Response of U.S. Air Carriers to On-Time Disclosure Rule

ROBERT SHUMSKY

The On-Time Disclosure Rule, implemented by the U.S. Department of Transportation in 1987, makes reports of the on-time performance of the major U.S. carriers available to the public. The purpose of the disclosure rule was to create incentives for the carriers to improve their on-time performance by either reducing the amount of time to complete a flight or lengthening the amount of time scheduled for a flight. The evidence shows that although actual flight times have fluctuated, scheduled flight times have increased significantly since 1988. The largest increases occurred just after the disclosure rule went into effect, and a regression analysis shows that since 1988, on-time performance has been a significant factor in the scheduling decisions of many carriers. Then two scheduling strategies are presented that are designed to improve on-time performance. The strategies are implemented, and their performance is compared with the performance of the carriers' schedules. The comparison highlights the challenges that carriers face when designing schedules with the disclosure rule in mind.

According to a U.S. Department of Transportation (DOT) study in 1987, fewer than half the flights operated by the major carriers in and out of eight major hub airports during 1986–1987 arrived on time, where an on-time flight is defined to be an arrival within 15 min of schedule (M. Langelan, DOT, 1987, unpublished data). In response to the 1987 DOT study and many consumer complaints, DOT adopted regulations in 1987 designed to encourage the carriers to improve the percentage of on-time flights. These regulations included the On-Time Disclosure Rule (OTDR) (*1*).

Rather than establish performance goals, the OTDR makes available to the public information about the carriers' on-time performance. The carriers are now required to submit data on their on-time performance to DOT, and DOT publishes cumulative results in the monthly *Air Travel Consumer Report*. The statistics are quoted in the media (*2*), carriers who perform well cite the results in their advertising (*3*,p.15), and the computer reservation systems used by most travel agents display summaries of the statistics. DOT's goal was to allow the marketplace to pressure the carriers to improve their on-time performance. Whether the *Air Travel Consumer Report* contains the most appropriate statistics, and whether consumers make effective use of the information published in the report, has recently been questioned by Cunningham and Brand (*4*) and by consumer advocates (*5*). This paper will focus on the behavior of the carriers under the OTDR.

According to DOT, by March 1988 more than 81 percent of flights flown by the major U.S. carriers were arriving on time, indicating a marked improvement over the results of DOT's 1987 study (*6*). This improvement might be attributable to factors beyond the carriers' control, such as reduced congestion in the national airspace system, but major carriers may have improved their on-time performance by operating their flights in a more timely manner or by changing their posted schedules to more accurately reflect the lengths of their flights. In this paper we investigate whether the carriers implemented these strategies, gauge the success of their efforts, and evaluate the effectiveness of the OTDR itself.

Operations Research Center, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

DATA

Under the OTDR, the major U.S. passenger carriers submit performance data on all domestic flights to DOT. A flight is a scheduled entity that flies regularly between a pair of airports (i.e., USAir Flight 427 from Boston to Washington National). An operation is an instance of a flight on any given day. In the original data, a flight generates between 1 and 31 operations each month. The performance data for each operation include the destination and arrival airport codes, the carrier code, the scheduled departure and arrival times, and the actual departure and arrival times. One month of data contains information on approximately 150,000 operations representing about 14,000 flights. From this information one can calculate the scheduled and actual transit times for each operation. Scheduled transit time is the time from the scheduled departure from the gate to the scheduled arrival at the destination gate; actual transit time is the time from the scheduled departure to the actual arrival. Actual transit time includes time spent holding at the origination gate, taxiing, queueing at the head of the runway, flying to the destination, and taxiing to the destination gate. A flight is on time if its actual transit time is less than 15 min longer than its scheduled transit time. A flight is late if it is not on time.

We obtained DOT data for March of each year from 1988 to 1991. This was more information than could be processed, so we extracted from the 1988 data the operations of a random 5 percent sample of the flights. We then searched the data in subsequent years for flights considered to be equivalent to flights in the 1988 sample.

Specifically, the 1988 sample was selected by sorting all 1988 flights by departure airport code, arrival airport code, and flight number. We then chose every 20th flight that operated at least five times a week. We ignored all Eastern Airlines flights, because Eastern was crippled by a strike during the time covered by our data. This method was simple and easily reproducible, and it resulted in a random sample of 698 flights. The number of flights was large enough to ensure reasonably small standard errors for the estimates that will be described in the next two sections.

