United States General Accounting Office

GAO

Report to Congressional Requesters

April 1991

AIRLINE COMPETITION

Effects of Airline Market Concentration and Barriers to Entry on Airfares

143865

RESTRICTED——Not to be released outside the General Accounting Office unless specifically approved by the Office of Congressional Relations.

GAO

United States
General Accounting Office
Washington, D.C. 20548

Resources, Community, and Economic Development Division

B-242870

April 26, 1991

The Honorable Jack Brooks
Chairman, Committee on the Judiciary
House of Representatives

The Honorable James L. Oberstar
Chairman, Subcommittee on Aviation
Committee on Public Works and Transportation
House of Representatives

The Honorable John C. Danforth
Ranking Minority Member
Committee on Commerce, Science, and Transportation
United States Senate

Rising fares on certain routes and a wave of mergers and bankruptcies have raised concerns that the airline industry has become less competitive than envisioned when the Airline Deregulation Act of 1978 was passed. This report is one in a series of GAO reviews on competition in the nation's airline industry. (See Related GAO Products.) In an earlier report, we identified the presence of factors that inhibit airlines from entering a market, often called "barriers to entry," and discussed their potential effects on competition.[1] This report presents estimates of how several other factors, such as an airline's market share and airport congestion, as well as barriers to entry, affect air fares. It also discusses the policy implications of our analysis.

To conduct this analysis, we developed an econometric model that examines how several competitive conditions influence an airline's fare and market share on a route. The model also includes other factors, such as distance and traffic volume, that were considered likely to influence fares and market shares. This letter focuses on the model results that are particularly significant for setting federal policy for the airline industry. Our model tests the influence of factors that we believe directly influence fares and factors that indirectly influence fares by affecting an airline's market share on a route. For a more complete discussion of model results, see appendix I.

[1] Airline Competition: Industry Operating and Marketing Practices Limit Market Entry (GAO/RCED-90-147, Aug. 29, 1990).

Results in Brief

In terms of direct impacts, our model results indicate that fares are higher on routes when any one of three barriers to entry is present. When airport limitations on take-offs and landings, known as slot restrictions, were present at an airport, fares were, on average, 4 percent higher. The presence at an airport of a majority-in-interest clause, which often gives an airline veto power over airport expansion, was associated with, on average, 3-percent higher fares on routes to and from that airport. When a major airline had a code-sharing agreement, which involves joint marketing with a commuter airline, the major airline's fares were, on average, 2 percent higher on routes involving that airport. Two or more of these barriers are often present together on a route. In such cases, fares were 5 to 9 percent higher.

Several factors other than barriers to entry also directly influence fares. Airlines with higher market shares generally charge higher prices. For example, a 65-percent increase in an airline's market share on a route was associated with 6-percent higher fares on the route. In addition, the higher the operating costs of the least-cost airline on a route, the higher were the fares of all airlines serving the route.[2] Finally, fares were higher on routes involving more congested airports.

In terms of indirect impacts, our model results indicate that some competitive factors influence fares by increasing an airline's market share on a route. For instance, airlines having a large share of gates at the airports on a route, or a dominant computerized reservation system in the cities on a route, tend to have greater market shares and, therefore, higher fares.

The effects of several factors on fares are not uniform. We found substantially different impacts, depending on the distance of the route, the price-sensitivity of the passenger, and the size of the airline. For example, slot restrictions affect fares more on short routes than on long routes.

Because no single factor has a large impact on fares, a policy designed to affect any single factor or entry barrier is not likely to have a large impact on fares across all routes. However, such policies may have a substantial effect on certain kinds of routes or passengers. Policies

[2]We define the least-cost airline on a route to be the airline with the lowest operating costs of any airline serving the route. Our results show that, for example, when a very low-cost airline serves a route, the prices charged by all airlines serving the route will be lower.

Airline Competition: Impact of Computerized Reservation Systems (GAO/RCED-86-74, May 9, 1986)

Airline Takeoff and Landing Slots: Department of Transportation's Slot Allocation Rule (GAO/RCED-86-92, Jan. 31, 1986)

Deregulation: Increased Competition Is Making Airlines More Efficient and Responsive to Consumers (GAO/RCED-86-26, Nov. 6, 1985)

Related GAO Products

Airline Competition: Weak Financial Structure Threatens Competition (GAO/RCED-91-110, Apr. 15, 1991)

Airline Competition: Fares and Concentration at Small City Airports (GAO/RCED-91-51, Jan. 18, 1991)

Airline Competition: Passenger Facility Charges Represent a New Funding Source for Airports (GAO/RCED-91-39, Dec. 13, 1990)

Airline Deregulation: Trends in Airfares at Airports in Small and Medium-Sized Communities (GAO/RCED-91-13, Nov. 8, 1990)

Airline Competition: Industry Operating and Marketing Practices Limit Market Entry (GAO/RCED-90-147, Aug. 29, 1990)

Airline Competition: Higher Fares and Reduced Competition at Concentrated Airports (GAO/RCED-90-102, July 11, 1990)

Effects of Airline Entry Barriers on Fares (GAO/T-RCED-90-62, Apr. 5, 1990)

Airline Competition: DOT and Justice Oversight of Eastern Air Lines' Bankruptcy (GAO/RCED-90-79, Feb. 23, 1990)

Barriers to Competition in the Airline Industry (GAO/T-RCED-89-65, Sept. 20, 1989)

Airline Competition: DOT's Implementation of Airline Regulatory Authority (GAO/RCED-89-93, June 28, 1989)

Airline Service: Changes at Major Montana Airports Since Deregulation (GAO/RCED-89-141FS, May 24, 1989)

Factors Affecting Concentration in the Airline Industry (GAO/T-RCED-88-65, Sept. 22, 1988)

Airline Competition: Fares and Service Changes at St. Louis Since the TWA-Ozark Merger (GAO/RCED-88-217BR, Sept. 21, 1988)

Competition in the Airline Computerized Reservation System Industry (GAO/T-RCED-88-62, Sept. 14, 1988)

directed at easing entry into slot-controlled airports, for example, may lower fares more on short-haul routes than on longer routes.

Background

Before 1979, the Civil Aeronautics Board, which regulated prices of air travel, limited the creation of new airlines and the entry of existing airlines into new markets, so that prices were generally above competitive levels, according to industry analysts. In the initial years following deregulation, many new airlines began service, existing airlines entered new markets, and competition increased on many routes. However, since 1985, rising fares on certain routes and a wave of mergers and bankruptcies have raised concerns that the industry has become less competitive than originally envisioned.

Industry analysts have argued that barriers to entry might account for this decline in competition. Generally, barriers to entry are practices or conditions that may impede a firm's ability to enter a market. These conditions are considered entry barriers if they allow firms presently in the market to charge prices higher than what would be charged in a more competitive environment. Without these entry barriers, additional firms would enter the market, bidding prices down to competitive levels.

We developed an econometric model to estimate the impact of various factors, including several entry barriers, on an airline's fares and market share on a particular route. An econometric model allows the analyst to examine how each of a set of factors independently influences some other factor. In our analysis, the model estimates how each of several factors independently affects an airline's fares and market shares, while holding constant all other factors in the model.

The model consists of two equations. The first equation estimated the effects of several factors on an airline's fares on a route. The second equation estimated the effects of several factors on an airline's market share on a route. Because we hypothesize that factors included in the market share equation have an indirect effect on price, market share is also one of the factors included in the price equation. The decision about which factors, particularly the competitive conditions, should be included in each equation was based on our judgment of whether the factor primarily affected fares directly or primarily affected fares indirectly through its effect on market share.

Key factors in the fare equation included several barriers to entry, an airline's market share on a route, the level of operating costs of the least-cost airline on a route, and the level of airport congestion.[3] The barriers to entry in this equation included the following:

- Slot restrictions. Limitations on the number of take-offs and landings at the four airports under the High Density Rule—Washington National, New York's Kennedy and LaGuardia, and Chicago O'Hare.[4]
- Majority-in-interest clauses. Provisions in an airport's general use agreement with an airline that typically give those airlines performing a majority of the operations at the airport veto power over airport expansion when those airlines would be responsible for paying the cost of that expansion.
- Code-sharing. An agreement between a jet airline and a commuter airline to market services jointly by sharing the jet airline's two-letter airline code.
- Noise restrictions. Limits on the amount of airplane noise generated at an airport.
- Airport expansion constraints. Institutional or physical conditions that, according to an airport's officials, hamper that airport's ability to expand.

The competitive conditions in the market share equation include the combined enplanement share[5] of an airline at the airports on a route, the enplanement share of its largest competitor, the dominance of an airline's computerized reservation system at the endpoint cities, and an airline's share of gates at the endpoint airports.

We performed our analysis using third-quarter 1988 data, the most current data available at the time of our analysis. We believe that these data are representative of current airline pricing practices. Our sample consisted of over 1,600 routes. Our analysis produced estimates of each factor's average effect on fares across the sample. In addition, we conducted a series of analyses to determine if these effects differed depending on the distance of the route, the price sensitivity of the passenger, or the size of an airline flying the route.

[3] Other potential entry barriers that could affect fares or market shares, such as frequent flyer plans, were not included in the model because of data availability problems. Frequent flyer plans may have a significant effect on fares. Our survey of travel agents indicated, for example, that business travelers usually choose their airline on the basis of these plans.

[4] 14 C.F.R. Sec 93, Subpart 5.

[5] An enplanement is a passenger boarding a flight. This includes passengers beginning their trip at an airport, as well as passengers who are connecting from another flight. Enplanement share equals the average of a carrier's share of enplanements at each of the endpoint airports of a route.

Several Factors Directly Influence Fares

Several factors of policy interest to the Congress directly influence fares, according to the results of our model. These factors include certain barriers to entry, an airline's market share on a route, the level of operating costs of the least-cost airline on a route, and the level of airport congestion.

Fares were higher on routes when any one of three barriers to entry was present. In particular, fares were, on average, 4 percent higher when slot restrictions were present at an airport on a route. The presence at an airport of a majority-in-interest clause was associated with, on average, 3-percent higher fares on routes involving that airport. A major airline's fares were, on average, 2 percent higher on routes when the airline had a code-sharing agreement at either endpoint of the route.[6] Neither noise restrictions nor airport expansion impediments had a statistically significant impact on fares.

Two or more of these barriers are often present together on a route. For example, two or more of the barriers that influence fares were present on 844, or 51 percent, of the routes in our sample, affecting 64 percent of the passenger trips in the sample. In these situations, because the effects of these barriers are additive in this model, fares were higher on average by 5 to 9 percent.

Other factors also directly influence fares. Airlines with higher market shares charged higher prices. For example, our model indicated that a 65-percent increase[7] in an airline's market share on a route was associated with, on average, 6-percent higher fares. In addition, the higher the systemwide operating costs of each route's least-cost airline, the higher are the fares of <u>all</u> airlines serving the route. For every 15-percent increase in the costs of the lowest-cost airline on a route, fares were, on average, 3 percent higher for all airlines serving the route. Finally, fares are higher on routes involving more congested airports. If the airport congestion on a route (as measured by the number of take-offs and landings per runway at the two endpoint airports) was 30 percent greater than on another route, fares on the more congested route were, on average, 2 percent higher than on the less congested route.

[6] There is a 95-percent probability that the estimated effect of code-sharing on fares was not due to chance alone. There is a 99-percent probability that the estimated effects of each of the other variables did not occur by chance.

[7] Each percentage increase used in our illustrations in this letter is "typical" for that variable. We used the statistical concept of the "standard deviation," which measures the typical variation around the average value of the variable, to calculate typical percentage changes.

Factors that indirectly affect fares are of policy interest as they relate to airline domination at hub airports and limitations in airport capacity. These factors include high airport enplanement shares, high gate shares, and dominant computerized reservation systems.

Particular policies designed to limit an airline's ability to exercise market power on routes from airports at which they are dominant could benefit consumers without reducing the efficiencies provided to the airline by its hub-and-spoke system. For example, a policy designed to revise the Department of Transportation's (DOT) rules governing computerized reservations systems (CRS) could improve competitive conditions in the industry.[11] Greater airport capacity could also ease entry conditions by making it less likely that a particular airline will control most traffic and facilities at an airport.

Policies that enhance competition may improve the financial health of certain weaker airlines by enabling them to enter additional profitable markets that are now dominated by other airlines. However, policies to enhance competition must also be sensitive to possible adverse effects on some airlines' financial health.

It is important to note that policies directed at the competitive factors discussed in this report will not necessarily eliminate the entire estimated fare differentials that our model found were associated with those factors. For example, even though the model finds that routes affected by slot controls are associated with 4-percent higher fares, a policy that modifies or eliminates slot limitations may not reduce fares on these routes by the full 4 percent. This is because some of the fare differential on routes affected by slot controls is due to a scarcity of capacity at these airports. This scarcity would exist even in the absence of any formal slot allocation system. Thus, policies should be formulated recognizing that a particular factor's effect on competition may be directly attributable to that factor or may simply reflect other costs or competitive problems.

This letter presents an overview of the econometric model and its basic results. We discuss the technical issues of the modeling effort in the

[11]On Mar. 26, 1991, DOT issued a notice of proposed rulemaking that would extend current CRS rules (14 C.F.R. 255) with revisions. These revisions, according to DOT, are designed to enhance competition in the airline and CRS industries. They include allowing travel agents to use equipment obtained from suppliers other than a CRS vendor and to use a single terminal for access to all CRSs.

sensitive passengers, such as business travelers.[9] For example, while highly price-sensitive passengers pay only 1-percent higher fares when the level of congestion increases by 30 percent, less price-sensitive passengers pay 4-percent higher fares for the same increase in congestion.

Finally, some factors, particularly airline market share and the systemwide operating cost level of the least-cost airline, affect the fares of the three largest airlines differently from the other airlines.[10] For instance, the three largest airlines appear better able to translate higher market shares into higher prices than the other airlines. If one of the largest airlines increased its market share on a route by 65 percent, its fares were, on average, 15 percent higher. By comparison, a 65-percent increase in the market share for one of the other airlines did not lead to a statistically significant increase in fares.

