INDUSTRY STUDIES ASSOCIATION

INDUSTRY STUDIES ASSOCATION
WORKING PAPER SERIES

An Analysis of Passenger Delays Using Flight Operations and Passenger Booking Data

By

Stephane Bratu
Manhattan Associates
Burlington, MA 01803

Cynthia Barnhart
Global Airline Industry Program
Massachusetts Institute of Technology
Cambridge, MA 02139

2004
Industry Studies Association
Working Papers

WP-2004-20
http://isapapers.pitt.edu/

Submitted to the *Air Traffic Control Quarterly* for possible publication

An analysis of passenger delays using flight operations and passenger booking data

Stephane Bratu (sbratu@manh.com)
Manhattan Associates, Burlington MA 01803

Cynthia Barnhart (cbarnhar@mit.edu)
Massachusetts Institute of Technology, Cambridge MA 02139

Abstract

The performance metric used to evaluate on-time performance in the US airline industry is flight-based, measuring the number of flight legs with arrival delay of 15 minutes or more. We analyze airline passenger operations and schedule performance and conclude that this flight-based performance metric does not accurately reflect delays to passengers, primarily because it does not recognize the long passenger delays resulting from flight leg cancellations and missed connections. Using passenger bookings and flight operations data from a major US airline, we develop a *Passenger Delay Calculator* to compute passenger delays and to establish relationships between passenger delays and cancellation rates, flight leg delay distributions, load factors, and flight schedule design. Using the insights gained in our analysis, we define new passenger-centric metrics to address the shortcomings of existing flight-based metrics and more accurately evaluate schedule reliability.

1 Introduction

Although flight schedules and ticket prices have proven to be the main drivers of airline profitability (Gopalan and Talluri [GoT98]), studies show that on-time performance and service reliability are important to achieving long-term profitability. Heskett et al. [HLS94] show that customer satisfaction and loyalty drive long-term corporate profitability and growth. This point is illustrated by the America West Airlines experience in 1999. America West "found itself at the bottom of the DOT Consumer Report," and "... business load fell 2 points in the second half of 2000 as high-yield travelers decamped to other airlines. This contributed to a 98% fall in annual profits" (Flint [Flin00]).

Because the airline industry is a highly competitive business, service reliability can serve as a major competitive advantage to attract and retain passengers; especially business passengers who are time sensitive and particularly important to airline profitability (Belobaba and Simpson, [BeS82]). Moreover, low service reliability can result in decreasing average fares (Janusewski [Jan02]). Janusewski cites LaGuardia airport as an example of this, noting that airlines charged lower fares in 2000 when flight delays at LaGuardia were particularly high.

Airlines are able to schedule as many flights as they desire at all but four major airports in the domestic US. O'Hare International Airport (ORD), Reagan National Airport (DCA), Kennedy International (JFK) and LaGuardia (LGA) airports are the only slot constrained airports in the US, that is, the only airports at which the number of scheduled arrivals and departures are constrained by the government. The growing economy of the 90's led airlines to increase service, and to sometimes schedule more flights than airports could handle, even in optimal weather conditions. According to the Bureau of Transportation Statistics (BTS), the number of flights operated by commercial airlines in the United States from 1990 to 2000 grew from 6.6 million in 1989 to 9.0 million in 2000, a 36% increase. In that same period, airport capacity increased by only 1% (Federal Aviation Administration, [FAA01]). As more and more flights were scheduled at congested airports, ground congestion became a major impediment to efficient operations, and taxi-out times at major US airports increased sharply. According to the BTS ([Mea01]), the number of flight legs with taxi-out times exceeding one hour increased by 165% from 1995 to 2000, from 17,331 to 45,993. Moreover, the disproportionate increase in flights relative to total airport capacity resulted in severe system congestion and numerous flight delays and cancellations, adversely affecting the traveling public.

From 1995 to 2000, the number of passenger complaints recorded by major US airlines (source: DOT Air Travel Consumer Reports) and the number of news articles reporting poor performance in airline service reliability (source: Lexis-Nexis) rose dramatically (Table 1-1).

	1995	2000	Ratio (2000/1995)
Number of articles in US newspapers	22	101	4.6
Complaints (Per 100,000 passengers)	0.76	2.98	3.9

Table 1-1: Airline passenger dissatisfaction

Although a growing number of dissatisfied travelers in this period were able to file their complaints more readily via emails, Mitra [Mit01] argues that this factor alone does not account for the severe increase in passenger complaints. According to the DOT Air Travel Consumer Report, the most common complaints involved flight problems, defined as "flight cancellations, delays, or any other deviations from schedule, whether planned or unplanned".

The deterioration of schedule reliability was so severe that in 2000, public concerns and media coverage resulted in the introduction of the Shuster Passenger Bill of Rights, requiring airlines to:

- Provide reasons of flight schedule disruptions to the passengers; and
- Compensate passengers if they waited for more than two hours on the runways prior to take-off or after landing.

Although this bill was not passed, airlines responded with *Customer First Plans* that recognized the importance of notifying customers of known delays, cancellations and diversions. For more details on the Customer First Plans see http://www.customers-first.org/commitment1.html.

1.1 On-time performance measurements in the US

US regulators and airlines measure on-time performance using the *15 minute on-time performance (15OTP) metric*, also called the *airline dependability statistic.* With this metric, a flight leg is considered to be on time if it arrives less than 15 minutes after its scheduled arrival time, and a canceled flight is classified as a delayed flight. Since 1987, the general public has had access on a monthly basis to flight delay statistics published in the *Air Travel Consumer Report* and maintained in the *Airline Service Quality Performance (ASQP)* database. The *major airlines* in the US (defined as airlines generating revenues of $1 billion or more annually) are required by federal law to provide regulators with flight operation information, including actual departure time, arrival time, and cancellation and diversion data, for each domestic US flight leg serviced by jet aircraft. In Figure 1-1 (source: ASQP), we depict on-time performance from 1995 to 2000, as measured by the industry 15OTP yardstick of airline on-time performance; and; the percentage of flight legs canceled or arriving later than scheduled. With these two flight delay statistics alone, it is difficult to explain the sharp increase in passenger complaints and negative press reports shown in Table 1-1.

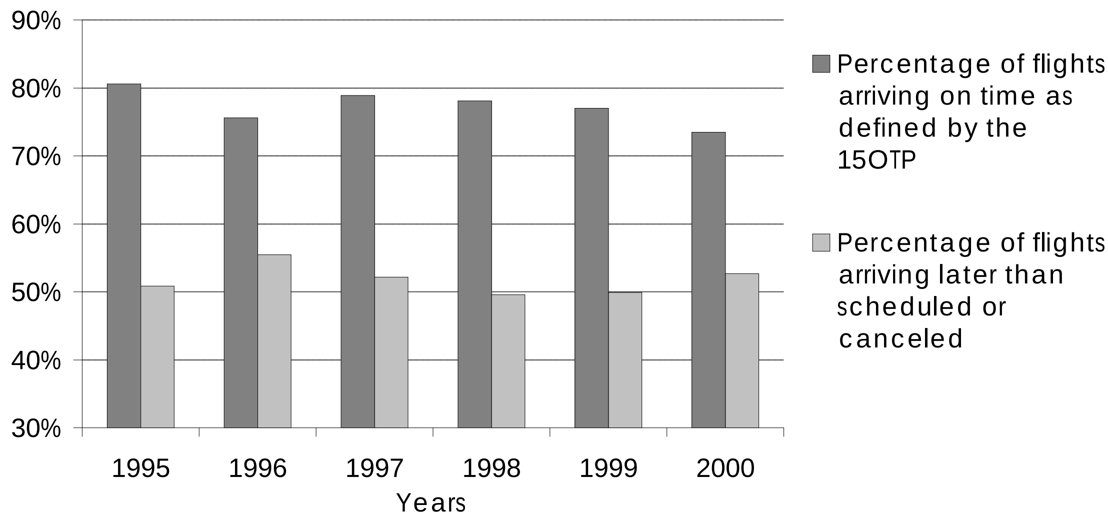

Figure 1-1: Trends in flight schedule performance

Shumsky [Shu93] and Hall [Hal99] report that US carriers have responded to the 15OTP metric by increasing planned block times and/ or scheduled gate-to-gate times for flight legs. Between 1973 and 1994, planned block times increased significantly, with continued growth at a slower rate between 1994 and 1999. This increase is a response to two factors:

- *The introduction of the 15OTP in 1987*. In order to improve on-time performance, airlines increased their block times; and

- *The increase in airspace and airport congestion*. The number of jet-operated flights scheduled by US major airlines in the domestic US at Boston Logan Airport (BOS), for example, increased by 16.1% from 1995 to 1999 (source: ASQP), while the airport capacity remained unchanged.

