How to Please Most of the People Most of the Time

Matt Davison
Dave Diller
David Jessup
University of Dayton
Dayton, OH 45469

Advisor: R.C. Steinlage

Summary

We consider the dissatisfaction with delay to which flights are subjected. We develop a function to represent a combination of these dissatisfactions from the airline's and the passengers' points of view. We then use this function to rate five different systems of ordering takeoffs, by simulating a day at an airport and calculating the total dissatisfaction for all departing flights.

These same systems are also analyzed separately over the same simulated day, on the basis of the number of passengers who miss connecting flights due to delays in takeoff, the number who get fed up with the length of delay, and the distribution of passenger delays.

The dissatisfaction ratings give the same rank ordering as the analysis does. Further, the relative difference between dissatisfactions appears to reflect the amount by which one system is better than another. Most importantly, the best system turns out to be the system that tries to minimize the increase in dissatisfaction between one takeoff and the next.

1. Restatement of the Problem

At an airport, when an aircraft leaves the gate ("pushes back"), it calls the control tower and is added to the list of flights waiting to take-off. We interpret the problem to be to develop a model of airline and passenger dissatisfaction which can be used to evaluate the order in which aircraft take off. We assume that the control tower has a fast online database with access to the following information about each aircraft:

- Scheduled pushback time
- Actual pushback time
- Number of passengers aboard
- Number of passengers aboard who need to make a connection, as well as the time of that connection
- Scheduled time of arrival at the next stop.

2. Assumptions

1. There are seven sizes of aircraft, which range in seating capacity from 100 to 400 in steps of 50.
2. We can use a database such as the one described above.
3. The Dayton International Airport provides an adequate testing ground for systems ordering aircraft takeoff.
4. No airline is to be given special preference over any other airline or passenger.
5. The aircraft may take off in any order we see fit (i.e., the aircraft are waiting side by side, and any aircraft could use the runway next).
6. Airline dissatisfaction is a result of passenger dissatisfaction.
7. The following are independent for each flight: trip length, aircraft size, lag between scheduled pushback and actual pushback, and the percentage of seats occupied on the flight.
8. The actual air time required for a flight is the difference between scheduled pushback and arrival times.
9. There are only two types of passengers, and the dissatisfaction of all passengers of the same type is calculated the same way (passengers of a given type are all generic).

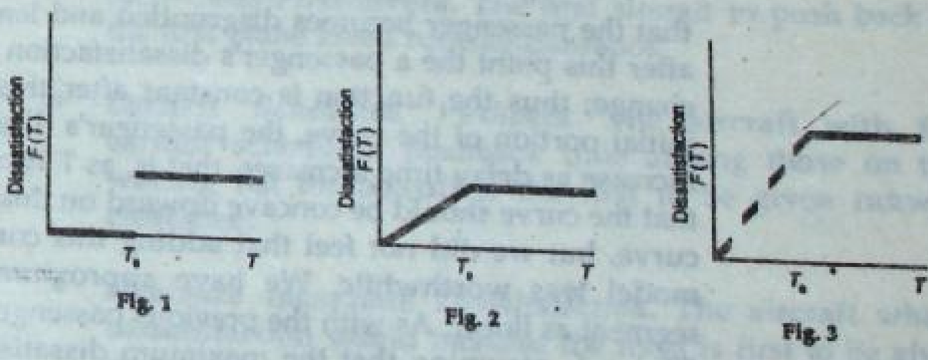
3. Justification of Our Approach

Airports are supported by airlines. We assume that the goal of an airport is to satisfy the airlines that it serves. The airlines' dissatisfaction is assumed to be proportional to their passengers' dissatisfaction. The reason for this assumption is that an approximately constant fraction of dissatisfied passengers will file complaints or quit using the airline. We assume that the airline has no other source of dissatisfaction besides the dissatisfaction of its passengers.

A passenger's dissatisfaction f depends on the amount of delay d and the actual air time t. It seems likely that f is dependent on the ratio of the two, that is, $f(d,t) = f(T)$, where $T = d/t$. Henceforth T will be called *travel-adjusted delay*. This assumption implies that a passenger will be equally dissatisfied with a one-hour wait for a two-hour flight as with a six-hour wait for a twelve-hour flight.

We break passengers into two general groups. The first are trying to reach connecting flights at another airport (Figure 1); the second do not need to catch a connecting flight (Figure 2).