Policy Implications

The results of our model suggest that several factors of policy interest to the Congress are directly associated with higher fares. These factors include airport slot restrictions, code-sharing agreements, majority-in-interest clauses, and congestion. Because none of these factors individually has a large influence on fares, a policy designed to affect any one of these factors is not likely to have a large impact on fares across all routes.

However, this does not mean that policies targeting these factors would be ineffective. Because the effects of each factor are not uniform across routes, policies to address a particular factor may have a substantial effect on certain kinds of routes or passengers. Policies directed at easing entry into slot-controlled or congested airports, for example, by reallocating slots to entering airlines, will have more of an influence on fares for short-haul routes and business travelers than on fares for other routes and passengers generally. In addition, because the effects of the factors in this model are additive, a policy directed at more than one factor will have a more substantial effect on fares.

[9]Because highly price-sensitive passengers, such as tourists, generally pay lower fares, we used the 25th percentile fare on a route to represent the fares paid by more price-sensitive passengers. Less price-sensitive passengers, such as business travelers, are less likely to qualify for lower fares because of their less flexible schedules. We used the 75th percentile fare to represent prices paid by less price-sensitive travelers.

[10]The three largest airlines (based on national revenue passenger miles) are American, Delta, and United.

Some Factors Indirectly Influence Fares by Affecting Market Share

Some factors influence an airline's market shares and, thus, indirectly affect prices, according to the results of our model. For example, our model estimated that a 65-percent increase in an airline's average enplanement shares at the airports on a route was associated with 21-percent higher market shares. At the same time, a 45-percent increase in the enplanement share for a given airline's largest competitor at the airports on a route was associated with a 6-percent decline in the given airline's market share on the route. In addition, an airline's market share tended to be higher when that airline owned a computerized reservation system that was dominant in the endpoint cities, and when the airline had a large share of gates at the endpoint airports of the route.[8]

Because these factors lead to higher airline market shares, they indirectly lead to higher prices for the airline as well. For example, if an airline's market share was 23 percent higher because of a rise in enplanement shares as described above, fares on the route involving those airports would be expected to be, on average, 2 percent higher.

Effects of Factors Not Uniform

The effects of several factors on fares and market shares differed depending on the distance of the route, the price-sensitivity of the passenger, and the size of the airline.

Fares on short-haul routes (less than 1,000 miles) were more strongly influenced by slot restrictions, the level of operating costs of the least-cost airline, and airport congestion than were fares on long-haul routes. For example, short-haul routes affected by slot restrictions had fares that were an estimated 11 percent higher than fares on other short-haul routes, as compared with 4-percent higher fares for slot-restricted routes overall. On the other hand, fares on long-haul routes affected by slot restrictions were not higher than fares on other long-haul routes to a statistically significant degree.

Several factors—including slot restrictions, an airline's market share, and airport congestion—had stronger impacts on the fares of less price-

[8]Because both an airline's enplanement share and gate share at an airport indicate the degree to which the airline dominates operations at the airport, they were included alternatively in two different specifications of the model. For further details, see app. I.

appendixes that follow. Appendix I presents in detail the model structure and results, including our estimation methodology and sensitivity analyses. Appendix II describes data sources, discusses data quality issues, and provides a formal definition of each variable included in the model. Appendix III explains the methods for selecting a set of airline routes and processing airline data.

Throughout the model's development, we consulted with expert economists, including Dr. David Belsley of Boston College and Dr. Theodore Keeler of the University of California at Berkeley.

As agreed with your office, we did not obtain agency comments on a draft of this report. However, we did present preliminary model results to DOT officials.

Unless you publicly announce its contents earlier, we do not plan to distribute the report until 30 days after the date of this letter. At that time we will send copies to the Secretary of Transportation and other interested parties.

If you have any questions about this report, please contact me at (202) 275-1000. Major contributors to this report are listed in appendix IV.

K. M. Mead

Kenneth M. Mead
Director, Transportation Issues

Contents

Abbreviations

ASM	Available Seat Mile
CAB	Civil Aeronautics Board
CRS	Computerized Reservation System
DOT	Department of Transportation
IFAPA	International Foundation of Airline Passenger Associations
OAG	Official Airlines Guide
SH&E	Simat, Helliesen, and Eichner
TACO	travel agent commission override

GAO Airline Model

The purpose of our study was to determine whether several competitive factors, including certain barriers to entry, influence airlines' market shares and prices on individual routes. We examined these relationships by specifying models of market share and price, and estimating the parameters of those models to determine whether and to what extent these competitive conditions independently influence the dependent variables. In addition to the primary analysis of the determinants of market share and price, we tested the sensitivity of the model across certain subsamples of the data set and alternative specifications of the dependent price variable.

This appendix (1) discusses the theory of airline markets; (2) discusses the structure of the GAO model; (3) presents the specification of the base-case GAO model; (4) discusses estimation methodology; (5) presents base-case model results; and (6) discusses extensions to the base case and sensitivity-analysis issues.

Theory of Airline Markets

One of the primary justifications for deregulating the airline industry in the late 1970s was the expectation by many industry analysts that the ease of entry into individual route markets would lead to lower prices. Some empirical evidence suggested, for example, that, prior to the deregulation of the interstate airline industry, prices on intrastate routes, not under the jurisdiction of the Civil Aeronautics Board (CAB), had lower prices than similar interstate routes.[1] Thus, proponents of deregulation believed that eliminating the strict CAB controls over entry and fares would allow new airlines to begin service and existing airlines to enter new markets, leading to vigorous price competition in airline markets.

In particular, some of these analysts believed that airline markets might have the characteristics of a "contestable" market, in which market entry has no unrecoverable costs. When this is the case, an incumbent firm (even if a monopolist) is constrained from charging a high price by the mere threat that charging such a price will attract new firms to enter the market and compete with, or replace incumbent firms. Airline markets were believed to be easy to enter and exit because the capital an airline devotes to a route—the airplane—is highly mobile and can be readily reassigned to alternative routes.

[1] See Stephen Breyer, Regulation and Its Reform (Cambridge, Mass. Harvard University Press, 1982), p. 201.

After deregulation, as was predicted, many new airlines did begin service and existing carriers entered and exited routes as they reconfigured their route networks to the highly efficient hub-and-spoke system. As a result, price competition became very intense on many routes. By the mid-1980s, however, the industry began to become more concentrated as a result of a series of mergers and bankruptcies. While competition remained vigorous on many routes, certain routes became less competitive. Yet the benefits of deregulation remain clear: A study by Brookings Institution economists found, for example, that travelers receive an annual savings of approximately $6 billion because of deregulation.

Even with these benefits of deregulation, it has become apparent that the conditions of a contestable market do not apply to all airline routes. For example, although airplanes themselves are highly mobile, an airline needs other capital resources to serve a route and can find it difficult or costly to obtain these necessary resources at endpoint airports. In particular, airport resources such as gate space and landing slots can be difficult to obtain and cannot be reassigned from one airport to the next. The difficulty of obtaining these resources can act as a barrier to entry into certain airline routes.

Additionally, even if an airline can obtain the necessary resources to enter a particular route, the carrier may be unable to provide such service efficiently unless the route in question is "connected with," or integrated into that carrier's network of routes. Thus, while the relevant market in which prices are set is an individual route, there are important joint costs incurred in the production of a <u>set of routes</u>, so that entry into a particular route can be impractical in the absence of an established route network or simultaneous entry into several routes.

In addition to a realization by many analysts that the theoretical requirements of contestability may not apply to all airline routes, empirical studies have also shown that contestability theory is not always applicable in the airline industry. Several studies in the early 1980s tested the contestability of airline markets by examining whether market structure indicators (such as concentration ratios or actual market entry) were related to prices. These studies generally found that <u>actual</u> market structure was, indeed, related to pricing, indicating that the mere threat of entry was not constraining price to competitive levels on many routes.[2] More recently, Borenstein showed that a firm's share of originating passengers at the endpoint airports of a route influence the airline's prices. His work provides further evidence that ease of entry cannot be counted on to restrain prices in airline markets and points to the dominance of airlines at hub airports as a source of market power on routes to and from a dominated airport.[3]

Structure of the GAO Airline Model

Our model examines whether several competitive conditions, including certain barriers to entry, inhibit airline market contestability. By examining whether these factors influence how airlines compete and set prices on routes, we will shed light on some of the reasons that entry, or the threat of entry, does not appear to restrain prices on many airline routes. We are primarily interested in examining the effects of barriers to entry on airline prices. Some of these factors are likely to affect price directly, and others are more likely to affect price indirectly through their influence on an airline's ability to attract traffic on a route. Therefore, we developed a two-equation model to estimate the determinants of both price and market share on individual airline routes.[4]

The Nature of Barriers to Entry

In a well-functioning competitive market, free entry and exit of firms in response to price signals guarantees that, over time, the price of a good approximates its cost of production. However, markets will not be competitive if some characteristic or attribute of the market acts as a barrier to entry, so that firms are not able to enter markets easily in response to price signals that imply profitable opportunities for entering firms. Because entry barriers can insulate incumbent firms from competitive pressure, these barriers may allow only a few firms (or perhaps even one) to dominate a market and exercise market power by raising price above the level that would be set in a competitive market.

[2]See, for example, Elizabeth E. Bailey, David R. Graham, and Daniel P. Kaplan, Deregulating the Airlines (Cambridge, Mass.: The MIT Press, 1985), and Gregory D. Call and Theodore E. Keeler, "Airline Deregulation, Fares and Market Behavior: Some Empirical Evidence" Analytical Studies in Transport Economics, Andrew F. Daughety ed. (Cambridge, Eng.: Cambridge University Press, 1985), pp. 221-247.

[3]See Severin Borenstein "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry," Rand Journal of Economics, Vol. 20 No. 3 (Autumn 1989), pp. 344-365.

[4]Our price model is, in many respects, similar to that developed by Severin Borenstein in "Hubs and High Fares." Borenstein's model contains a more complete description of certain airline operations (by including, for example, variables for load factor, average plane size, and non-change of plane stops). Our model, on the other hand, includes several variables describing conditions related to entry barriers. In addition, we explicitly specify an equation for market share.

In our model we include several factors that have been claimed to affect entry conditions on routes and at airports. We test whether these factors have a detrimental effect on competition by estimating whether prices are higher on routes affected by these conditions. In some cases we test whether the factor affects prices directly, while in other cases we test whether the factor affects prices indirectly through its influence on market share.

Carrier-Route Focus

This work focuses on market share and price determination for a particular airline on a particular route. For example, one observation in the model will include information on United Airlines' price and market share on the route between Boston and San Diego, while another observation will be American Airlines' price and share of passengers on that route, and a third will be for United's operations on some other route. We refer to each observation as containing "carrier-route" data. In contrast to the carrier-route focus used here, most earlier work on the pricing effects of airline market structure examined price determination at the market level—that is, on a route-by-route basis—and analyzed the determinants of average price on a route without distinguishing between fares charged by different carriers.[5]

We believe that the lower level of aggregation involved in a carrier-route model provides a more refined description of market structure. First, we are able to model differences among carriers serving a given route. Second, we can model more directly many of the factors that are the focus of this analysis. In particular, several of the variables in the model describe attributes of specific airlines at specific airports and should not be expected to affect <u>all</u> carriers serving routes from those airports.[6]

Specification of the Base-Case Model

The base-case model presented in this section represents our judgment of the most useful single model for policy analysis purposes. Later in this appendix we present other specifications as alternatives to the base case.

[5]One exception is the work of Severin Borenstein, "Hubs and High Fares." A carrier-route focus is used in that analysis.

[6]For example, if an airline has a code-sharing agreement with a regional or commuter carrier at a particular airport, we would not expect that agreement to affect the pricing of <u>other</u> airlines at that airport. Without a carrier-route focus, we could not specify the code-sharing variable correctly.

In the case of the market-share equation, we present two estimations that differ with respect to the measurement of a key concept: endpoint dominance. In one estimation we use a carrier's enplanement (passenger boardings) share at endpoint airports, and in the second estimation we use the carrier's share of leased gates at the endpoint airports.

For the price equation, we present alternative estimations that differ with respect to underlying theoretical considerations. While most of the decisions and assumptions we made in specifying these models were fairly straightforward, others were more difficult because economic theory and policy analysis did not imply a single, logical choice regarding some aspect of model specification. In particular, theory implies that endpoint dominance at an airport should increase a carrier's prices on routes from that airport, yet it is not clear whether endpoint dominance affects prices directly in addition to its effect on price through an influence on a carrier's route market shares. Thus, we view the decision about whether endpoint dominance is appropriately included in the price equation as ambiguous. Consequently, we present two estimations for the base-case price equation: Version A includes variables describing the given carrier's endpoint dominance, and version B excludes those variables.

The following are some key aspects of the base-case specification:

- Sample. The base-case estimates are for the full set of 3,331 carrier-route observations drawn as a stratified random sample. The sample includes routes as short as 150 miles and as long as over 2,700 miles. We include observations for the operations of 12 major and national carriers on the sampled routes.[7] Appendix III discusses sample selection further.
- Market Definition. A route is defined as airline service between two airports. This includes nonstop, direct, and connecting service as long as the connecting service entails no more than one change of planes along a one-way segment of a flight. We call these routes airport-pairs. Alternatively, we could define the market as city-pair routes, in which all

[7]Nine of these carriers were classified as "major" carriers in 1988 with annual operating revenues over $1 billion. Three carriers were classified as "nationals" with annual operating revenues of between $100 million and $1 billion.

traffic for a given carrier between two cities (as opposed to individual airports) is included in a single observation.[8]

- Median Price. We use the median price charged by each carrier on each route as the base-case dependent price variable because it represents the price paid by the "typical passenger."[9]

Next we discuss the specific factors that we believe influence, or help to determine, the level of the dependent, or endogenous variables. Taking market share and price in turn, we discuss the factors that we believe influence each variable, and provide some justification for our expectations. We will also briefly define each of the independent, or exogenous variables. Appendix II offers a more detailed description of data sources and construction of variables.