While longer planned block times can improve on-time performance, they result in greater operating costs (for example, crew costs increase) and in reduced productivity (for example, aircraft utilization decreases). Average actual block time has increased by 5 minutes from 1995 to 1999 for the 100 *routes* with the highest *frequency* in 1999. (A route is the sequence of one or more flight legs serving an origin/destination airport pair. By convention, flight legs from an origin airport *o* to a destination airport *d* constitute a route, while the return flight legs from *d* to *o* represent a different route, called the *opposite route*. We define the frequency of route *r* to be the number of times it is flown.) The incremental operating costs associated with increased block times are estimated at $1 billion, assuming an average direct operating cost of $1,800/hour and ignoring opportunity costs (that is the potential passenger revenue gain using 1995 block times).

Caulkins et al. ([CBL93]) examine this trade-off between on-time performance and its related costs. They argue that airlines operating at congested airports are disadvantaged by the 15OTP metric. They propose alternative approaches to estimate schedule reliability that compare airline schedule performance at an airport with average on-time performance of all airlines at that airport.

The 15OTP metric and the approaches suggested by Caulkins et al. are "*flight-centric*" measures of schedule reliability, each measuring delays to *aircraft*. In this paper, we present "*passenger-centric*" metrics aimed at measuring passenger delays and expressing schedule reliability as a function of passenger experiences. We demonstrate, for period 1995 to 2000, that there is a discrepancy between passenger perceptions of delay and delay as described by flight-based delay statistics. Note that we have chosen not to present more recent statistics (from 2001 to 2003) because we believe they represent anomalies in air transportation trends due to the September 11, 2001 terrorist attack and the economic recession.

1.2 Outline

In section 2, we describe our *passenger delay calculator* to quantify passenger delays, given passenger booking and flight leg delay and cancellation data. We show, in Section 3, that existing flight-based delay metrics are not accurate surrogates of passenger delays for hub-and-spoke airlines. Using our passenger delay calculator, we demonstrate that existing metrics are inadequate because they do not capture the effects of passenger disruptions caused by missed connections and flight leg cancellations. Using our delay calculator, we establish relationships between passenger delays and cancellation rates, flight leg delay distributions, load factors, and flight schedule design. Based on our findings, we propose in Section 4, new flight-based metrics to measure schedule performance.

2 Quantifying passenger delay

Knowledge of aircraft delays is critical to many airline functions. Without this information, for example, aircraft maintenance compliance and crew salary calculations cannot be performed. Passenger delays, however, have not been recorded historically, primarily because they are not indispensable to airline operations. As a result, passenger-based delay metrics are not commonly used in the industry. In this section, we develop the Passenger Delay Calculator algorithm (*PDC*) to:

- Estimate passenger delay in order to better measure passenger schedule reliability in the airline industry; and
- Recover disrupted passengers in real-time.

2.1 Definitions

To facilitate our description of *PDC*, we introduce the following notation and definitions. A non-stop *flight leg f*, also referred to more succinctly as *flight f*, is defined by a flight number, an origin airport, a destination airport, a *planned departure time*, PDT(f), and a *planned arrival time*, PAT(f). In operations, AAT(f) represents the *actual arrival time* of flight leg *f* at the gate and ADT(f), the *actual departure time* from the gate. The *flight arrival delay* of *f*, denoted FAD(f), equals max(AAT(f) - PAT(f); 0) and the *flight departure delay,* denoted FDD(f), equals max(ADT(f) - PDT(f); 0).

A *scheduled itinerary* is a sequence of scheduled flights serving a group of passengers. The group of passengers on a given scheduled itinerary is called a *scheduled passenger type*. If a scheduled itinerary contains only one flight leg then passengers are referred to as *local*, otherwise they are *connecting*. For a given day of operations, passengers are operated on a sequence of flight legs called the *actual itinerary*. A passenger is *disrupted* if:

- One or more of the flights in his/her scheduled itinerary is canceled; or
- The time between consecutive flights in his/her scheduled itinerary is less than the *Minimum Connecting Time* (MCT); that is, the minimum time required to walk between the arrival and departure gates of the consecutive flight legs.

Hence, scheduled and actual itineraries of *disrupted* passengers are different. Alternatively, *non-disrupted* passengers have the same scheduled and actual itineraries. Thus, the set of passengers P can be partitioned into subsets: D and ND, corresponding respectively to the set of disrupted and non-disrupted passengers. The queue of passengers to be re-accommodated by the airline is denoted *DQ*. Let DT(p) be the *disruption time* of passengers of type *p* and, d(p) the *passenger arrival delay* for passenger type *p*, computed as the maximum of zero and the difference between *p*'s actual arrival time and scheduled arrival time. Hence, letting L(p) denote the last flight in *p*'s actual itinerary, d(p) = max(AAT(L(p)) - PAT(L(p)); 0). We denote the maximum passenger delay as *MPD*.

2.2 Passenger Delay Calculator

Inputs to *PDC* include: 1) the planned flight schedule with given aircraft routings, 2) for each scheduled itinerary and corresponding passenger type, the number of booked passengers and their show-up rates (that is, the fraction of passengers of a given type who book seats and show up for their flight legs); and 3) the actual flight leg departure and arrival delays for each operated leg in the schedule; and 4) each canceled flight leg.

The steps of the *PDC* algorithm are illustrated in Figure 2-1.

