M&SOM MANUFACTURING & SERVICE OPERATIONS MANAGEMENT
Vol. 14, No. 3, Summer 2012, pp. 423–440
ISSN 1523-4614 (print) | ISSN 1526-5498 (online)

http://dx.doi.org/10.1287/msom.1120.0379

The Impact of Airline Flight Schedules on Flight Delays

Vinayak Deshpande
Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, vinayak_deshpande@kenan-flagler.unc.edu

Mazhar Arıkan
School of Business, University of Kansas, Lawrence, Kansas 66045, mazhar@ku.edu

Airline flight delays have come under increased scrutiny lately in the popular press, with the Federal Aviation Administration data revealing that airline on-time performance was at its worst level in 13 years in 2007. Flight delays have been attributed to several causes such as weather conditions, airport congestion, airspace congestion, use of smaller aircraft by airlines, etc. In this paper, we examine the impact of the scheduled block time allocated for a flight, a factor controlled by airlines, on on-time arrival performance. We analyze empirical flight data published by the Bureau of Transportation Statistics to estimate the scheduled on-time arrival probability of each commercial domestic flight flown in the United States in 2007 by a major carrier. The structural estimation approach from econometrics is then used to impute the overage to underage cost ratio of the newsvendor model for each flight. Our results show that airlines systematically "underemphasize" flight delays, i.e., the flight delay costs implied by the newsvendor model are less than the implied costs of early arrivals for a large fraction of flights. Our results indicate that revenue drivers (e.g., average fare) and competitive measures (e.g., market share) have a significant impact on the scheduled on-time arrival probability. We also show that the scheduled on-time arrival probability is not positively affected by the total number of passengers on the aircraft rotation who could be affected by a flight delay, or the number of incoming and outgoing connecting passengers on a flight. Operational characteristics such as the hub and spoke network structure also have a significant impact on the scheduled on-time arrival probability. Finally, full-service airlines put a higher weight on the cost of late arrivals than do low-cost carriers, and flying on the lowest fare flight on a route results in a drop in the scheduled on-time arrival probability.

Key words: flight delays; flight schedules; newsvendor model; forecasting
History: Received: September 10, 2008; accepted: December 27, 2011. Published online in *Articles in Advance* May 4, 2012.

1. Introduction

Flight delays have been a growing issue and have reached an all-time high in recent years. According to the U.S. Department of Transportation (DOT), a flight is considered as delayed if it arrives at the destination gate 15 minutes or more after its scheduled arrival time. The statistics show that there were 1,804,028 arrival delays out of a total of 7,455,458 commercial flight operations in the United States in 2007. Furthermore, in 2007, the airlines' overall on-time performance was at its worst level since 1995 when the DOT first started to collect detailed on-time performance data.

Flight delays have a significant impact on the U.S. economy. A recent report by the Joint Economic Committee of the U.S. Congress, chaired by Senator Charles Schumer (Schumer and Maloney 2008), estimated the total cost of flight delays to the U.S. economy at as much as $41 billion in 2007. This includes an estimated $19 billion in operating costs to the airlines, as well as $12 billion in passenger delay costs. The report also estimated that flight delays resulted in consumption of 740 million additional gallons of jet fuel, costing an additional $1.6 billion in fuel costs, and releasing an additional 7.1 million metric tons of climate-disrupting carbon dioxide in the atmosphere.

Flight delays have been attributed to several causes such as weather conditions, airport congestion, airspace congestion, use of smaller aircraft by airlines, etc. Airlines typically blame a number of external factors that are out of their control. In this paper, we examine the impact of the scheduled block time allocated for a flight, a factor controlled by the airlines, on on-time arrival performance. We combine empirical flight data published by the Bureau of Transportation Statistics (BTS) and the Federal Aviation Administration (FAA), and use the structural estimation approach from econometrics and the newsvendor framework from operations literature.

Each flight scheduled by an airline consists of a scheduled block time from the scheduled departure time to the scheduled arrival time. This schedule is determined typically more than six months in advance and is based on the estimate of the time it takes to complete each flight. Actual travel/block time for a flight is the difference between the actual arrival time and the scheduled departure time. Actual block time is uncertain because of various reasons such as weather conditions and airport/airspace congestion, and it is not known at the time of the scheduling decision. This creates a "too much" versus "too little" trade-off in the scheduling decision, typical of a newsvendor model: If too much scheduled block time is allocated to a flight, it creates leftover inventory (overage) costs for the airline such as idle aircraft, pilot compensation, etc. If too little scheduled block time is allocated for a flight, then the actual block time is likely to exceed the scheduled block time resulting in flight delays and shortage (underage) costs for the airline such as dealing with unhappy customers, overtime costs, etc. Airlines typically make the scheduled block-time decision in a two-step process (see §1.1 for a detailed description): (1) a probabilistic distribution of the actual travel time is constructed from past data and (2) a target on-time arrival probability is set for the flight. The target (scheduled) on-time arrival probability for the flight in conjunction with the actual travel time distribution determines a flight's scheduled block time. In this paper, we first construct a method for estimating the scheduled on-time arrival probability of a flight. We then focus on understanding the drivers of variation in this scheduled on-time arrival probability.

We use the "structural estimation" approach from econometrics, which assumes that the decision maker (airline scheduler in our context) is rational and makes a decision (in our case the flight schedule) that optimizes his objective. The structural estimation approach then uses observed decisions to estimate the underlying parameters of the decision model (in our case the relative weight on costs of late arrival versus early arrival) for which this decision is rational. The implied relative weights put on late versus early arrivals can be imputed from the scheduled on-time arrival probability using the newsvendor framework as explained in detail in §1.1.

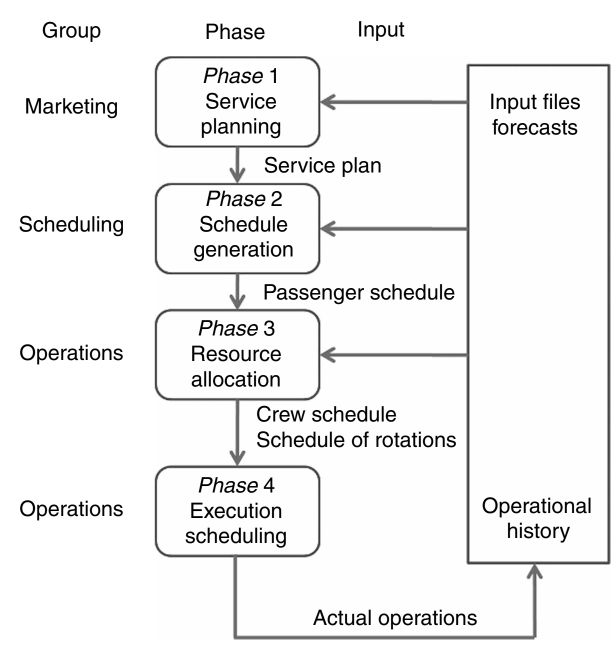
The goal of our paper is to answer the following questions: How is on-time performance affected by the scheduled block time? What is the relative emphasis put on late arrivals versus early arrivals by airline schedulers, as implied by the imputed overage to underage cost ratio (or the scheduled on-time arrival probability)? How is the overage to underage cost ratio (or the scheduled on-time arrival probability) affected by operational, competitive, revenue, and cost factors (i.e., what are some of the drivers of the heterogeneity in airline scheduling decisions)? What are some implications of the heterogeneity in the scheduled on-time arrival probability?

Our research questions have implications for policy planners, airline managers, and passengers. For policy planners, our research assesses the impact of alleviating congestion on on-time arrivals, as well as estimates the impact of competition on the quality of service to customers. For airline managers, we quantify the impact of operational factors such as congestion and network structure on the actual block-time distribution. Also, our models estimate how these factors affect the scheduled on-time arrival probability. Finally, our model provides a tool for passengers to evaluate the on-time arrival probability of a flight. Our analysis also assesses the impact of the choice of carrier (low cost versus full service) and fares on the on-time arrival probabilities. A complete discussion of the implications of our research for different stakeholders is provided in §7. We begin by first describing the airline schedule development process.

1.1. Airline Schedule Development and Block-Time Decisions

The airline schedule development process consists of four phases: (1) service planning, (2) schedule generation, (3) resource allocation, and (4) execution scheduling. We briefly describe this process here as depicted in Figure 1 (details can be found in Grandeau 1995). The service planning phase is conducted by the marketing group with the goal of creating a set of services that an airline will offer in each market. This usually consists of the frequency of flights offered in each market and also usually includes desired time windows (e.g., 5 P.M.–6 P.M.) and aircraft types (e.g., wide body, narrow body, long

Figure 1 Four Phases of Airline Schedule Development



Source. Grandeau (1995).

range, etc.). The scheduling group takes this service plan and develops the actual passenger schedules by considering aggregate constraints such as the total number of available aircraft and flight crew. Note that each passenger schedule includes the exact departure and arrival time of each flight, and hence the scheduled block-time decision for each flight is made at this stage (i.e., schedule generation phase). The passenger schedules then become an input to various specific resource allocation decisions that are usually the responsibility of the operations group (i.e., resource allocation phase). For example, aircraft with specific tail numbers are assigned to appropriate aircraft rotations by taking into consideration various constraints such as maintenance requirements. Finally, the execution scheduling phase involves implementing the developed schedule by taking schedule deviations (irregular operations) into account.

Airlines track and report five segments of travel time for each of their flights to the FAA: (1) departure delay, (2) taxi-out, (3) air time, (4) taxi-in, and (5) arrival delay. Figure 2 displays this segmentation. We use this information (publicly available through the BTS) to determine the scheduled block time of each flight.

Based on Figure 2, we first define different terms that constitute the travel time segments: computerized reservation system (CRS) departure/arrival time is the scheduled departure/arrival time of the flight, wheels off is the time when the wheels of the aircraft leave the ground at the origin airport, and wheels on is the time when the wheels of the aircraft touch the ground at the destination airport. The departure delay of an aircraft is the difference between the actual departure time and the CRS departure time of the flight. Arrival delay equals actual arrival time minus the scheduled arrival time.