The Market-Share Equation

We expect that a carrier's ability to establish market share on a route is affected by a variety of factors related to (1) the carrier's degree of endpoint presence at the airports defining a route, (2) characteristics of the route, and (3) characteristics of the given carrier relative to other carriers providing service on the route.

The following variables related to endpoint presence are included in the market-share equation:

- Endpoint dominance. To indicate when a carrier is dominant at an endpoint (which could affect entry conditions at the airport), two variables are alternatively included in the market-share equation. First, enplanement share equals the average of the carrier's share of enplanements, or passenger boardings, at each of the endpoint airports of a route. We expect that the higher a carrier's share of enplanements, the higher the carrier's market share on a route should be, since a large value for enplanements indicates the carrier has a large commitment to serving routes from an airport. Alternatively, we include a carrier's average gate share (the percentage of gate space at an airport leased by

[8]In an alternative version of the model, we defined a route to include traffic between two cities (such as Washington, D.C. to New York), as opposed to traffic between two airports (such as Washington National to LaGuardia). We performed an analysis using this broader market definition, since an airport definition could be too narrow if alternative service over airport-pairs within a city-pair are, in fact, highly substitutable. Overall, the results for the city-pair model are very similar to results for the airport-pair case, and we do not present them in this report.

[9]Alternatively, we could use the mean, or average price. In an alternative estimation using the mean price, results were, for the most part, similar to those for the median price specification. However, we found airport congestion (slots and congestion) and endpoint dominance (hub and code-sharing) had stronger effects on mean prices.

the given carrier) at the two endpoint airports. Similarly, we expect that a higher gate share will be associated with higher market shares on routes.

- Computerized reservation system share. Computerized reservation systems (CRS) are owned by airlines and used by travel agents to make airline bookings for their clients. The variable, based on revenues booked through carriers' CRSs, measures the degree to which the given carrier has a CRS advantage over other carriers serving the route. We expect that when a carrier has a larger than average share of the CRS market in the two endpoint cities, it will be able to achieve higher market shares on a route because travel agents using that carrier's CRS are more likely to book their clients on the flights of that airline.[10, 11] The effect of CRS dominance may make entry for other carriers into routes from those cities more difficult.
- Scheduled service. This variable measures the capacity that a carrier has devoted between two airports on the basis of its shares of available seat miles and scheduled connecting flights. We expect that when a carrier devotes a large percentage of the total capacity available between two airports, it will be offering more convenient service than its rivals, enabling it to gain a higher market share on the route.
- Enplanements of others. This variable equals the largest enplanement share of any carrier other than the given carrier at either endpoint airport of a route. This variable measures the degree to which some other airline is in position to have a large market share on a route. We expect that larger enplanement shares for the largest other airline, will lead to smaller market shares for the given carrier.

The following variables related to characteristics of routes are expected to affect a carrier's market share:

- Distance. This variable equals the one-way, straight-line mileage of a route. Longer routes are expected to be associated with lower market shares because such routes will be served efficiently by a greater

[10]"Study of Airline Computer Reservation Systems," May 1988, U. S. Department of Transportation, Office of the Secretary of Transportation, DOT-P-37-88-2. DOT finds that CRS dominance leads to a "halo" effect whereby carriers gain larger revenue shares on routes to and from a city because of the use of their CRSs by local travel agents.

[11]In some estimations of the model we also include a variable (that is somewhat related to the CRS variable), which we call "TACOs" (travel agent commission overrides). The TACO variable represents monetary bonuses paid by airlines to travel agents who book a large volume of business with the particular airline. Because DOT information on TACOs is only available for the five carriers that owned a CRS system in 1986, we can only analyze the effects of TACO payments for these five carriers. When we included the TACO variable in the market-share equation, its coefficient is negative and significant. In the price equation, the coefficient on the TACO variable is positive and significant.

number of carriers offering connecting service through intermediate hub airports.

- Traffic volume. This variable is the total number of passenger directional trips on a route.[12] Since some minimum scale on a route might be needed for a carrier to serve the route efficiently, a large passenger base offers an opportunity for a greater number of carriers to be in a position to profitably serve the route. Therefore, routes with a greater traffic volume will generally be characterized by lower market shares.
- Route direct traffic. Route direct traffic measures the proportion of traffic on a route (for all carriers serving the route) that is direct rather than connecting.[13] The degree of direct traffic on a route is largely related to the distance and the traffic volume of the route. When a route is mostly served with direct service, connecting service is likely to be considered a poor substitute, and a carrier will need to offer direct service to compete effectively. To provide direct service efficiently, a carrier will usually need to have some significant degree of presence at one of the endpoint airports so that the route is integrated into the carrier's route network on a direct service basis. Thus, we expect higher market shares on primarily direct service routes because such routes will likely fit into fewer carrier's systems.

The following variables are included in the market share equation to describe characteristics of carriers:

- Relative directness. This variable measures the given carrier's percentage of direct traffic on a route compared with the percentage of direct traffic offered on the route by all carriers. The variable indicates any advantage in gaining passengers that a carrier may have because it offers higher quality service. If a carrier is offering more direct service than its rivals on a route, we expect that carrier to gain a higher market share.
- Relative costs. This variable equals the weighted average of the cost per available seat mile of all carriers serving a route divided by the cost per available seat mile of the given carrier.[14] Cost per available seat mile

[12] A single directional trip represents one passenger flying a one-way segment of an itinerary. Therefore, a single person flying round-trip would count as two directional trips. Two people flying a one-way segment would also count as two directional trips.

[13] Passengers fly direct itineraries if they fly from their origin to destination point on a single plane. Nonstop flights and flights that stop but do not require the passenger to change planes are both considered to be direct flights.

[14] For cost information, we use the distance-adjusted cost per available seat mile for a carrier. See app. II for a discussion of our derivation of these data.

represents a measure of carrier cost that is systemwide and not necessarily indicative of costs incurred in serving any particular route. Holding constant such factors as distance, traffic volume, enplanements, and enplanements of others, the given airline may be in a better position to gain market share if it is competing with higher-cost airlines than with lower-cost airlines. Thus, the lower a carrier's costs relative to its rivals serving a particular route, the higher its market share is expected to be.[15]

- Relative preferences. Using information from a survey by the International Foundation of Airline Passenger Associations in which passengers identified their most preferred airline, we derived a relative preference variable that is equal to the given carrier's preference rating divided by the average rating for all carriers serving the route. We expect that a carrier with a high relative preference ranking will gain a higher market share, and therefore, the variable should have a positive coefficient.

Table I.1 lists the variables included in the market-share equation.

Table I.1: Variables Included in Market-Share Equation

Type of factor	Specific variable
Degree of endpoint dominance	Endpoint dominance Enplanements Gates CRS share Scheduled service Enplanements of other
Route characteristics	Distance Traffic volume Route direct traffic
Carrier characteristics	Relative directness Relative costs Relative preferences

The Price Equation

For the price equation, we identified four categories of factors influencing price: (1) factors related to the costs of serving a route, (2) factors related to characteristics of the airports or cities at the endpoints of a route, (3) factors related to market structure and degree of endpoint dominance, and (4) characteristics of the given carrier.

[15]Since the relative cost variable is equal to the weighted average costs of all carriers on the route divided by the costs of the given carrier, the expectation is that this variable will have a positive coefficient.

The following are the cost variables included in the price equation:

- Distance. Distance measures the one-way, straight-line mileage of a route. Because an airline incurs higher costs (for example, fuel costs) in serving a longer route, increased distance should be associated with higher prices.[16]
- Minimum Costs. This variable equals the cost per available seat mile of the least-cost carrier serving a route. The higher the cost of the "least-cost carrier," the higher we expect prices for all carriers serving the route. This is because the presence of a low-cost carrier should discipline the pricing behavior of all carriers serving a route.[17]

Variables describing airport or city characteristics include:

- Multiple-airport service. This variable measures the proportion of traffic served on alternative airport-pairs within the city-pair. If the airport-pair route for an observation is St. Louis to Chicago O'Hare Airport, for example, the multiple-airport service variable will indicate whether and to what extent there is also service on the St. Louis to Chicago Midway Airport airport-pair route. Routes for which this form of competing service is available are expected to have lower prices.
- Tourism. This variable equals the weighted average value of hotel revenues per capita in the two endpoint cities.[18] Routes with a tourist destination at an endpoint are expected to have lower prices since a greater proportion of consumers traveling these routes are highly price-sensitive.
- Noise restrictions. The noise variable is a dummy variable[19] equal to one if either endpoint airport has any of several noise restrictions. Noise restrictions can raise an airline's cost of serving an airport by requiring, for example, that the most modern aircraft (quiet but expensive) be

[16]We recognize that we make a strong assumption in specifying distance to have a constant elasticity. However, since the most appropriate specification of distance is not a primary focus of this analysis, we chose to use this simple specification.

[17]Our cost variable is continuous. Therefore, we do not really identify routes as having or not having a "low-cost carrier." Instead, it is more accurate to think of this variable as indicating the cost level of whichever carrier serving the route happens to have the lowest cost.

[18]The weights are based on the proportion of traffic originating at the opposite city of the route. That is, for the Las Vegas to Cleveland route, the value of the Las Vegas tourism index is weighted by the percent of passengers on the route originating in Cleveland, and the value of the Cleveland tourism index is weighted by the percent of passengers originating in Las Vegas.

[19]A dummy variable takes a value of either zero or one, depending on whether a particular condition holds. For example, in the case of the noise dummy variable, observations for which neither endpoint airport has any noise restrictions will have a zero value for this variable, while observations with an airport that has some noise restrictions will have a value of one.

used. Since some airlines do not have as large a fleet of newer aircraft as other airlines, noise restrictions can act as a barrier to certain airlines serving certain airports.

- Congestion. This variable measures the degree of runway congestion at the endpoint airports of a route. Specifically, the variable equals the average of the two endpoint airports' operations (takeoffs and landings) per runway. Serving congested airports can be more costly for an airline in a variety of ways. For example, personnel and equipment cannot be applied to some alternative use when these resources are tied up because of delays at a congested airport. In addition, prices on those routes may be higher because of the high demand, relative to capacity, at congested airports. That is, higher fares at congested airports in part reflect the scarcity of resources at those airports. Thus, we expect that routes involving more congested airports will have higher prices.
- Slots. Slots is a dummy variable that indicates if the route involves one of the four slot-controlled airports at which the number of takeoffs and landings is restricted under the High Density Rule.[20] Serving a slot-controlled airport can raise the cost of service because an airline must obtain or hold the transferable slot (which can be purchased only at a substantial cost) in order to operate. Since an airline must obtain a slot to serve these airports, slot restrictions can be a barrier to entry at these airports. Therefore we expect the slots variable to have a positive coefficient.
- Majority-in-interest agreements. This is a dummy variable indicating whether either endpoint airport has a majority-in-interest clause in the use agreement between the airport and airlines serving the airport. These clauses give airlines a voice in airport planning decisions, such as expansion plans, where the cost of expansion would be recovered from fees charged to the airlines. This could give the signatory airlines the power to inhibit expansion at an airport. We expect that a majority-in-interest clause can indicate that an airport may have been unable to respond fully to rising demand for service.[21] The power of incumbent airlines to influence airport expansion could make entry by other airlines more difficult and could therefore lead to higher prices.

[20]Under the High Density Rule (14 C.F.R. Part 93 Subpart K), scheduled airline service is limited to a specified number of takeoffs and landings (i.e., slots) per hour or half-hour period. Airlines wanting to fly into or out of a slot-controlled airport must reserve a slot in advance for the appropriate time period.

[21]In the case of both the majority-in-interest and the expansion variables, our dummy variables indicate which airports, in responding to our survey, stated that these conditions apply. We do not know, however, whether these factors actually constrained airport behavior in the sense that airport expansion was actually necessary based on current demand for air service. Therefore, it is not clear that prices would be affected by these conditions.

- Expansion. Expansion is a dummy variable equal to one if either endpoint airport indicated (in responding to the GAO airport survey) that it would have difficulty expanding airport facilities for any of a variety of reasons. Because it may be difficult for airlines to enter routes involving airports unable to expand their facilities, we expect prices on affected routes to be higher.

We include the following market structure variables in the price equation:

- Market share. This variable equals the share of traffic on the route handled by the given carrier. We expect that market share indicates market power, and thus larger market shares will be associated with higher prices.
- Herfindahl index. The Herfindahl index measures overall concentration in a market.[22] Given the level of market share, we expect carriers will be able to get a higher price when the overall concentration in a market is higher.
- Traffic volume. This variable measures the total number of origin-to-destination passenger directional trips on a route. We expect that heavily traveled routes will have lower prices because the potential for entry is greater (since more firms can achieve a minimum efficient scale in serving the route), and because larger traffic volumes enable carriers to achieve higher load factors and/or to use larger aircraft, which should lower the cost of service.

Version B of the base-case price equation also includes two variables that provide information about a carrier's endpoint presence and operations:

- Code-sharing.[23] Code-sharing is a dummy variable indicating that the given carrier has a code-sharing agreement with a regional or commuter airline at either endpoint airport of a route. Under code-sharing agreements, the regional or commuter carrier will generally use the larger

[22]The Herfindahl index equals the sum of the squared market shares of all firms serving a market.