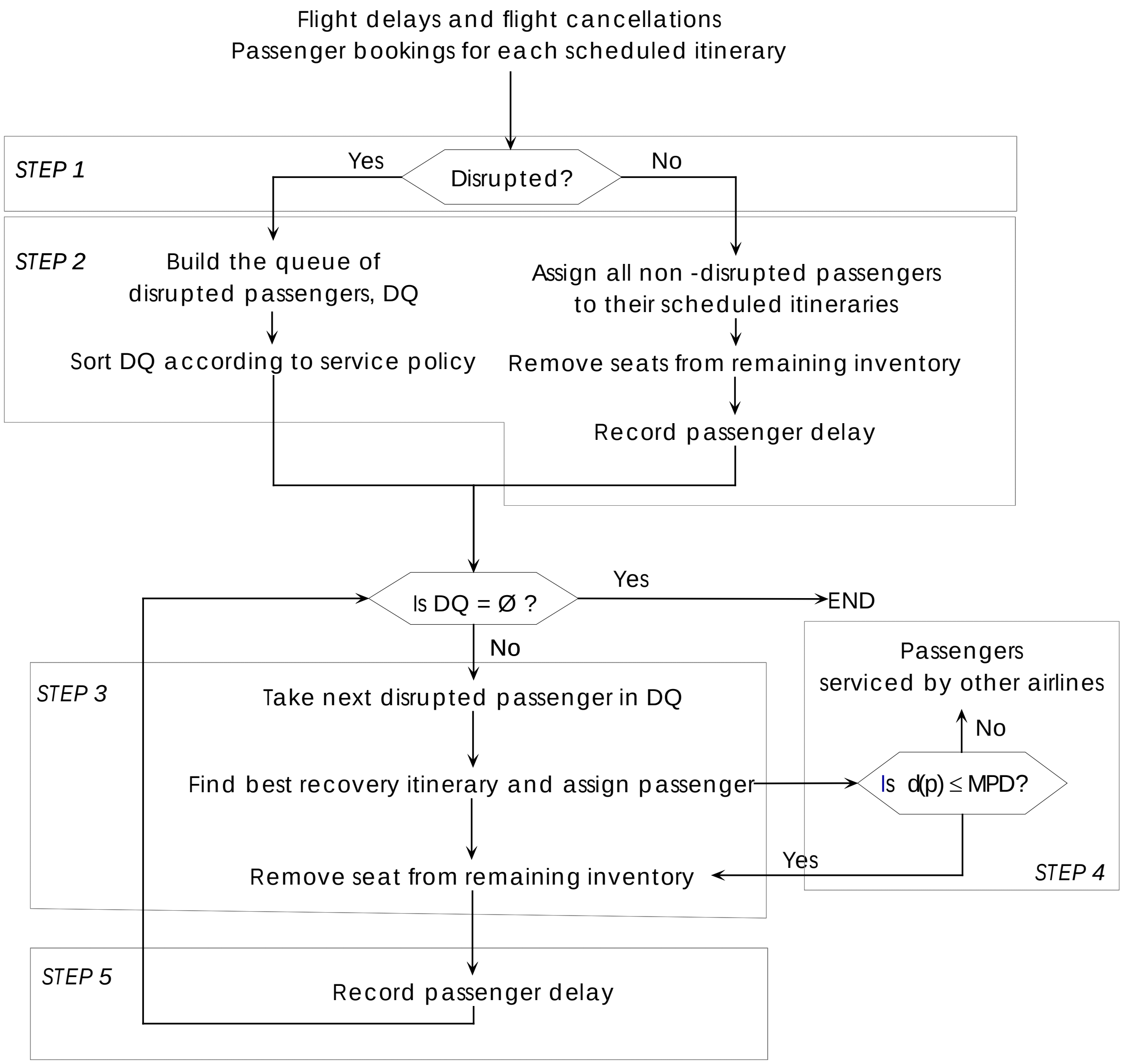

Figure 2-1: Passenger Delay Calculator schematic

In STEP 1, for each itinerary in the schedule, we determine the set of disrupted itineraries and their associated passenger types. In Step 2, disrupted passengers are placed in a recovery queue according to a specified recovery policy. In STEP 3, each disrupted passenger is re-accommodated on the *recovery itinerary* with seat availability that arrives earliest at the desired destination. In STEP 4, disrupted passengers, for whom there are no efficient recovery itineraries within a specified time frame, are assumed recovered on other airlines. In STEP 5, passenger delays and schedule reliability statistics are computed.

Details of the steps of PDC are as follows:

✧ <u>STEP 1: Identify disrupted and non-disrupted itineraries</u>

Consider each passenger type *p* on an itinerary, containing for example flight leg f(1) followed by f(2). The following sequential algorithm is executed to determine if passengers are disrupted:

❖ Is f(1) canceled?
⇒ If yes, passengers are disrupted, $p \in D$, and $DT(p) = PDT(f(1))$
❖ Else, is f(2) canceled?
⇒ If yes, passengers with PI are disrupted and $DT(p) = PDT(f(2))$
❖ Else, is $ADT(f(2)) - AAT(f(1)) < MCT$?
⇒ If yes, passengers are disrupted, $p \in D$, and $DT(p) = AAT(f(1))$
❖ Otherwise, passengers of type p are not disrupted, $p \in ND$, and $d(p) = Max(0; AAT(f(2)) - PAT(f(2)))$.

Figure 2-2: Passenger disruption evaluation

In *PDC,* consistent with industry practice, we remove seats assigned to non-disrupted passengers from the list of those available, ensuring that disrupted passengers are not assigned these seats. Available seat inventories are computed by subtracting the number of non-disrupted passengers on a given flight leg from the total number of assigned seats.

✧ STEP 2: Order the disrupted passengers

In the second step of PDC, we build the disrupted passenger queue. We sort disrupted passengers, D, according to a selected airline recovery policy. Various service policies are possible, including: 1) re-accommodating passengers using a *first-disrupted-first-recovered* policy; 2) re-accommodating passengers in order of decreasing fare class value; or 3) re-accommodating passengers in order of decreasing frequent flyer status. Whichever policy is selected, disrupted passengers are sorted and processed in order. When two different passenger types are disrupted at the same time, we randomly select whom to re-accommodate first, or use other characteristics, such as generated revenue, to rank passengers in the recovery list.

✧ STEP 3: Re-accommodate disrupted passengers

For each passenger, the next step in the PDC algorithm is to find the recovery itinerary commencing at the airport where the passenger is located and arriving the earliest at the passenger's desired destination. Each recovery itinerary must be *operationally feasible*. Operational feasibility requires that all of the flight legs in the recovery itinerary are operated; the time between the arrival of a flight leg and the departure of the next flight leg in the itinerary is greater than *MCT*; the itinerary's first flight leg departs later than the disrupted passenger's ready time; and there is at least one available seat on each of the flight legs in the recovery itinerary. For each disrupted passenger, we refer to the operationally feasible itinerary that arrives earliest at the disrupted passenger's destination as the *best itinerary*.

Assuming that the passenger is of type *p,* two lists of recovery itineraries are generated: the *Direct Itinerary List* (DIL(p)) for which itineraries have one flight leg only, and the C*onnecting Itinerary List*, CIL(p,H), for which itineraries have multiple flight legs and connect through hub airport *H*. DIL(p) and CIL(p,H) are merged and sorted in increasing arrival time, and then type *p* passengers are re-assigned to the earliest arriving itineraries with seat availability. In Section 3, the recovery itinerary search algorithm is described and an example is provided.

✧ STEP 4: Recover severely delayed passengers on other airlines

In this module, disrupted passengers with delays exceeding the *Maximum Passenger Delay (MPD)* threshold (of 15 hours, see Bratu, [Brat03], for analysis supporting this selection) are re-accommodated on another airline. Because the other airlines' schedules are not known to *PDC*, we do not include these passengers in our delay statistic calculations.

✧ STEP 5: Generate outputs

The output of the PDC algorithm is a vector of passenger delay statistics, including average delays for each passenger type, and sizes of different passenger groups including: disrupted; non-disrupted; disrupted and recovered the same day; and disrupted overnight.

3 Example: recovery of disrupted passengers

To illustrate how disrupted passengers are aggregated into clusters, how the recovery list is built and how passengers are re-assigned to recovery itineraries in *PDC*, consider two scheduled passenger types, *SPT 1* and *SPT 2,* with the following planned connecting itineraries:

SPT	Number of passengers	Flight leg	From	To
1	10	f	A	H1
		h	H1	C
2	5	g	B	H1
		h	H1	C

Table 3-1: Example of disrupted itineraries

Assume that the airline operates 3 hubs, *H1, H2* and *H3*, and that flight leg *h*, scheduled to depart at 7:30PM and to arrive at 9:00PM, is canceled. Consequently, passengers belonging to *SPT* 1 and 2 are disrupted at *H1*.

3.1.1 Disrupted passenger clustering

Passenger types 1 and 2 can be aggregated within the disruption queue because:

- They are disrupted at the same time (e.g., 7:30PM),
- They are destined to the same airport (e.g., airport C) and,
- They were originally scheduled to arrive at the same time (e.g., 9:00PM).

By aggregating these two *scheduled passenger types* into a new, single type *p,* we generate only a single recovery list for passengers disrupted from *both* itineraries. This clustering technique is particularly useful when a flight leg departing a hub airport is canceled. In this case, most passengers with a common destination can be clustered into the same passenger type, thereby facilitating the decision-making process for operations controllers and speeding up the re-accommodation process.