We then searched the March 1989, 1990, and 1991 data for flights that were equivalent to flights from the 1988 sample. To match a 1988 flight, a flight in a later year must be operated by the same carrier, flown between the same airports, and scheduled to depart within ½ hr of the original 1988 departure time. For example, Delta's Flight 1719 was scheduled to leave Los Angeles (LAX) for Seattle (SEA) at 6:55 a.m. in 1988. A matching Delta flight was found in 1989 that left LAX for SEA at 6:30 a.m. No matching flights were found in 1990 or 1991, so this flight appears for 2 years in our sample. Of the original sample of 698 flights, 296 appeared in all 4 years. Of these, the schedules of 219 were found in the *Official Airline Guide* for March 1987 (*7*). We will refer to these 219 flights as the "5-year sample." Because the OTDR was instituted in September 1987, the schedules of the 5-year sample allow the comparison of scheduling practices before and after the passage of the legislation.

TRENDS IN TRANSIT TIMES

Transit Time Changes in 5-Year Sample

Because the 5-year sample contains the same set of flights in every year, the mean scheduled and actual transit times for each year may be compared. The middle line in Figure 1 shows a strong upward trend in scheduled transit times from 1987 to 1991, with the largest increase, one of 4.8 min, occurring between 1987 and 1988, when the OTDR was implemented. Note that only schedule times, and no performance data, were available for 1987. Actual transit times were consistently higher than scheduled transit times throughout the period; the average gap between actual and scheduled times peaked at 11.8 min in 1989. From this observation, one would expect that on-time performance suffered in 1989 and improved as scheduled times lengthened and actual times dropped. The lowest line in Figure 1, which displays the on-time performance of the 5-year sample, confirms this expectation.

To quantify the degree to which the changes in scheduled transit times affected on-time performance, we calculated the fraction of operations that would have arrived on time had the original 1987 schedule not been altered. Figure 2 shows that the fraction of on-time operations would have been 0.5 in 1990, rather than the 0.73 achieved with the schedule changes.

The original sample of 698 flights was a random sample of all 1988 flights, so its characteristics mirror the characteristics of the parent population. Therefore, statistics derived from the original sample will not be biased with respect to the parent population. The same cannot be said about the 5-year sample, because many of its characteristics differ significantly from the original sample, and therefore from the parent population. For example, the original sample and the 5-year sample were significantly different in the percentage of flights allocated to each carrier. TWA flights accounted for 6 percent of the original sample and only 1 percent of the 5-year sample. Other statistically significant differences between the two

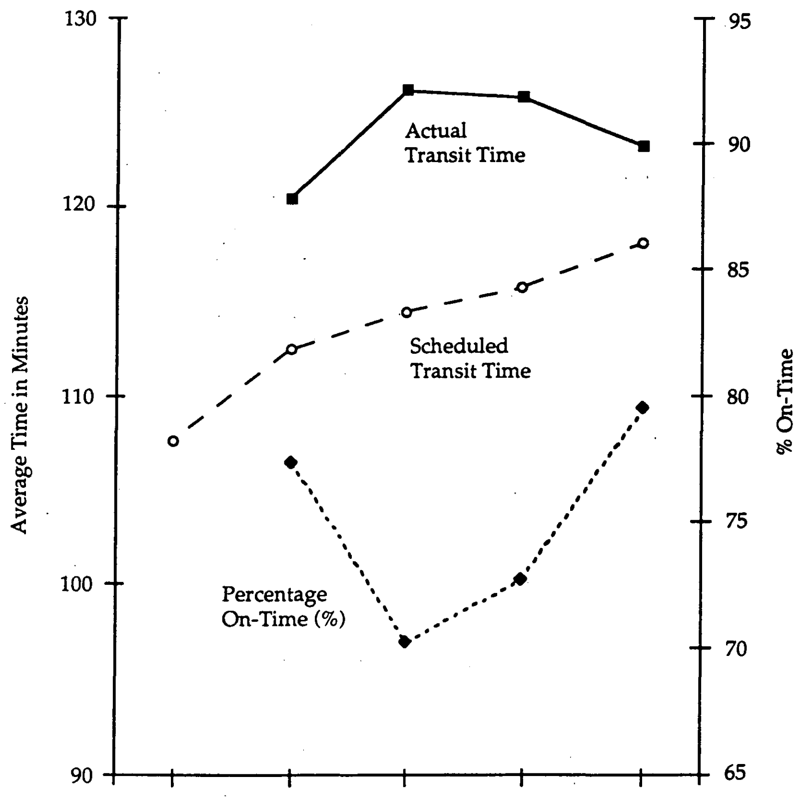

FIGURE 1 Transit times and on-time performance in 5-year sample.