[23]The code-sharing variable is not necessarily an endpoint-dominance variable. However, since we have specified code-sharing as a dummy variable, and since much of code-sharing occurs at carrier hubs, we believe that in our model, and as specified, code-sharing operates much as an endpoint dominance variable. In fact, in one specification of the model we included two separate variables to account for the different contexts of code-sharing: one variable was a dummy for code-sharing if it occurs at an airline's hub airport, and another one was a dummy for code-sharing when it occurs at nonhub airports. We found that only code-sharing at a hub airport was price-increasing. Moreover, a statistical test indicated that the effects of these two variables on price were, indeed, statistically different from one another.

carrier's two-letter airline designation code in its own flight listings so that connecting flights between the two airlines at that airport will be viewed as online (same carrier) connections. We expect that code-sharing agreements will be associated with higher prices on routes to and from airports where carriers have such agreements. This is because the major carrier is obtaining additional connecting traffic onto its primary flights from connecting code-sharing flights, thus raising load factors and reducing the amount of price discounting needed to fill its planes.

- Hub.[24] Hub is a dummy variable indicating whether either endpoint airport is a hub for the given carrier.[25] Hub is a barrier-to-entry factor related to endpoint dominance. We expect that, generally, it is more difficult for entering carriers to compete with an incumbent carrier at its hub airport.[26] Fares on routes involving a hub airport are therefore expected to be higher.

The following variables are included in the model to indicate characteristics of carriers:

- Carrier size. This variable equals the total domestic revenue passenger miles of the carrier divided by the total domestic passenger miles of all the carriers in our sample. We expect that carriers operating a large

[24]A hub dummy variable is a fairly generic indicator of endpoint presence. The estimated effect of this variable could be capturing the pricing effects of any number of airport-dominance factors. For example, the variable could be picking up an effect related to a passenger's membership in a dominant carrier's frequent flyer program (which we did not include in the model because of data availability problems), or it could represent a premium paid for nonstop service that is more prevalent from hub airports.

[25]To define hub airports we use the set of airports so defined in "Hub Operations: An Analysis of Airline Hub and Spoke Systems Since Deregulation," prepared for the Air Transport Association by Simat, Helliesen, and Eichner (SH&E), Inc., (May 1989). Because some of the airports denoted as hubs in the study did not, in our opinion, represent airports at which carriers engage in a significant degree of interconnections among flights, we made some modifications to the list of hub airports used in the SH&E study.

[26]The hub variable is similar to enplanements or gates used in the route-share equation in the sense that it indicates endpoint dominance. The reason that we choose to use the hub variable (rather than enplanements or gates) in the base-case price equation is because of certain characteristics of the code-sharing dummy variable included in the price equation. As mentioned earlier, code-sharing agreements usually occur in two types of situations: at a carrier's hub airports and at other large airports (such as those in Boston, New York and Los Angeles). Because most carriers have code-sharing at their hub airports, and because we treat code-sharing as a simple dummy variable, we were concerned that an estimated code-sharing effect cannot be distinguished from a hub airport effect. By including a hub dummy in the estimation, we believe that our code-sharing variable now can indicate any additional effect—over and above a hub effect—that code-sharing is providing carriers.

national network may be able to receive higher prices because passengers are familiar with these carriers and because these carriers' frequent flyer programs are more extensive. Additionally, since larger airlines serve many of the same markets, larger carriers may achieve higher prices in part because of possible cross-market collusion.

- Preferences. Passengers' preferences for airlines may be important in explaining carrier pricing. However, larger airlines tend to be more preferred by passengers because passengers are more familiar with these airlines. Because the carrier size variable accounts for pricing effects related to size differences among carriers, the preference variable is adjusted for size-related preference differences. Preference is specified as a 0-1-2 categorical dummy variable indicating how preferred a carrier is relative to its national size. If an airline was significantly more preferred than its relative national size, the preference value for that airline is equal to two. If the carrier's preference ranking was about the same as its relative national size, its preference value is equal to one, while a carrier with a lower preference ranking than its relative national size would have a preference value equal to zero.[27] A higher preference ranking is expected to be associated with higher fares.

- Relative directness. This variable equals the carrier's percent of direct traffic relative to the percent of direct traffic offered by all carriers on the route. A carrier offering better quality service on average than its rivals should be able to obtain higher prices.

Table I.2 lists the variables included in the price equation.

[27]This specification assumes that pricing effects of a change from a zero value of preference to a value of one is the same as a change from a value of one to a value of two. In an alternative specification, we estimated the separate effects of having at least an average (size-adjusted) preference value from the effects of having a better than average (size-adjusted) preference value. We found that the latter effect was the stronger of the two, but other results in the model were not affected by this change in specification.

Table I.2: Variables Included in Price Equation

Type of factor	Specific variable
Costs	Distance Minimum costs
Airport or city characteristics	Multiple airports Tourism Noise restrictions Airport congestion Slots Majority-in-interest Expansion difficulties
Market structure	Market share Herfindahl Traffic volume
Version B (Endpoint dominance included)	Hub Code-sharing
Carrier characteristics	Carrier size Preferences Relative directness

Estimation Methodology

The parameters of the price and market-share equations are jointly determined because a carrier's market share is itself one of the factors hypothesized to influence a carrier's price on the route. Therefore, the coefficients presented in tables I.3 and I.4 must be estimated using simultaneous-equations techniques. We used two-stage least squares, which is an appropriate estimation method for this situation.[28] The first-stage regression determines "fitted" (or estimated) market share values, for use in estimating the effects of market share on price. In the first-stage equation, market share is regressed on all predetermined variables in the price or market-share equations, as well as other exogenous variables related to airline operations, demand conditions, and geographic indicators that are not included as part of the structural equations for either price or market share.[29]

We do not include price, or any other variable jointly determined with market share, as part of the structural explanation of market share.

[28] As an alternative estimation approach, we estimated the model using three-stage least squares, which accounts for the possibility that error terms are correlated across equations. However, because the residuals from the two equations were not highly correlated and the three-stage results did not provide any meaningful differences from the two-stage results, we did not use this procedure as our primary method of estimation.

[29] We include additional exogenous variables in our first-stage regression for route share based on the assumption that these additional variables, although not part of the structural explanation of either market share or price, are part of the larger (and unanalyzed) system related to airline demand, costs, and operations. A list and description of these variables is found in app. II.

Therefore, technically, the market-share equation is estimated using ordinary least squares. However, some of the variables included in the market-share equation, most notably scheduled service and relative directness are jointly determined to some extent with a carrier's market share on a route.[30] We treat these variables as exogenous because, given hub-and-spoke technology, an airline's decisions about the capacity and service on an individual route are based not only on how many origin-to-destination passengers the carrier handles on that route, but also on how many passengers the airline can expect to carry on many other routes on which passengers are carried jointly with the route in question.

To estimate the model, we used 3,331 carrier-route observations drawn as a stratified random sample of domestic airline routes for the third quarter of 1988. These observations represent carrier service on 1,667 airport-pair routes. (For details of our sample selection, see app. III.) We used a multiplicative model and, therefore, all continuous variables are expressed in natural logarithms.[31] Thus, the estimated coefficients for the continuous variables can be interpreted as elasticities[32] . In the case of dummy variables, coefficients are interpreted as percentage changes (in decimal form) resulting from the existence of the condition described by the dummy variable.

We tested for and found evidence of heteroskedasticity in both our market share and price equations. Heteroskedasticity occurs when error terms do not have a common variance. For the estimation of the market-share equation, t-statistics are obtained from the White heteroskedasticity-consistent variance-covariance matrix.[33]

[30] A variable measuring the systemwide (as opposed to route-specific) operating costs of the given carrier, compared with other carriers serving the route, is included in the market-share equation because this information provides a general representation of the competitive pressure provided by other airlines serving the route. This variable is essentially acting as a proxy for a relative price variable that was not feasible to include in the model. Because the relative cost variable is based on systemwide costs (and not route-specific), we believe it is appropriately treated as exogenous.

[31] Because the natural logarithm of zero is undefined, we added one to the observed value of any continuous variable that can take the value of zero before we took the natural logarithms of these variables.

[32] An elasticity equals the percentage change in the value of one variable resulting from a given percentage change in the value of another variable. In the case of our price equation, a coefficient for one of the explanatory factors expresses the percentage change in price associated with a percentage change in the value of that explanatory factor.

[33] See H. White, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica, Vol. 48, (1980), pp. 817-838.

For the price equation, we applied a heteroskedasticity correction because the model generally provided less accurate predictions of prices on short-distance routes. Thus, the heteroskedasticity correction had the effect of giving less weight in the regression to shorter routes.[34]

We also examined the degree of collinearity in the model. Although we did find evidence of collinearity among the independent variables in each of our equations, the problem was not severe. In most cases, the presence of collinearity did not harm our ability to make the statistical inferences that were most important in this model.

Base-Case Results

Tables I.3 and I.4 present results estimated from the base-case specification of the model. In general, the estimation results suggest that several of the competitive conditions included in the model influence market shares and prices. These results are generally consistent with our prior expectations about the mechanics of the airline industry's competitive structure and pricing and with findings of previous studies.[35] Additionally, most other factors related to conditions on routes and characteristics of carriers appear to affect market shares and prices as we expected.

[34]Specifically, the correction was based on iteratively-determined weights generated from a regression of the absolute value of the two-stage least squares residuals on the natural log of distance and a constant term. The predicted values of this regression were used to weight the dependent and explanatory variables in the price equation. Under an alternative specification of the heteroskedasticity correction, our results were largely unchanged.

[35]See, for example, Elizabeth E. Bailey, David R. Graham, and Daniel P. Kaplan, Deregulating the Airlines (Cambridge, Mass.: The MIT Press, 1985); Gregory D. Call and Theodore E. Keeler, "Airline Deregulation, Fares and Market Behavior: Some Empirical Evidence" Analytical Studies in Transport Economics, Andrew F. Daughety ed. (Cambridge, Eng.: Cambridge University Press, 1985), pp. 221-247; D. R. Graham, D. P. Kaplan, and D. S. Sibley, "Efficiency and Competition in the Airline Industry," The Bell Journal of Economics, Vol. 14, No. 1, (Spring 1983); and Severin Borenstein "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry," Rand Journal of Economics, Vol. 20, No. 3 (Autumn 1989), pp. 344-365.

Table I.3: Base-Case Market-Share Equation

Explanatory variables	Coefficients Case 1	Case 2
Intercept	4.136[a] (25.823)	5.727[a] (37.496)
Scheduled service	0.156[a] (20.579)	0.170[a] (20.895)
Distance	−0.106[a] (−7.998)	−0.160[a] (−11.542)
Route direct percent	0.078[a] (9.547)	0.112[a] (13.134)
Relative directness	0.119[a] (13.990)	0.158[a] (18.525)
CRS	0.082[a] (10.440)	0.080[a] (9.371)
Relative preferences	−0.038[a] (−4.211)	−0.022[b] (−2.282)
Traffic volume	−0.179[a] (−20.639)	−0.213[a] (−23.758)
Enplanements of others	−0.134[a] (−7.000)	−0.238[a] (−12.159)
Relative costs	0.562[a] (12.023)	0.583[a] (11.779)
Endpoint dominance		
Enplanements	0.323[a] (22.578)	
Gates		0.065[a] (6.092)
Summary statistics		
R^2	.57	.50
n	3,331	3,331

Notes: Dependent variable is expressed in logarithmic form.

The first estimation includes enplanement shares as a measure of endpoint dominance and the second estimation uses gate shares.

t-statistics in parentheses.

All variables are expressed in natural logarithmic form.

[a]Statistically significant at the 1-percent level.

[b]Statistically significant at the 5-percent level.

Table I.4: Base-Case Price Equation, With and Without Endpoint Dominance

Explanatory variables	Coefficients Version A	Coefficients Version B
Intercept	1.302[a] (7.389)	1.201[a] (6.569)
Market share[b]	0.087[a] (5.202)	0.135[a] (8.479)
Herfindahl index[b]	−0.026 (−1.625)	−0.046[a] (−2.850)
Slots	0.038[a] (3.874)	0.040[a] (4.007)
Distance[b]	0.342[a] (45.279)	0.338[a] (43.419)
Code-share	0.017[c] (1.808)	
Hub	0.059[a] (5.121)	
Relative directness[b]	−0.011[a] (−2.455)	−0.015[a] (−3.176)
Multiple airport service[b]	−0.003 (−1.394)	−0.004[c] (−2.277)
Tourism[b]	−0.035[a] (−8.118)	−0.037[a] (−8.264)
Preference ranking	0.046[a] (9.347)	0.042[a] (8.254)
Airline size[b]	0.075[a] (10.814)	0.084[a] (11.949)
Traffic volume[b]	−0.032[a] (−8.626)	−0.021[a] (−6.173)
Noise restrictions	0.010 (1.340)	0.008 (1.104)
Majority-in-interest	0.029[a] (3.923)	0.032[a] (4.173)
Expansion difficulties	−0.001 (−0.113)	−0.006 (−0.550)
Minimum cost[b]	0.228[a] (9.490)	0.235[a] (9.494)
Airport congestion[b]	0.066[a] (5.146)	0.072[a] (5.336)
Summary statistics[d]		
R^2	.54	.52
n	3,331	3,331

Notes: Dependent variable is expressed in natural logarithmic form.

Version A includes variables related to endpoint dominance, while version B deletes those variables.

t-statistics in parentheses.

[a]Statistically significant at the 1-percent level.

[b]Variable expressed in natural logarithmic form.

[c]Statistically significant at the 5-percent level.

[d]R^2 is calculated as the square of the correlation between the dependent variable and its predicted value.

Market Share and Pricing Effects of Endpoint Dominance

For the most part, the variables that describe the scale of an airline's presence at airports influence both market share and price as anticipated. In the market-share estimation, enplanement shares—or alternatively gate shares—have a positive influence on a carrier's market share on a route; larger CRS market shares are also associated with higher market shares for a carrier; and larger enplanement shares by some other carrier are associated with lower market shares for the given carrier. The estimation results for the price equation indicate that a carrier hub at an endpoint airport is associated with higher prices for the carrier with the hub, and that a code-sharing agreement at an endpoint airport might enable carriers to achieve a small increase in price.

