*	-0.11*	-0.10*	-0.14*	-0.17*	-0.13*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Trans World	-0.17*	-0.16*	-0.15*	-0.13*	-0.19*	-0.13*	-0.12*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
United	0.16*	0.16*	0.18*	0.18*	0.17*	0.13*	0.10*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
US Air	-0.19*	-0.18*	-0.19*	-0.16*	-0.19*	-0.19*	-0.16*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SouthWest	-0.05*	-0.04*	-0.01	0.01	-0.04*	-0.06*	0.05*
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)

Note: See Table 3B for explanations of the specification in each column, no. of observations, and function values.

Table 3B: Cost Parameter Estimates from Different Specifications -- 1999

Cost Variables	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Constant 1	1.07*		1.07*	1.29*	0.85*	0.88*	0.81*
	(0.03)		(0.03)	(0.03)	(0.03)	(0.05)	(0.04)
Distance 1	0.26*		0.26*	0.28*	0.26*	0.26*	0.23*
	(0.00)		(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Connection 1	-0.06*		-0.06*	-0.02	0.01	-0.08*	-0.08*
	(0.02)		(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Constant 2	1.61*		1.61*	1.92*	1.38*	1.38*	1.30*
	(0.04)		(0.04)	(0.04)	(0.04)	(0.06)	(0.05)
Distance 2	0.09*		0.09*	0.10*	0.09*	0.10*	0.07*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Connection 2	-0.09*		-0.10*	-0.07*	-0.02	-0.10*	-0.10*
	(0.02)		(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
HubMC	-0.02†		-0.02	-0.08*	-0.03*	0.00	0.00
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SlotMC	0.08*		0.08*	0.11*	0.08*	0.09*	0.06*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cost Carrier Dummy							
Other Carriers	-0.03†		-0.02†	-0.04*	-0.03*	-0.04*	-0.02
	(0.01)		(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
American West	-0.22*		-0.22*	-0.26*	-0.20*	-0.22*	-0.20*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Continental	-0.03*		-0.03*	-0.02*	-0.02*	-0.03*	-0.03*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Delta	-0.10*		-0.10*	-0.12*	-0.09*	-0.10*	-0.09*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NorthWest	-0.02†		-0.02†	-0.04*	-0.01	-0.01	-0.01
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Trans World	0.02		0.02	0.01	0.03*	0.00	0.02*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
United	-0.05*		-0.05*	-0.07*	-0.05*	-0.03*	-0.04*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
US Air	-0.08*		-0.08*	-0.11*	-0.08*	-0.06*	-0.07*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SouthWest	-0.12*		-0.12*	-0.19*	-0.12*	-0.10*	-0.10*
	(0.01)		(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Function Value	49.37	46.54	49.44	40.22	51.15	42.90	44.89
Observations	214.8k	214.8k	214.8k	214.8k	238.5k	147.3k	214.8k

Note: See Table 1 for the variable definitions. Column one is the base case. Column two does not use the markup condition. Column three adds delays to demand. Column four groups nearby airports. Column five and six use a finer and a rougher set of fare bins, respectively. Column seven includes 25 airport dummies. * (†) denotes significance at the 5% (10%) confidence level. Standard errors are in parentheses.