3.1.2 Recovery list generation

We generate the recovery list of itineraries from *H1* to *C* that depart later than the *passenger ready time* (*PRT*). *PRT* is the sum of the planned departure time of canceled flight leg *h* and the *MCT* for passengers to walk to the gate of the first flight leg in their recovery itinerary.

The Direct Itinerary List for *p*, DIL(p), is easy to generate as it contains only flights from *H1* to *C* departing later than passenger *p*'s ready time *PRT(p)*. We sort these flights in increasing order of their arrival time at C.

Next, we build the Hub *H2* Connecting Itinerary List, CIL(p,H2), for passengers *p*. We begin by sorting all flight legs from *H2-C* in increasing order of arrival times. We denote the sorted list as g(1), g(2), ..., g(n). We also sort the flight legs from hub *H1* to hub *H2* that depart later than *PRT(p)*. We denote this sorted list of flight legs as f(1), f(2), ..., f(m). Then, beginning with *g(1)*, we look among all the flight legs f(1), f(2), ..., f(m) for feasible connections with *g(1)*. (The set of feasible connections is illustrated in Figure 2-3.) {f(1),g(1)} is operationally feasible, f(1) has 7 available seats and g(1) has 10 available seats, hence {f(1),g(1)} enters CIL(p,H2) with 7 seats available. Because the next itinerary {f(2),g(1)} is not feasible, all remaining itineraries with g(1) are not feasible. The next step, then, is to consider flight leg g(2). Both {f(1),g(2)} and {f(2),g(2)} are feasible itineraries with available seats, and hence they are included in CIL(p,H2). The same procedure is repeated to generate connecting itinerary lists for hub *H3*. To generate 3-flight leg recovery itineraries, we build upon the 2-flight leg approach. First, we build all feasible 2-flight leg itineraries from any hub to the destination, and then we expand these by adding any feasible leg from the location of the disruption to the origin of the 2-flight leg itinerary. Generation of recovery itineraries is terminated when the cumulative number of seats for the itineraries in CIL(p,H2) is at least as great as the number of disrupted type *p* passengers.

Figure 3-1 illustrates the generation process. The sequence of arrows in the picture represents the order in which itineraries enter CIL(p,H2).

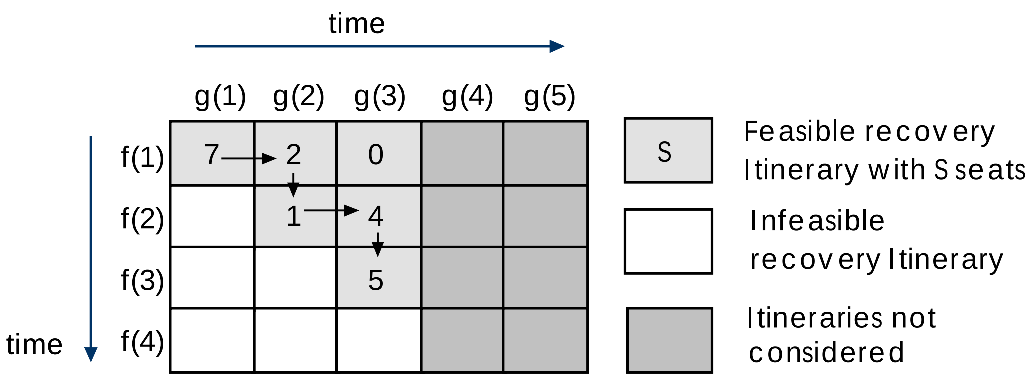

Figure 3-1: Example of connecting recovery itinerary generation

3.1.3 Recovery itinerary assignment

Recovery itinerary lists, DIL(p) and CIL(p,H), are merged for all hub airports and the aggregated list is sorted in order of increasing arrival time. The passengers in the recovery queue are then processed in order of increasing disruption time and assigned to the first recovery itinerary in the list with available seats. For every assigned passenger, the number of available seats on each flight leg in the assigned itinerary is decremented by one.

As shown in Table 3-2, the 15 disrupted type *p* passengers in this example are assigned to all seats on *Recovery Itineraries* (*RI*) 1 and 2 and to 5 of the 10 available seats on *RI* 3.

RI#	Path	Number of seats	Arrival delay (minutes)	Number of passengers assigned
1	H1-C	3	146	3
2	H1-H2-C	7	386	7
3	H1-H3-C	10	407	5

Table 3-2: List of feasible recovery itineraries

3.2 PDC assumptions

Underlying PDC are assumptions that lead to approximations, usually *underestimates*, of *actual* passenger delays. These assumptions are summarized as follows:

- *Perfect information*: we assume that at any point in time we know future operations exactly. Consequently, disrupted passengers assigned to recovery itineraries will not be disrupted again. We also assume that the airline has perfect knowledge of the number of seats available for each flight. In actuality, airlines do not know this information, as some passengers do not show-up for their flights. For the major US airline we study, 15.4% of booked passengers are no-shows on average, with higher no-show rates for business passengers who often have fully refundable tickets.

- *Instantaneous information*: we assume that a disrupted passenger is instantaneously rebooked on the recovery itinerary that provides at least the minimum connection time between the time of the disruption and the itinerary departure, and arrives earliest at the desired destination. Because we do not know the *disruption time* of a canceled flight, we set its disruption time equal to its scheduled departure time.

- *No bumped passengers*: when there are disrupted passengers who are severely delayed, we do not account for the possibility of recovering seats for them by *bumping* passengers, that is, by enticing non-disrupted passengers to fly later and give up their seats for the disrupted passengers.

4 Delay analysis at a major hub-and-spoke airline

In this section, we compute and compare passenger and aircraft delay statistics for a major US airline using the airline's passenger reservation information and no-show rates for the month of August 2000. The airline operates a hub-and-spoke network in which three hubs serve 74 airports in the domestic United States. Typical of a hub-and-spoke carrier, the passenger mix is 65% local and 35% connecting, with almost all passenger connections occurring at one of the three hub airports.

Using the booking and no-show data covering 307,675 planned itineraries and 2.56 million passengers for August 2000, we derived the average passenger demand, that is, the number of passengers showing up, for each flight leg. Then, given operational delay and cancellation data from the ASQP database, we applied *PDC* to construct new itineraries for disrupted passengers, and to compute delay statistics for all passengers. In computing these statistics, we assume the following within *PDC*:

- *Disrupted passengers* are re-accommodated in increasing order of the time they are disrupted, that is, we employ the *first-disrupted-first-recovered* service policy. When a flight is canceled, the order in which we add the associated disrupted passengers to the disruption queue is random.

- Although transfer times differ for each connecting passenger type at each airport because of differences in the distances between gates and disembarking times, we set *MCT* to 10 minutes based upon a sensitivity analysis presented in Bratu [Brat03].

4.1 Passenger and flight statistics

We limit our analysis to jet-operated flight legs (of which there are 33,730, with 303 assigned jet aircraft) because the ASQP database includes only jet operated flight leg information. For each itinerary *j* containing one or more flight legs operated by other than jet aircraft, we reduce the seat capacity of each jet-operated flight leg in *j* by the number of itinerary *j* passengers. Passengers for whom we do not have all flight leg information are thus, assumed to be non-disrupted. This assumption is supported by our analysis of the airline's operations showing that fewer than 4% of the passengers are disrupted.

Using ASQP and airline data for August 2000, we compare the average performance of major US airlines with the airline we study (referred to as *our airline)* in Table 4-1 (source: ASQP). Compared to the major airlines in August 2000, our airline cancelled fewer flight legs, experienced shorter delays on average for operated flight legs, and achieved better on-time performance.

	Our airline (August 2000)	All major US airlines (2000 average)
15 minutes On Time Performance (15OTP)	78.0%	73.6%
Percentage of flights delayed by more than 45 minutes	11.7%	9.0%
Average delay of operated flights (minutes)	9.0	10.5
Percentage of canceled flights per day	2.2%	3.7%

Table 4-1: Flight operations statistics for our airline versus industry average

4.2 Discrepancy between passenger and flight delays

Using ASQP and airline data as inputs to *PDC*, we conclude that flight leg delays severely underestimate passenger delays. For August 2000, the average passenger delay of 25.6 minutes is 1.7 times greater than the average flight leg delay of 15.4 minutes, as illustrated in Table 4-2.