The primary difference in results when hub and code-sharing are excluded from the price equation (version B of the base case) is that the market-share coefficient is considerably larger. That is, if we do not include endpoint-dominance variables directly in the price equation, market-share (in part determined by endpoint dominance) is more important in explaining price. If endpoint presence is included in the price estimation, the magnitude of the market-share effect is diminished.

Effects of Certain Barrier-To-Entry Factors at Airports

Several variables were expected to influence prices because they represent barrier-to-entry factors at airports. We find that the existence of slot restrictions at an airport on a route is associated with higher prices, as expected. A majority-in-interest agreement between airlines and an airport also appears to have a positive influence on price. However, we did not find any independent pricing effect related to either noise restrictions or limited expansion opportunities at airports.

Market Share and Pricing Effects of Quality and Reputation

Several of the variables in the model characterize the quality or reputation of carriers. The results for the market-share estimation indicate that a carrier offering a greater percentage of direct service than its rivals will achieve a higher market-share. However, in the price equation, the relative directness variable has a negative coefficient, indicating that offering more direct service is associated with lower prices.[36]

[36]See later discussion of airline subsamples for a further analysis of these findings.

Another unexpected result occurs for the relative preference variable, which has a negative and significant coefficient in the market-share equation, indicating that more preferred carriers actually have lower market shares. Yet, the coefficients on both the preference and size variables included in the price equation are positive and significant, indicating that larger and more preferred airlines are able to charge higher prices. There appears, therefore, to be some evidence that factors related to major carrier reputation and size influence market shares and prices, yet not all results for these variables were in keeping with our expectations.

Other Model Results

Estimated parameters for key route characteristics such as distance, traffic volume, and tourism indicate that these variables influence market share and price as we expected. Other results include the following:

- A carrier that has lower than average costs (on a systemwide basis) among the carriers serving a route will obtain a higher market share,
- The level of operating costs of the least-cost airline will tend to affect prices of all carriers serving a route,
- The coefficient on the multiple-airport service variable is not statistically significant in version A of the base case but is significant in version B of the base case,[37]
- The coefficient on the market-share variable is positive and significant as expected,[38]
- Overall concentration in a market, measured by the Herfindahl index, has an unexpectedly negative and significant coefficient in some specifications,[39] and

[37]We also used a more general definition of the multiple airport service variable. As an alternative, we used a dummy variable if either endpoint airport on a route was within 100 miles of another airport that was classified by FAA as a medium or large airport. This specification was not significant in any version of the model.

[38]Although actual route-shares must lie between 0 and 100 percent, estimated route shares are not so bounded. As an alternative specification of the route-share variable, we used a logit transformation to constrain its range between zero and 100 percent. Model results were not sensitive to this specification of market share.

[39]This result, although not in keeping with our expectations, is not uncommon in empirical studies examining the determinants of profitability where both market share and a measure of overall industry concentration are included as explanatory variables. See, for example, David J. Ravenscraft "Structure-Profit Relationships at the Line of Business and Industry Level," The Review of Economics and Statistics, Vol. 65, (Feb. 1983), and Dennis C. Mueller, Profits in the Long Run, (Cambridge, Eng. Cambridge University Press, 1986), pp. 83 and 103.

- Prices tend to be higher on routes from congested airports.[40]

Extensions of the Base Case and Sensitivity Analysis

Of necessity, particular assumptions and decisions were made in specifying the base-case model. Because economic theory does not always lead to an obvious choice in how to specify certain aspects of the model, examining the sensitivity of the estimations to alternative specifications can be useful and important in assessing the robustness of the model. Additionally, because there are significant variations within the sample used for the base-case estimation in terms of the types of carriers and routes included, we believe that, as part of our sensitivity analysis, it is appropriate to examine the stability of model estimates over subsamples of the full data set.

The model does appear to be sensitive to the particular sample used for estimation, since statistical tests indicate that the determinants of airline competition and prices are different across subsamples of various route and carrier characteristics. Additionally, there are some important differences in model estimates when the price variable used as the dependent variable is a high-end fare (such as the 75th percentile price) as opposed to a low-end fare (such as the 25th percentile price). Differences in estimations over subsamples and for alternative price specifications conform, for the most part, to our prior expectations.

In this section we present some results based on the following alternatives to the base-case model: (1) subsamples of the full data set based on distance, (2) subsamples based on carrier size, and (3) estimations in which the dependent price variable is alternatively a high (75th percentile) or low (25th percentile) price.

[40]Measuring congestion is highly complex. The operations per runway variable that we created is a crude proxy for congestion because it does not account for differences in weather, runway configuration and many other factors affecting airport congestion. Further, some of the cost-raising effects of congestion may be due to general airspace congestion that will not be indicated in an airport-based measure. Because the runway variable is only an approximation of true congestion, we also measured congestion by creating a dummy variable based on an FAA list of airports subject to large amounts of delay. Using FAA information on hours of delay at each airport, we classified an airport as "delayed" if it had more than 20,000 hours of delay in 1987 as reported by FAA in its 1989 report Airport Capacity Enhancement Plan 1989, U.S. Department of Transportation, Federal Aviation Administration, (DOT/FAA/CP-89-4, May, 1989). When we replace the runway variable with the delay measure, we find that delay is also price-increasing, and other model results are not sensitive to the change in how congestion is measured.

Subsamples Based on Distance Categories

If a route begins or ends at a hub airport, the carrier with the hub may be able to exercise market power on the route if other carriers cannot serve the route directly from their own hub or offer connecting service through an intermediate hub. If a route is likely to be served via connecting service through hub airport(s), we would expect that such routes would be fairly competitive and that dominance at an endpoint airport does not necessarily imply market power.

Because large endpoint presence may not confer significant market power to a dominant carrier on routes for which other carriers can provide service on the route through one of their intermediate hub airports, it might be useful to determine if the effect of endpoint dominance is the same for all types of routes. We viewed the distance of a route to be a reasonable proxy for whether a route is likely to be competitive because of the viability of connecting service through intermediate hub airports—longer routes are more likely to pass over a greater number of hub airports. To examine whether competitive factors affect long and short routes differently, we split the sample at 1,000 miles (about halfway between the mean and the median distance in the sample) and estimated the model for both subsets of observations.[41]

Our findings, shown in table I.5, suggest that the determinants of price on long routes are different than on short routes. A statistical test indicated that the estimated coefficients for the two distance subsets are not generated by the same model structure.[42] The coefficients of several variables are statistically different across the distance subsamples.

[41] We used alternative distance partitions, and the results were robust across these alternative breaks.

[42] A Chow test is a statistical test of the equivalence of regression coefficients between two sets of data. The test generates a statistic, which is compared to a critical value based on the number of regression coefficients and the number of observations. In the case of the price equation, the test statistic for the distance subsets was 14.4, which exceeds the critical value for a 5-percent test with (18,3295) degrees of freedom. This means that we reject the hypothesis that the two sets of coefficients are equivalent.

Table I.5: Price Equation, Subsamples of Short and Long Routes

	Coefficients	
Explanatory variables	**Short routes**	**Long routes**
Intercept	1.065[a] (2.878)	1.319[a] (6.117)
Market share[b]	0.068[a] (2.040)	0.099[a] (5.064)
Herfindahl index[b]	−0.011 (−0.342)	−0.030[c] (−1.647)
Slots	0.109[a] (5.520)	0.003 (0.276)
Distance[b]	0.271[a] (14.756)	0.405[a] (26.345)
Code-share	0.010 (0.450)	0.011 (1.025)
Hub	0.090[a] (4.028)	0.045[a] (3.228)
Relative directness[b]	−0.008 (−0.801)	−0.013[a] (−2.542)
Multiple airport service[b]	−0.016[a] (−4.201)	0.000 (0.186)
Tourism[b]	−0.047[a] (−4.351)	−0.029[a] (−6.038)
Preference ranking	0.064[a] (5.991)	0.042[a] (7.558)
Airline size[b]	0.055[a] (4.232)	0.082[a] (9.686)
Traffic volume[b]	−0.056[a] (−6.913)	−0.019[a] (−4.415)
Noise restrictions	0.014 (0.862)	0.016[c] (1.998)
Majority-in-interest	0.015 (0.833)	0.029[a] (3.442)
Expansion difficulties	−0.009 (−0.434)	−0.008 (−0.600)
Minimum cost[b]	0.466[a] (10.172)	0.062[c] (2.108)
Congestion[b]	0.106[a] (3.657)	0.041[a] (2.766)
Summary statistics[d]		
R^2	.43	.41
n	1,583	1,748

Notes:Dependent variable is expressed in natural logarithmic form.

Long routes are defined to be those exceeding 1,000 miles, and short routes are those less than 1,000 miles.

t-statistics in parentheses.

[a]Statistically significant at the 1-percent level.

[b]Variable expressed in natural logarithmic form.

[c]Statistically significant at the 5-percent-level.

[d]R^2 is calculated as the square of the correlation between the dependent variable and its predicted value.

These results suggest that long routes are, indeed, more competitive than short routes. Prices on long routes are more influenced by distance—a primary cost factor. Yet, as expected, prices on long routes are less influenced by endpoint dominance (hubs) since connecting service can mitigate the effects of hub dominance. Further, since longer routes generally have more carriers than shorter routes, prices are less affected by factors that mitigate the exercise of market power. For example, the presence of a low-cost carrier, a high degree of tourist traffic on the route, and additional competition provided over alternative airport-pairs are less important in determining prices of longer routes. Also, the costs incurred in serving congested airports (indicated by the slots and congestion variables) appear to be borne by passengers on short (and probably less competitive) routes.

Subsamples Based on Carrier Size

We include 12 carriers in the base-case estimations. These carriers vary significantly in terms of national size, route networks, strategies, and reputations. By pooling observations for all carriers into the same estimations, we assume that model coefficients are the same across all types and sizes of carriers. That assumption may be incorrect. Because the data are firm-specific, it is important to test whether such pooling is appropriate.

In order to examine this issue, we tested the stability of model estimates for the three largest carriers (American, Delta, and United) compared with all other carriers in the model.[43] Our findings suggest that the determinants of price across these carrier groupings are different, because a statistical test, which tests the hypothesis that the sets of coefficient estimates across these two subsets are generated by the same

[43]We also tested the stability of model estimates for the three smallest carriers (America West, Midway, and Southwest were all "national" carriers in 1988) compared to all other carriers in the model. The subsampling results suggest that the three smallest carriers are different from the rest of the industry. Several variables (such as market share, slots, and the hub dummy) are not significant in the national carrier subsample, but they are significant in the major carrier subsample (the nine remaining carriers) as well as the full carrier sample. Prices for the national carriers seem to be determined primarily by distance, density, minimum costs, and carrier size. These results suggest that smaller carriers may set prices primarily by formula and are not generally as able as larger carriers to influence price and exclude competition.

model structure, is rejected.[44] Results for this subsampling are presented in table I.6.

Table I.6: Price Equation, Airline Subsamples

	Coefficients	
Explanatory variables	Large airlines	Other airlines
Intercept	1.176[a] (3.846)	1.369[a] (5.492)
Market share[b]	0.237[c] (9.400)	−0.024 (−1.042)
Herfindahl index[b]	−0.118[a] (−5.013)	0.057[a] (2.637)
Slots	0.035[a] (2.562)	0.0246[c] (1.728)
Distance[b]	0.345[a] (30.093)	0.343[a] (33.390)
Code-share	0.006 (0.470)	0.0243 (1.577)
Hub	0.071[a] (4.334)	0.054[a] (3.130)
Relative directness[b]	−0.042[a] (−6.348)	0.0144[c] (2.244)
Multiple airport service[b]	−0.003 (−1.131)	−0.002 (−0.692)
Tourism[b]	−0.042[a] (−7.075)	−0.023[a] (−3.597)
Preference ranking	0.0574[a] (4.801)	0.0386[a] (3.311)
Airline size[b]	0.099 (1.584)	0.059[a] (4.576)
Traffic volume[b]	−0.011[c] (−2.030)	−0.0441[a] (−8.724)
Noise restrictions	0.012 (1.115)	0.002 (0.188)
Majority-in-interest	0.021[c] (1.893)	0.020[c] (1.985)
Expansion difficulties	0.009 (0.597)	0.002 (0.152)
Minimum cost[b]	0.090[c] (2.243)	0.300[a] (9.482)

(continued)

[44]For the price equation, the test statistic for the test comparing the largest three carriers with the rest was 4.6, which exceeds the critical value for a 5-percent test with (18,3295) degrees of freedom. In the case of the route-share equation, the test statistic for the test comparing the largest three carriers with the rest was 31.0. This exceeds the critical value for a 5-percent test with (11,3309) degrees of freedom.

	Coefficients	
Explanatory variables	**Large airlines**	**Other airlines**
Airport congestion[b]	0.102[a] (5.643)	0.027 (1.465)
Summary statistics[d]		
R^2	.39	.56
n	1,458	1,873

Notes:Dependent variable is expressed in natural logarithmic form.

Large carriers are defined to be American, Delta, and United.

t-statistics in parentheses.

[a]Statistically significant at the 1-percent level.

[b]Variable expressed in natural logarithmic form.

[c]Statistically significant at the 5-percent level.

[d]R^2 is calculated as the square of the correlation between the dependent variable and its predicted value.

Several interesting differences in the results for the large carriers compared with all other carriers are indicated in table I.6. In particular:

- While the coefficient on the relative directness variable for the largest three carriers is negative and significant, the coefficient on this variable is positive and significant (as expected) for the other nine carriers. We believe that this occurs because of the large route networks operated by the largest carriers—these carriers can gain passengers and avoid discounting prices on routes even when they are offering lower quality connecting service.[45]
- The largest carriers had a larger coefficient for market share, indicating that these carriers can turn a given market share into higher prices more effectively than is the case for other carriers.