The airline scheduler has to pick a scheduled block time (Q) for a flight, whereas the actual travel/block time (D) for the flight is uncertain. The scheduled block-time decision and the actual block time is defined as

$$\text{scheduled block time } (Q) = \text{CRS arrival time} - \text{CRS departure time}, \quad (1)$$

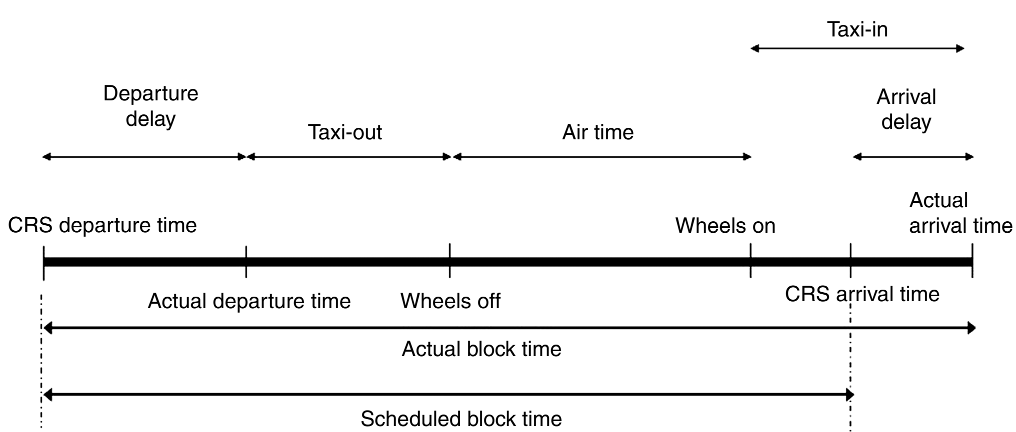
$$\text{actual block time } (D) = \text{actual arrival time} - \text{CRS departure time}. \quad (2)$$

From the above definitions, it is clear that if a flight's actual block time exceeds its scheduled block time, then the flight is delayed, whereas if its actual block time falls short of its scheduled block time, then the flight arrives early. For any scheduled flight i, airlines first estimate the actual block-time distribution by analyzing historical (typically over the previous two to three years) travel time data. It is important to note that airline schedulers typically remove late aircraft delay (LAD) from the actual block time when estimating the distribution of the actual block time. LAD is the portion of departure delay attributed to an aircraft arriving late from its earlier flight, and is tracked in the BTS on-time data set for every flight. Thus, truncated block time, which is used for making scheduled block-time decisions, is defined as follows:

$$\text{truncated block time } (TD) = \text{actual block time } (D) - \text{late aircraft delay } (LAD). \quad (3)$$

The reason airlines typically remove LAD from the actual block time is that the scheduled block-time decisions are made in the schedule generation phase (phase 2 of the schedule development described above) at which point the aircraft rotation (i.e., the sequence of flights operated by a particular aircraft on the same day) is unknown. Hence, it is difficult for an airline scheduler to estimate the actual block time for a flight, which includes LAD. Another advantage of using truncated block time instead of actual

Figure 2 **Main Segments of Air Travel Time**



block time is that it reduces the error term correlation between flights flown by the same aircraft in a given rotation, which may be induced by late aircraft delays. Finally, from the perspective of our research questions, it makes no difference to exclude LAD from the block time. This is because all flights look similar in terms of LAD due to unknown rotations, and our objective is to explore the drivers of *heterogeneity* in scheduling decisions.

The cumulative distribution function (CDF) of the truncated block time, $F_i(\cdot)$, becomes an important input in the block-time decision Q_i for flight i. Airline schedulers pick a scheduled block-time Q_i such that the on-time arrival probability for flight i meets a target service level criterion SL_i; i.e., Q_i is set to the SL_ith percentile of the truncated block-time distribution $F_i(\cdot)$.

To summarize, the scheduled block-time decision made by airline schedulers can be described as a three-step process: (1) estimate $F_i(\cdot)$ from historical observations over the last two to three years; (2) pick a scheduled (target) on-time arrival probability, SL_i, for flight i; and (3) set $Q_i = F_i^{-1}(SL_i)$.

Although the above approach used by airline schedulers to make block-time decisions is not exactly identical to the newsvendor approach (for additional details, see Cachon and Terwiesch 2006), it does have similarities. A decision maker facing an uncertain demand D has to make a quantity decision Q before demand is realized. Let $F(\cdot)$ be the CDF of the probability distribution of demand. If the realized demand exceeds the quantity, then the newsvendor incurs a shortage (underage) cost of C_u per unit of shortage. If the realized demand is less than the quantity decision, then the newsvendor incurs a leftover (overage) cost of C_o per unit of leftovers. The goal of the newsvendor is to pick the quantity Q that minimizes the total of expected leftover and shortage costs. The optimal quantity is then given by

$$F(Q^*) = \frac{C_u}{C_u + C_o} = \frac{1}{1+\gamma}, \tag{4}$$

where $\gamma = C_o/C_u$.

For the airline scheduling context, the underage cost, C_u, consists of operating costs to airlines due to flight delays (such as fuel costs, pilot, and flight attendant wages); goodwill costs to passengers due to flight delays; costs associated with accommodating passengers with missed connections; as well as the cost impact of a flight delay on subsequent flights flown by the same aircraft. The overage cost, C_o, consists of the opportunity costs of an early arrival such as an idle aircraft that does not generate revenue, lost revenue due to passenger perception that the flight takes too long to travel, lost revenue because flight time may change to a less desirable time slot, costs associated with managing gates, ground crew and other resources at the destination airport due to early arrivals, etc. minus the passenger goodwill benefits of on-time arrivals.

Note that $F(Q^*)$ is the scheduled on-time arrival probability (SL) with respect to the truncated demand (TD), associated with the optimal quantity decision Q^*. Airline schedulers typically do not calculate the cost of underage, C_u, and the cost of overage, C_o, explicitly for the purpose of making block-time decisions. But, defining target service levels SL_i for a flight i and setting $F(Q_i^*) = SL_i$ is equivalent to using an *implied* overage to underage cost ratio $\gamma_i = 1/SL_i - 1$ and then applying the newsvendor model with this γ_i. Thus, the newsvendor framework is useful to understand the drivers of the variation in the implied overage to underage cost ratio (γ_i) imputed through the target on-time probability SL_i.

1.2. Our Approach

We obtained detailed data from the BTS's website on each flight flown in the United States in the years 2005–2007. Other information that was collected includes aircraft type and registration information, airline fare information, load factor information, airport congestion, etc. Our approach mirrors the approach used by airline schedulers. We first estimate the truncated block-time distribution $F_i(\cdot)$, for each flight i flown in 2007, using 2005–2006 data (a detailed description of the forecasting model is provided in §5). We then plug in the observed quantity decision (scheduled block time) Q_i to estimate the "scheduled on-time arrival probability," $SL_i = F_i(Q_i)$, for each flight scheduled in 2007. Throughout this paper, scheduled on-time arrival probability, unless mentioned otherwise, is calculated based on truncated block time. The overage to underage cost ratio (γ_i) is then set equal to $1/SL_i - 1$. Next, this overage to underage cost ratio or, equivalently, the scheduled on-time arrival probability SL_i is used to test several hypotheses. Note that our approach is similar to model N1 described in Olivares et al. (2008).

Our paper is organized as follows. The hypotheses are stated in §2. We conduct a literature review in §3. Section 4 explains the data and the variables used in the model. Section 5 introduces the model and assumptions. In §6, we present the results, and §7 contains the conclusions and policy implications of our research.

2. Hypotheses Development

Based on our conversations with airline schedule planners and managers, as well as academic literature, we identified several factors that can affect the scheduled on-time arrival probability of a flight. We group our hypotheses on the drivers of scheduled on-time arrival probability into four general economic factors: (1) operational, (2) competition, (3) cost minimization, and (4) revenue maximization.

2.1. Hypotheses Related to Operational Factors

Periods of high congestion usually occur when the demand for travel is elevated. For example, airlines usually schedule a large number of flights in the 5 p.m.–6 p.m. departure time block to capture business travel demand, which results in increased congestion. On the one hand, the cost of overage to an airline is higher during these periods of high congestion because an idle aircraft loses potential revenue. Also, adding minutes to block time may result in changing the flight departure time to a less desirable time slot from a passenger's perspective, thus losing revenue. Mayer and Sinai (2003a) also hypothesized that as congestion increases, delays increase at hubs because the network benefits at hubs outweigh the marginal cost of delays. Therefore, C_o should potentially increase with airport congestion. So, we test the following:

Hypothesis 1A. *Increasing departure and arrival congestion (normalized by airport capacity) results in an increased overage to underage cost ratio, leading to negative effects on scheduled on-time arrival probability.*

On the other hand, one can also argue that, as there could potentially be more business customers during congested hours of the day, having more business customers dissatisfied due to delays could result in a higher goodwill cost, which raises C_u. Hence, our alternate hypothesis regarding normalized congestion is as follows:

Hypothesis 1B. *Increasing departure and arrival congestion (normalized by airport capacity) results in a decreased overage to underage cost ratio, leading to positive effects on scheduled on-time arrival probability.*

Our next hypothesis tests whether or not airlines try to minimize the cost (and inconvenience) to passengers due to flight delays. The higher the number of expected passengers on a flight, the larger is the cost impact of a delayed flight on passengers. Also, the cost to an airline of accommodating passengers who get affected by a delayed flight, increases in the number of total delayed passengers. Hence, all else equal, flights with higher number of expected passengers should potentially have a higher underage cost (C_u) because of the possibility of the total passenger inconvenience it may create in case of a delay. Besides, airlines schedule several consecutive flights on an aircraft before the aircraft takes an overnight halt or is taken out for maintenance. So, a delay on a flight not only affects the passengers on that flight but can also potentially affect other passengers scheduled to fly on subsequent flights with the same aircraft. As a result, the cost of a flight delay is larger if the total number of potential passengers who are affected by a flight delay is large. The overage cost (C_o) is typically not affected by the number of passengers on a flight. Hence, we test the following:

Hypothesis 2. *As the total number of passengers who could be potentially affected by a flight delay increases, the overage to underage cost ratio decreases, resulting in a higher scheduled on-time arrival probability.*

We refine the above hypothesis by also looking at the number of passengers making a connection for flights arriving at (departing from) a hub airport. Not all passengers on a flight traveling to a hub airport take a connecting flight. For some passengers, a hub airport may be their final destination. If the percentage of passengers, or the number of passengers, who will take a connecting flight decreases, that flight becomes more like a spoke flight instead of a hub flight. From an airline's perspective, the underage cost (C_u) is potentially higher for a flight with larger number of passengers making connections. The reason for this is that a delay of a flight with more connecting passengers leads to more unsatisfied passengers (and higher cost to an airline) than that of a flight with less connecting passengers. On the other hand, the overage cost (C_o) is not affected by the number of connecting passengers. Thus, we hypothesize the following:

Hypothesis 3. *A flight's scheduled on-time arrival probability increases with the number of passengers who are connecting to (from) another flight at destination (origin) airport.*

2.2. Hypotheses Related to Competition

The relationship between quality of services and competition in the literature is generally inconclusive. One strand of the literature claims that more competition leads to better service. Depending on the values of parameters for the cost and demand structure, Banker et al. (1998) found that increased competition can lead to higher quality levels. Similarly, studies considering different industries, such as healthcare (Kessler and McClellan 2000), education (Hoxby 2000), and local phone services (Economides et al. 2008) have indicated a positive effect of competition on service quality. In addition, there exist studies with similar results in which on-time performance from the airline industry was used as the quality measure. For example, Mazzeo (2003) found that delays are significantly greater in less competitive routes and Rupp and Holmes (2006) indicated that more competitive routes have lower flight cancellation rates. In terms of the newsvendor model, flights experiencing more competition have a higher C_u because, as shown by

Suzuki (2000), passengers may switch from an airline with larger delays to another airline with lower delays, i.e., the potential loss of revenue on flights with larger delays results in an increase in the underage costs. Another measure of quality is flight frequency: a larger variety of departure times offers more convenience to customers. Given capacity, an airline can improve this measure of quality by increasing plane rotation, at the expense of running a tighter schedule. A monopolist in a route may have fewer incentives to offer more frequent flights and would perceive a lower C_o, resulting in a lower overage to underage cost ratio. Therefore, we test the following:

Hypothesis 4A. *The overage to underage cost ratio decreases with increasing flight competition, hence airlines schedule higher on-time arrival probability on flights facing more competition.*