The total dissatisfaction of an aircraft is then the sum of the dissatisfactions of all the passengers on board (Figure 3).



Figures 1–3. Dissatisfaction of a passenger needing a connection, of one not needing a connection, and of the aircraft as a whole.

4. The Model

4.1. Dissatisfaction of a passenger needing a connection (Figure 1)

Since the travel time of an individual is constant, T varies only with time of delay. T_0 is the point at which the time of delay causes the individual to miss a connecting flight. We have assumed that as long as a passenger makes the connecting flight, the passenger is completely satisfied. Once the passenger has missed the connecting flight, an increase in the time of delay does not affect the passenger's dissatisfaction.

It could be argued that the second portion of the curve should slope upward, or that there could be a second discontinuity representing another missed flight that could have taken the place of the connecting flight; but these differences are not reflected in our model. We have arbitrarily set $f(T_0) = 1$ for use in our model. In the end, it is only the differences in dissatisfaction that matter. For that reason it does not matter what $f(T_0)$ is so long as it remains consistent.

4.2. Dissatisfaction of a passenger not needing a connection (Figure 2)

Here T_0 represents the point at which the delay gets so long that the passenger becomes disgruntled and leaves. Of course, after this point the a passenger's dissatisfaction is not going to change; thus the function is constant after this point. On the initial portion of the curve, the passenger's dissatisfaction will increase as delay time increases, that is, as T increases. It seems that the curve should be concave upward on this segment of the curve, but we did not feel that adding this complexity to the model was worthwhile. We have approximated the initial segment as linear. As with the previous passenger group, we set $f(T_0) = 1$, assuming that the maximum dissatisfaction of both types of passengers is the same. For purposes of evaluation, we set $T_0 = 2$ for passengers not needing to make a connecting flight. For passengers who needed a connecting flight, of course, T_0 would vary from passenger to passenger.

4.3. Total dissatisfaction on an aircraft (Figure 3)

The total dissatisfaction function $F(T)$ is just the sum of the individual dissatisfaction functions $f(T)$ of the passengers aboard the aircraft. Each discontinuity represents one or more of the passengers missing a connecting flight. The initial segment of the curve is piecewise-linear, because the individual dissatisfaction curves are linear. The point T_0 represents the point at which either every passenger has left or the flight is cancelled. It is worth noting that $F(T)$ is not necessarily continuous at T_0. From Figure 2, $T_0 = 2$, and the maximum dissatisfaction is equal to the number of passengers aboard the aircraft.

5. Testing the Model

This model of a dissatisfaction function F is now calculated for "actual" flights, that is, all the flights of one computer-simulated day. During the simulation, dissatisfaction totals are kept for several different takeoff system algorithms.

These same takeoff systems are also evaluated by using some pertinent statistics generated during the simulated day, from which we can determine if the dissatisfaction model is giving us lower dissatisfactions for systems which we find more satisfactory.

The following takeoff systems are considered:

- *First-Come-First-Served*. The first aircraft to push back is the first plane given runway clearance.
- *Earliest Scheduled Pushback*. The aircraft with the earliest scheduled pushback time among those on the waiting list for takeoff is the first to be given runway clearance.
- *Minimize Immediate Dissatisfaction*. The aircraft whose dissatisfaction would increase the most is first to be given runway clearance.
- *Three-Deep, Best*. The three aircraft with the earliest scheduled pushback times are considered. All possible sequences of permitting these three aircraft to take off are considered; the sequence that minimizes total

dissatisfaction is chosen. The first aircraft in this sequence is given runway clearance, the other two aircraft are returned to the waiting list; and the process is repeated.

- *Three-Deep, Worst.* This system is identical to Three-Deep, Best, except that the sequence chosen maximizes dissatisfaction. This choice is not being considered in the hope that it would be useful in practice! Rather, it is intended to help evaluate our model of dissatisfaction by determining whether a large dissatisfaction really corresponds to poor service.

To develop a simulated airport day, we generated all of the information that the database would contain during that day. To obtain scheduled pushback times, arrival times, and destinations of the flights (which we shall use later in this development), we entered the Dayton International Airport's daily schedule [Wood 1989] into a personal computer.

Figure 4 represents our feeling of what the probability density function of having left by time t should look like. The characteristic negative-exponential shape suggests a truncated normal distribution, in which the probability of leaving early is zero. We use a binomial distribution to approximate the normal distribution, since we need to know the actual departure time only to the nearest minute.