[45]A similar result is found in the route-share equation subsample on the three largest carriers when compared with the other nine. In that equation, the coefficient on the preference variable is negative and significant for the largest three carriers (as it was in the full sample base case), while it is positive and significant for the other nine carriers (as expected). Again, we believe that this result obtains because of the large networks of the largest carriers—their preference ranking is high, but they have small market shares on many routes because of their widely operated networks. It is also interesting to note that in a full sample specification for the market-share equation that included a set of carrier dummy variables to take account of carrier-specific information, the preference variable had a positive and significant coefficient. We believe that in such a specification carrier dummies control for the larger carriers' wider networks and thus allow the preference variable to pick up a market-share-enhancing effect related to preferences per se, rather than absorbing carrier-specific characteristics more generally.

- Prices of the largest carriers appear to be less influenced by the presence of low-cost carriers since the minimum cost variable has a lower coefficient for the three largest carriers.[46]

Alternative Dependent Price Variable

Airlines practice a good deal of differential pricing, or yield management—identifying passengers of different price-sensitivities and charging them different prices. Highly price-sensitive passengers (primarily tourist/leisure travelers) will be charged low prices since they will choose to travel less (or will travel to different locations) in response to high prices, while less price-sensitive passengers (mostly business travelers) will be charged higher prices because they are less likely to change their travel plans in response to high prices. Differential pricing is a relatively easy practice for airlines to engage in because airline service is not easily transferable from one consumer to another, and because the price-sensitivity of passengers can be identified by advance purchase and day-of-the-week travel restrictions for lower-priced tickets.

It is likely that less price-sensitive passengers, who tend to pay high prices, bear a larger share of the pricing effects of carrier market power and other factors related to the competitive conditions on routes. In order to examine whether the model indicates that determinants of high-end prices are different than low-end prices, we reestimated version A of the base-case model twice: once defining the dependent variable to be the 75th percentile price (a high price) and second defining the dependent variable to be the 25th percentile price (a low price). As expected, some differences do exist (see table I.7):

- The coefficients on both the market share and the hub variables are lower for the 25th percentile price, indicating that carrier market power does not influence low-end prices as much as high-end prices,
- The coefficients on size and preference are higher in the 75th percentile price estimation, indicating that it may be primarily high-paying passengers who are willing to pay for reputation and convenience,
- Higher coefficients on both slots and congestion in the 75th percentile price estimation imply that congestion costs are primarily passed on to high-paying passengers.

[46]Similarly, in the market-share equation, the coefficient on the relative cost variable is smaller for the largest three carriers indicating that, again, they can face a lower-cost competitor and not be at as much of a disadvantage in gaining market share as is the case for the other nine carriers when facing lower-cost competitors on a route.

Table I.7: Price Equation, Alternative Dependent Variables

Explanatory variables	Coefficients: 25th percentile	Coefficients: 75th percentile
Intercept	1.037[a] (7.097)	1.501[a] (5.677)
Market share[b]	0.039[a] (2.839)	0.130[a] (5.225)
Herfindahl index[b]	−0.002 (−0.186)	−0.082[a] (−3.417)
Slots	0.031[a] (3.940)	0.074[a] (4.994)
Distance[b]	0.411[a] (64.011)	0.275[a] (26.320)
Code-share	0.026[a] (3.305)	0.034[c] (2.254)
Hub	0.034[a] (3.591)	0.129[a] (7.559)
Relative directness[b]	−0.003 (−0.802)	−0.021[a] (−3.041)
Multiple airport service[b]	−0.009[a] (−5.603)	−0.001 (−0.501)
Tourism[b]	−0.025[a] (−7.137)	−0.085[a] (−12.471)
Preference ranking	0.029[a] (7.093)	0.063[a] (8.437)
Airline size[b]	0.051[a] (8.764)	0.107[a] (10.614)
Traffic volume[b]	−0.037[a] (−12.157)	−0.045[a] (−7.881)
Noise restrictions	0.002 (0.272)	0.011 (1.007)
Majority-in-interest	0.032[a] (5.274)	0.030[a] (2.578)
Expansion difficulties	−0.001 (−0.159)	−0.008 (−0.487)
Minimum cost[b]	0.207[a] (10.257)	0.373[a] (10.857)
Airport congestion[b]	0.040[a] (3.785)	0.136[a] (6.892)
Summary statistics[d]		
R^2	.68	.48
n	3,331	3,331

Notes:Dependent variable is expressed in natural logarithmic form.

The alternative specifications of the dependent variables are the 25th percentile price and the 75th percentile price.

t-statistics in parentheses.

[a]Statistically significant at the 1-percent level.

[b] Variable expressed in natural logarithmic form.

[c] Statistically significant at the 5-percent level.

[d] R^2 is calculated as the square of the correlation between the dependent variable and its predicted value.

Conclusions

As expected, we found that several key factors were important in determining a carrier's market share and price on a route. Market shares, for example, are influenced by the distance and traffic volume of a route. In addition, prices are influenced by distance, the degree of tourist traffic, and whether a low-cost airline serves the route. Additionally, larger and more preferred airlines appear able to charge higher prices.

Our model also indicates that several competitive factors affect market structure and pricing in airline markets. In particular, factors related to endpoint dominance influence airlines' market shares and in turn their pricing on routes. Certain barrier-to-entry factors such as majority-in-interest agreements, code-sharing agreements and slot controls also appear to be associated with higher prices on routes from airports with these conditions.

Our findings indicate that these factors do not affect the dependent variables in the same way on all routes, for all airlines, or for all passengers. In particular, prices on long routes appear to be less influenced by some competitive factors, such as those indicating endpoint dominance, because connecting service on longer routes mitigates the pricing effects of a dominant endpoint operation. The determinants of price across various carrier groups are shown to have statistically significant differences as well. Finally, prices paid by business travelers appear to be more influenced by certain competitive factors than are prices paid by more price-sensitive tourist travelers.

Data Sources, Adjustments and Variable Definitions

This appendix describes how we obtained our data and how we created our data base. We had to manipulate the data significantly in order to derive the variables we use in the model. At times we identified significant problems related to data quality and reliability. This appendix also outlines the steps we took to address these problems.

Data Sources

DOT/FAA Data Bases

DB1A is a quarterly 10-percent sample of actual passenger tickets submitted by airlines to DOT and processed by the agency. This data base provided us with information on prices and quantities, as well as several other variables related to carriers' service on our sampled routes. Appendix III discusses our use of DB1A in more depth.

DB6 is a DOT ranking of domestic city-pair routes by passenger traffic volume and includes information on the straight-line distance of each route. We used this data source to choose our stratified (by route traffic volume) sample of routes and to obtain route distance information. Appendix III discusses issues related to our use of DB6 for sample selection.

Study of Airline Computer Reservation Systems is a DOT report that presents a compilation of information from travel agents on their subscriptions to computerized reservation systems. As printed, this 1988 report only contained information on CRS market shares in large and medium-size cities. By obtaining additional data from DOT, we were able to derive airlines' CRS market shares in all cities in our sample.[1] These data were for 1986; other surveys have suggested that CRS market shares did not change greatly from 1986 to 1988.

The DOT Service Segment data base provides a variety of operational and financial information about airlines. This data source provided the necessary information for several of our variables, including enplanements, available seat miles, revenue passenger miles, and airline costs.

FAA operations and runway data were obtained from FAA. We were provided with information on operations (takeoffs and landings) and

[1]This data source also contained the information on travel agent commission override payments made by CRS vendor airlines in each of the cities.

the number of runways at the airports in our sample in 1988. We generated a measure of congestion for each airport, defined as the number of operations per runway.

GAO Airport Survey

The GAO survey of airports provided the basis for several variables measuring conditions at airports and relationships between airlines and airports. These variables include gate shares, "locked" gate shares (gates under long-term exclusive use leases), the incidence of noise restrictions, the existence of majority-in-interest clauses, and expansion difficulties. Generally, responses to the survey reflect circumstances at airports in 1988.

U.S. Department of Commerce and U.S. Department of Labor Data

For each of our cities, we obtained U.S. Department of Commerce and U.S. Department of Labor data on unemployment, income, population, and hotel revenues. For most cities, we used the Consolidated Metropolitan Statistical Area or Metropolitan Statistical Area data. For some smaller localities, we used county-level data. Data on income, population, and unemployment are for 1987. Information on hotel revenues are from the 1982 Census of Services.[2]

International Foundation of Airline Passengers Associations

The International Foundation of Airline Passengers Associations (IFAPA) surveyed travelers on their most preferred airlines in 1987. The survey included responses from nearly 30,000 frequent flyers worldwide.[3] We used the results for the respondents in North America to derive two variables reflecting consumer preference patterns across different airlines.

Official Airline Guide

The Official Airline Guide (OAG) provides a guide to airline schedules. We used the August 15, 1988, edition to obtain information that served as part of our measure of each airline's scheduled service on a route. We also used the OAG to identify flights that involved an airline's code-sharing (commuter) partner.

[2]We also use population data from 1982 in order to derive a tourism measure equal to hotel revenues per capita.

[3]According to IFAPA, the responses more generally reflect preferences of business travelers than tourist passengers.

Annual Report of the Regional Airline Association

We used the Annual Report of the Regional Airline Association, 1988 edition, to determine the cities in which each of the major airlines had a code-sharing agreement with a commuter airline.

Air Transport Association (Simat, Helliesen, and Eichner, Inc.)

We used Hub Operations: Analysis of Airline Hub and Spoke Systems Since Deregulation prepared for the Air Transport Association as the basis for our classification of hub airports.

Variable Descriptions

This section describes how we calculated each of the variables in the model. For purposes of presentation, we divide the variables into three groups: dependent variables, independent variables included in the market share and/or the price equations, and other exogenous variables not included in the two structural equations but included in the first-stage regression for market share.

Although our model requires information at a route level, only some of our variables naturally describe route attributes. For instance, all information derived from DB1A is naturally expressed on a route basis: A carrier's price and traffic level are defined for a particular route. However, some information, such as a carrier's enplanements, measures or describes conditions at airports or cities at a route's two endpoints. In these cases, our description of the variable will include how we constructed a route variable by relating or linking information from the two endpoints together in some way (for example, by taking the mean of the two endpoint values).

Appendix III discusses more fully how ticket information from DB1A was collapsed into carrier-route observations for the model. Without discussing this process for each individual variable here, we note that for the set of variables derived from DB1A, each carrier-route variable reflects the information contained in the set of underlying tickets for a particular carrier on a particular route. In some cases, a variable reflects a summation of information contained in the underlying tickets (for example, traffic volume), and sometimes a variable might be an average of the information contained in underlying tickets (as with price information).

Here we describe how each variable is constructed and where the data were obtained. The function of each variable in the model is discussed in appendix I.

Dependent Variables

Price. A carrier's price on a route comes from the ticket information in DB1A. We use the median price as the primary specification of price for each carrier-route observation. We also use the 75th and the 25th percentile fares to represent high- and low-end prices, respectively.

Market share. Market shares, or route shares, were derived from ticket information in DB1A. First, route traffic volume and carrier traffic volume were calculated from the information in DB1A. Route shares were calculated from these traffic volumes by dividing the carrier's traffic volume on a route by the traffic volume for all carriers on the route.[4]

Independent Variables in Market Share and Price Equations

Code-sharing. We defined a dummy variable that equals 1 if the given carrier has a code-sharing agreement at either of the endpoint airports of a route. Data are obtained from The 1988 Annual Report of the Regional Airline Association.

Congestion. We used operations per runway at the endpoint airport as a proxy measure of congestion. We took a simple average of the endpoint values to create a route variable for the model. All data were obtained from FAA.

Relative CRS share. This variable measures the degree to which a carrier has a CRS advantage on a route. We used information about each airline's CRS revenue market share (available from DOT) in each of the two endpoint cities and derived a weighted average CRS market share for the airline where the weights are based on originating passengers.[5] We then took an average of each carrier's CRS market shares on a particular route

[4]Since we included interline tickets in the total route traffic volume, market shares calculated reflect interline tickets in the base. Using interline tickets in the base of the market share calculations has some benefits but some drawbacks as well. We chose to include these tickets in the base because we did not want to overstate a carrier's market share when many of the passengers on a route are traveling with an interline itinerary. However, because a portion of any interline itinerary could be on the given airline, we are likely understating market share in some cases.

[5]That is, when calculating the weighted average of the endpoint CRS values, the CRS share at one endpoint is weighted by the percentage of the passengers on the route that originated at that endpoint.

to get the average CRS market share for all the carriers on the route. We then defined a carrier's relative CRS share as its CRS share on the route divided by the route average CRS—that is, divided by the average of all carriers' CRS shares on that route. On routes where no carrier had a positive CRS market share, we defined each carrier's relative CRS share to equal one.

We defined the CRS variable in this manner because we believe that the importance of CRS systems, in terms of how dominance of CRSs will affect market shares, is primarily a function of how large a carrier's CRS dominance is relative to those of its rivals on a route.

Traffic volume. This "size of the market" measure is calculated as the total number of passenger directional trips on each route provided by all carriers (and on an interline basis) during the quarter of our analysis. Information for this variable is obtained from DB1A.

Distance. This variable is the one-way, straight-line miles (great circle distance) between the two endpoints of a route. This information is obtained from the DB6 data base.

Endpoint dominance. We have two alternative measures of endpoint dominance that we alternatively include in the market-share equation. The first is based on an airline's share of the enplanements (passenger boardings) at the endpoint airports of a route, and the second is based on an airline's share of gates at those airports.

Information on enplanement shares of each airline at each airport was first derived from enplanement data obtained from DOT's Service Segment Data. To derive a variable in "route" form for the model, we took a weighted average of the enplanement share values for the two endpoint airports, where the weights are the number of originating passengers in the two endpoint airports (see footnote 5, this appendix).

The second measure of endpoint dominance is a gate share variable. This information was obtained from the GAO survey of airports. The construction of the gate variable parallels that of the enplanement variable. First, gate shares at each airport were calculated, and then a route version of the variable was derived by taking a weighted average of each airline's gate shares at the two endpoint airports. The weighting scheme was the same as that used with the enplanement measure.