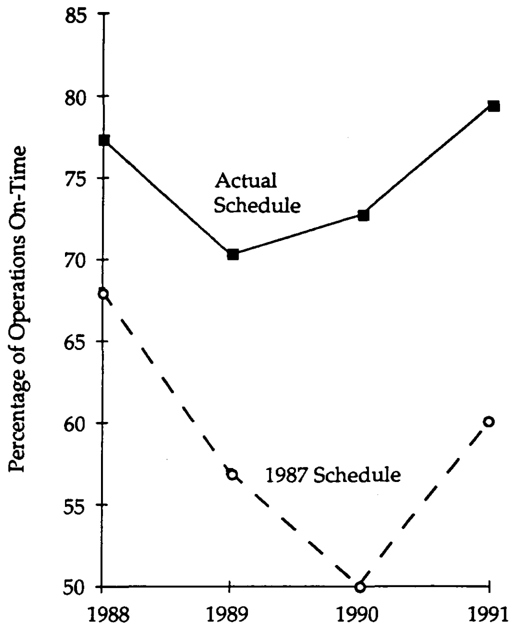

FIGURE 2 Percentage of operations on time under actual and 1987 schedules.

samples included a difference in the types of airports served: the 5-year sample included a higher percentage of flights serving the larger, busier airports.

Transit Time Changes for All Flights

Because the compositions of the 5-year sample and the original random sample have significant differences, we investigated whether the flights left out of the 5-year sample also display increases in scheduled transit times. The subsample of those flights that are not 5-year flights will be called the intermittent sample, because the flights appear intermittently from 1987 to 1991. A reasonable method for measuring schedule changes in 1989, 1990, and 1991 for the intermittent sample is to calculate the mean change in scheduled transit times from the "baseline" transit time established for each flight in 1988. Note that the mean change for 1989 is the mean over a group of flights, which differs from the groups averaged in 1990 and 1991 because some flights "disappear" in 1989 and reappear in 1990 or 1991 and others appear in 1989 but disappear from the sample in a later year. The following table displays these means, as well as the mean change from the 1988 baseline for the 5-year flights and the p-values:

	1988–1989	*1988–1990*	*1988–1991*
5-year sample (min)	2.0	3.3	5.7
Intermittent sample (min)	1.6	3.0	4.2
p-value	0.53	0.71	0.04

These schedule changes were approximately normally distributed, and t-tests were performed on the hypothesis that the means of each of the underlying populations were equal. We have no reason to reject the hypothesis that the mean scheduling increases in 1989 and 1990 were the same in the two samples. Although the difference in 1991 is statistically significant at a .05 level, the table indicates that the scheduled transit times in the intermittent sample have grown at a rate similar to that of the 5-year sample. This implies that the overall trends observed in the 5-year sample may also be observed in the population of flights as a whole.

SCHEDULE CHANGES OF INDIVIDUAL CARRIERS

In the previous section we saw that the carriers as a group have consistently increased the scheduled transit times of flights between 1987 and 1991. The chance to improve on-time performance may motivate carriers to lengthen their scheduled flight times, but other factors restrain the carriers from adding slack. Carriers compete for customers, who prefer short flight times. In addition, contracts for many carriers specify that the flight crew be compensated for each operation on the basis of a maximum of the scheduled and actual transit times. The total time allocated to all flights may be viewed as a scarce resource and the allocation of those minutes among flights as a constrained optimization problem. In this section we investigate how individual carriers solved this problem.

Adjustments in Scheduled Transit Times

Within the 5-year sample, there are large differences between each carrier's schedule adjustments. Table 1 displays the adjustments by year and carrier. The scheduled transit times of American Airlines flights in the 5-year sample increased an average of 17 min between 1987 and 1991, but those of United Airlines increased by an average of just 5 min. American experienced its largest mean increase, of 12 min, in the first year.