	Average delay (minutes)
All passengers	25.6
Flight legs	15.4
Ratio: passenger/flight leg delays	166%

Table 4-2: Average passenger and flight delay, August 2000

For the ten days in August 2000 with the lowest average flight leg delays and smallest number of canceled flights, we estimate the average passenger delay to be 1.6 times greater than the average flight leg delay of 15.4 minutes. On these days, 85.7% of non-disrupted passengers arrive within one hour of their scheduled arrival time, experiencing an average delay of 16 minutes, close to the average flight leg delay. Passengers disrupted by flight leg cancellations (causing 71% of the disruptions) or missed connections (causing the remaining 29%), however, experience an average delay of 303 minutes, accounting for 39% of all passenger delay minutes. The significantly longer than average *flight leg* delays experienced by disrupted passengers underscore the inadequacy of conventional flight-based metrics, such as 15OTP, in measuring schedule performance as experienced by passengers.

4.3 Passenger disruption analysis

Recognizing the significance of disrupted passengers in understanding passenger delays and schedule performance, we further investigate characteristics that influence the degree of disruption and the effects of disruption.

4.3.1 Connecting and local passengers

As shown in Table 4-3, connecting passengers, scheduled to connect between 2 or more flights, are on average 2.8 times more likely to be disrupted than local passengers and 1.5 times more likely to be disrupted by a flight cancellation than local passengers.

Passenger group	All	Connecting (C)	Local (L)	%(C/L)
Percent of disrupted passengers	3.20%	5.46%	1.97%	277%
Percent of passengers on canceled flights	2.32%	2.95%	1.97%	150%
Percent of passengers missing connections	0.88%	2.51%		

Table 4-3: Disruption risk for local and connecting passengers

Consequently, while only 35% of passengers are connecting, they represent 49% of the disrupted passengers. Among them, 52% are disrupted because of a flight cancellation and 48% because of a missed connection.

4.3.2 Time of passenger disruptions

In Figure 4-1, we depict average arrival delays in August 2000 for flight legs scheduled to arrive in each one-hour time window during the day at the airline's major hub. The average arrival delay, for example, of flight legs scheduled to arrive between 7:30 PM and 8:30 PM is 21 minutes. (Note that *8 PM* corresponds to *20* in the figure). We observe that flight leg delays generally increase as the day progresses. Moreover, the largest number of flight legs canceled per hour occurs between 6:30 PM and 7:30 PM, leaving limited time to re-accommodate the resulting group of disrupted passengers.

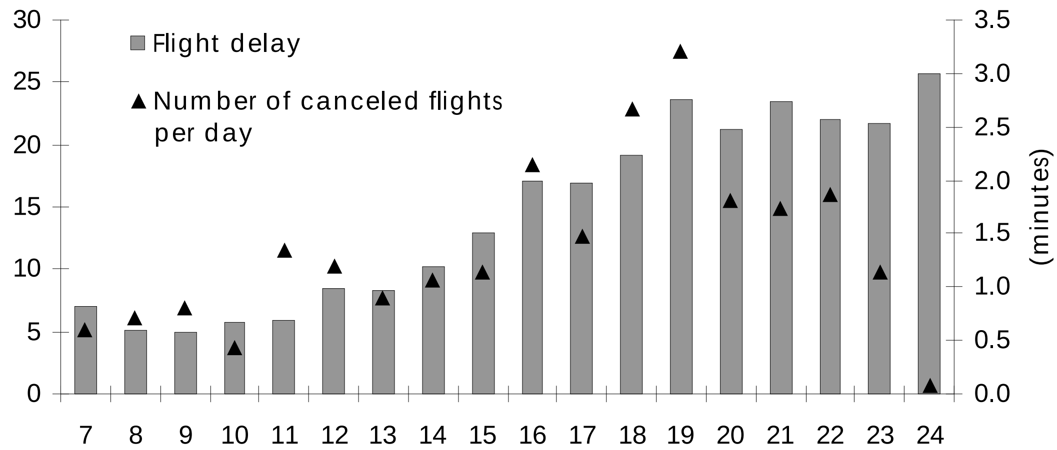

Figure 4-1: Flight arrival delays and number of canceled flights per day for flights scheduled to arrive in each hour time window

Growing delays as the day progresses and late-in-the-day cancellations provide an explanation for why half of the disrupted passengers are originally scheduled to arrive after 6 PM, and why the numbers of overnight passengers (those arriving at least one day late) increase each hour as the day progresses (Figure 4-2).

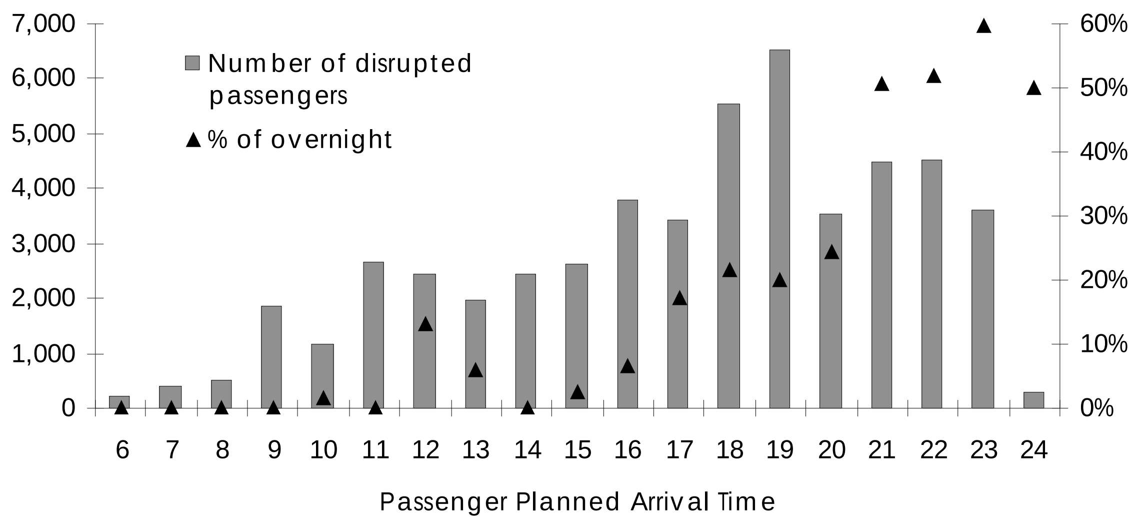

Figure 4-2: Number of disrupted passengers and percentage of overnight passengers per hour

Figure 4-3 summarizes the average delay of disrupted passengers for selected groups where delays are in hours. E*arly* corresponds to before noon and *late* to after 6PM. *Low frequency* corresponds to less than 3 flights per day and *high frequency* to more than 8 flights per day.

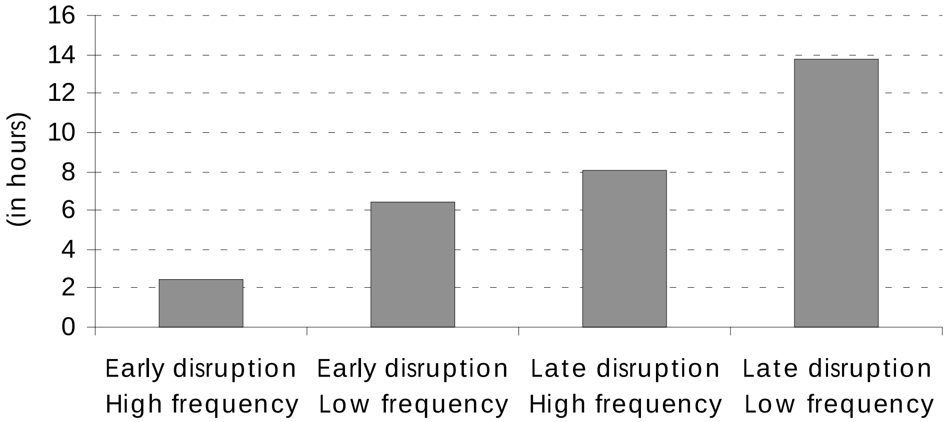

Figure 4-3: Average delay in hours for different groups of disrupted passengers

Clearly, low service frequency and disruptions late in the day contribute significantly to the delay of disrupted passengers. Consistent with these results, we plot average delay of disrupted passengers as: 1) frequency increases from the point of disruption to the passenger's destination (Figure 3-5, graph #1); and 2) disruption time changes (Figure 3-5, graph 2). As indicated in the graphs, delay is more sensitive to disruption time than to route frequency.