Table 4A: Demand Parameter Estimates from Different Specifications -- 2006

Demand Variables	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Fare 1	-1.05*	-1.49*	-1.06*	-1.13*	-1.09*	-0.96*	-1.04*
	(0.01)	(0.15)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Connection 1	-0.59*	-0.33*	-0.56*	-0.46*	-0.48*	-0.72*	-0.62*
	(0.02)	(0.04)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Constant 1	-5.68*	-4.50*	-5.61*	-4.85*	-5.87*	-5.44*	-6.06*
	(0.12)	(1.04)	(0.12)	(0.14)	(0.11)	(0.17)	(0.11)
Fare 2	-0.10*	0.00	-0.10*	-0.10*	-0.10*	-0.09*	-0.10*
	(0.00)	(0.02)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Connection 2	-0.51*	-0.60*	-0.53*	-0.67*	-0.50*	-0.52*	-0.55*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant 2	-8.60*	-9.16*	-8.55*	-8.39*	-8.64*	-8.40*	-8.85*
	(0.19)	(0.82)	(0.19)	(0.19)	(0.18)	(0.28)	(0.16)
No. Destination	0.27*	0.20*	0.29*	0.26*	0.27*	0.25*	0.43*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
No. Departures	0.11*	0.10*	0.11*	0.13*	0.09*	0.12*	0.12*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance	0.53*	0.55*	0.53*	0.41*	0.52*	0.55*	0.58*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Distance^2	-0.08*	-0.08*	-0.08*	-0.05*	-0.08*	-0.09*	-0.09*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Tour	0.36*	0.37*	0.35*	0.41*	0.34*	0.37*	0.34*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Slot-Control	-0.18*	-0.18*	-0.17*	-0.10*	-0.18*	-0.19*	-0.13*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Delay			-0.82*				
			(0.04)				
lambda	0.72*	0.67*	0.72*	0.63*	0.72*	0.72*	0.77*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
gamma	0.63*	0.49	0.63*	0.60*	0.63*	0.65*	0.61*
	(0.07)	(0.43)	(0.07)	(0.08)	(0.07)	(0.10)	(0.06)
Demand Carrier Dummy							
Other Carriers	0.06*	0.13*	0.04*	0.14*	0.03*	0.07*	0.11*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Continental	0.07*	0.13*	0.09*	0.14*	0.09*	0.06*	0.11*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Delta	-0.21*	-0.24*	-0.23*	-0.21*	-0.19*	-0.29*	-0.22*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NorthWest	0.07*	0.08*	0.04*	0.11*	0.06*	0.07*	0.08*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
United	0.08*	0.14*	0.09*	0.14*	0.09*	0.06*	0.03*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
US Air	0.06*	0.11*	0.02*	0.13*	0.07*	0.02†	0.06*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
JetBlue	0.39*	0.55*	0.38*	0.56*	0.24*	0.53*	0.46*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
SouthWest	0.08*	0.19*	0.08*	0.11*	0.10*	0.06*	0.14*
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

Note: see Table 4B for explanations of the specification in each column.

Table 4B: Cost Parameter Estimates from Different Specifications -- 2006

Cost Variables	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Constant 1	1.16*		1.16*	1.30*	1.02*	1.22*	1.07*
	(0.03)		(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
Distance 1	0.19*		0.19*	0.22*	0.19*	0.21*	0.17*
	(0.00)		(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Connection 1	0.07*		0.10*	0.25*	0.14*	-0.03	0.05*
	(0.02)		(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Constant 2	1.59*		1.58*	1.73*	1.40*	1.73*	1.44*
	(0.04)		(0.04)	(0.04)	(0.03)	(0.05)	(0.04)
Distance 2	0.04*		0.04*	0.06*	0.04*	0.06*	0.04*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Connection 2	0.06*		0.10*	0.27*	0.16*	-0.07*	0.05†
	(0.03)		(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
HubMC	-0.05*		-0.05*	-0.07*	-0.06*	-0.06*	-0.05*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
SlotMC	0.03*		0.03*	0.06*	0.03*	0.03*	0.02*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Cost Carrier Dummy							
Other Carriers	-0.22*		-0.22*	-0.22*	-0.22*	-0.27*	-0.22*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Continental	-0.19*		-0.18*	-0.11*	-0.18*	-0.22*	-0.20*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Delta	-0.15*		-0.15*	-0.15*	-0.13*	-0.19*	-0.15*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
NorthWest	-0.04*		-0.04*	-0.06*	-0.03*	-0.04*	-0.05*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
United	-0.06*		-0.06*	0.00	-0.04*	-0.09*	-0.07*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
US Air	-0.11*		-0.10*	-0.03*	-0.10*	-0.13*	-0.12*
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
JetBlue	-0.32*		-0.30*	-0.16*	-0.32*	-0.36*	-0.39*
	(0.03)		(0.03)	(0.03)	(0.02)	(0.04)	(0.03)
SouthWest	-0.19*		-0.18*	-0.19*	-0.21*	-0.18*	-0.19*
	(0.02)		(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Function Value	58.07	49.59	57.38	40.48	60.84	43.86	52.20
Observations	226.5k	226.5k	226.5k	226.5k	257k	146k	226.5k

Note: See Table 1 for the variable definitions. Column one is the base case. Column two does not use the markup condition. Column three adds delays to demand. Column four groups nearby airports. Column five and six use a finer and a rougher set of fare bins, respectively. Column seven includes 25 airport dummies. * (†) denotes significance at the 5% (10%) confidence level. Standard errors are in parentheses.