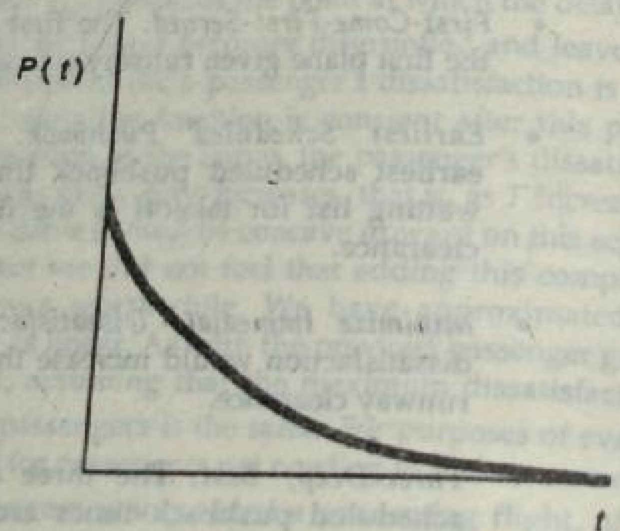


Figure 4. Probability density function for late pushbacks.

Once a plane reaches its scheduled pushback time, its probability of pushing back is $p = .17$ every minute. That is, the probability that an aircraft pushes back in a given minute, given that it has not yet pushed back and that t is greater than the scheduled pushback time, is 17%. Consequently, the probability of having pushed back within 15 min of schedule is 94%, which is in keeping with the information that about 92% of all planes push back within 15 min of schedule [Piedmont Information Desk 1989].

In order to prevent any fouling of the simulation, we assumed that if a plane has not pushed back within 24 min after its scheduled departure time, it then pushes back (24 min is the most that an aircraft can delay pushback). Since the probability of pushing back within 24 min of schedule is over 99%, this assumption should not have any significant effect on our results.

To obtain the number of passengers on each flight, we first need to determine the aircraft's carrying capacity.

The histogram in **Figure 5** is for frequency of occurrence vs. aircraft size. Our sample space consists of approximately every 15th aircraft listed in the Delta System Timetable [Delta Airlines 1989, 412-417]. The aircraft tend to fall into seven size categories, which fact helps to justify the assumption that the airport deals with only seven sizes of aircraft. We use **Table 1** to translate real aircraft sizes to the model aircraft sizes.

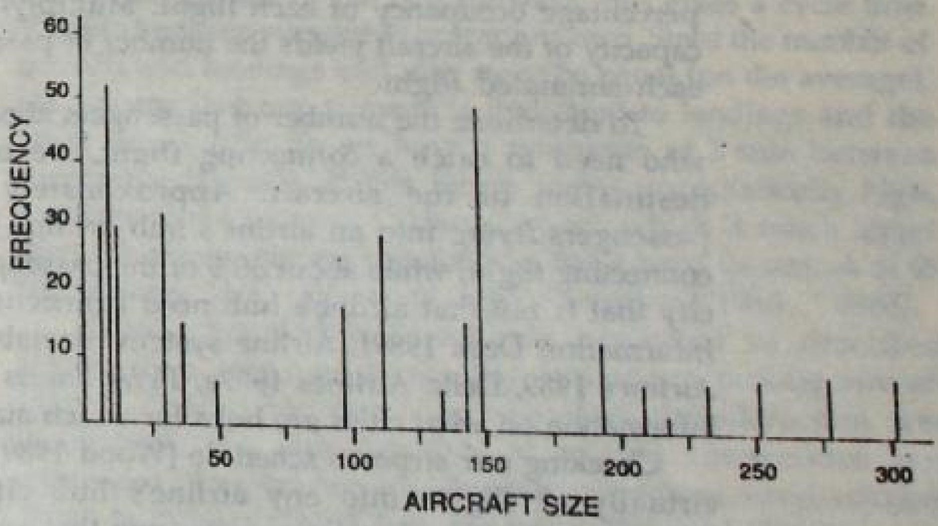


Figure 5. Frequency vs. aircraft capacity.

Table 1.
Conversion from actual aircraft size to model size.

Actual Aircraft Size	Size in Model	Percentage
15–20	100	32.4
30–50	150	18.0
95–115	200	13.6
125–150	250	20.4
180–215	300	6.2
230–260	350	4.4
270 +	400	5.0

The percentages in **Table 1** yield a probability distribution for aircraft size, which we use in generating the aircraft size for each individual flight. These percentages also tend to agree with the information that approximately half of all flights are "small" flights [Piedmont Information Desk 1989].