On the other hand, the other relevant stream of literature supports the hypothesis that more competition leads to worse service. Among these, Kranton (2003) demonstrated that if consumers are not able to observe the service quality prior to their purchase, high-quality production cannot be sustained in equilibrium when firms compete for market share without collusion. Using an oligopoly model with firms choosing both of their quantity and quality of products, Gal-Or (1983) showed that additional entry to a market reduces the average quality when the distribution of consumers is uniform. Using the data provided by UK hospitals, Propper et al. (2004) found that the impact of competition is to reduce quality in healthcare. Likewise, for the airline industry, Mayer and Sinai (2003b) found that competition leads to worse on-time performance, and Brueckner and Spiller (1991) showed that although competition in one city-pair market benefits the passengers in that market, the welfare of all passengers in all other city-pair markets is reduced. Based on the above literature, our alternate hypothesis is as follows:

Hypothesis 4B. *The overage to underage cost ratio increases with increasing flight competition, hence airlines schedule lower on-time arrival probability on flights facing more competition.*

2.3. Hypothesis Related to Revenue Maximization

The airline industry has been a pioneer in the practice (and research) of revenue management. Airline fare pricing is an important area of study for revenue management and airlines try to price their seats to maximize revenues. In this respect, airlines should carefully assess the effects of a possible delay on each route they serve, because it also affects potential revenue for an airline. For instance, Forbes (2008) found that airline prices fall by \$1.42 on average for direct passengers, and by \$0.77 on average for connecting passengers for each additional minute of delay. It is a common practice to use yield (i.e., average revenue per mile) as one measure for airline pricing (Windle and Dresner 1999, Oliveira and Huse 2009). Also, one important dimension of airline travel is the price elasticity difference between business and leisure passengers (i.e., leisure passengers have higher absolute price elasticity) (Brons et al. 2002). Surveys show that air travelers are willing to pay more for better on-time performance, with business travelers valuing on-time performance more than leisure travelers (Prince and Simon 2009). Because the average fare per mile is higher for flights with a larger fraction of business travelers who have higher sensitivity to delays, we posit that for any route, as the average fare per mile increases, the underage cost (C_u) goes up, whereas the overage cost (C_o) is typically not affected by the fare per mile.

Hypothesis 5. *The overage to underage cost ratio for a flight decreases with the average fare per mile, and, hence, the average fare per mile has a positive effect on a flight's scheduled on-time arrival probability.*

2.4. Hypotheses Related to Cost Minimization

After the deregulation of the airlines in the United States in 1978, carriers faced intense competition that forced them to continuously find ways to minimize their costs. The deregulation also gave rise to the emergence of so-called "low-cost carriers." These carriers took away market share from full-service airlines by offering considerably low fares. We want to explore the relationship between on-time service quality and being the lowest fare carrier on a specific route. One way for an airline to offer a lower fare on a market is through cutting cost—particularly its scheduled block-time cost. This relationship can shed light on to what extent lowest fare carriers forgo their on-time performance to be able to offer the minimum fare. Thus, low-fare carriers have higher overage costs (C_o). Besides, these airlines typically carry passengers who are more sensitive to price and less sensitive to delays (e.g., leisure passengers) (Dresner 2006, Proussaloglou and Koppelman 1995). Hence, low-fare carriers have lower underage costs (C_u). Thus, we hypothesize the following:

Hypothesis 6. *On a route served by more than one airline, flying the lowest fare carrier results in an increase in the overage to underage cost ratio and, hence, results in lower scheduled on-time arrival probability.*

In their study, Rhoades and Waguespack (2005) showed that although low-cost carriers are not as good as full-service carriers in the area of service quality (e.g., complaint rate per departure), both groups of airlines are not significantly different in terms of safety rates. Similarly, Lapré and Tsikriktsis (2006) and Tsikriktsis (2007) grouped the airlines into two groups: (1) focused airlines and (2) full-service airlines. The justification for this type of grouping is that the variation of the types of aircraft in the airlines' fleet was significant between the groups and not significant within the groups. They also mentioned considerable differences in terms of fleet utilization of the two groups. In addition to these, the Massachusetts Institute of Technology Global Airline Industry Program's Airline Data Project reports airline industry data by categorizing carriers into three distinct subgroups: (1) network carriers, (2) low-cost carriers, and (3) other. We employed a similar grouping strategy by clustering the 19 airlines in our data into four distinct groups: (1) full-service airlines (i.e., American, United, Delta, Northwest, Continental, and US Airways); (2) commuter/subsidiary airlines (i.e., Comair, American Eagle, Atlantic Southeast, Mesa, ExpressJet, and Skywest); (3) low-cost airlines (Southwest, AirTran, JetBlue, and Frontier); and (4) leisure airlines (Alaska, Hawaiian, and Aloha). Because of their network strategy, the types of routes they serve, fleet utilization, etc., we expect that the average on-time performance differs between the groups. Also, considering the different mix of customers (leisure travelers versus business travelers) across groups, we hypothesize that the overage to underage cost ratio is different for these different airline groups. Hence, we test the following:

Hypothesis 7. *The overage to underage cost ratio is decreasing* (*and*, *hence*, *scheduled on-time arrival probability is increasing*) *for the four airline groups in the following order*: *leisure*, *low-cost*, *commuter/subsidiary*, *and full-service airlines*.

3. Background and Literature Review

Flight delays can be attributed to several reasons such as weather conditions, origin/destination airport or airspace congestion, aircraft mechanical problems, scheduling complications, etc. Airlines began reporting information on the causes of delays in June 2003 to the BTS. The airlines report the causes of delays in five broad categories: (1) air carrier, (2) extreme weather, (3) National Aviation System (NAS), (4) late-arriving aircraft, and (5) security.

Many possible causes of flight delays have been analyzed by airline industry analysts, including the use of smaller airplanes with more flights scheduled per day (McCartney 2007a), increased congestion both at the airports and in the sky, slow cruising speeds due to very high fuel costs, and lack of modern equipment for the air traffic controllers (McCartney 2007b). Airlines have an incentive to reduce their scheduled flight times to cut operating costs, and the trade-off with this decision is increased delays, higher customer waiting times, and customer dissatisfaction. Trottman and Carey (2007) reported that the number of passengers per full-time employee has been steadily increasing since 2001, during when airline delays have also been steadily going up.

Academic literature on flight delays can be classified into three streams: (1) statistical models exploring the variations in various components of total travel time, (2) econometric models analyzing the economic drivers of flight delays, and (3) operations management models.

Several researchers have developed statistical models for forecasting different components of air-travel time. Highly stochastic nature of air transportation has encouraged researchers to analyze the different aspects of flight scheduling issues. Shumsky (1995) contributed to the literature of airline scheduling performance analysis by generating accurate forecast models for gate delays and aircraft roll-out times. There are several studies that model aircraft departures. See Mueller and Chatterji (2002), Tu et al. (2008), Hebert and Dietz (1997), and Willemain et al. (2003) for models of different components of the actual block time such as the departure process, air time, and taxi-in process. Our contribution to this stream of statistical models lies in how we model the uncertainty in the *total* travel time. We model the total travel time without dividing it into individual segments such as departure delays or taxi-out times. In addition, we develop a model of total travel time for *all* flights flown in the United States at *all* airports, as reported in the BTS data set, instead of focusing on specific airports or short time periods as done in prior studies.

Many econometric research papers have tried to explain the reasons of flight delays. Among these studies, Mayer and Sinai (2003a) developed models to estimate the effect of airline hub size and airport concentration on air travel time. Mayer and Sinai (2003b) hypothesized that wage cost minimization and aircraft utilization maximization lead to tight schedules of flights. Another stream of econometric models have looked at the impact of competition on airline service quality. Mazzeo (2003) showed that the prevalence and duration of flight delays are significantly greater on routes where only one airline provides direct service. Rupp and Holmes (2006) examined the determinants of flight cancellations such as revenue, competition, aircraft utilization, and airline network. Forbes (2008) estimated the effect of air traffic delays on airline fares.

A key difference between previous econometric models and our model is as follows: Most prior econometric models conduct a "reduced form" analysis by regressing observed flight delays against explanatory variables, whereas we use a "structural estimation" approach from econometrics. There are several advantages of the structural estimation approach over the reduced form approach, as highlighted by Olivares et al. (2008). A key advantage is the ability to disentangle whether a specific factor affects the quantity decision (Q_i), through the distribution of the random demand variable (D_i), and/or through the overage to underage cost ratio (γ_i). Indeed, the focus of our research is on understanding the variation in the overage to underage cost ratio used by the decision maker to set target service levels (on-time arrival probability). The overage to underage cost ratio provides a perspective on the relative emphasis placed on late arrivals versus early arrivals by the airline schedulers. Also, we are interested in how this relative weight changes with different operational, competitive, revenue, and cost factors. The structural estimation approach is well suited to answer these questions. See Reiss and Wolak (2006) for more on the structural estimation approach.

Although our paper is most closely related to Mayer and Sinai (2003b), there are important differences in methodology (reduced form versus structural estimation), statistical results, and findings. We add several statistical refinements by using an individual flight as the unit of analysis instead of the flight number, which makes our results statistically robust. A key finding of Mayer and Sinai (2003b) is that the scheduled block time is close to the mean travel time and increased standard deviation results in a decrease in the scheduled block time, i.e., the "z" value implied by their model is constant and negative across all flights. Our results show that the z value is positive for approximately 66% of flights, which would result in an increase in the scheduled block time if standard deviation were to increase. This result does not support the key finding in Mayer and Sinai (2003b). In addition, we test the impact of revenue drivers such as average fare, number of passengers affected by delays and passengers making connections, and cost drivers on scheduled on-time arrival probability, which were not examined in Mayer and Sinai (2003b).

Operations management literature exploring the operational aspects of running an airline and its impact on flight delays has been sparse. Ramdas and Williams (2009) were the first to combine an operations management (i.e., queueing) perspective with econometrics to investigate the effects of aircraft utilization on flight delays. Our approach is very closely related to Olivares et al. (2008), who used a structural estimation model to impute the overage to underage cost ratio associated with scheduling operating room capacity at hospitals. Scheduling flight block times is similar to scheduling block times of surgeries in operating rooms. Our findings are also similar: airlines—like hospitals—place more emphasis on the tangible costs of having idle capacity.

4. Data and Variables

We obtained the data for our study from several sources within the BTS website.[1] The main data are the Airline On-Time Performance data for all domestic flights flown in the United States by major carriers that account for at least 1% of domestic scheduled passenger revenues. We downloaded three years of on-time performance data from 2005 to 2007. This "on-time" data set contains actual/scheduled arrival/departure times, arrival/departure delays, flight numbers, taxi-in and taxi-out times, air carrier names, tail numbers of aircraft for each commercial flight flown by major airlines in the United States, etc. The tail number or the N number of an airplane is an alphanumeric string starting always with the letter N, which uniquely identifies each aircraft. Although the on-time performance data from the BTS is very useful for analysis, it lacks some important aircraft-specific information such as number of seats and year of manufacture. There exists another database called the Aircraft Registry Database,[2] where one can find aircraft-specific information for all registered aircraft in the United States. By using unique tail numbers in the on-time data set, we merged the on-time performance data with the aircraft registry data set. We also used DOT's Domestic Airline Fares Consumer Report (Table 6),[3] which contains information regarding the average fares for all city-pair markets that average at least 10 passengers each day. In addition to these, we collected *load factor* data from T-100 Domestic Market data from the BTS website. Our data set consists of 21,735,733 observations (one observation for each flight flown) across three years. Of these, 7,140,596 flights were for the year 2005, 7,141,922 flights were from 2006, and 7,453,215 flights were from 2007. Similar to the methodology in Ramdas and Williams (2009), we eliminated some bad data from our data set. See the online appendix (available at http://msom.journal.informs.org/) for the details on how we cleaned the data. As a result, our number of observations dropped down to 20,681,160 flights across three years covering 294 U.S. airports.