It may be that American anticipated or reacted to the OTDR in a more timely manner than its competitors. On the other hand, American's scheduled transit times in 1987 may have been shorter than the transit times of the other carriers, so that American was compelled to add a greater amount of time to its schedule in order to achieve reasonable on-time performance. To evaluate the latter hypothesis, we found the scheduled transit times of the major carriers along 16 popular routes in the March 1987 *Official Airline Guide* (7). On average, American's scheduled times were 7 min shorter than the times of each of the other carriers. Before the OTDR, American underrepresented the transit times of those flights to a greater extent than its competitors.

TABLE 1 Average Change in Scheduled Transit Times by Carrier

	Average Change (min.)				
Year	AA	CO	DL	NW	UA
1987-88	12	5	4	2	1
1988-89	-1	5	4	7	2
1989-90	3	-1	0	1	1
1990-91	3	4	3	-2	1

AA=American, CO=Continental, DL=Delta, NW=Northwest, UA=United

Schedule Changes and On-Time Performance

A carrier's decision to lengthen the scheduled transit time of a particular flight may depend on myriad factors, including competitive pressures, the departure times of connecting flights, and on-time performance history. If on-time performance does enter the decision, we would expect a positive correlation between the fraction of operations that were late in a year and the number of minutes added to that flight's schedule in the next year. Figure 3 displays such a relationship for Continental flights that operated in both 1988 and 1989. Each point represents one such flight, and the point is plotted with the fraction of operations late in 1988 as the abscissa and the subsequent schedule increase as the ordinate. For many of these flights, poor on-time performance led to longer scheduled transit times.

If we assume that the relationship between on-time performance in one year and scheduling adjustments in the next is linear, we may model each carrier's behavior with the following equation:

$$y_i = c + \beta_d x_{di} + \varepsilon_i \qquad (1)$$

where

- y_i = change in minutes of scheduled transit time of Flight i,
- c = change in scheduled transit time applied to all flights,
- β_d = constant for value of (schedule change in minutes)/(percentage late),
- x_{di} = percentage of operations of Flight i late in previous year, and
- ε_i = change in scheduled transit time of Flight i not explained by other variables.

For each carrier we performed a least-squares fit for flights that operated in both 1988 and 1989 in order to obtain the estimates $\hat{c}$ and $\hat{\beta}_d$ that are listed in the first two rows of Table 2. The numbers in parentheses are the standard deviations of the estimates around their true values if we assume normality of the error term ε_i. According to these results, if a Continental flight experienced no late operations in 1988, we would expect its scheduled transit time to decrease by about 4 min in the next year. On the other hand, its schedule would increase by (0.2 min)/(1 percent late), or about 12 sec, for each percentage point of operations late in the previous year.

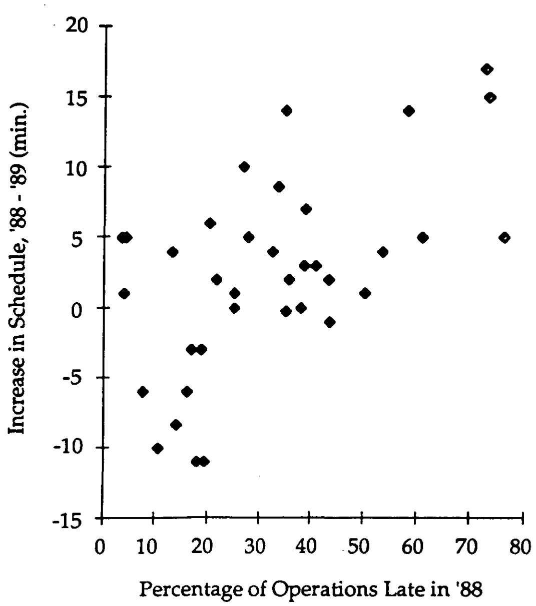

FIGURE 3 Percentage late versus schedule increase for Continental flights.

To test hypotheses about the coefficient estimates in Table 2, we must assume the normality and constant variance of ε_i. We tested these assumptions by examining standard regression diagnostics, such as normal probability plots of the residuals. For all the sample carriers except United, for which there were only 33 sample flights, we found no reason to reject the assumptions. We may therefore construct confidence intervals for our parameters. For example, the true β_d for Continental is within the region $(\hat{\beta}_d - 2\sigma_{\beta_d}, \hat{\beta}_d + 2\sigma_{\beta_d}) = (0.1, 0.3)$ with a probability of approximately 0.95.