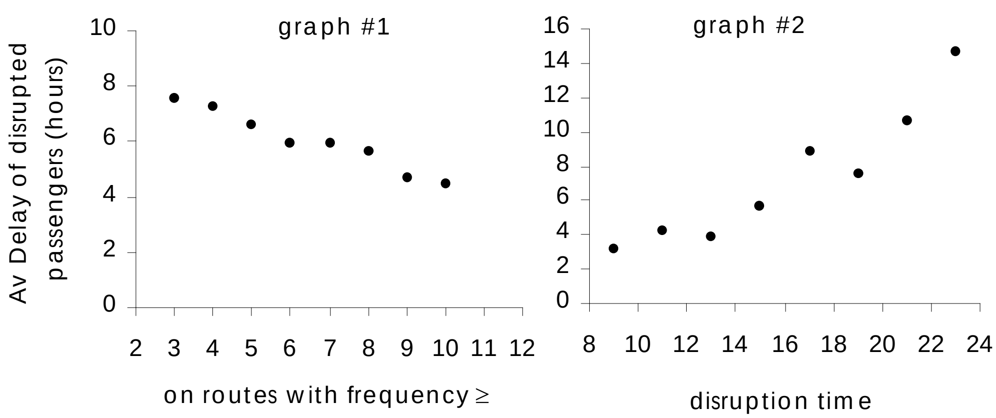

Figure 4-4: Average delay of the disrupted passengers versus route frequency and time of disruption

4.3.3 Seat availability

Due to lack of seat availability, 31% of disrupted passengers are not re-accommodated on their best itineraries (Table 3-5). If the disruption is caused by a flight cancellation rather than a missed connection, the situation worsens, with just over half (55.3%) of the disrupted passengers re-accommodated on their best itineraries. This occurs because each cancellation results in an average of 81.0

disrupted passengers, all competing for seats. In contrast, only 3.4 passengers on average are disrupted per missed connection.

	Percentage of All disrupted passengers	Percentage of disrupted passengers caused by:	
		Flight cancellations	Missed connections
Best itinerary	69.0%	55.3%	94.2%
Other itineraries	31.0%	44.7%	5.8%

Table 4-4: Statistics on disrupted passenger recovery itinerary

We conduct an experiment in which we run *PDC* with all aircraft assumed to have unlimited capacity, the *infinite capacity scenario*. We then contrast its solution, in which all passengers are re-accommodated on their best itineraries, with the *PDC* solution in which finite, actual aircraft capacities are specified. We find that limited seat capacity most adversely affects local passengers, who are all disrupted by a flight leg cancellation and likely, have fewer efficient recovery itineraries than connecting passengers. Moreover, from the infinite capacity scenario, we compute that 25% of the delays experienced by disrupted passengers are due to lack of seat availability and 75% to schedule design and airline operations.

4.3.4 Flight schedule design

The structure of the airline network and the scheduling of the flight legs are critical determinants of passenger disruption. To illustrate, consider the airline we investigate in which at the largest hub airport, the airline schedules 10 *complexes* or *banks*. A complex, or bank, is a set of arriving flight legs scheduled closely with a set of departing flight legs to allow passenger connections between arriving and departing flight legs. If passengers have the same probability of being disrupted at all complexes, the percentage of overnight passengers should be about 10%, much less than the 22% calculated using *PDC*. We believe this difference is due to propagation effects in the network. A delayed or canceled flight leg causes downstream delays to aircraft, crews and passengers, resulting in growing numbers of flight delays and cancellations as the day progresses.

Higher average flight leg delays and cancellations, however, do not always translate into longer delays for disrupted passengers. As shown in Table 4-5, we compare two days of operation, denoted *day 1* and *day* 2. *Day 1* has longer aircraft delays and more cancelled flight legs than *day 2*, but average delays for disrupted passenger are less on *day 2* than on *day 1*.

Day	Average delay of disrupted passengers (minutes)	Percent of flight legs canceled	Average flight delay (minutes)
1	495	1.0%	9.5
2	334	8.1%	40.4

Table 4-5: Illustration of the importance of disrupted itineraries

We explain this apparent contradiction through a simple example. Consider an arriving complex with flight legs *f(1)* and *f(2)*, and a departing complex with flight legs *f(3)* and *f(4)*. Assume that one passenger plans to connect on each connecting pair of flight legs, that is, *f1-f3*, *f1-f4*, *f2-f3* and *f2-f4*. We consider two flight delay scenarios, illustrated in Table 4-6, where delays are in minutes:

Scenario 1			
Inbound flight leg	Arrival delay	Outbound flight leg	Departure delay
f(1)	0	f(3)	0
f(2)	15	f(4)	15
Scenario 2			
Inbound flight leg	Arrival delay	Outbound flight leg	Departure delay
f(1)	10	f(3)	10
f(2)	10	f(4)	10

Table 4-6: Flight delay scenarios

Assume that disrupted passengers experience 4 hours of delay. Although, the average flight delay for Scenario 1 is lower than for Scenario 2 (7.5 versus 10 minutes), connecting passengers experience more than 6 times as much delay in Scenario 1 (64 minutes) than in Scenario 2 (10 minutes). The difference results because the passenger on itinerary {f(2),f(3)} in Scenario 1 is disrupted, and none of the passengers in Scenario 2 is disrupted.

Using *PDC,* we estimate that the number of misconnecting passengers in August 2000 would have increased by 38.1% (12,362 passengers) had the flight legs departed on time from hub airports. This hypothetical case is, of course, unrealistic but it nonetheless indicates that downstream departure delays can benefit connecting passengers. We conclude that flight leg delay *differences* are better indicators of missed connections than flight leg delays alone.

4.3.5 Load Factor Analysis

The *Average Load Factor* (AvLF) is defined as the ratio of the total number of booked passengers to the total number of seats supplied in the schedule. From 1994 to 2000, the average load factor in the US airline industry increased noticeably due to:

- A 4% per annum increase in passenger traffic,
- A 1% per annum decrease in average seat capacity per flight leg (according to the US Department of Transportation Form 41 data), and
- An increase in competition from 1996 to 2002 from the low cost carriers, whose market share increased from 12.6% to 17.1%, forcing yields down by 9% and causing break-even load factors to increase during that period (source: Air Traffic Association).

In this section, we present results that quantify the impact of load factors and flight schedule disruptions on passenger delays. We begin by generating a hypothetical *demand scenario* by multiplying the number of passengers on each itinerary by the same demand factor *d*. Changing the value of *d,* we create additional demand scenarios. Then, for each of these demand scenarios, we estimate passenger delays by solving *PDC* with August 2000 flight delay and cancellations data.

In our *base case,* AvLF is 40%, similar to the infinite capacity scenario described in Section 4.3.3, with almost all disrupted passengers recovered on their best itineraries. Re-running *PDC* for each demand scenario, we find that when loads increase, expected average delay of disrupted passengers increases exponentially, as illustrated in Figure 4-5.