4.1. Drivers of Truncated Block Time

The main purpose of the initial data analysis is to develop a model to estimate the block time by characterizing the drivers, and to use this model to forecast delays and on-time arrival probability. Note that, unlike prior research that used a flight number as

[1] http://www.transtats.bts.gov/databases.asp?Mode_ID=1&MODE_Desc=Aviation&Subject_ID2=0 (last accessed April 12, 2012).

[2] http://www.faa.gov/licenses_certificates/aircraft_certification/aircraft_registry/releaseable_aircraft_download/ (last accessed April 12, 2012).

[3] http://ostpxweb.dot.gov/aviation/x-50%20Role_files/consumerairfarereport.htm (last accessed April 12, 2012).

Table 1 Description of Variables

Variables	Description
Drivers of truncated block time (OLS1 variables)	
Route	An origin-destination airport pair combination (captures all the fixed effects on a particular route).
Carrier	The airline carrier that flew the flight (23 different airlines exist in our data set).
Origin	The origin airport of the flight.
Destination	The destination airport of the flight.
Dep-Congestion	Number of flights scheduled to depart between 45 minutes before and 15 minutes after the scheduled departure time of the observed flight, normalized by airport departure capacity.
Arr-Congestion	Number of flights scheduled to arrive between 45 minutes before and 15 minutes after the scheduled arrival time of the observed flight, normalized by airport arrival capacity.
Aircraft-Age	Age of the aircraft that flew the flight.
Avg-Passengers	The expected number of passengers on the aircraft.
Month	The month of the flight.
Day-of-Week	The day of week of the flight.
Dep-Time-Block	One-hour time block based on the scheduled departure time (e.g., 6 A.M.–7 A.M.).
Arr-Time-Block	One-hour time block based on the scheduled arrival time.
Additional drivers of scheduled on-time arrival probability (additional variables in OLS2)	
Org-Hub (*Dest-Hub*)	{0, 1} variable indicating whether a flight departed (arrived) from (to) a hub airport.
Total-Passgrs-Affected	Expected number of passengers on the observed and all subsequent flights of the same aircraft in the same rotation.
From-Connecting-Passgrs	The expected number of incoming passengers with connections at the origin airport.
To-Connecting-Passgrs	The expected number of passengers who will be catching a connection at the destination airport.
Market-Share	The number of seats provided by an airline on a route divided by the total number of seats available on that route.
Num-Major-Carrier	The total number of distinct major carriers on a route (airline subsidiaries and regional partners were merged with their parent companies for the purpose of measuring competition).
Rank	The competitive rank of an airline on a route was assigned by measuring the number of seats available on a route by each airline.
Fare-per-Mile	The average market fare divided by the distance flown.
Min-Fare-Flight	{0, 1} variable indicating whether or not the observed flight belongs to the lowest fare carrier on that route.
Carrier-Group	We clustered the 19 airlines (in 2007) of our data into four distinct groups: (1) full-service airlines (i.e., American, United, Delta, Northwest, Continental, and US Airways); (2) commuter/subsidiary airlines (i.e., Comair, American Eagle, Atlantic Southeast, Mesa, ExpressJet, and Skywest); (3) low-cost airlines (i.e., Southwest, AirTran, JetBlue, and Frontier); and (4) leisure airlines (Alaska, Hawaiian, and Aloha).

the unit of analysis, we use an individual flight as the unit of analysis. For example, although a Monday morning flight from Boston to Chicago may share the same flight number with a Saturday morning flight from Boston to Chicago, and would be treated homogeneously for the purpose of estimating travel time distribution if the unit of analysis was a flight number, we treat these flights separately because the travel time distributions are different on Mondays and Saturdays. To understand the drivers of travel time, we analyzed the following data fields: (1) route, (2) carrier, (3) origin airport, (4) destination airport, (5) congestion at the origin airport, (6) congestion at the destination airport, and (7) aircraft-specific variables. We describe these variables in Table 1.

4.2. Drivers of Scheduled On-Time Arrival Probability

We obtained several more variables that can affect scheduled on-time arrival probability. We grouped these variables into four categories: (1) operational (i.e., hub airports, total number of passengers potentially affected by a flight delay, and number of passengers with connections); (2) competition; (3) revenue maximization (i.e., average fare per mile); and (4) cost minimization (i.e., minimum fare flight and airline group). We describe these additional variables also in Table 1.

Table 2 Descriptive Summary Statistics After Cleaning the Data

Variable	*N*	Mean	SD	Median
Air-Time[a]	13,227,945	101	68	82
Distance[a]	13,227,945	705	555	550
Taxi-in[a]	13,227,945	6	5	5
Taxi-out[a]	13,227,945	16	11	13
Arr-Congestion[a]	13,227,945	0.365	0.244	0.353
Dep-Congestion[a]	13,227,945	0.412	0.272	0.400
Aircraft-Age[a]	13,227,945	10	7	8
Avg-Passengers[a]	13,227,788	91	52	94
Arr-Delay[a]	13,227,945	8	35	−1
Dep-Delay[a]	13,227,945	9	32	0
Actual Block Time (*Demand*)[a]	13,227,945	132	77	112
Scheduled Block Time (*Quantity*)[a]	13,227,945	124	69	105
Truncated Block Time[a]	13,227,945	127	74	108
Fare-per-Mile[b]	6,828,914	0.386	0.359	0.270
From-Connecting-Passgrs[b]	6,708,652	14	23	2
To-Connecting-Passgrs[b]	6,708,652	14	24	2
Total-Passgrs-Affected[b]	6,881,303	297	283	214
Market-Share[b]	7,022,036	0.740	0.298	0.869

[a]2005–2006.
[b]2007.

Summary statistics of the continuous variables used in our analysis is shown in Table 2. An important observation from the descriptive statistics in Table 2 is that there is a considerable difference between the average actual block time and the average scheduled block time. This is a good illustration of how airlines set very aggressive schedules. The data show that airlines set their scheduled block times on average eight minutes less than what it takes to complete a flight (a significant difference when measured over 13 million flights). The arrival and the departure delays are also consistent with the above inference (average arrival delay of eight minutes matches the difference between average actual and scheduled block time).

5. Estimating the Truncated Block-Time Distribution and Scheduled On-Time Arrival Probability

In this section, we model truncated demand, TD_i, as defined in §1.1 as a random variable. We assume that TD_i's are random variables from a common family of distributions $\{F(\cdot, \boldsymbol{\theta})\colon \boldsymbol{\theta} \in \Theta\}$, where $\boldsymbol{\theta}$ is a vector parameter from the parameter space Θ. The distribution of TD_i is given by $F(\cdot, \theta_i)$. We let this distribution depend on a vector of covariates X_i, such as arrival/departure congestion, route, etc., as defined in §4. The goal of this section is to analyze some classes of distributions F that provide a good fit to the empirical flight data.

Our initial analysis showed that a log-normal distribution may provide a good fit for the truncated block-time distribution. In addition, our initial data analysis reveals a very high kurtosis value, suggesting that a log-Laplace distribution may provide a better fit. Hence, we decided to test both of these block-time distributions. In Model 1 we assume that TD_i has a log-normal distribution, whereas in Model 2 we fit a log-Laplace distribution. For both types of distributions, we assume that the distribution of TD_i is characterized by the following equation:

$$\ln(TD_i) = X_i^T \boldsymbol{\beta} + \varepsilon_i, \tag{5}$$

where $\boldsymbol{\beta}$ is a $K \times 1$ column vector of the coefficient of the regressors and ε_i is the disturbance associated with the ith observation, where $\varepsilon_i \sim \text{Normal}(0, \sigma_i^2)$ for Model 1 and $\varepsilon_i \sim \text{Laplace}(\gamma, 2b_i^2)$ for Model 2.

The distribution F is characterized by the parameter set (μ_i, σ_i^2) and $(\mu_i, \gamma, 2b_i^2)$ for Model 1 and Model 2, respectively. We want to estimate the parameter vectors $(\hat{\mu}_i, \hat{\sigma}_i)$ and $(\hat{\mu}_i, \hat{\gamma}, \hat{b}_i)$ so that we can obtain a predicted on-time arrival probability for each flight.

The first estimator, $\hat{\mu}_i$, can be found by regressing $\ln(TD_i)$ on X_i. This regression gives an estimate of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$, which can be used in

$$\hat{\mu}_i = X_i^T \hat{\boldsymbol{\beta}}. \tag{6}$$

We call this regression OLS1. The estimation model of our analysis is

$$\begin{aligned} \ln(TD)_i &= \beta_0 + \beta_1 \times Route_i + \beta_2 \times Origin_i + \beta_3 \times Destination_i \\ &+ \beta_4 \times Carrier_i + \beta_5 \times Month_i + \beta_6 \times Day\text{-}of\text{-}Week_i \\ &+ \beta_7 \times Dep\text{-}Time\text{-}Block_i + \beta_8 \times Arr\text{-}Time\text{-}Block_i \\ &+ \beta_9 \times Arr\text{-}Congestion_i + \beta_{10} \times Dep\text{-}Congestion_i \\ &+ \beta_{11} \times Aircraft\text{-}Age_i + \beta_{12} \times Avg\text{-}Passengers_i + \varepsilon_i. \end{aligned} \tag{7}$$

A summary of the results of OLS1 estimation is provided in Table 3.

The second estimator we need is $\hat{\sigma}_i^2$ or $2\hat{b}_i^2$. We denote the variance of the disturbances of the two models we analyze by σ_i^2 and $2b_i^2$, respectively, for Model 1 and Model 2. Note that since $\sigma_i^2 = 2b_i^2$, Model 2 estimated disturbance variance can be found by $\hat{b}_i^2 = \hat{\sigma}_i^2/2$. Hence, below we only show the estimation of $\hat{\sigma}_i^2$, which can easily be converted to $2\hat{b}_i^2$.