Table 2 also contains the R^2-statistic, sometimes called the coefficient of determination. R^2 measures the proportional reduction of the variation in y_i when the model is used. Many of the values in Table 2 are small, indicating that the linear model does not greatly reduce the variation in y_i. This is not unexpected, since it would be difficult to imagine that the carriers rely on a formula such as Equation 1 to construct their schedules.

Despite the low R^2 scores, the estimates of the model parameters allow us to examine the strength of the relationship between on-time performance and the schedule changes for each carrier. In addition, we may use R^2 to test the hypothesis that these two variables are independent. We assume that y_i and x_{di} are jointly normally distributed and test the following hypothesis:

H_0: y_i and x_{di} are independent

H_a: y_i and x_{di} are not independent

If n is the number of flights in our sample and R is the square root of R^2, then under H_0, the statistic

$$t^* = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}} \qquad (2)$$

follows a student-t distribution with $n - 2$ degrees of freedom (*8*). We used t^* to test H_0 for each of the carriers, and the p-values are shown in Table 2. H_0 may be rejected for all airlines except American with a 0.05 level of confidence, and even for American the p-value is only 0.07. Note that this test is equivalent to testing the hypothesis that the constant β_d is equal to 0 for each carrier. Taken together, these results supply strong evidence that the schedule changes are correlated with on-time performance in the previous year.

Schedule Changes and Marginal Gain

Besides the relationship demonstrated in the last section, specific scheduling practices with respect to on-time performance may vary among carriers, so any detailed analysis of these

TABLE 2 Least-Squares Estimates and Diagnostics for Schedule Change Model, 1988–1989

	AA	CO	DL	NW	UA
$\hat{c}(\sigma_c)$	-2 (0.9)	-4 (1.8)	0 (0.9)	4 (1.2)	-1 (1.9)
$\hat{\beta}_d(\sigma_{\beta_d})$	.08 (.05)	.20 (.05)	.14 (.03)	.11 (.04)	.09 (.07)
R^2	0.04	0.33	0.19	0.14	0.08
p-value	0.069	0.001	0.001	0.007	0.041

practices should focus on each carrier as a distinct decision maker. An investigation of the practices at American Airlines may be particularly interesting because American is known for its sophisticated decision-support techniques.

For example, in 1991 American reduced the scheduled transit time of Flight 330 by 4 min from the previous year, even though 69 percent of the operations of this flight were late in 1990. The reason for this decision may be found in the actual transit times of each operation of the flight: 27 percent of the operations arrived more than 35 min after the scheduled arrival time, so a substantial increase in scheduled transit times would not have improved on-time performance. Even if the scheduled transit times had been 10 min shorter, 69 percent of the operations would still have been on time. Using 1990 performance as a guide, American's schedulers may have realized that the gain in on-time performance realized by increasing Flight 330's scheduled transit time was likely to be quite low. To quantify the effect of this factor on American's scheduling decisions, we calculate the marginal gain for a particular flight by finding the percentage of operations that would have been on time had 10 min been added to the schedule, subtracting the percentage that would have been on time had 10 min been subtracted from the schedule, and dividing by 20. In Figure 4, 40 percent of the operations of American's Flight 25 arrived on-time (less than or equal to 14 min after schedule). If the schedule had been 10 min longer, 60 percent would have arrived on-time and if the schedule had been 10 min shorter, 8 percent would have been on-time. The marginal gain, then, is (60 − 8)/20 = 2.6 percent/min.

Again, we will approximate the relationship between performance in one year and schedule changes in the next with a linear function. For this model, American's schedule changes will depend on both on-time performance and marginal gain in the previous year:

$$y_i = c + \beta_d x_{di} + \beta_g x_{gi} + \varepsilon_i \qquad (3)$$

where β_g is a constant for the value (schedule change in minutes)/[marginal gain (%)], and x_{gi} is the marginal gain of Flight i in the previous year.