Table 5: Robustness Check

	Markets w/o LCC Entry		Markets Longer than 1.5k Miles	
Demand Variables	1999	2006	1999	2006
Fare 1	-0.85*	-1.20*	-0.88*	-1.60*
	(0.01)	(0.02)	(0.02)	(0.06)
Connection 1	-0.48*	-0.53*	-0.53*	-0.54*
	(0.01)	(0.03)	(0.03)	(0.04)
Constant 1	-5.68*	-5.38*	-3.44*	-1.74*
	(0.22)	(0.26)	(0.62)	(0.56)
Fare 2	-0.07*	-0.11*	-0.06*	-0.12*
	(0.00)	(0.00)	(0.00)	(0.00)
Connection 2	-0.36*	-0.48*	0.00	-0.55*
	(0.01)	(0.02)	(0.09)	(0.03)
Constant 2	-8.55*	-8.75*	-8.88*	-7.36*
	(0.51)	(0.41)	(3.37)	(0.65)
No. Destination	0.38*	0.26*	0.39*	0.27*
	(0.02)	(0.01)	(0.04)	(0.02)
No. Departures	0.05*	0.12*	0.02*	0.07*
	(0.00)	(0.00)	(0.01)	(0.00)
Distance	0.31*	0.51*	-0.16*	-0.14*
	(0.02)	(0.02)	(0.01)	(0.01)
Distance^2	-0.05*	-0.08*		
	(0.00)	(0.00)		
Tour	0.18*	0.18*	0.51*	0.63*
	(0.01)	(0.01)	(0.02)	(0.01)
Slot-Control	-0.15*	-0.18*	-0.16*	-0.17*
	(0.00)	(0.00)	(0.01)	(0.01)
lambda	0.75*	0.72*	0.75*	0.67*
	(0.00)	(0.00)	(0.01)	(0.01)
gamma	0.71*	0.63*	0.85†	0.57*
	(0.15)	(0.15)	(0.51)	(0.28)
Cost Variables				
Constant 1	1.09*	1.24*		
	(0.04)	(0.03)		
Distance 1	0.29*	0.23*		
	(0.01)	(0.01)		
Connection 1	0.04*	0.09*		
	(0.02)	(0.03)		
Constant 2	1.66*	1.68*	3.13*	2.22*
	(0.06)	(0.06)	(0.09)	(0.07)
Distance 2	0.10*	0.05*	0.04*	0.06*
	(0.01)	(0.01)	(0.01)	(0.01)
Connection 2	0.04	0.09*	-0.12*	0.10*
	(0.02)	(0.04)	(0.05)	(0.04)
HubMC	-0.08*	-0.07*	-0.01	-0.13*
	(0.01)	(0.01)	(0.04)	(0.02)
SlotMC	0.10*	0.05*	0.10*	0.04*
	(0.01)	(0.01)	(0.01)	(0.01)

Note: column one and two only use markets that did not experience LCC entry between 1999 and 2006. Column three and four use markets longer than 1500 miles that are less likely to be affected by the regional jets.

Table 6: Percentage Changes in Demand When Product Attributes Change

	1999	2006
No. Destination Doubles	11%	9%
Add One Daily Departure	6%	16%
Distance up 10%	-1%	-1%
Tour Dummy Changes from 0 to 1	32%	39%
Slot Changes from 0 to 1	-22%	-22%
Carrier Dummy Changes from 0 to 1		
Other Carrier	-20%	8%
American West	-20%	
Continental	-24%	9%
Delta	-15%	-24%
NorthWest	-17%	9%
Trans World	-19%	
United	22%	11%
US Air	-20%	7%
JetBlue		58%
SouthWest	-5%	10%

Note: the top panel displays the percentage change in market demand when the relevant product attribute is changed as specified. For example, in 2006, adding one departure to all products increases the market demand by 16% on average. The bottom panel reports changes in demand for the relevant carrier. For example, in 2006, changing Continental's carrier dummy from 0 to 1 increases its average market demand by 9%.