If the variances of the disturbances across all observations were to stay constant (i.e., homoscedastic),

Table 3 Summary of OLS1 Estimation

Dependent variable: ln(*Truncated Block Time*)		
Variable	Df	OLS1 parameter estimate
Intercept ($\hat{\beta}_0$)		5.1315*** (0.1918)
Route ($\hat{\beta}_1$)	4,244	
Origin ($\hat{\beta}_2$)	293	
Destination ($\hat{\beta}_3$)	290	
Carrier ($\hat{\beta}_4$)	21	
Month ($\hat{\beta}_5$)	11	
Day-of-Week ($\hat{\beta}_6$)	6	
Dep-Time-Block ($\hat{\beta}_7$)	18	
Arr-Time-Block ($\hat{\beta}_8$)	18	
Arr-Congestion ($\hat{\beta}_9$)	1	0.1263*** (0.0005)
Dep-Congestion ($\hat{\beta}_{10}$)	1	0.1123*** (0.0004)
Aircraft-Age ($\hat{\beta}_{11}$)	1	0.0015*** (1.19×10^{-5})
Avg-Passengers ($\hat{\beta}_{12}$)	1	0.0002*** (2.64×10^{-6})
Number of observations		13,227,718
R-square		0.8747

Note. Standard errors are in parentheses.
***$p < 0.0001$.

then an unbiased estimator of σ^2 (or $2b^2$) would be easily calculated using $\hat{\sigma}^2 = (\hat{\boldsymbol{\varepsilon}}^T\hat{\boldsymbol{\varepsilon}})/(n-K)$, where $\hat{\boldsymbol{\varepsilon}}^T\hat{\boldsymbol{\varepsilon}} = 1/(n-1)(\sum_{i=1}^{n}(\ln(TD_i)-\hat{\mu}_i)^2)$. However, considering the variation across flights, across airlines, and across routes in different months, we allow the variance of travel time to be heteroscedastic. We assume that the error term of a flight i flown on day d of month m can be written as the sum of three shock terms: origin airport (org) shock on the day (d) of the flight, destination airport (dest) shock on the day (d) of flight, and a carrier (c) dependent idiosyncratic shock for flight i. So,

$$\varepsilon_{i,d,m,c} = \xi_{\text{org},d,m} + \xi_{\text{dest},d,m} + \xi_{i,c}, \tag{8}$$

where $\mathbb{E}(\xi_{\text{org},d,m}) = \mathbb{E}(\xi_{\text{dest},d,m}) = \mathbb{E}(\xi_{i,c}) = 0$, $\text{Var}(\xi_{\text{org},d,m}) = \sigma^2_{\text{org},m}$ is the (daily) variance of the error term at the origin airport in month m, $\text{Var}(\xi_{\text{dest},d,m}) = \sigma^2_{\text{dest},m}$ is the (daily) variance of the error term at the destination airport in month m, and $\text{Var}(\xi_{i,c}) = \sigma^2_c$ is the (daily) variance of the error term due to the carrier related issues (i.e., idiosyncratic shock).

Also, we assume independent airport and carrier shocks, i.e., $\mathbb{E}(\xi_{\text{org},d,m}\,\xi_{\text{dest},d,m}) = 0$, $\mathbb{E}(\xi_{\text{org},d,m}\,\xi_{i,c}) = 0$, and $\mathbb{E}(\xi_{\text{dest},d,m}\,\xi_{i,c}) = 0$.

To estimate the variance terms, we followed the four-step procedure listed below:

Step 1. Let $\mathbb{S}_{a,d}$ be the set of all flights that flew either from or to airport a on day d; also let $\varepsilon^a_{j,d,m,c}$ and $\varepsilon^a_{k,d,m,c}$ denote the residuals from OLS1 of jth and kth flights of day d operated by a carrier c, which has airport a as destination or origin. Now, we define a set $\mathbb{P}_{a,d}$ ($\mathbb{P}_{a,d} \subset \mathbb{S}_{a,d}$) for all pairs of flights that satisfy

$$\begin{aligned}\mathbb{P}_{a,d} = \big\{&\forall\, j,k \in \mathbb{S}_{a,d}\\ &\text{s.t. } [(j,\text{org}) \neq (k,\text{org}) \text{ or } (j,\text{dest}) \neq (k,\text{dest})]\\ &\text{and } [(j,\text{org}) \neq (k,\text{dest}) \text{ or } (j,\text{dest}) \neq (k,\text{org})]\big\}.\end{aligned} \tag{9}$$

Then, the destination airport shock and the origin airport shock in month m can both be estimated using

$$\hat{\sigma}^2_{a,m} = \frac{\sum_{d\in\{\text{days in month } m\}} \sum_{\forall j,k\in\mathbb{P}_{a,d}} \varepsilon^a_{j,d,m,c}\varepsilon^a_{k,d,m,c}}{\sum_{d\in\{\text{days in month } m\}} |\mathbb{P}_{a,d}|}. \tag{10}$$

Step 2. Use the following equation to estimate σ^2_c:

$$\hat{\sigma}^2_c = \frac{\sum_{\forall i,d,m,c}(\varepsilon^2_{i,d,m,c} - \hat{\sigma}^2_{\text{org},m} - \hat{\sigma}^2_{\text{dest},m})}{N_c}, \tag{11}$$

where N_c is the total number of observations of carrier c in the data set.

Step 3. For each observation, calculate

$$\hat{\sigma}^2_{i,d,m,c} = (\hat{\sigma}^2_{\text{org},m} + \hat{\sigma}^2_{\text{dest},m} + \hat{\sigma}^2_c) \tag{12}$$

using Equations (10) and (11).

Step 4. By using the relation of $\hat{b}^2_i = \hat{\sigma}^2_i/2$, estimate $\hat{b}_{i,d,m,c} = \hat{\sigma}_{i,d,m,c}/\sqrt{2}$.

Thus, our model assigns a different estimate of variance to each flight depending on which carrier, route, and month the flight is flown. There are considerable differences in variances across airlines, with Frontier and Southwest having the lowest variance estimates (see Table 5 in the online appendix). Also, our analysis shows that airports in the northeast such as Newark, Philadelphia, and Boston have very high variance estimates, whereas airports in the southwest such as Phoenix and Tucson have low variance estimates (see Table 6 in the online appendix).

5.1. Model 1: Log-Normal Demand Distribution

In this subsection we assume that $\ln(TD_i)$ is normally distributed. Using the estimated $\hat{\mu}_i$ and $\hat{\sigma}^2_i$, one can calculate the scheduled on-time arrival probability, SL_i, for each flight i as follows:

$$\begin{aligned}\widehat{SL}_i &= F(Q_i;\hat{\theta}_i) = \Pr(D_i \le Q_i)\\ &= \Pr(\ln(TD_i) - X_i^T\hat{\boldsymbol{\beta}} \le \ln(Q_i) - X_i^T\hat{\boldsymbol{\beta}}),\end{aligned} \tag{13}$$

where $\ln(TD_i) - X_i^T\hat{\boldsymbol{\beta}} = \hat{\varepsilon}_i$ and $\varepsilon_i \sim \text{Normal}(0,\sigma^2_i)$. Thus, the estimated scheduled on-time arrival probability for flight i is given by

$$\widehat{SL}_i = F(Q_i;\hat{\theta}_i) = \Phi\left(\frac{\ln(Q_i)-\hat{\mu}_i}{\hat{\sigma}_i}\right) = \Phi(z), \tag{14}$$

where Φ denotes the cumulative distribution function for the standard normal distribution.

5.2. Model 2: Log-Laplace Demand Distribution

This model is very similar to the log-normal distribution, except that it provides a better fit when data show very high kurtosis. In this model, we assume that TD_i has a log-Laplace distribution (i.e., $\ln(TD_i)$ has a Laplace distribution and $\varepsilon_i \sim \text{Laplace}(\gamma, 2b^2_i)$). The probability function of the Laplace(γ,b) distribution is $f(x\mid\gamma,b) = (1/2b)\exp(-|x-\gamma|/b)$, where γ is a location parameter and b is a scale parameter. The maximum likelihood estimator of γ, $\hat{\gamma}$, is the median of all ε_i's we have in our data set. For the estimation of b, we used the relation $\hat{b}^2_i = \hat{\sigma}^2_i/2$, where $\hat{\sigma}^2_i$ is found from Equation (12).

Using the estimated $\hat{\mu}_i$, $\hat{\gamma}$, and $\hat{b}^2_i$, one can calculate the scheduled on-time arrival probability for each flight i as follows:

$$\begin{aligned}\widehat{SL}_i = F(Q_i;\hat{\theta}_i) &= \mathscr{L}\left(\frac{\sqrt{2}(\ln(Q_i) - X_i^T\hat{\boldsymbol{\beta}} - \hat{\gamma})}{\hat{\sigma}_i}\right)\\ &= \mathscr{L}\left(\frac{(\ln(Q_i) - X_i^T\hat{\boldsymbol{\beta}} - \hat{\gamma})}{\hat{b}_i}\right),\end{aligned} \tag{15}$$

where $\mathscr{L}$ is the CDF of a Laplace distribution with parameters $(0,1)$, where $\mathscr{L}(x) = 0.5[1+\text{sgn}(x)(1-\exp(-|x|))]$.

5.3. Fitting Demand Distributions

We fitted both the log-normal and log-Laplace truncated block-time models, as described in the earlier subsections, using the data for 13,227,718 flights flown in 2005 and 2006. The result of the log-block-time regression on explanatory variables (OLS1) (given by Equation (7)) is available in Table 3. The R^2 value for this regression is 0.8747. As expected, the airport congestion variable estimates are positive indicating that as airports become more congested, travel time increases. Route, airport, airline, month, day of week, and time of day fixed effects are also significant in the regression. The expected number of passengers and the age of the aircraft have positive and significant effect on the travel time indicating that larger and/or older aircraft require more time to complete a flight.

We first estimated Model 1 (log-normal truncated block-time distribution). Hence, we initially tested for normality of errors. We used the normality tests Kolmogorov–Smirnov, Cramer–von Mises, and Anderson–Darling. All of these tests revealed that the disturbances are not normally distributed because the p-value for these tests are all less than the critical value $\alpha = 0.05$. Therefore, the null hypothesis that the data are normally distributed must be rejected. It should be noted that these tests are not usually used alone to test the normality of the data. Besides, the power of any test (i.e., the ability to reject the null hypothesis) increases with the sample size; in other words, if the sample size is large, any small deviations from normality can be easily detected.

Next, we also tested Model 2 (log-Laplace truncated block-time distribution). Figure 6 in the online appendix displays the empirical CDF of the disturbances obtained from OLS1 and the parametric Laplace CDF. The maximum difference between the cumulative predicted value of parametric log-Laplace distribution and the actual cumulative value is 0.069 (i.e., 6.9%), while this difference is 12.51% for the log-normal distribution. However, the Kolmogorov–Smirnov test still rejects the hypothesis that the data have a log-Laplace distribution. Our data set is very large (with more than 13 million observations), and it is unlikely that any fitted distribution will pass a distribution test such as the Kolmogorov–Smirnov test. This is because even small deviations between the empirical and fitted CDF will be detected.

The above analysis shows that the log-Laplace distribution provides a better fit to the truncated block-time data than does the log-normal distribution. From a practical perspective, one can measure the maximum difference between the empirical and fitted CDF. In particular, we measure the difference between the empirical and fitted CDF in the region of practical interest, i.e., on-time arrival probability between 35% and 85% (more than 90% of flights fall in this region). In this area of practical interest, the log-Laplace distribution shows a very good fit. The maximum difference between the empirical and fitted log-Laplace distribution in this region is 1.20%, with an average difference of only 0.47%, i.e., the average difference between the predicted CDF using log-Laplace distribution and the actual CDF is less than 0.5%. Hence, we assumed log-Laplace truncated block-time distribution (instead of log-normal) while estimating the parameters necessary to compute the on-time arrival probability for each flight. All results in §6 are based on the log-Laplace distribution.

6. Hypotheses Tests, Results, and Implications

Using the procedure described in §5, we estimated the overage to underage cost ratio ($\hat{\gamma}_i$) for 6,651,779 flights flown in 2007. The median of $\hat{\gamma}_i$ value was 0.7232, and its mean was 0.965. We also found that a large number of flights (2,379,862 or 35.7%) had an estimated γ value of greater than one, i.e., the imputed cost of a flight delay was less than the cost of an early arrival for these flights. This shows that airlines systematically "underemphasize" flight delays for a large number of flights. A histogram of the imputed γ values is provided in Figure 3.