To determine the number of passengers on each plane, all that remains is to determine the percentage of seats occupied on each flight. The actual range of 35% to 95% occupancy [Piedmont Information 1989] suggests using a distribution with a mean of 65% and a standard deviation of 10%. A normal distribution with these parameters was used to determine the percentage occupancy of each flight. Multiplying this by the capacity of the aircraft yields the number of passengers aboard each simulated flight.

To determine the number of passengers aboard each flight who need to catch a connecting flight, we considered the destination of the aircraft. Approximately 70% of the passengers flying into an airline's hub on that airline need a connecting flight, while about 30% of the passengers flying to a city that is not that airline's hub need connections [Piedmont Information Desk 1989]. Airline system timetables [American Airlines 1989, Delta Airlines 1989a, *TWA Routes* 1989] provide information on what cities are hubs for which airlines.

Checking our airport's schedule [Wood 1989], we find that virtually all flights into any airline's hub city are on the airline in question. Therefore we assume that any flight into a city that is a hub for some airline is a flight on that airline. The cities designated as hubs are Chicago, Cincinnati, Nashville, and St. Louis. The percentage of passengers on each flight who need to make a connection is assumed to have a normal distribution with a standard deviation of 5% and a mean of 70% if the destination is a hub, or with a mean of 30% if the destination is not a hub.

The concept of a hub city is used again to determine the number of connecting flights that passengers are trying to catch. This quantity is assumed to be variable, since not all passengers have the same final destination. The number of connecting flights is taken to be a Poisson variable with mean 5 if the destination is a hub, or with mean 2 if the destination is not a hub. The passengers are then distributed among the available connecting flights, by giving each passenger who needs a connecting flight an equal chance of needing each of the available connecting flights.

Each available connecting flight is then given a different time at which the passengers needing that connection will miss it. Since most airlines recommend providing at least an hour to make a connection, the distribution is designed to leave at least an hour in most cases. The time by which each connecting flight will be missed is assumed to have a normal distribution with a standard deviation of 20 min and a mean of 80 min after the scheduled arrival time of the feeder flight.

The last piece of information required for the simulation is the length of time that must elapse between successive takeoffs. The airport we studied generally handles a maximum of one flight operation—that is, one takeoff or landing—every 90 sec, using two runways [Piedmont Information Desk 1989]. With alternate use of the two runways, this gives a cycle time of 3 min between successive operations each. Since the number of takeoffs and landings each day must be equal (on the average), we assume that one runway is dedicated to landings and the other to takeoffs. So we have a minimum of 3 min between successive takeoffs. If this figure seems unrealistically high, consider that Chicago's O'Hare Airport, which is much larger than the airport we are considering, has a legal minimum of 45 sec between flight operations in good weather ["FAA..." 1988].

Three airport days, randomly generated as described above, were subjected to the five systems of picking aircraft takeoff order. For each day, the total dissatisfaction was calculated, the total number of missed connections was computed, the total number of passengers whose travel-adjusted delay is greater than 2 hrs was computed, and histograms of frequency of delay time were prepared.

The dissatisfaction ratings were compared in light of the rest of the information for each day and system to determine if lower dissatisfaction ratings actually reflect "better" systems. To assist in the determination of which systems are "better," they were also tested under adverse conditions (which were intended to cause pushback to come to shove): thunderstorms at peak traffic hours, simulated by prohibiting takeoffs 8:15-8:45 A.M. and 5:15-5:45 P.M.

In our testing of the model, we recorded a missed connection for a passenger whenever the difference between scheduled pushback and actual takeoff times was greater than the difference between time of connection and scheduled arrival time. The delay time histograms represent " runway delay time," that is, the difference between actual pushback time and takeoff time. This quantity is used rather than actual delay time because it is felt that in the analysis of the system we should not include any delay that the system cannot prevent.

We present here the results of only one of the test days, both with and without thunderstorms (Table 2) This day was quite representative of the three days. [EDITOR'S NOTE: For space reasons, 10 accompanying histograms are omitted].

6. Results

We present the systems from best to worst, with comments.