Table 7A: Elasticity Estimates from Different Specifications -- 1999

Price Elasticity	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Type One	-5.01	-7.81	-5.01	-5.64	-4.90	-4.77	-4.40
Type Two	-0.44	-0.65	-0.44	-0.46	-0.42	-0.48	-0.43
Both Types	-1.96	-2.16	-1.96	-2.35	-1.95	-1.63	-1.62
Aggregate Price Elasticity	-1.55	-1.69	-1.55	-1.68	-1.53	-1.38	-1.37
Connection Semi-Elasticity							
Type One	0.75	0.73	0.75	0.78	0.69	0.79	0.74
Type Two	0.55	0.64	0.55	0.59	0.51	0.63	0.58
All	0.66	0.68	0.66	0.71	0.61	0.71	0.66
Percentage of Passengers							
Type One	0.59	0.47	0.59	0.64	0.57	0.58	0.54
Type Two	0.41	0.53	0.41	0.36	0.43	0.42	0.46

Table 7B: Elasticity Estimates from Different Specifications -- 2006

Price Elasticity	Base Case	No MC	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Type One	-6.55		-6.57	-8.09	-6.41	-6.66	-6.10
Type Two	-0.63		-0.63	-0.70	-0.61	-0.63	-0.60
Both Types	-2.10		-2.15	-2.94	-2.15	-1.97	-1.89
Aggregate Price Elasticity	-1.67		-1.70	-2.01	-1.63	-1.66	-1.58
Connection Semi-Elasticity							
Type One	0.80	0.63	0.79	0.77	0.74	0.86	0.80
Type Two	0.75	0.83	0.76	0.88	0.75	0.76	0.76
All	0.77	0.76	0.77	0.83	0.74	0.80	0.77
Percentage of Passengers							
Type One	0.51	0.47	0.52	0.59	0.48	0.55	0.48
Type Two	0.49	0.53	0.48	0.41	0.52	0.45	0.52

Note: the aggregate price elasticity measures the percentage change in total demand when all products' prices increase by 1%. Connection semi-elasticity measures the percentage change in product j's demand when it switches from a direct flight to a connecting flight, fixing other products' attributes.

Table 8A: Marginal Cost and Markup from Different Specifications -- 1999

Marginal Cost ($)	Base Case	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Connecting Flights	160	160	190	153	141	125
Direct Flights	149	149	170	126	132	120
All Products	156	156	183	142	138	123
% Markup						
Connecting Flights	0.60	0.60	0.53	0.60	0.69	0.69
Direct Flights	0.66	0.66	0.61	0.68	0.78	0.74
All Products	0.63	0.63	0.56	0.63	0.72	0.71

Table 8B: Marginal Cost and Markup from Different Specifications -- 2006

Marginal Cost ($)	Base Case	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
Connecting Flights	167	173	229	165	157	149
Direct Flights	138	137	158	120	147	124
All Products	155	158	199	145	153	139
% Markup						
Connecting Flights	0.56	0.54	0.41	0.54	0.60	0.60
Direct Flights	0.66	0.66	0.60	0.66	0.69	0.69
All Products	0.60	0.59	0.49	0.60	0.64	0.64

Table 9: Carrier Profit and Revenue Per Market

		Profit ($100k)			Revenue ($100k)		
Year		All Fares	Bottom 90% Fares	Top 10% Fares	All Fares	Bottom 90% Fares	Top 10% Fares
1999	All flights	17.80	11.77	6.03	26.38	19.79	6.60
	Direct	14.95	10.17	4.77	21.90	16.62	5.29
	Connecting	2.86	2.14	0.72	4.48	3.64	0.84
2006	All flights	14.46	12.19	2.27	23.92	20.72	3.19
	Direct	12.53	11.03	1.50	20.53	18.31	2.23
	Connecting	1.94	1.62	0.32	3.38	2.93	0.45

Table 10: Carrier Profit and Revenue Per Market for Different Counter-Factual Scenarios: Connecting Flights

	Profit ($100k)			Revenue ($100k)		
Different Scenarios	All Fares	Bottom 90% Fares	Top 10% Fares	All Fares	Bottom 90% Fares	Top 10% Fares
1999 Base Case	2.86	2.14	0.72	4.48	3.64	0.84
2006 Base Case	1.94	1.62	0.32	3.38	2.93	0.45
1999 Demand Parameters	2.47	1.97	0.50	4.05	3.43	0.63
1999 Demand Parameters and ξ	2.45	1.91	0.54	3.95	3.27	0.68
1999 MC Parameters	2.02	1.64	0.38	3.51	2.99	0.52
No LCC Expansion	2.01	1.62	0.39	3.51	2.97	0.54
All Factors	2.59	2.04	0.56	4.15	3.45	0.69

Note: we use 2006 product attributes for all counter-factual exercises. In each row, we solve for a new vector of the optimal prices that satisfy the first order conditions incorporating the parameter changes as specified.