We also calculated the scheduled on-time arrival probability ($\widehat{SL}_i$) for flights flown in 2007. The average scheduled on-time arrival probability across all flights was 56.51%, and the median was 57.88%. Note that, our definition of scheduled on-time probability, SL_i, excludes LAD as defined in §1.1. Including LAD would result in a drop in the average scheduled on-time probability to below 50%. Another interesting observation from our analysis is that scheduled on-time probability, SL_i, which excludes LAD, steadily decreases as the day progresses until around 6 P.M. and begins to improve thereafter. The average on-time arrival probabilites for different one-hour departure time blocks is displayed in Figure 4. This suggests that the drop in on-time performance as the day progresses is not completely explained by flights delayed earlier in the day.

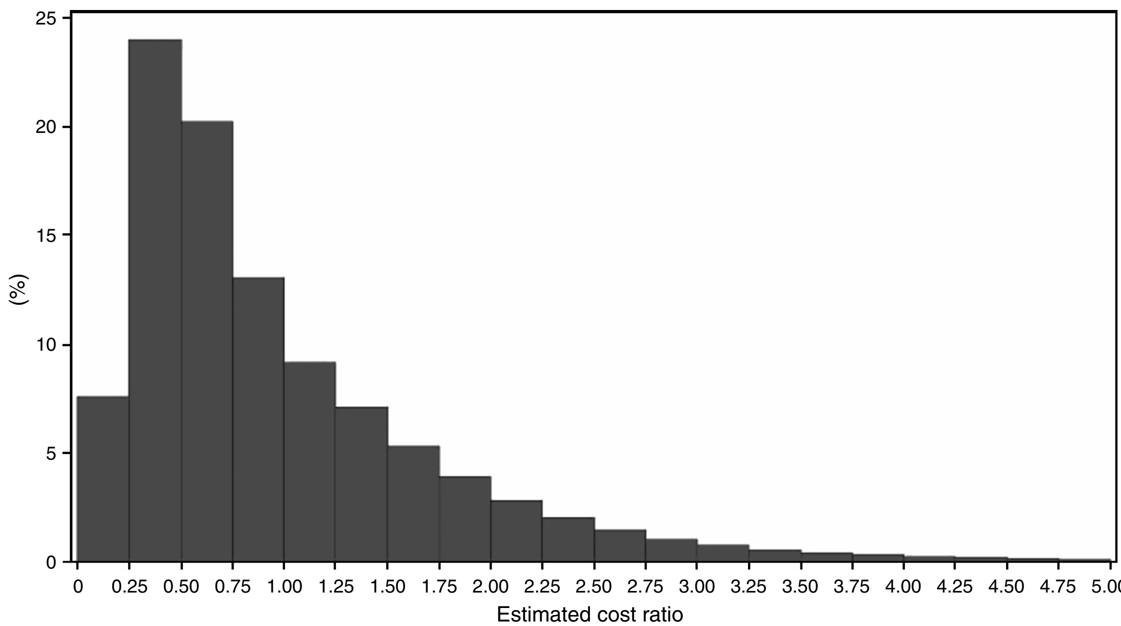
Let Z_i be the vector of covariates that impacts the overage to underage cost ratio and $\boldsymbol{\alpha}$ be the vector of the coefficients of these covariates. To test how the imputed overage to underage cost ratio ($\hat{\gamma}_i$) is affected by the covariates Z_i, we ran a regression of the form

$$-\ln \hat{\gamma}_i = Z_i \boldsymbol{\alpha} + \zeta_i. \qquad (16)$$

Note that this approach can be shown to be equivalent to model N1 in Olivares et al. (2008). We now expand Equation (16) to list the covariates used in our hypothesis tests:

$$\begin{aligned} -\ln \hat{\gamma}_i &\\ &= \alpha_0 + \alpha_1 \times \textit{Arr-Congestion}_i + \alpha_2 \times \textit{Dep-Congestion}_i \\ &\quad + \alpha_3 \times \textit{Total-Passgrs-Affected}_i \\ &\quad + \alpha_4 \times \textit{From-Connecting-Passgrs}_i \end{aligned}$$

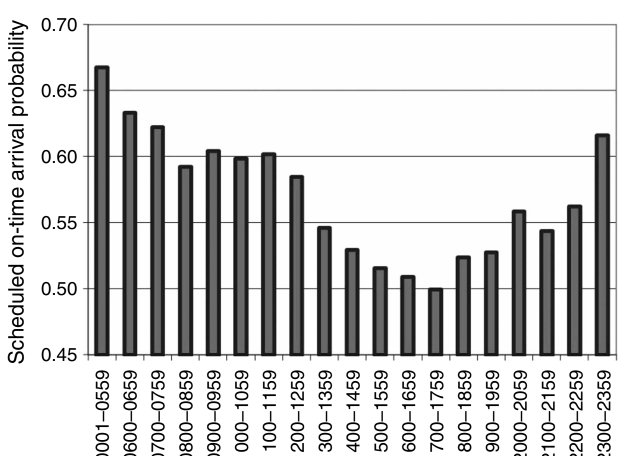
Figure 3 **Histogram of Imputed Gamma Values**



$$\begin{aligned}
&+\alpha_5 \times \textit{To-Connecting-Passgrs}_i + \alpha_6 \times \textit{Market-Share}_i \\
&+\alpha_7 \times (\textit{Num-Major-Carrier} \times \textit{Rank})_i \\
&+\alpha_8 \times \textit{Fare-per-Mile}_i + \alpha_9 \times \textit{Min-Fare-Flight}_i \\
&+\alpha_{10} \times \textit{Carrier-Group}_i + \alpha_{11} \times \textit{Org-Hub}_i \\
&+\alpha_{12} \times \textit{Dest-Hub}_i + \alpha_{13} \times \textit{Route}_i + \alpha_{14} \times \textit{Origin}_i \\
&+\alpha_{15} \times \textit{Destination}_i + \alpha_{16} \times \textit{Month}_i \\
&+\alpha_{17} \times \textit{Day-of-Week}_i + \alpha_{18} \times \textit{Dep-Time-Block}_i \\
&+\alpha_{19} \times \textit{Arr-Time-Block}_i + \alpha_{20} \times \textit{Aircraft-Age}_i + \zeta_i
\end{aligned} \tag{17}$$

The results of the above regression (labeled OLS2) are provided in Table 4.

Figure 4 **Scheduled On-Time Arrival Probability by Departure Hour of the Day**



Our first hypothesis tests whether the overage to underage cost ratio is increasing in the normalized departure/arrival congestion. We test if $\alpha_2 < 0$ through a t-test, where α_2 is the coefficient of the normalized departure congestion variable in OLS2. The estimated coefficient value is -0.8382, with a t-value -441.21 ($p < 0.0001$). Hence, we find support for the hypothesis that $\alpha_2 < 0$, indicating that the overage to underage cost ratio is increasing (and therefore, the scheduled on-time arrival probability is decreasing) in the normalized departure congestion. A hypothesis test on the normalized arrival congestion variable shows similar results. The mean of the *Dep-Congestion* (*Arr-Congestion*) variable in our data set is 0.412 (0.365). At the mean, an increase in the normalized departure (arrival) congestion by 1% would result in a 0.346% (0.396%) increase in the overage to underage cost ratio, resulting in a 0.28% (0.33%) decrease in scheduled on-time arrival probability. At a higher value of the *Dep-Congestion* (*Arr-Congestion*) variable corresponding to one standard deviation above the mean, an increase in the normalized departure (arrival) congestion by 1% would result in a 1.2% (1.5%) increase in the overage to underage cost ratio, resulting in a 2.2% (3.7%) decrease in scheduled on-time arrival probability.

Our above result on congestion has implications for policy planners who are interested in knowing the impact of alleviating congestion on on-time arrivals. There are several ways of alleviating congestion such as capacity expansions at congested airports or investing in new technologies (e.g., modernized landing and tracking technologies, En Route Automation Modernization program (ERAM), etc.) assessed

Table 4 Summary of OLS2 Estimation

Dependent variable: $-\ln($*Imputed Overage-to-Underage Cost Ratio*$)$

Variable	Df	Level	OLS2 parameter estimate
Intercept ($\hat{\alpha}_0$)			1.6875*** (0.0999)
Arr-Congestion ($\hat{\alpha}_1$)	1		−1.0829*** (0.002)
Dep-Congestion ($\hat{\alpha}_2$)	1		−0.8382*** (4.54×10^{-5})
Total-Passgrs-Affected ($\hat{\alpha}_3$)	1		-7.5×10^{-6} *** (0.0038)
From-Connecting-Passgrs ($\hat{\alpha}_4$)	1		−0.0041*** (2.49×10^{-5})
To-Connecting-Passgrs ($\hat{\alpha}_5$)	1		−0.0036*** (0.0012)
Market-Share ($\hat{\alpha}_6$)	1		−0.2350*** (0.0015)
Num-Major-Carrier × *Rank* ($\hat{\alpha}_7$)	16	(See online appendix)	
Fare-per-Mile ($\hat{\alpha}_8$)	1		0.3145*** (2.78×10^{-5})
Min-Fare-Flight ($\hat{\alpha}_9$)	1	0	—
		1	−0.1087*** (0.0028)
Carrier-Group ($\hat{\alpha}_{10}$)	3	Leisure	—
		Low cost	0.4973*** (0.003)
		Commuter/subsidiary	0.0196*** (0.0029)
		Full service	0.6396*** (0.0231)
Org-Hub ($\hat{\alpha}_{11}$)	1	0	—
		1	−0.0054*** (0.0015)
Dest-Hub ($\hat{\alpha}_{12}$)	1	0	—
		1	0.0196*** (0.0702)
Route ($\hat{\alpha}_{13}$)	3,924		
Origin ($\hat{\alpha}_{14}$)	245		
Destination ($\hat{\alpha}_{15}$)	240		
Month ($\hat{\alpha}_{16}$)	11	(See online appendix)	
Day-of-Week ($\hat{\alpha}_{17}$)	6	(See online appendix)	
Dep-Time-Block ($\hat{\alpha}_{18}$)	18	(See online appendix)	
Arr-Time-Block ($\hat{\alpha}_{19}$)	18		
Aircraft-Age ($\hat{\alpha}_{20}$)	1		−0.0119*** (1.17×10^{-6})
Number of observations		5,992,040	
R-square		0.4650	

Note. Robust standard errors are in parentheses.
***$p < 0.0001$.

by the FAA under its Air Traffic Control (ATC) and Next Generation (NextGen) modernization program (*ATC Modernization and NextGen: Near-Term Achievable Goals* 2009). It is obvious that reduced congestion will lead to shorter travel times (as verified by OLS1). However, this may result in airlines cutting their scheduled block times, so its net impact on on-time arrival performance is not immediately clear. Our structural estimation approach allows us to separate out the impact of congestion on actual block times and scheduled on-time arrival probability. Our result shows that reducing congestion will lead to an increase in the scheduled on-time arrival probability; hence, it is important for policy planners to focus on reducing airport congestion. Our model also provides policy planners with a tool for estimating the magnitude of the improvement in on-time performance by reducing congestion.