We obtained estimates $\hat{c}$, $\hat{\beta}_d$, and $\hat{\beta}_g$ by performing a least-squares fit on three sets of data: flights operated by American in both 1988 and 1989, in both 1989 and 1990, and in both 1990 and 1991. The results are shown in Table 3, where the numbers in parentheses are the standard deviations of the estimates around the true values if we assume the normality of the error term ε_i. If the marginal gains for two flights in 1988 were 0 percent/20 and 15 percent/20, and if all other factors were equal, we would expect American to increase the scheduled transit time of the second flight by $\hat{\beta}_g$ (15 − 0) = (2.8)(15/20) = 2.1 min more than the first flight.

As we did for our first linear model, we used the R^2-statistic to test whether schedule changes are independent of the other factors in the linear model. Specifically, we test the following hypothesis:

H_0: y_i is independent of x_{di} and x_{gi}

H_a: y_i is not independent of x_{di} and x_{gi}

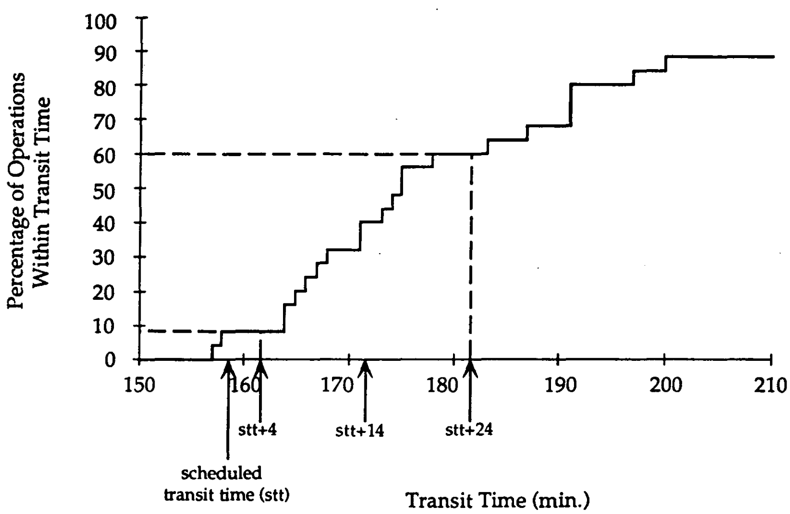

FIGURE 4 Marginal gain of American Flight 25 in 1990.

TABLE 3 Least-Squares Estimates for Schedule Change Model with Marginal Gain

	1988-89	1989-90	1990-91
$\hat{c}(\sigma_c)$	-3.9 (1.0)	-5.4 (2.2)	-0.6 (1.9)
$\hat{\beta}_d(\sigma_{\beta_d})$	0.00 (0.05)	0.11 (0.06)	0.03 (0.05)
$\hat{\beta}_g(\sigma_{\beta_g})$	2.8 (0.8)	3.0 (1.1)	2.3 (1.3)
R^2	0.18	0.20	0.09
p-value	0.0005	0.0008	0.08

When H_0 is true, and when y_i, x_{di}; and x_{gi} are distributed as multivariate normal random variables, then the statistic

$$F^* = \frac{R^2}{1 - R_2}\left(\frac{n-3}{2}\right) \quad (4)$$

follows an F-distribution with 2 and $n - 3$ degrees of freedom (*8*). Table 3 displays p-values from the tests with this statistic. We may reject the null hypothesis of independence for 2 out of the 3 years with a .05 level of confidence, and the test for the third year obtains a p-value of 0.08.

Note that the inclusion of marginal gain among those factors influencing schedule decisions is an artifice of DOT's definition of being on time as being within 15 min of schedule. If DOT reported the mean number of minutes off schedule for all of a carrier's flights, rather than the percentage more than 15 min late, then any increase in scheduled transit time for any flight would improve the carrier's on-time performance. All flights would have the same marginal gain.

In this analysis we have focused on American and have seen that both the mean and the distribution of actual transit times may influence changes in scheduled transit times. Flights with a large amount of variability in actual transit times or large "right tails" in their histograms offer little marginal gain; these flights require a large investment of schedule minutes to achieve small gains in on-time performance. Carriers whose transit times are extremely variable have a difficult choice: either they implement large increases in scheduled transit times, or they accept relatively poor on-time performance.