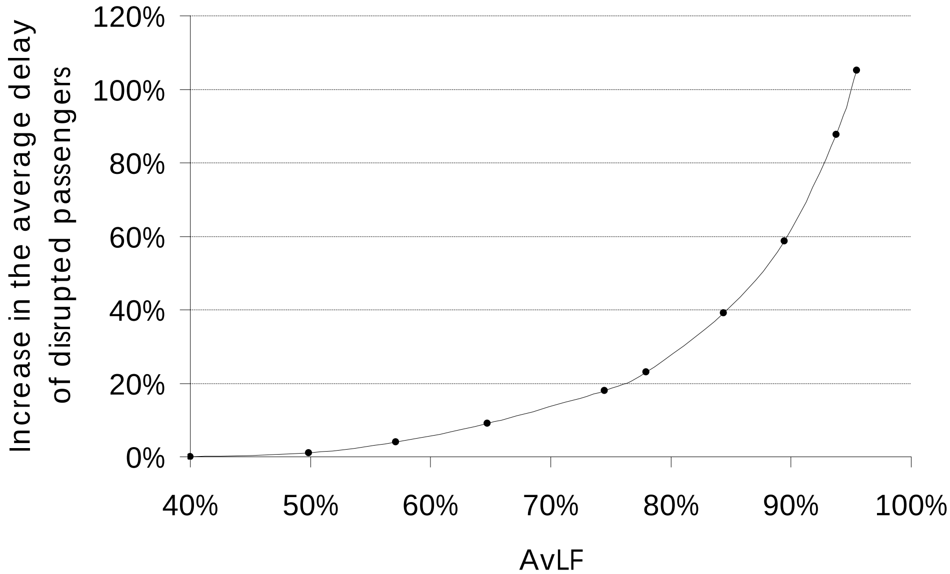

Figure 4-5: Average delay of disrupted passengers versus average load factor

When loads are low (that is less than 60%), a small number of disrupted passengers results and there are enough empty seats to re-accommodate most of them on their best itineraries. When the average load factor increases, however, the number of disrupted passengers increases and fewer seats are left unoccupied. As more disrupted passengers compete for fewer empty seats, the average delay increases rapidly. For average load factors in excess of the average load factor in August 2000 (73%), delays increase sharply. As loads increase, there are more overnight passengers because there are no seats available to re-accommodate them the same day. For low load factors, 80% of the disrupted passengers are recovered the same day, while for load factors between 90% and 95%, only 65% of the disrupted passengers are re-accommodated the same day.

In Figure 4-6, we quantify the effects of load factors on the average delay experienced by passengers disrupted by: 1) flight cancellations; and 2) missed connections. We observe that for low load factors, the cause of disruption does not impact the average delays experienced by passengers, and most disrupted passengers are re-assigned to their best itineraries. Passengers disrupted because of a flight cancellation become increasingly more difficult to re-accommodate as load factors increase, for the reasons previously discussed in Section 4.3.3. Load factors, however, do not greatly affect the airline's ability to re-accommodate passengers who miss their connections due to a delayed flight.

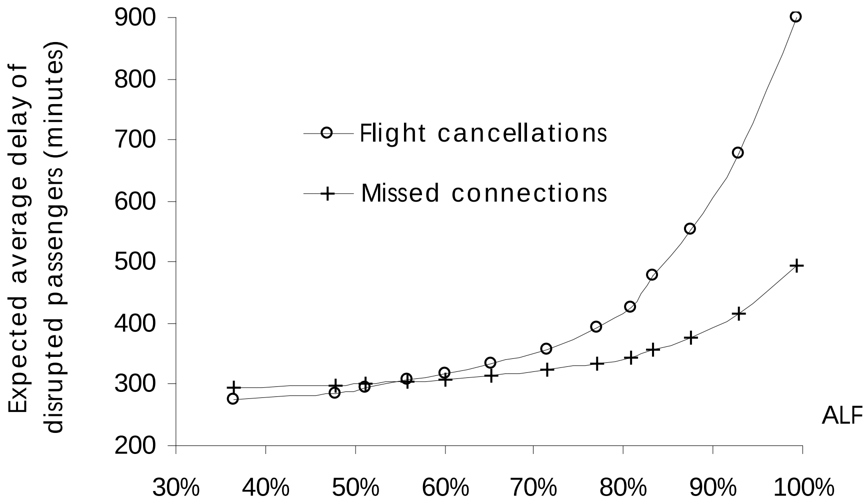

Figure 4-6: Average delay of passengers disrupted because of flight cancellations or missed connections versus average load factor

In Figure 4-7, we contrast the percentages of passengers re-accommodated on their best recovery itineraries for varying load factors when disruptions are caused by flight cancellations and missed connections.

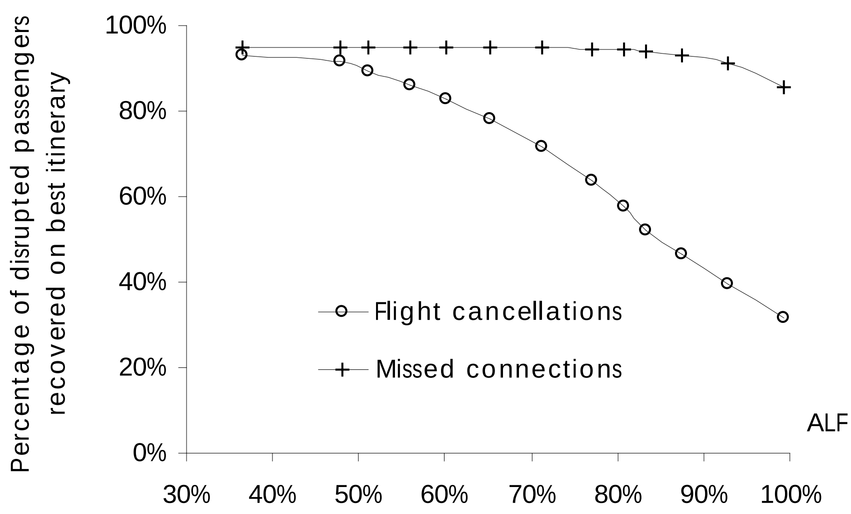

Figure 4-7: Percentage of disrupted passengers re-assigned to their best recovery itinerary versus average load factor

4.3.6 Disruption intensity

In this section, we present an analysis of the combined effects of higher levels of flight schedule disruption and average load factor on the average delay of disrupted passengers.

We sort the days of August 2000 into two lists, *L1* and *L2*, one list sorted in increasing order of the percentage of flights severely delayed and the second list in increasing order of the daily flight cancellation rate. We assign a weight of *1* to each day in the top third of the lists, a weight of *2* to each day in the middle third and a weight of *3* to each day in the bottom third of lists *L1* and *L2*. We then average the two weights assigned to each day and sort the days in increasing order of their average weight. The days in the first third of this sorted list form *LD*, the set of days with *low levels of flight schedule disruptions*. The days in the next third of the sorted list form *MD*, the set of days with *moderate levels of flight schedule disruptions,* and the remaining days in the sorted list form the *days with high levels of flight schedule disruptions* (*HD*).

Using *PDC*, we find that the expected delay of disrupted passengers for days in *HD* is double that for days in *LD*. As depicted in Figure 4-8, for low average load factors, enough seats remain unoccupied to re-accommodate the relatively low number of disrupted passengers.