Our second hypothesis states that the overage to underage cost ratio is decreasing with increasing total number of passengers on the aircraft rotation who are potentially affected by a delayed flight. So, we test if $\alpha_3 > 0$. The coefficient estimate for the total number of affected passengers ($\hat{\alpha}_3$) is -7.5×10^{-6} and is significant at 99.99% significance level. Hence, we reject our second hypothesis. The third hypothesis looks at the number of incoming/outgoing connecting passengers at a hub airport. We check if the coefficients of the variables *From-Connecting-Passgrs* (α_4) and *To-Connecting-Passgrs* (α_5) are greater than zero. However, the estimated coefficients are $\hat{\alpha}_4 = -0.0041$ and $\hat{\alpha}_5 = -0.0036$ and statistically significant. Hence, we reject Hypothesis 3 that a flight's scheduled on-time arrival probability is increasing in the number of passengers making connections. To further support our analysis, we also consider our hub control variables: *Org-Hub* and *Dest-Hub* (i.e., *Org-Hub* $= 1$ if departure airport is a hub, and *Dest-Hub* $= 1$ if destination airport is a hub). The coefficient estimate for *Org-Hub* variable ($\hat{\alpha}_{11}$) is -0.0054, and the coefficient estimate for *Dest-Hub* variable ($\hat{\alpha}_{12}$) is 0.0196 (both are significant at 99.99% level). Our model projects that a 1% increase in these three variables (*Total-Passgrs-Affected*, *From-Connecting-Passgrs*, and *To-Connecting-Passgrs*) from their mean values would result in 0.0002%, 0.057%, and 0.054% increase, respectively, in their overage to underage cost ratios, resulting in a 0.001%, 0.034%, and 0.029% decrease, respectively, in the scheduled on-time arrival probability.

This set of results has implications for airline schedule managers. The above analysis shows that although airlines do schedule higher on-time arrival probability for flights with hub airports as destinations, the total number of passengers who might be making a connection does not positively affect scheduled on-time arrival probability. The total number of passengers who are potentially affected by a flight delay also does not positively affect the scheduled on-time arrival probability. To the extent that minimizing passenger inconvenience (and delay cost) is one of the goals for airlines, their scheduling policies seem to be inconsistent with that goal. The data imply that the number of passengers inconvenienced by a flight delay is inadvertently ignored in the current scheduling decisions. If reducing total passenger inconvenience is a desired objective for airline schedule managers, then they need to incorporate both the total number of passengers on a flight and the number of passengers making connections into their scheduling decisions.

Next, we test the effects of different competition measures on the scheduled on-time arrival probability. OLS2 results show that the route market share has a negative impact on the scheduled on-time arrival probability. The parameter estimate for the *Market-Share* variable ($\hat{\alpha}_6$) is −0.2350 (significant at 99.99% level), which means that the higher the market share of an airline on a route, the higher the overage to underage cost ratio and hence, the lower the scheduled on-time arrival probability. Our model estimates that a 1% increase in market share at its mean value would result in a 0.17% increase in the overage to underage cost ratio, resulting in a 0.12% decrease in the scheduled on-time arrival probability. Also, at a value of one standard deviation above the mean, a 1% increase in market share will result in a 0.41% increase in the overage to underage cost ratio and, hence decrease scheduled on-time arrival probability by 0.35%. Our other competition variable *Num-Major-Carrier* × *Rank* variable also shows a similar result for routes with two and three carriers, i.e., the coefficient estimate for carriers that have the rank 1 is lower than the coefficient estimate for airlines with rank 2 or 3. As a result, we find support for the hypothesis that as competition decreases, the overage to underage cost ratio increases and hence, the scheduled on-time arrival probability goes down; we reject Hypothesis 4B.

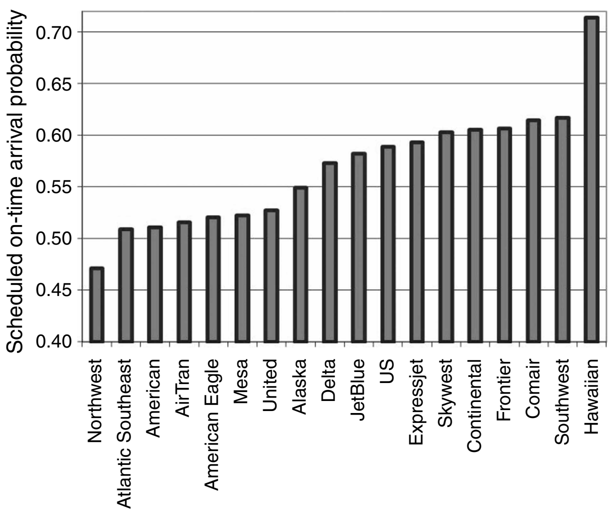
The above result on competition has implications for policy planners interested in understanding the impact of airline mergers on service quality to consumers. Currently, the U.S. airline industry is going through a profound transformation with the mergers of Delta and Northwest Airlines, Continental and United Airlines, and Southwest Airlines and AirTran Airways. On one hand, these mergers can create positive network externalities for airlines because of economies of scale. On the other hand, our above results suggest that, all else equal, on overlapping routes where a merger leads to an increase in the market share of one airline, there is a likelihood of observing worse on-time performance. Although we have not conducted a route by route analysis to understand the impact of mergers on on-time performance, our model provides policy planners a tool to conduct such analysis.

We next examine the revenue related variables. The parameter estimate of *Fare-per-Mile* variable ($\hat{\alpha}_8$) is equal to 0.3145 and is significant at 99.99% significance level, which gives support to our fifth hypothesis, i.e., as the average fare per mile increases, the imputed overage to underage cost ratio decreases, and hence, the scheduled on-time arrival probability increases. Our model projects that, at the mean value, a 1% increase in fare per mile would result in a 0.12% decrease in the overage to underage cost ratio, and hence increase the scheduled on-time arrival probability by 0.06%. Moreover, at a value of one standard deviation above the mean, a 1% increase in fare per mile would result in a 0.44% decrease in the overage to underage cost ratio, and hence increase the scheduled on-time probability by 0.16%. Our next hypothesis tests whether being the lowest fare carrier has a negative effect on the scheduled on-time arrival probability. The OLS2 results support the hypothesis that lowest fare carriers provide worse service because the estimated coefficient is $\hat{\alpha}_9 = -0.1087$ (significant at 99.99% level). Hence, we find support for the Hypothesis 6.

The above set of results has implications for passengers. These results suggest that there is a trade-off between fare paid for a flight and on-time performance. Passengers need to carefully assess this trade-off when booking their itineraries. Passengers who are very sensitive to flight delays may prefer to focus on the on-time arrival performance, whereas passengers less sensitive to delays may direct their attention to finding cheaper flights when making their booking decisions. The above result also has implications for travel websites that provide fare and on-time performance information. Most travel websites currently provide on-time performance summary of a flight, which is an aggregate average of observed performance over the past three–four months. Our analysis shows that such estimates are not accurate in projecting the on-time performance of a future flight. Our model can help travel websites improve the transparency and accuracy of expected on-time performance of a future flight, and thus help passengers make more informed decisions.

Last, we examine if the overage to underage cost ratio, or the scheduled on-time arrival probability, is affected by the airline group that flies the flight. We first display some summary statistics on how the scheduled on-time arrival probability varies by airline. Figure 5 plots the average of $\widehat{SL}_j$, where the average is computed over all flights flown in 2007 for each airline j. This figure shows that there exist significant differences across airlines in their scheduled on-time arrival probability. For example, the best airline in this data set (Hawaiian) has a scheduled on-time arrival probability of 71%, which is significantly higher than the airline with the worst scheduled on-time arrival probability (Northwest) at around 47%. Among the

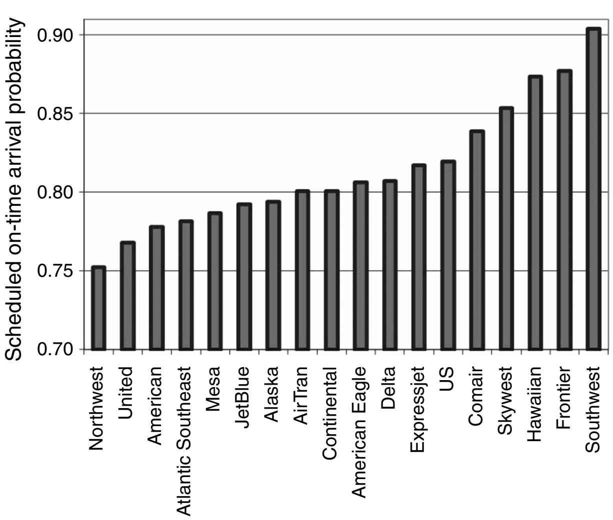
Figure 5 **Scheduled On-Time Arrival Probability by Airline**



full-service airlines, Continental and US Airways rank at the top of scheduled on-time arrival probability, whereas Northwest and American rank at the bottom.

Figure 6 plots the average scheduled on-time arrival probability based on the DOT on-time definition for each airline. As seen in these graphs, the airline scheduled on-time arrival probability rankings depend on the metric used for the ranking. For example, Southwest ranks the highest on the DOT scheduled on-time arrival probability metric but not on our scheduled on-time arrival probability metric. This suggests that Southwest may be managing its performance based on the DOT service metric. Some airlines consistently rank at the bottom of the rankings, independent of which metric is used. Northwest, for example, is at the bottom of both graphs, suggesting poor schedule planning. Finally, the graphs show that the 15 minutes buffer provided by the DOT on-time definition significantly inflates on-time arrival probability. Although the BTS reported average scheduled on-time arrival probability may be as high as 83%, the true average scheduled on-time arrival probability is closer to 56.5%. This suggests that the DOT on-time definition may be driving the scheduling performance, where the average scheduled block time is lower than the average actual block time.

Figure 6 **Scheduled On-Time Arrival Probability by Airline (DOT Definition)**



To formally test our hypothesis, we conduct an F-test to determine if $\alpha_{10,\,\text{leisure}} < \alpha_{10,\,\text{low-cost}} < \alpha_{10,\,\text{commuter}} < \alpha_{10,\,\text{full-service}}$. Although the F-test rejects the above hypothesis, we do find support for an alternate hypothesis $\alpha_{10,\,\text{leisure}} < \alpha_{10,\,\text{commuter}} < \alpha_{10,\,\text{low-cost}} < \alpha_{10,\,\text{full-service}}$. Thus, excluding commuter/subsidiary airlines, we do find support for our hypothesis regarding the impact of airline groups on the overage to underage cost ratio.

Our results suggest that to the extent one sees better on-time performance for low-cost airlines (e.g., Southwest), this can be attributed to other factors such as choice of airports, network congestion, etc., rather than their scheduling performance. All else equal, full-service airlines do place higher emphasis on the cost of flight delays and, hence, have a higher target on-time arrival probability. Therefore, passengers need to evaluate their choice of carrier carefully by understanding the trade-offs that might be involved.

7. Conclusions

The goal of our paper was to answer the following questions: How is on-time performance affected by the scheduled block time? What is the relative emphasis put on late arrivals versus early arrivals by airline schedulers, as implied by the imputed overage to underage cost ratio (or the scheduled on-time arrival probability)? How is the overage to underage cost ratio affected by operational, competitive, revenue, and cost factors (i.e., what are some of the drivers of heterogeneity in airline scheduling decisions)? What are some implications of the heterogeneity in the scheduled on-time arrival probability?

The modeling contribution of this paper to the current literature on flight delays are: First, we collected data from several public sources to model the total travel time distribution. We measured several explanatory variables of total travel time in an innovative fashion, and the multivariate regression model with these variables explained 88% of the variation in total travel time. Second, by modeling the total air travel time distribution using empirical flight data and using the observed scheduled block times, we provide a method for forecasting scheduled on-time arrival probability for each *individual* scheduled domestic flight in the United States, which captures the heterogeneity in various factors that affects the on-time arrivals.