APPROACHES FOR SCHEDULING TRANSIT TIMES

In the previous section we saw that the decision to alter a scheduled transit time may depend on the on-time performance of a flight as well as the distribution of the flight's actual transit times. This section explores the opportunities and difficulties that confront the carriers when they rely on past performance to predict future transit times. We examine the effectiveness of the carriers' solutions by comparing their performance with the performance of two strategies of our own design.

Fifty-five of the flights operated by American appear in all 4 years of the sample. The average increase in scheduled transit times for these flights between 1990 and 1991 was 3.4 min, so that American distributed a total of 187 extra min among these flights when it constructed its 1991 schedule. In this section we develop and test two specific strategies for distributing such a budget of extra minutes. To determine the effectiveness of the carriers' scheduling decisions, we test our strategies on the 4-year flights of American, Delta, and United and compare the performance of our strategy with the performance of the carriers.

The first strategy is the simplest possible: take the budget and distribute it uniformly among all flights. For example, to generate American's 1991 schedule for its 4-year flights, we add 3.4 min to the 1990 scheduled transit time of each American flight in the 4-year sample. This strategy is called the uniform strategy.

The second strategy will attempt to allocate the budget in a manner that optimizes the on-time performance of the flights in 1991. An obvious method would be to look ahead at actual 1991 transit times and then design the schedule to capture as many flights as possible within the 15-min on-time limit. However, we would then be using more information than was available to the carriers; they did not know how individual flights would perform in the future when they designed the schedule.

On the other hand, we might use only the actual transit times from 1988 through 1990 to estimate the distributions of 1991 transit times and then use the estimated distributions to design the schedule. We would not allow ourselves to use any 1991 data at all, but this may be too much of a handicap, for the carriers would have had some knowledge of general trends in transit times. For example, the mean actual transit times for all 4-year flights were 121.7, 126.5, 127.4, and 125.2 min in 1988, 1989, 1990, and 1991, respectively. The carriers, by forecasting the economic and political situation, may have been able to anticipate the downward trend for 1991.

Our second strategy, then, compromises between total use and no use of 1991 data by first centering the actual transit times from 1988, 1989, and 1990 by 3.5 min, −1.3 min, and −2.2 min, respectively, so that the aggregate mean for each year is equal to the 1991 mean. We then use these centered 1988–1990 actual transit times to estimate the cumulative distribution function of all actual transit times [for more details on the density estimation procedure, see elsewhere (*9*)]. The distribution functions are used to allocate the budget in a way that maximizes the expected number of on-time operations. We formulate this problem as a mathematical program in the following.

In the formulation, the decision variables are x_i, the proposed changes in scheduled transit times for Flights i, $i = 1 \ldots n$ from the 1990 schedule. The objective function is the sum of the functions $g_i(x_i)$, which are the estimated cumulative distribution functions for the flight's actual transit time, shifted along the horizontal axis so that they are plotted with respect to a change in the 1990 scheduled transit time. In other words, $g_i(x_i)$ represents the expected proportion of operations that would arrive on time if a given change x_i in the schedule were implemented.

The first two constraints in the formulation limit the total number of minutes added to the schedule to the budget of minutes set by the carrier (i.e., 187 min for American's 4-year flights). The last two constraints limit each flight's transit time increase and decrease to 24 and 18 min, respectively.

These constraints match the carriers' own behavior and keep the optimization problem to a manageable size.

Maximize

$$\sum_{i=1}^{n} g_i(x_i) \tag{5}$$

Subject to

$$\sum_{i=1}^{n} x_i \leq B \tag{6}$$

$$B = \sum_{i=1}^{n} t_i \tag{7}$$

$$x_i \leq 24 \text{ for all } i \tag{8}$$

$$x_i \geq -18 \text{ for all } i \tag{9}$$

where

x_i = change in scheduled transit time of Flight i,
$g_i(x_i)$ = expected proportion of operations on time if scheduled transit time were adjusted by x_i, and
t_i = carrier's change in scheduled transit time of Flight i.

To solve this problem, the functions $g_i(x_i)$ were approximated with piecewise linear functions so that a mixed-integer program (MIP) formulation could be used. We will say that schedules generated with this formulation were generated under the MIP strategy.