Table 11: Carrier Profit and Revenue Per Market for Different Counter-Factual Scenarios: Direct Flights

	Profit ($100k)			Revenue ($100k)		
Different Scenarios	All Fares	Bottom 90% Fares	Top 10% Fares	All Fares	Bottom 90% Fares	Top 10% Fares
1999 Base Case	14.95	10.17	4.77	21.90	16.62	5.29
2006 Base Case	12.53	11.03	1.50	20.53	18.31	2.23
1999 Demand Parameters	10.97	9.62	1.35	18.08	16.28	1.80
1999 Demand Parameters and ξ	15.06	11.72	3.34	22.11	17.67	4.44
1999 MC Parameters	11.99	10.41	1.58	19.85	17.48	2.36
No LCC Expansion	12.81	11.20	1.61	20.85	18.49	2.36
All Factors	14.80	11.46	3.34	22.03	17.55	4.48

Note: we use 2006 product attributes for all counter-factual exercises. In each row, we solve for a new vector of the optimal prices that satisfy the first order conditions incorporating the parameter changes as specified.

Table 12A: Percentage of Profit Changes Explained by Different Counter-Factual Scenarios -- Connecting Flights

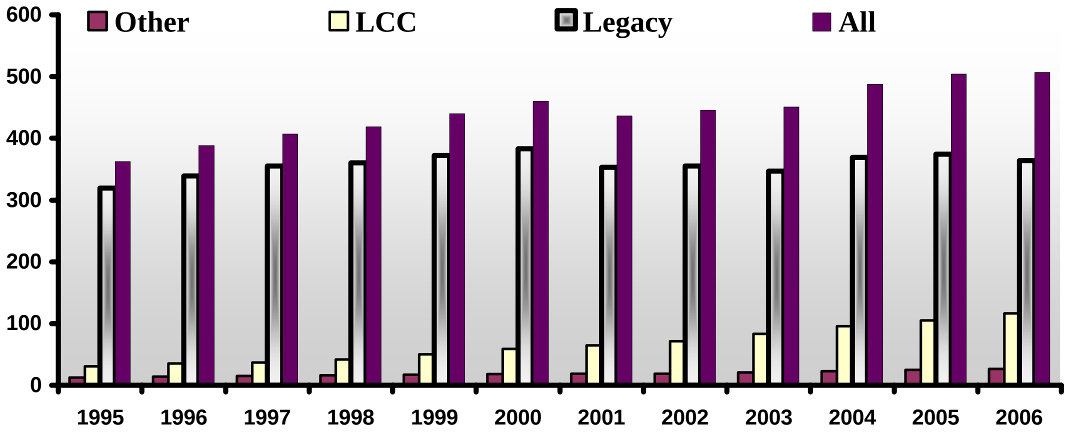
Different Scenario	Base Case	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
1999 Demand Parameters and ξ	0.56	0.46	0.49	0.56	0.53	0.54
1999 MC Parameters	0.09	0.14	0.33	0.14	0.11	0.20
No LCC Expansion	0.08	0.08	0.08	0.08	0.08	0.06
All Factors	0.72	0.66	0.87	0.76	0.70	0.77

Table 12B: Percentage of Profit Changes Explained by Different Counter-Factual Scenarios -- Direct Flights

Different Scenario	Base Case	Delay	Combine Airports	Small Bin	Large Bin	Airport Dummy
1999 Demand Parameters and ξ	1.05	1.02	0.85	1.29	0.70	0.87
1999 MC Parameters	-0.22	-0.23	-0.20	-0.16	0.18	0.07
No LCC Expansion	0.12	0.11	0.18	0.14	0.08	0.08
All Factors	0.94	0.90	0.77	1.26	0.90	0.98