Our results indicate that the on-time arrival probability depends significantly on the definition used. If a flight is measured to be late if it arrives after its scheduled arrival time (including LAD), the overall on-time arrival percentage for the airline industry drops to 56.5%. This value is significantly lower than the 82% on-time arrival percentage for the year 2007 based on the DOT's definition (recall that the BTS reports a flight to be late only if it arrives 15 minutes or more after its scheduled arrival time). An interesting question, left for future research, is if the DOT definition of flight delays incentivizes airlines to cut their flight schedules below the mean total travel time. If this is true, a policy implication of our analysis is to change the definition of a flight delay in the BTS official reports so that "on time" *really* means on time.

We estimated the overage to underage cost ratio ($\hat{\gamma}_i$) for 6,651,779 flights flown in 2007. The median $\hat{\gamma}_i$ value was 0.7232, and the mean was 0.965. We also found that a large number of flights (2,379,862 or 35.7%) had an estimated γ value of greater than one, i.e., the imputed cost of flight delay was less than the cost of an early arrival for these flights. This shows that airlines systematically underemphasize flight delays for a large number of flights.

We show that increased congestion results in an increased overage to underage cost ratio for airlines, resulting in decreased scheduled on-time arrival probability. However, an increase in the number of expected passengers on a flight and subsequent flights on the aircraft rotation does not result in a decrease in the overage to underage cost ratio, and hence does not result in an increase in the scheduled on-time arrival probability. We also show that the number of passengers making connections does not positively affect the on-time arrival probability. These results show that policy planners should focus on reducing congestion to improve on-time performance, and airlines need to incorporate the total number of passengers affected by flight delays in their scheduling decisions, if reducing passenger inconvenience is an objective.

We also show that increasing market share results in an increase in the overage to underage cost ratio for airlines, and hence a reduction in the scheduled on-time arrival probability. Also, airlines with the highest market share on routes served by multiple airlines have a higher overage to underage cost ratio and, hence, set lower on-time arrival probability targets. These results have implications for evaluating the impact of airline mergers on service quality in air travel industry.

We show that a significant variable that drives the overage to underage cost ratio for airlines is the fare per mile flown. Increasing fare per mile flown results in a reduction in the overage to underage cost ratio for an airline, and hence the scheduled on-time arrival probability is higher for flights with higher fare per mile. Finally, flying the lowest fare carrier on a route results in an increase in the overage to underage cost ratio, and hence lowers the scheduled on-time arrival probability. We also show that full-service carriers do place a higher weight on the cost of flight delays than do leisure or low-cost carriers. This shows that passengers need to carefully assess the trade-off between fares—also their choice of carrier—and on-time arrival performance while making their booking decisions.

Although our paper primarily took a "descriptive" approach in analyzing the flight delays, our framework can also be used for prescriptive purposes. An airline manager can look at the overage to underage cost ratio implied by airline's schedules to see if it confirms intuition. For instance, if the management believes that the underage costs are 75% higher than overage costs, and they find that the flight schedule implies overage costs almost equal to underage costs (as our results suggest), then this might signal a mismatch between management objectives and actual scheduling behavior. Our analysis suggests that there exist a large number of flights for which the overage to underage cost ratio is either very high or very low. Airline managers need to look at the implied overage to underage cost ratio for each flight to do some reality checks.

Structural estimation is a useful tool for prescriptive analysis of the system. One advantage of the structural estimation approach is that it disentangles whether a specific factor affects the observed block-time decision, through the distribution of the random travel time, or through the overage/underage cost ratio. For example, often it might be easier for the airlines to change factors that affect the cost ratio (e.g., fares) than to change factors that affect travel-time. It can also help policy planners (e.g., the FAA) to understand the impact of their decisions on on-time arrivals. Also, because congestion affects both the travel time and the overage to underage cost ratio, our model is able to disentangle these effects. For example, our model can be used to estimate the impact of easing congestion (e.g., by adding to airport capacity) on travel time, and also on on-time arrivals.

The newsvendor model is also useful in suggesting remedies for improving airline scheduled on-time performance. The newsvendor model suggests that scheduled on-time arrival probability would increase if the underage cost for the airlines were to increase. There are several ways by which underage costs can be increased for airlines, such as passing a passenger bill of rights, charging landing fees based on congestion, or restricting the number of scheduled flights based on the capacity of an airport. Our model suggests a way of estimating the impact of these policy changes on on-time arrivals through the impact of policy changes on the overage to underage cost ratio.

Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://msom.journal.informs.org/.

Acknowledgments

The authors are grateful for the thoughtful and constructive feedback provided by the editor, Steve Graves; the associate editor; and three anonymous referees. The authors thank organizers and participants of the Wharton Empirical Research Workshop, in particular Chris Ittner, for their helpful comments. The authors thank seminar participants at Cornell University, Indian School of Business, University of British Columbia, University of Michigan, University of Illinois at Urbana–Champaign, University of Cincinnati, University of North Carolina at Chapel Hill, and Purdue University. The authors also thank airline operations research practitioners and participants of the Airline Group of the International Federation of Operational Research Societies (AGIFORS) 2008 Montreal Conference for providng a practical perspective.

References

ATC Modernization and NextGen: Near-Term Achievable Goals. 2009. 111th Congress Subcommittee on Aviation Hearing. U.S. Government Printing Office; Washington, DC.

Banker, R. D., I. Khosla, K. K. Sinha. 1998. Quality and competition. *Management Sci.* **44**(9) 1179–1192.

Brons, M., E. Pels, P. Nijkamp, P. Rietveld. 2002. Price elasticities of demand for passenger air travel: A meta-analysis. *J. Air Transport Management* **8**(3) 165–175.

Brueckner, J. K., P. T. Spiller. 1991. Competition and mergers in airline networks. *Internat. J. Indust. Organ.* **9**(3) 323–342.

Cachon, G., C. Terwiesch. 2006. *Matching Supply with Demand: An Introduction to Operations Management*. McGraw-Hill, New York.

Dresner, M. 2006. Leisure versus business passengers: Similarities, differences, and implications. *J. Air Transport Management* **12**(1) 28–32.

Economides, N., K. Seim, V. B. Viard. 2008. Quantifying the benefits of entry into local phone service. *RAND J. Econom.* **39**(3) 699–730.

Forbes, S. J. 2008. The effect of air traffic delays on airline prices. *Internat. J. Indust. Organ.* **26**(5) 1218–1232.

Gal-Or, E. 1983. Quality and quantity competition. *Bell J. Econom.* **14**(2) 590–600.

Grandeau, S. C. 1995. The process of airline operational control. Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge.

Hebert, J. E., D. C. Dietz. 1997. Modeling and analysis of an airport departure process. *J. Aircraft* **34**(1) 43–47.

Hoxby, C. M. 2000. Does competition among public schools benefit students and taxpayers? *Amer. Econom. Rev.* **90**(5) 1209–1238.

Kessler, D. P., M. B. McClellan. 2000. Is hospital competition socially wasteful? *Quart. J. Econom.* **115**(2) 577–615.

Kranton, R. E. 2003. Competition and the incentive to produce high quality. *Economica* **70**(279) 385–404.

Lapré, M. A., N. Tsikriktsis. 2006. Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Sci.* **52**(3) 352–366.

Mayer, C., T. Sinai. 2003a. Network effects, congestion externalities, and air traffic delays: Or why not all delays are evil. *Amer. Econom. Rev.* **93**(4) 1194–1215.

Mayer, C., T. Sinai. 2003b. Why do airlines systematically schedule their flights to arrive late? Working paper, University of Pennsylvania, Philadelphia.

Mazzeo, M. J. 2003. Competition and service quality in the US airline industry. *Rev. Indust. Organ.* **22**(4) 275–296.

McCartney, S. 2007a. Frequent flying: Small jets, more trips worsen airport delays. *Wall Street Journal* (August 13) A1.

McCartney, S. 2007b. Why flights are getting longer. *Wall Street Journal* (May 29) D1, D5.

Mueller, E. R., G. B. Chatterji. 2002. Analysis of aircraft arrival and departure delay characteristics. *Proc. AIAA Aircraft Tech., Integration, and Oper. Forum at Aviation Week's Aerospace Expo, Los Angeles*.

Olivares, M., C. Terwiesch, L. Cassorla. 2008. Structural estimation of the the newsvendor model: An application to reserving operating room time. *Management Sci.* **54**(1) 41–55.

Oliveira, A. V. M., C. Huse. 2009. Localized competitive advantage and price reactions to entry: Full-service vs. low-cost airlines in recently liberalized emerging markets. *Transportation Res. Part E: Logist. Transportation Rev.* **45**(2) 307–320.

Prince, J. T., D. H. Simon. 2009. Multimarket contact and service quality: Evidence from on-time performance in the US airline industry. *Acad. Management J.* **52**(2) 336–354.

Propper, C., S. Burgess, K. Green. 2004. Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market. *J. Public Econom.* **88**(7–8) 1247–1272.

Proussaloglou, K., F. Koppelman. 1995. Air carrier demand: An analysis of market share determinants. *Transportation* **22**(4) 371–388.

Ramdas, K., J. Williams. 2009. An empirical investigation into the tradeoffs that impact on-time performance in the airline industry. Working paper, London Business School, London.

Reiss, P. C., F. Wolak. 2006. Structural econometric modeling: Rationales and examples from industrial organization. J. J. Heckman, E. E. Leamer, eds. *Handbook of Econometrics*, Vol. 6. North-Holland, Amsterdam, 4277–4415.

Rhoades, D. L., B. Waguespack. 2005. Strategic imperatives and the pursuit of quality in the US airline industry. *Managing Service Quality* **15**(4) 344–356.

Rupp, N. G., G. M. Holmes. 2006. An investigation into the determinants of flight cancellations. *Economica* **73**(292) 749–783.

Schumer, C. E., C. B. Maloney. 2008. Your flight has been delayed again: Flight delays cost passengers, airlines, and the U.S. economy billions. Report, Joint Economic Committee Majority Staff of U.S. Senate, chaired by Senator Charles E. Schumer and vice-chaired by Rep. Carolyn B. Maloney. Accessed April 12, 2012, http://jec.senate.gov/public/?a=Files.Serve&File_id=47e8d8a7-661d-4e6b-ae72-0f1831dd1207.

Shumsky, R. A. 1995. Dynamic statistical models for the prediction of aircraft take-off times. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.

Suzuki, Y. 2000. The relationship between on-time performance and airline market share: A new approach. *Transportation Res. Part E: Logist. Transportation Rev.* **36**(2) 139–154.

Trottman, M., S. Carey. 2007. Unfriendly skies: As pays fall, airlines struggle to fill jobs. *Wall Street Journal* (May 16) A1, A19.

Tsikriktsis, N. 2007. The effect of operational performance and focus on profitability: A longitudinal study of the US airline industry. *Manufacturing Service Oper. Management* **9**(4) 506–517.

Tu, Y., M. O. Ball, W. S. Jank. 2008. Estimating flight departure delay distributions: A statistical approach with long-term trend and short-term pattern. *J. Amer. Statist. Assoc.* **103**(481) 112–125.

Willemain, T. R., H. Ma, N. V. Yakovchuk, W. A. Child. 2003. Factors influencing estimated time en route. Technical report, The National Center of Excellence for Aviation Operations Research (NEXTOR). Accessed April 12, 2012, http://www.nextor.org/pubs/Factors.pdf.

Windle, R., M. Dresner. 1999. Competitive responses to low cost carrier entry. *Transportation Res. Part E: Logist. Transportation Rev.* **35**(1) 59–75.