Enplanements of others. Using DOT's Service Segment enplanement data, we examined carrier enplanement shares at the two endpoint airports and identified the largest enplanement share at each airport by some carrier other than the given carrier. We then chose the larger of these two enplanement shares to represent a large presence by a competing airline. The "largest other" airline in terms of enplanements did not necessarily compete on the route in question.

Hub. This is a dummy variable indicating that one of the endpoint airports serves as a hub for the given airline. This information was based on a study prepared for the Air Transport Association.

Expansion limited. This is a dummy variable indicating that at least one endpoint airport would have difficulty or face delays in expanding airport facilities. This information was obtained from the GAO airport survey.

Herfindahl Index. This index of overall concentration in a market equals the sum of the squared market shares. Data are derived from DB1A.

Multiple-airport service. Multiple-airport service denotes the percentage of a city-pair's traffic accounted for by airport-pairs other than the airport-pair of the given observation.

Using DB1A data, the variable was calculated by dividing the total traffic volume of an airport-pair route by the traffic volume of the associated city-pair. We subtracted this value from one, so that the value of the variable would decline as the airport-pair in question approached the city-pair level (for example, if neither airport is in a multiple-airport city, the value of the variable is zero). Thus, the value of the variable is larger if traffic on alternative airport-pairs (within the relevant city-pair) accounts for most of the city-pair traffic.

Majority-in-interest clauses. This is a dummy variable indicating that at least one of the endpoint airports of a route has a majority-in-interest clause in the use agreement between airlines and the airport. This clause gives the airlines a voice in airport planning decisions, such as expansion plans, where the cost of expansion would be recovered from fees charged to the airlines. This could give the signatory airlines the power to inhibit expansion at an airport. This information was obtained from the GAO survey of airports.

Minimum costs. We calculated the cost per available seat mile (ASM) for each airline in this analysis using data available from DOT's Service Segment Data base. We adjusted the cost per ASM for distance so that data across carriers with different average flight distances were comparable.[6]

For each route, we identified the carrier with the lowest cost level and denoted this value as the "minimum costs" for the route. This value was then assigned to all carriers on the route as the "costs of the least-cost competitor."

Noise. This is a dummy variable indicating that at least one endpoint airport of a route responded to the GAO airport survey that it had at least one of several different types of noise restrictions. Data are based on the GAO airport survey.

Preference (in the price equation). This is a 0-1-2 categorical dummy variable indicating how preferred an airline is (according to the results of a passenger survey) relative to its size. The results of a survey of North American travelers provided information on the percent of passengers citing each airline as their most preferred. Since larger airlines were more likely to be cited than smaller airlines as preferable, we used DOT Service Segment Data to compare these preference rankings according to each carrier's share of national revenue passenger miles. We subtracted each airline's size (its percent of total revenue passenger miles) from its preference ranking value. If an airline's preference ranking was considerably higher than its relative size, we assigned the carrier a preference value of 2. If an airline had similar values for the survey on preferences and relative size, the preference value was 1. Finally, if the carrier was significantly less preferred than its relative size, we assigned a preference value of 0.

Preferences (in the market-share equation). This preference ranking variable is different from the variable included in the price equation described earlier, although it is based on the same passenger survey of preferences. We used these data to derive, for each route, the average "preference" ranking for the route by taking the average preference ranking of all carriers serving the route. For each airline on the route,

[6]Each airline's total operating expenses during the quarter was divided by its ASMs. We then ran a regression to adjust for the distance effect on costs per ASM. Because two of the airlines included in the model were strong outliers—that is, were highly unusual—in terms of their cost per ASM relationship, we used a procedure to give less weight to these outlier observations that was suggested by F. J. Anscombe, as discussed in G. S. Maddala, Econometrics (New York: McGraw-Hill, 1977), p. 309.

we then calculated the ratio of that carrier's preference ranking to the average for the route.

Relative cost. The relative cost variable uses the same distance-adjusted cost per ASM data described in the section on minimum costs. For this variable we took the ratio of the weighted average (weighted by passenger traffic volumes) of the cost per ASM of all airlines serving the route to the cost per ASM of the given airline in order to derive a cost of the airline relative to the costs of all the carriers serving the route.

Relative directness. This variable indicates the degree to which the given airline is offering a greater percentage of its service as direct service (as opposed to connecting) than is generally offered on the route. We use this variable to indicate the product quality of the given airline relative to others on the route. Data for this variable came from DB1A.

Route directness. This variable measures the proportion of a route's traffic that is served on a direct basis. Data for the construction of the variable came from DB1A.

Scheduled service. This variable indicates the commitment of the given carrier to serving the route in question. We used both the OAG and DOT's Service Segment data base to construct this variable. From Service Segment, we obtained a carrier's ASMs—a measure of a carrier's capacity on a nonstop basis between two points—on each route. We used this information to derive each carrier's share of nonstop service on each route. Next, we used OAG information to obtain scheduled connecting flights (and direct flights with stops) for each carrier on each route in our sample. Again, we used these data to calculate each carrier's share of such flights. We then took a weighted average of each of these share values where the weights were the proportions of direct and connecting traffic on the route. That is, ASM shares are weighted by the proportion of direct traffic, and connecting (plus direct flights with stops) shares are weighted by the proportion of connecting traffic.

Size. This variable gives information about each carrier's share of domestic revenue passenger miles, expressed on an industrywide basis. Necessary data were obtained from Service Segment Data. In deriving this measure, we included only the airlines in our analysis.[7]

[7]Calculations of relative size were made, however, prior to the decision to exclude Pan Am and Braniff from the analysis.

Slots. This is a dummy variable indicating that one of the endpoints of the route is one of four slot-controlled airports.

Tourism. We defined a tourism value for each city equal to hotel revenues per capita. To derive a route variable from the endpoint information, we took a weighted average of the two city tourism values where the weights for each were the percentage of passengers on the route destined for that city.[8] All information is obtained from the Bureau of the Census.

Independent Variables in First-Stage Regression for Route Share

This section defines some additional variables that we believe are related to airline operations, costs, and demand conditions. They may be thought of as part of a broad system of factors influencing airline industry and route characteristics, even though they are not included in either the price or route-share equations. For example, some of these variables may not affect the firm's pricing decision directly, although they influence the larger market context in which those prices are set. Thus, we use these variables as instrumental variables (or preliminary regressors) in the first stage of the two-stage regression.

Airline cost per ASM. This variable measures the given carrier's costs (distance-adjusted) per ASM. It is derived from the same data used for the relative and minimum cost variables.

East/West. These variables are a pair of geographically based dummy variables. East indicates that both endpoints of a route are east of the Mississippi River, and West indicates that both endpoints are west of the Mississippi River.

Income. The data for 1987 personal income in each locality are from the Department of Commerce. To derive a route variable for income we took the geometric mean of the values of the endpoint cities (statistical areas).

Locked gates. Information for this variable was obtained from the GAO airport survey. For each airport at the endpoints of a route, we calculated the percentage of all gates held by each carrier under long-term, exclusive-use leases, which we refer to as "locked gates." We derived a

[8]That is, the tourism value for Reno on the route between Reno and Cleveland will be weighted by the percent of traffic originating in Cleveland. We used this weighting scheme so that the tourism weight for a city is the percentage of passengers who were destined for that city.

route variable by taking the weighted average of the two endpoint values where the weights are originating passenger percentages.

Population. This variable is derived from 1987 Census Bureau information. We allocated the population in each multiple-airport city among the airports in the city according to the proportion of total city originating enplanements accounted for by each airport. The route variable is the geometric mean of the population values for the two endpoint cities (usually metropolitan statistical areas).

Unemployment. This variable is the geometric mean of the unemployment rate in the two endpoint cities, and is derived from 1987 Department of Labor information.

Data Quality Issues

1986 CRS Data

Our study examines determinants of pricing in the third quarter of 1988. However, the DOT data on CRS market shares (and on travel agent commission override payments) were for 1986. Between 1986 and 1988, there were changes in travel agent CRS subscriptions—and thus CRS market shares in various cities—as well as changes in the ownership structure of the CRS systems. We cannot correct for changes in travel agent subscriptions, but we did make adjustments to account for the CRS ownership changes.

Between 1986 and 1988, two CRS systems had changes in ownership structure.[9] Northwest Airlines became a part owner (along with TWA) in PARS. Additionally, the System One CRS—owned by Eastern Airlines in 1986—was under the corporate umbrella of Texas Air Corporation by 1988, which operated both Eastern Airlines and Continental Airlines. To account for these ownership changes, we assigned PARS market shares to both TWA and Northwest Airlines, and System One market shares to both Eastern Airlines and Continental Airlines.

Southwest Pricing Data

Southwest Airlines realized that its original reported price data underrepresented its lower-priced tickets and decided to refile its data with

[9]During 1988 USAir also became part owner of United's Apollo CRS. However, because this transaction occurred late in the summer of 1988 (during the period of our analysis) we did not make any adjustments for this change in ownership structure.

DOT for 1988 and earlier time periods.[10] Our information for Southwest prices is obtained from the resubmitted data.

Bad Fare Data

As we will discuss in appendix III, we used yield screens to identify tickets with fares that were likely in error so that we could exclude such fares when calculating an airline's price on a route. Originally, we had planned to exclude an observation if a large portion of the airline's tickets failed our yield screens, since a large percentage of failed fares might indicate generally unreliable price information. However, one major airline had a great deal of bad fare data for a number of its routes—accounting for as much as half of its tickets on many of its primary routes. Because this airline was a major competitor on many routes, we decided to include observations for an airline even though it had a very large percentage of failed fares on these routes. Essentially, we assume that whatever fares passed our yield screens were representative of all fares actually paid.

Misreporting of Airport Designations in Multi-Airport Cities

For certain carrier-route observations in the airport-pair model, we adjusted price and traffic figures obtained from DB1A data for the third quarter of 1988. We did this because certain reporting errors by two airlines involving airports in New York, Washington, and Houston rendered their DB1A data incorrect. While these reporting errors did not affect DB1A data aggregated to the city-pair level, an adjustment was necessary to obtain reasonably correct data for the airport-pair model.[11] In this section, we describe the nature of the reporting problem, its implications for our modeling effort, and finally, the adjustments we made to correct the data.

Misreporting of an airline's traffic levels on routes to and from individual airports were found in the following cities:

- New York, where NYC is a valid reporting code for sampled tickets reported in DB1A even though no actual airport is so designated. (Major airports in the New York area are Newark International [EWR], La Guardia [LGA], and John F. Kennedy International [JFK].) One major

[10]In part, this problem arose because Southwest does not file a 10-percent ticket sample with DOT as do other airlines. Instead, Southwest estimates traffic levels and price information for the routes it serves and files information based on these estimates to DOT.

[11]Because the airport-pair model defines route markets on the basis of traffic between specific airports, it is crucial to have reliable estimates of routes defined at the airport level, and not just at the city level.

carrier reported little traffic to La Guardia and Kennedy, but reported much more traffic to NYC.

- Houston, where some airline itineraries report tickets to HOU, which is Houston Hobby, even though the tickets were actually to IAH, which is Houston Intercontinental.
- Washington, where as in New York, there is a valid city reporting code (WAS), in addition to the codes DCA and IAD designating actual airports (Washington National and Dulles International, respectively). For many routes involving Washington, one airline often reported significant portions of traffic at WAS rather than DCA or IAD, while another typically underreported the proportion of its Washington traffic at Dulles, and overreported the proportion of its Washington traffic at National.

In addition to creating problems in interpreting the affected airline's traffic levels, misreporting also results in inaccurate prices at the airport-pair route level for the affected airline and distorts measurements of total traffic volume on the airport-pair route and market share data.

After determining which routes were affected, we used the 1989 first quarter DB1A filing—with more accurate reporting—as the basis for adjusting the 1988 third quarter data.[12] We assumed that, in the third quarter of 1988, the airlines reported accurate city-level information, even if they reported inaccurate allocations of traffic between airports within the city. We then calculated what proportion a particular airport-pair route was of total city-pair traffic in the first quarter of 1989, and used this information to allocate 1988 third quarter city traffic levels to the airport level.[13]

For traffic-level variables, we computed the misreporting airline's total city-pair traffic in the third 1988 quarter, and the airline's city-pair traffic as well as the proportions accounted for by each airport-pair in the 1989 first quarter. Using these traffic proportions, we then allocated the airline's 1988 third quarter total city-pair traffic between the airport-pairs. We then recalculated the overall airport-pair traffic volume

[12]Some misreported observations did not need to be adjusted if the misreporting carrier's observation would be excluded from the data set on the basis of a low passenger count or route-share screen, and its misreporting did not lead to wildly inaccurate measurements of total route densities or route shares for other carriers.

[13]We assumed that between the third quarter of 1988 and the first quarter of 1989, the affected airlines did not materially change the allocation of traffic within the relevant cities. While there are seasonal traffic differences (for instance, there are probably more flights from the New York area to Florida vacation destinations in the winter than in the summer), we assumed that the shares of those flights leaving from La Guardia and Kennedy remained roughly the same.

by adding the misreporting airline's "adjusted" traffic to the traffic generated by the other airlines serving the airport-pair. This adjusted airport-pair route traffic volume was used to recompute route shares for all airlines serving the route.

For the median and mean price variables, we first calculated the misreporting airline's 1988 third quarter prices on the city-pair route. We also calculated this airline's 1989 first quarter prices for the city-pair route and each airport-pair route. We used the ratio of the airline's 1989 first quarter airport-pair price to its city-pair price to calculate the third quarter of 1988 airport-pair price, by multiplying the airport-pair/city-pair ratio for the first quarter of 1989 by the third quarter of 1988 city-pair price.