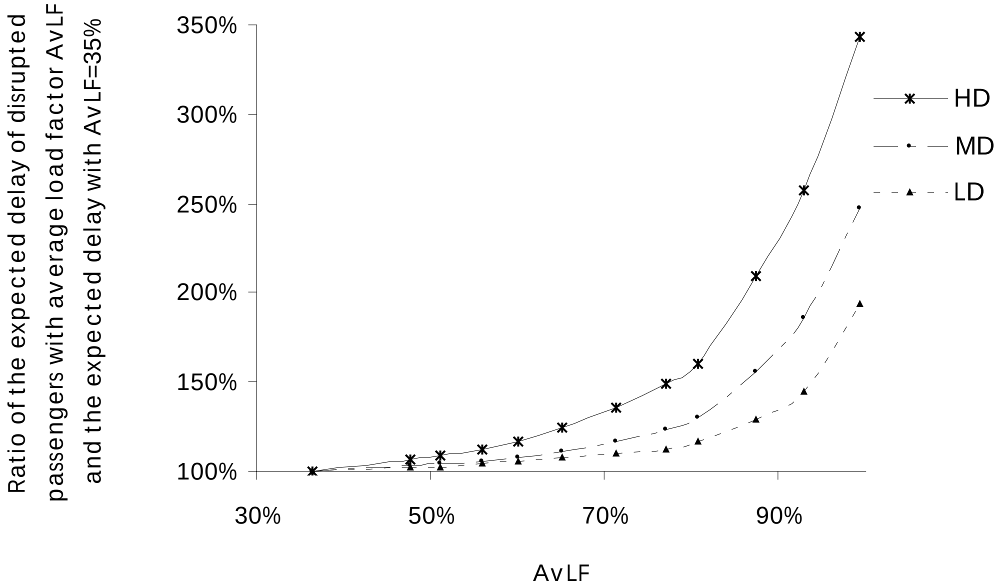

Figure 4-8: Average delay of the disrupted passengers for days with different disruption levels

As the average load factor increases, fewer seats are available to re-accommodate a growing number of disrupted passengers and the difference between delays for days in *HD* and *LD* widens. Note that the difference between *HD* and *LD* grows more sharply for AvLFs greater than about 75%, which is close to the average Load factor for 2000. Trends of increasing load factor (there were more days in which the load factor was above 85% in 2000 than in 1995), coupled with increasing numbers of flight schedule disruptions, amplified the delays experienced by passengers in the late 90's.

It is often the case in the service industry that quality of service deteriorates with increases in demand for service, with fewer resources available to correct service failures. The airline industry is not an exception and its ability to recover disrupted passengers deteriorates exponentially when the average load factor increases. This problem is exacerbated by increasing competition from low cost airlines in the domestic US market, causing major carriers to strive for even higher average load factors to breakeven. Passenger traffic in 2006 is forecasted to be about that of 2000, [FAA03], with similar system capacity. Hence, effectively managing passenger delays and schedule disruptions, critical to maintaining passenger loyalty and financial success, will be a major challenge for the airlines.

5 Alternative flight-based delay metrics

We suggest the following two alternative flight-based metrics, both designed to reflect the relative risk of passenger disruption and to allow ranking of airlines by schedule reliability:

- The percentage of operated flights that are delayed by more than 45 minutes (*45FD*) and,
- The percentage of flights that are canceled.

Like the *15OTP* metric, our metrics rely only on readily available flight delay and cancellation information from the ASQP database.

5.1 Percentage of operated flights delayed by more than 45 minutes

In Table 5-1, we observe a clear shift in the distribution of delayed flights from 1995 to 2000 (source: ASQP database).

	Flight delay window (minutes)		
Year	$\in$]0;15]	$\in$]15;45]	$>$45
1995	64.2%	25.4%	10.4%
2000	53.8%	28.0%	18.2%

Table 5-1: Trends in flight delay distribution

The percentage of flight legs with delays greater than 45 minutes increases from 10.4% in 1995 to 18.2% in 2000. This distribution shift from shorter to longer flight delays has a clear impact on connecting passengers, with flight legs delayed more than 45 minutes accounting for 45% of the total delay minutes in 1995 and for 61% of total delay in 2000.

Of the passengers with arrival delays in excess of 45 minutes, 68% missed their connections. In contrast, for passengers with arrival delays in excess of 15 minutes, 27% missed their connections. Hence, the number of passengers disrupted In August 2000 due to missed connections is much better correlated with *45FD* ($R^2 = 0.67$) than with 15OTP ($R^2 = -0.33$), with our least squares regression identifying the percentage of connecting passengers disrupted by missed connections as 0.92 multiplied by 45FD.

5.2 Percentage of flight legs canceled

The percentage of flights canceled has also risen dramatically, from 0.6% in 1995 to 3.5% in 2000, as illustrated in Table 5-2 (source: ASQP). The combined effect of longer delays and more cancellations is a substantial rise in the number of disrupted passengers, resulting in increased passenger dissatisfaction and perceptions of degraded service reliability.

	Number of canceled flights	Cancellation rate
1995	23,841	0.6%
2000	148,655	3.5%

Table 5-2: Trends in flight cancellation

The number of passengers disrupted because of flight cancellation is very well correlated with the cancellation rate ($R^2 = 0.98$), with the least square regression equating the percentage of passengers disrupted due to a canceled flight to 1.11 times the cancellation rate.

6 Summary

We summarize the key findings of our analysis as follows:

- In general, flight leg delays are not accurate surrogates of passenger delays for hub-and-spoke airlines. Average flight leg arrival delay can severely underestimate the average arrival delay of passengers. The chief contributor to this difference is passenger disruption, occurring when itineraries are disrupted by delayed flight leg arrivals and cancellations.

- Connecting passengers are almost three times more likely to be disrupted than local passengers. If connecting passengers miss their connections, however, they are often re-accommodated on their best itineraries. Alternatively, only about half of the local passengers, who are always disrupted by flight leg cancellations, are re-accommodated on their best itineraries.

- The inability to re-accommodate disrupted passengers on the best itineraries is exacerbated with increasing load factors, with average disrupted passenger delay growing exponentially with load factors.

- Alternative metrics measuring schedule performance, namely flight cancellation rates and the percentage of flights delayed by more than 45 minutes, are better indicators of passenger disruptions than *15OTP*.

7 Acknowledgments

We are grateful to the Alfred P. Sloan Foundation and the MIT Global Airline Industry Program for supporting our research.

References

[BeS82] Belobaba P., Simpson R. (1982) “Air Transportation Economics”, Massachusetts Institute of Technology Flight Transportation Laboratory, unpublished.

[Brat03] Bratu S., “Airline Passenger On-Time Schedule Reliability: Analysis, Algorithms and Optimization Decision Models,” Massachusetts Institute of Technology, PhD Thesis.

[CBL93] Caulkins J., Barnett A., Larkey P., Yuan Y., Goranson J. (1993) “The On-Time Machines: Some Analysis of Airline Punctuality”, Operations Research Vol. 41(4), pp. 710-720.

[FAA01] Federal Aviation Administration (2001), “Subcommittee on Aviation Hearing on FAA's Capacity Benchmarks.”

[FAA03] Federal Aviation Administration (2003), “FAA Aerospace Forecast for fiscal years 2003-2014.”

[Flin00] Flint P. (Nov. 2000) “Back on schedule”, Air Transport World, pp. 47-51.

[GoT98] Gopalan R., Talluri K. (1998) “Mathematical Models in Airline Schedule Planning: A Survey,” Annals of Operations Research 76, pp. 155-185.

[Hal99] Hall, W. (1999) "Efficient Capacity Allocation in a Collaborative Air Transportation System," Massachusetts Institute of Technology, PhD Thesis.

[HLS94] Heskett J., Jones T., Loveman G., Sasser E., Schlesinger L. (1994) "Putting the Service-Profit Chain to Work," Harvard Business Review March-April.

[Jan02] Janusewski S. (2002), “The Effect of Air Traffic Delays on Airline Prices”, Working Paper, Massachusetts Institute of Technology Department of Economics.

[Mea01] Mead K. (2001) “Flight Delays and Cancellations Continue as Major Sources of Customer Dissatisfaction,” Statement of the Inspector General, Department of Transportation.

[Mit01] Mitra T. (2001) “Measuring the causes of airline customer dissatisfaction,” S.M thesis, Massachusetts Institute of Technology, Dept. of Aeronautics and Astronautics.

[Shu93] Shumsky R. (1993) "The response of the U.S. air carriers to the DOT's On-Time Disclosure Rule," Massachusetts Institute of Technology SM Thesis.