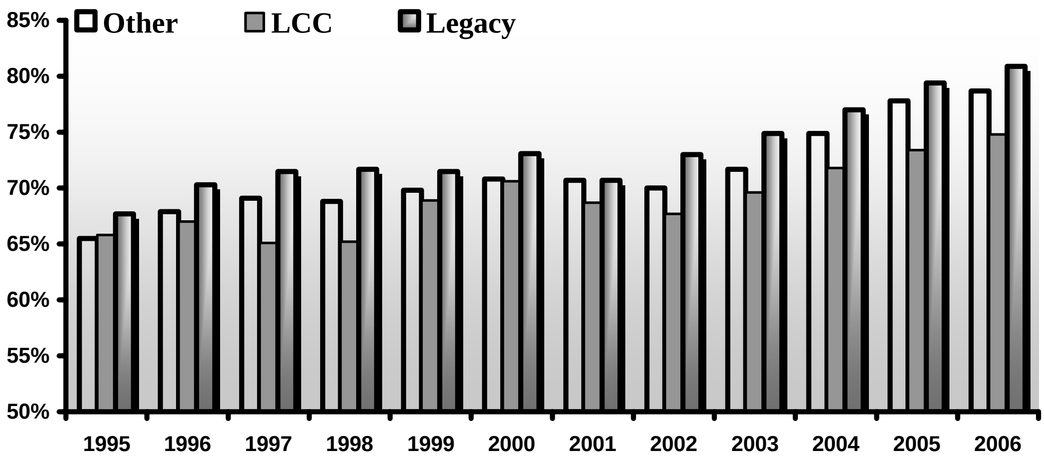
Note: we use 2006 product attributes for all counter-factual exercises. In each row, we solve for a new vector of the optimal prices that satisfy the first order conditions incorporating the parameter changes as specified.

Figure 1: U.S. Domestic Revenue Passenger Miles (Bill.)



Source: MIT Airline Data Project.

Figure 2: U.S. Airlines' System Load Factors



Source: MIT Airline Data Project.

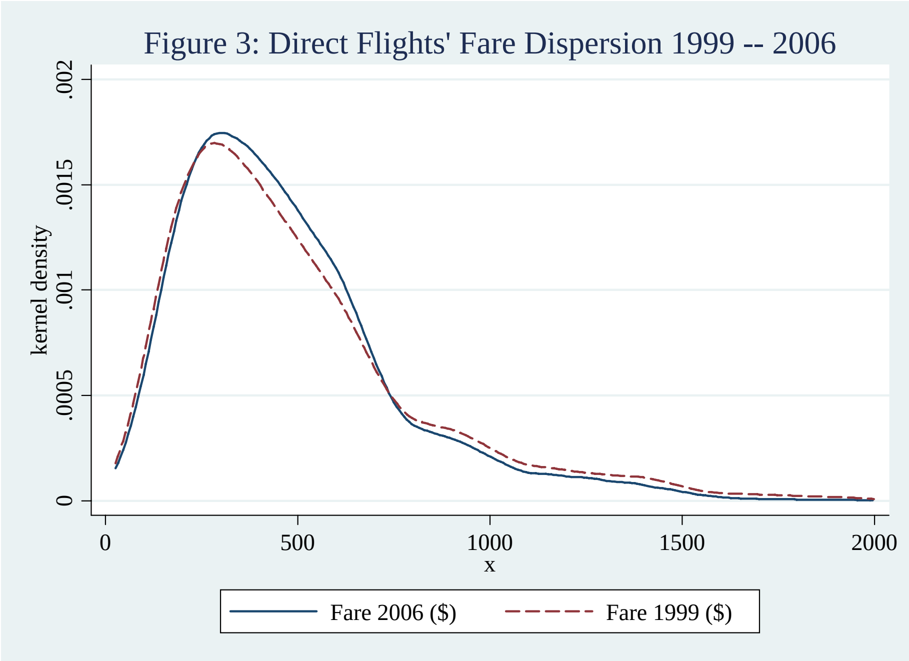


Figure 3: Direct Flights' Fare Dispersion 1999 -- 2006

Source: US DOT DB1B via BTS. Calculation based on the sample markets.