Route Sampling and Data Processing

Domestic airlines serve thousands of routes that vary tremendously in distance, traffic volume, and type of cities served. In order to capture the diversity of service in this industry, we include a broad range of routes in our econometric analysis of fare determination. However, not every route, or every carrier's service on a route, is appropriate to include in our analysis. In this appendix we discuss how we chose a sample of routes, and how we determined when carrier service on those routes should be included in our analysis. In the first section, we discuss our sample selection. In the next section, we discuss how we used DOT's data base of ticket information to form carrier-route observations for our model. In the final section, we note when and why certain observations that had been generated by our data processing were excluded from the data set used in our analysis.

Sample Selection of Routes

We analyzed prices on routes for which airlines provide primarily jet service in the continental United States. We do not include routes served primarily by regional or commuter carriers. Additionally, we excluded routes on which only a few passengers travel in any given time period. With this focus, we were able to include travel to both large and small cities, and on both heavily as well as less traveled routes.

Of the many thousands of domestic airline routes, a relatively small number of very heavily traveled routes account for the vast majority of all passenger trips.[1] Because these routes represent a disproportionate share of passenger trips, we chose to give more weight in our analysis to more heavily traveled routes. To accomplish this, we used a stratified sampling technique whereby more heavily traveled routes had a greater chance of being selected for our sample than did less heavily traveled routes.

Preparing the Sampling Frame

We used a DOT data base with information on city-pair traffic volume in order to select a sample of routes. For the purposes of selecting the sample, some adjustments to this DOT data base were necessary so that route traffic volumes were defined in a consistent manner. In addition, some routes were excluded from the universe of routes before we chose the sample.

[1]For example, there are over 40,000 domestic city-pair routes in the United States, but the 1,000 most densely traveled city-pairs account for over 70 percent of passenger trips.

Defining Consistent Traffic Volumes

The DOT data base we used for selecting our sample, DB6, did not define traffic volumes consistently across routes. In DB6, traffic to most multiple-airport cities is reported at a city-pair level; however, traffic to Los Angeles and San Francisco is reported at only an airport-pair level. In order to bring all routes to a consistent definition of traffic volume, we aggregated Los Angeles and San Francisco airport-pair routes to a city-pair level before selecting the sample.[2]

Excluding Certain City-Pair Routes

We attempt to exclude routes that are not likely to be served primarily by jet aircraft since we believe that the determinants of route shares and prices for such routes may be different from jet service routes.[3] Two screening criteria were used in an attempt to identify and exclude such routes:

- City-pair routes with fewer than 20 passengers per day.[4]

This exclusion rule helped to limit the focus to noncommuter routes. Additionally, on routes that are very lightly traveled—even if served by major carriers—competitive conditions may be different; or, in other words, carriers compete for traffic in different ways than they do on more heavily traveled routes. For example, major carriers on these routes may provide service primarily through commuter affiliates on

[2]In particular, we included the following airports as part of the Los Angeles- and San Francisco-area markets: Los Angeles included Los Angeles International, Burbank, John Wayne-Orange County, Ontario International, and Long Beach; San Francisco included San Francisco International, San Jose International, and Metropolitan Oakland. DB6 already contained aggregated traffic in other multiple airport cities. New York included Newark International, John F. Kennedy International, and La Guardia. Washington, D.C. included Washington National and Dulles International. Chicago included Chicago-O'Hare International and Chicago Midway. Dallas included Dallas-Ft. Worth International and Dallas Love Field. Houston included Houston Intercontinental and William P. Hobby. We made one additional change in these definitions. For New York we included two additional airports in the New York-area definition: Long Island-MacArthur (Islip) and Westchester County.

[3]We also exclude routes involving an endpoint in either Alaska or Hawaii. The cost conditions and competitive environment of routes to these destinations are likely to be considerably different from routes within the continental United States. We could not adequately address these differences within the context of this model.

[4]The sampling criteria did not eliminate all airport-pair routes of fewer than 20 passengers per day for 2 reasons. First, we used a 20-passenger-per-day selection criterion to choose city-pairs, but the more narrowly defined airport-pair routes contained within these city-pair definitions are often much less dense. (For example, if the city-pair route between Oklahoma City and Washington, D.C., has 100 passengers per day, 85 of those passengers could be Washington National passengers and only 15 Dulles passengers.) Second, the DB6 data base that was the source of traffic-level data used to select the sample included traffic for a full year of travel, although we used data on air fares for only 1 calender quarter for estimating the model. Therefore, a route was eligible for selection as long as it had at least 20 city-pair passengers per day on an annual basis, even though seasonal differences in the distribution of that traffic could mean that the traffic volume is fewer than 20 passengers per day during the quarter of our analysis.

small airplanes. Additionally, price and service data for very lightly traveled routes may be less reliable.

- City-pair routes less than 150 miles. Since travelers can easily drive a distance of less than 150 miles, excluding these routes helps to eliminate routes for which there would be important surface transportation alternatives available.[5] A short-route screen also helps to exclude commuter routes.

Sample Stratification

After aggregating the Los Angeles and San Francisco routes to the city-pair level and excluding the routes discussed above, we had a total of 3,231 city-pair routes from which to select our sample. We first grouped routes into four categories on the basis of traffic volume; we then randomly selected routes within each classification. The following discusses each group and explains our sampling rates:

- The highest traffic volume group contained routes with more than 500 passengers per day. A total of 255 routes were in this category, and we included all of them in our sample. The sampling rate for this stratum is 100 percent.
- The second category included routes of between 100 and 499 passengers per day. There were 830 routes in this category, from which we randomly selected 450. The sampling rate for this stratum is 54 percent.
- The third category contained routes of between 50 and 99 passengers per day. This category included a total of 678 routes from which we randomly selected 200. The sampling rate for this stratum is 29 percent.
- The final group included routes with between 20 and 49 passengers per day. There were 1,468 routes in this category, from which we randomly selected 200. The sampling rate for this stratum is 14 percent.

This stratified sampling procedure resulted in a sample of 1,105 city-pair routes. All airport-pairs contained within the chosen city-pairs were used for our primary analysis, which defines routes as airport-pair markets. That is, by choosing the Seattle-to-Chicago route for the city-pair model, we implicitly chose both the Seattle-to-O'Hare and the Seattle-to-Midway routes for the airport-pair model.[6]

[5]We realize this does not eliminate routes with alternative surface transportation available—particularly along the Eastern corridor. However, we had no clear criteria for determining which routes have viable competitive service from alternative modes of transportation, and any decisions we might have made would have been largely judgmental.

[6]Four small airports (out of 187 airports surveyed) were deleted from the analysis because they either did not respond to the survey or reported that they did not have regularly scheduled air service.

Deriving Carrier-Route Observations

Because our model examines the factors determining an individual carrier's fares and route share on a route, we needed to identify the carriers serving each route and to assemble information about each particular carrier's operations on the route—its prices and traffic volume, among other things.

Data on prices and traffic volumes are available in varying degrees of detail in several DOT data bases that are derived from the Origin to Destination Survey (O&D). The O&D data base with the most detailed information is DB1A. Because of its detailed level of information at the individual ticket level, we used DB1A to give us the greatest flexibility in defining the carrier-route variables needed for our model.

Using DOT's DB1A Ticket Data

The DB1A data base contains a 10-percent sample of air travel tickets sold during each quarter by U.S. certificated route air carriers. Each record in DB1A represents a passenger ticket (or a set of tickets)[7] and contains information about the set of airports that the passenger(s) passed through along the itinerary, as well as the price that was paid.

Each record also has information (generated by DOT) denoting the "directional breaks" of the itinerary. A directional break signifies that an airport in the itinerary was a destination point rather than one at which passengers merely changed planes. By using the directional break information, we were able to assign records to routes, and to classify tickets as one-way or round-trip.

A ticket with only one directional break is a one-way ticket. If a one-way ticket has only one passenger, it accounts for a single directional trip. A round-trip ticket will have two directional breaks, indicating two directional trips for a single passenger, where the one-way segment to the first destination point is the first directional trip and the return is the second directional trip.

Since we were interested in representing typical service on our sampled routes, we excluded some records if service seemed atypical, or if we had reason to believe the record contained errors. Specifically:

- We excluded a record when there was more than one connection on any directional trip segment. We assumed that it is generally atypical to

[7]DB1A records aggregate tickets when more than one passenger paid an identical fare for an identical itinerary.

have to change planes more than once to reach destination point (on routes served primarily by jet aircraft), and that such tickets would be likely to have unusual prices for the route.

- For calculating an airline's average price on a route, we excluded tickets that failed the GAO yield screens.[8] Although we exclude bad price data in average price calculations, we did not exclude these tickets from the calculation of quantity measures, such as passenger traffic volume and market shares.

We grouped DB1A records on each sampled route according to the identified airline. For each carrier serving a route, we derived variables measuring a carrier's price, number of passengers, and so forth. Each DB1A record was weighted in the calculation of carrier-route variables by the number of directional trips the record represented. The number of directional trips is determined by two factors:

- the number of passengers indicated on the record, which provides information about the number of underlying tickets each record contained (see footnote 7, this appendix); and
- whether the ticket was one-way or round-trip. We weighted round-trip records twice as much as one-way records because a round-trip is two directional trips while a one-way is a single directional trip. In order to treat a round-trip ticket as two directional trips, the round-trip fare was split in half, and one-half was applied to each directional trip.

Exclusion of Certain Observations

The data processing described above yielded an observation for each carrier that served a sampled route.[9] Although our processing generated an observation if a carrier had any service on a sampled route, many of the resulting observations were not appropriate to include in the model for various reasons. In this section we describe additional conditions under which we excluded certain carrier-route observations.

[8]Because of reporting errors, some of the fare data in the DB1A data base are incorrect, and it is necessary to screen out obviously bad fare data in order to have reliable price information. During earlier work, GAO developed fare screens for the purposes of identifying unusually high or unusually low fares for a particular route (based on the distance of the route). We developed these fare screens by examining available (or "listed") fares in the OAG and through discussions with industry experts. The determination of the highest possible fare for a route was based on OAG information, while the determination of the lowest possible fare for a route was somewhat more judgmental. For a presentation of the GAO yield screens, see Airline Competition: Higher Fares and Reduced Competition at Concentrated Airports (GAO/RCED-90-102, July 11, 1990), p. 18.

[9]In addition, we generated an observation for a route if there were any records (that is, tickets) on the route that involved travel on more than one carrier—such service is known as "interline" service.

Carrier "Presence" Criterion

We generated a very large number of carrier-route observations for which the carrier in question had only a handful of sampled passengers on the route during the quarter. We speculate that these tickets often exist because of coding errors or unusual itineraries and are not generally representative of scheduled airline service. We excluded observations for carriers with a very small presence on a route. We used two "screens" or elimination criteria to accomplish this goal:

- First, we deleted any observation for which the carrier did not have at least five passengers per day on a route during the quarter.
- Second, we used a screen based on a carrier's market share. We deleted any carrier-route observation if the carrier had less than 10 percent of the passengers on the route.

Carrier Screen

We focused our analysis of fare and market-share determination on the operations of the following major and national carriers that carried passenger traffic in the continental United States: American, America West, Continental, Delta, Eastern, Midway, Northwest, Piedmont, TWA, United, Southwest, and USAir. We did not treat carriers under joint ownership as a combined carrier. That is, although Eastern and Continental were both part of Texas Air, and although USAir and Piedmont were in the process of merging during the time period we studied, we treated each of these airlines as a separate entity.

These airlines were chosen from the set of major and national carriers that had operations in the third quarter of 1988, except for those that operated a significant portion of their system outside of the 48 contiguous states or that primarily served commuter routes. In addition, certain major carriers were excluded from the analysis:[10]

- Pan Am, because its domestic route structure is geared towards feeding into its international route system, and
- Braniff, because it was embarking on a major change in its network during the period we analyzed—pulling out of its hub in Dallas and establishing a new hub in Kansas City. Consequently, data in DB1A probably understates Braniff's commitment to its new routes and overstates its commitment to its old routes. Since observed market shares

[10] Although some airlines were excluded as observations in the model, their operations on a route were part of the data processing that calculated route traffic volume and other carriers' market shares.

may be misleading, the relationship between Braniff's market shares and prices may differ from the rest of the industry.

We classified any DB1A record on which more than one carrier provided service—known as "interline" service—into a separate group of its own—that is we treated such tickets as being provided by a mythical "interline" carrier. While the interline observations we derived on the basis of these tickets are themselves not included in our analysis, we included these tickets as part of "route" traffic when calculating market shares of each airline on the route and other quantity-based measures for which a carrier's operations on a route are compared with the route average.

Commuter/Code-Sharer Criterion

Despite the exclusion of commuter carriers and the elimination of very short and very lightly traveled routes, some of our observations could still represent travel on commuter carriers, since some commuter carriers have code-sharing agreements with large carriers and use the larger carrier's code for ticketing purposes. Because of the ticketing practices under code-sharing agreements, some commuter (code-sharing) records are included in the DB1A data base under the larger carrier code. We were unable—using DB1A data alone—to distinguish these records as code-sharing tickets. In order to identify code-sharing flights, we used information from the OAG to determine when service was provided by a carrier's code-sharing partner. We excluded any route for which most direct traffic was on a code-sharing carrier.

Sampling Error

We also deleted observations that had a high sampling error—defined as greater than 20-percent relative error around the mean price.

Major Contributors to This Report

Resources, Community, and Economic Development Division, Washington, D.C.

Francis P. Mulvey, Assistant Director
John V. Wells, Assignment Manager
Amy D. Abramowitz, Senior Economist
Stephen M. Brown, Senior Economist
Daren K. Sweeney, Evaluator-in-Charge
Jay R. Cherlow, Advisor
Thomas F. Noone, Senior Systems Analyst
Delores Parrett, Evaluator
Carol Herrnstadt Shulman, Reports Analyst

Office of the Chief Economist

Joseph Kile, Senior Economist

Ordering Information

The first five copies of each GAO report are free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20877

Orders may also be placed by calling (202) 275-6241.

United States
General Accounting Office
Washington, D.C. 20548

Official Business
Penalty for Private Use $300

First-Class Mail
Postage & Fees Paid
GAO
Permit No. G100