The 1991 schedules for the 4-year flights of American, United, and Delta were generated under our two strategies. We then examined how each of our strategies would have performed in 1991 year by computing the on-time performance of the flights had the alternative schedules been implemented. The results are given in the following table:

	On-Time Performance (%)		
Schedule Source	*American*	*Delta*	*United*
Uniform strategy	86.5	81.9	62.0
MIP strategy	87.9	83.0	60.9
Carrier	86.4	82.6	62.7

The most surprising result from the table is that the uniform strategy performed as well as the MIP strategy and the carriers' own schedules. The MIP strategy relies on past performance to predict transit times, and its failure stems from the variability in actual transit times from year to year. Even though we centered the data so that the overall mean of the 1991 transit times matched the mean from previous years, the mean transit time for each individual flight in 1991 varied widely from its mean in the past. For example, whereas the actual transit times of Delta Flight 961 averaged 95 min from 1988 to 1990, the transit times in 1991 had a mean of 103 min, a difference of 8 min that the MIP strategy did not anticipate. For all three carriers, these differences were approximately normally distributed and had an overall standard deviation of 8.3 min. However, United Airlines flights varied more widely than the flights of the other two carriers: the variations of United's flights had a standard deviation of 10.9 min. It is not surprising, then, that our MIP strategy performed particularly badly for United.

CONCLUSION

Our analysis of the data gathered under the OTDR demonstrates that between 1987 and 1991 the major carriers increased the scheduled length of their flights by an average of about 10 min, or about 10 percent (see Figure 1). From 1988 to 1991, actual transit times rose and then fell. If the carriers had not lengthened their schedules, the percentage of flights arriving on time would have been as much as 20 points lower (see Figure 2).

Of course, the schedule increases may have occurred in the absence of the OTDR. However, in Figure 1 we see that almost half of the increase occurred between 1987 and 1988, at the time the OTDR was implemented. In addition, our analysis of the behavior of individual carriers showed that many of the schedule increases were correlated with on-time performance as defined by the OTDR. Therefore, it is likely that the OTDR influenced the carriers' scheduling decisions.

Our attempts to create a schedule that was optimal with respect to on-time performance demonstrated that carriers with transit times that vary widely face particular scheduling difficulties. Because past performance may not provide reliable information about the future, only large schedule increases guarantee better on-time performance. The OTDR rewards carriers with stable transit times and carriers with unstable times that can afford to make such large increases.

Excellent on-time performance under the OTDR does not necessarily mean that an airline carries passengers efficiently; it indicates only that passengers arrived according to schedule. Because the primary goal of the OTDR is to protect consumers, we have seen that it has been successful in this respect; it appears to have prevented the carriers from developing two schedules: one that is advertised and one that is flown.

ACKNOWLEDGMENTS

This research was financially supported by the Graduate Research Award Program, which is sponsored by the FAA and administered by TRB. The author would like to thank Larry L. Jenney of TRB, Larry Cunningham of the University of Colorado at Denver, and Gerald McDougall of Wichita State University for their comments and encouragement. The suggestions of Marty Langelan of DOT were particularly helpful, and the advice and support of faculty advisors Arnold Barnett and Amedeo Odoni have been invaluable.

REFERENCES

1. U.S. Department of Transportation. *On-Time Disclosure Rule*. 14 C.F.R. Part 234, 1987.
2. The Best (and Worst) Airlines. *Consumer Reports*, Vol. 56, No. 7, July 1991, pp. 462–469.

3. Northwest Airlines paid advertisement. *Time*, Vol. 138, No. 12, Sept. 23, 1991.
4. L. Cunningham and R. Brand. Evaluating the Effectiveness of DOT's *Air Travel Consumer Report*. *Transportation Quarterly*, Vol. 43, No. 4, Oct. 1989, pp. 591–607.
5. E. McKenna. Air Travel Consumer Report Serving Airlines, Not Their Passengers. *Aviation Week and Space Technology*, Vol. 56, No. 7, April 1992, pp. 38–39.
6. *Air Travel Consumer Report*. Office of Consumer Affairs, U.S. Department of Transportation, May 1, 1988.
7. *Official Airline Guide*. Darryl J. Fisher, Oak Brook, Ill., March 1987.
8. J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Regression Models*. Richard D. Irwin, Homewood, Ill., 1983.
9. E. J. Wegman. Nonparametric Probability Density Estimation. *Technometrics*, Vol. 14, No. 3., Aug. 1972, pp. 533–546.