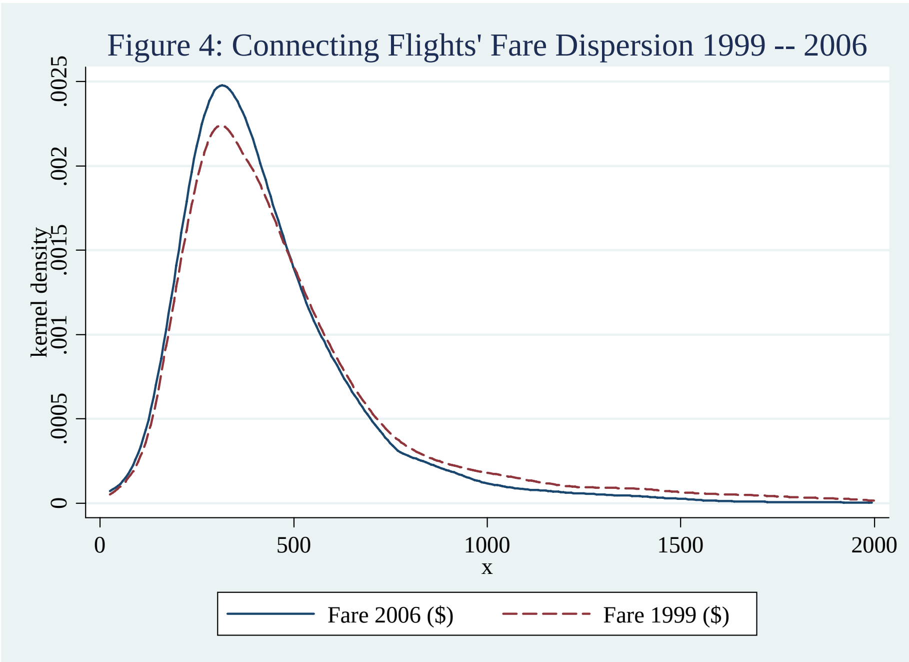
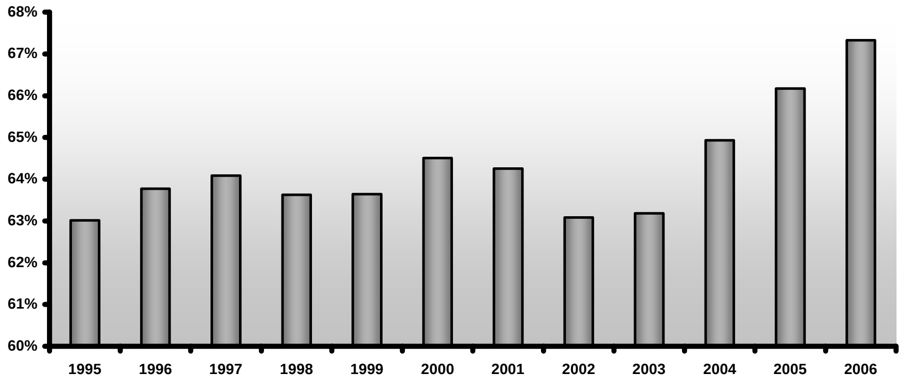


Figure 4: Connecting Flights' Fare Dispersion 1999 -- 2006

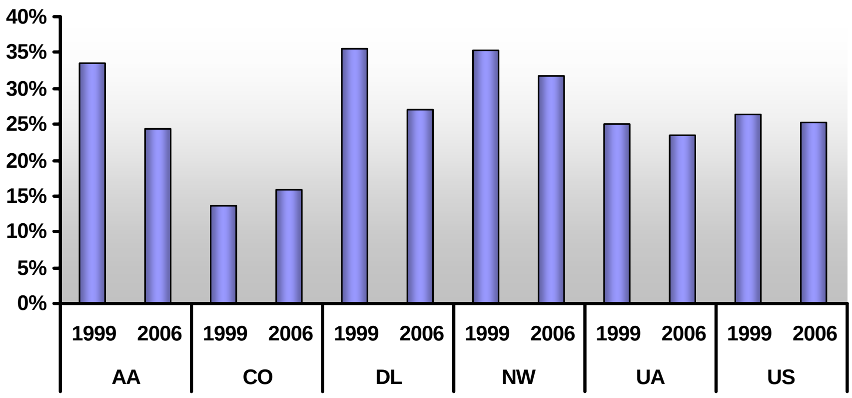
Source: US DOT DB1B via BTS. Calculation based on the sample markets.

Figure 5: Percentage of Direct Passengers in U.S.



Source: US DOT DB1B via BTS. Author's calculation.

Figure 6: Percentage of Connecting Passengers by Carrier -- 1999 and 2006



Source: US DOT DB1B via BTS. Calculation based on the sample markets.