1. *Minimize Immediate Dissatisfaction*. This system minimizes by a wide margin the total dissatisfaction, the number of missed connections, and the number of cancellations (i.e., the number of passengers for whom travel-adjusted delay is greater than 2 hrs). Also, this system allows more passengers to take off with less than 5 min runway delay than any other system. The only drawback is that this is the only system for which runway delays of more than 2 hrs occur. This probably happens because a small plane's dissatisfaction is likely to increase less than a large plane's, so we suspect this system is biased in favor of allowing larger planes to take off first.

2. *Three-Deep, Best*. This system allows more passengers to take off with less than 10 min runway delay, while minimizing (significantly over the three remaining systems) the total dissatisfaction, the number of missed connections, and the number of cancellations.

Table 2.
Results of simulations.

	Dissatisfaction	Missed Connections	Cancellations
Min. Immediate Dissatisfaction	2966	66	29
", with thunderstorms	3730	145	164
Three Deep, Best	3870	93	0
", with thunderstorms	4856	329	239
Earliest Flight Scheduled	4356	139	210
", with thunderstorms	5463	505	284
First-Come-First-Served	4381	159	255
", with thunderstorms	5538	577	364
Three Deep, Worst	5282	385	165
", with thunderstorms	6369	746	509

3. *Earliest Flight Schedule.* This algorithm for flight takeoff order beats First-Come-First-Served by a narrow margin on total dissatisfaction, number of missed connections, and number of cancellations, while giving about the same distribution of runway delay time. The consistency with which it is better than First-Come-First-Served suggests not only that this is a better system, but also that the dissatisfaction model is valid even for relatively narrow differences.

4. *First-Come-First-Served.* Analyzing the distribution of runway delay time in light of the fact that this is the system in use today reveals that we have made some faulty assumptions. In the model, more than half of all passengers have a runway delay of over 15 min; since size does not affect takeoff order, more than half of all planes in the model are leaving more than 15 min late. The actual fraction of aircraft that leave within 15 min of schedule is between 66% and 80% ["Airlines..." 1988].

5. *Three-Deep, Worst.* As expected, this system is dead last in all categories. These results help to justify the conclusion that a large total dissatisfaction actually reflects a poor system.

7. Strengths and Weaknesses

One of the most obvious weaknesses of our approach is that at least one of the assumptions we made in developing an airport day is wrong, as pointed out in the analysis of the First-Come-First-Served system.

It is very unlikely that such variables as aircraft capacity and trip length are independent of each other. Also, since percentage occupancy is dependent on the time of the week [Piedmont Information Desk 1989], this variable probably should have a high correlation between different flights.

Another potential problem is that the Dayton International Airport may not be a sufficient testing ground for our model.

The above weaknesses are unlikely to affect the validity of the model, although they could affect an individual airport's choice of a system of permitting aircraft to take off.

Weaknesses that could affect the validity of the dissatisfaction model are:

- Some of the assumptions made about the passengers' dissatisfactions may be inappropriate. Should passengers get fed up sooner? Should the maximum dissatisfaction be the same for both types of passengers?

- The model may ignore certain unusual biases in the systems. In our evaluations, the best system appears to be biased in favor of the larger aircraft.

The strength of our approach is the consistency of the rankings give by the dissatisfaction model with the rankings obtained by our analysis of the systems. Overwhelmingly, the systems with lower dissatisfaction ratings are judged, independently of this fact, to be the better systems. The dissatisfaction rating also appears to reflect how much better one system is than another.

References

"Airlines reach on-time performance record in April." 1988. *Aviation Week and Space Technology*. (June 13, 1988): 146.

American Airlines. 1989. *Domestic and International Timetable. (effective Jan. 31, 1989).*

Dayton International Airport. 1989. *Flight Lines.*

Delta Airlines. 1989a. *System Timetables (effective Feb. 1, 1989).*

Delta Airlines. 1989b. *Quick Reference Schedule (effective Feb. 1, 1989).*

"FAA imposes limit on airline traffic at Chicago O'Hare." 1988 *Aviation Week and Space Technology* (Oct. 10, 1988): 113.

Piedmont Airlines. 1989. *Quick Reference Schedule (effective Jan. 4, 1989).*

Piedmont Information Desk. 1989. Dayton International Airport.

TWA Routes (effective Jan. 31, 1989).

United City Time Table (effective Feb. 15, 1989).

US Air City Timetables. Dayton International Airport (effective Nov. 1, 1988).

Wood, James R. 1989. *Dayton International Airport Flight Guide (effective Feb. 1989).*