Managerial Uses of Causal Models of Subway On-Time Performance

GARY HENDERSON AND VENGAL DARAPANENI

On-time performance (OTP) indicators reflect the performance of employees and the effectiveness of policies and organizational structure. However in New York City's subway system the operating environment includes factors beyond the control of operating employees, such as mechanical reliability of subway cars, route merges, scheduled headways, construction projects, crowding, passenger behavior, and so on. These factors differ from route to route and complicate comparisons of routes. Planning decisions affecting ridership levels, route design, and the operating environment can benefit from precise estimates of impacts on OTP, especially to help in evaluating the trade-offs involved. The results of a statistical causal model of on-time performance are presented. A set of hypothetically important variables was developed from New York City Transit Authority documents and train movement records. Data for over 54,000 morning rush hour trains traveling from terminals to central business district stations during 1988 and 1990 were used. The model quantifies the effect of variables on the probability that a train will be on time. How the results of such a model can be used to make predictions of performance that control for the operating environment are also shown, allowing performance comparisons between routes with different characteristics. How different OTP goals can be set for different routes is suggested. Finally by converting the results to odds ratios, it is shown how small improvements on routes with OTP of more than 90 percent can provide large benefits from the perspective of riders and how OTP measures obscure that fact.

On-time performance (OTP) indicators reflect employee performance and the effectiveness of managers, operating policies, and organizational structure. However performance levels also reflect the operating environment, for example, the mechanical reliability of subway cars, the frequency of route merges, the spacing between trains (scheduled headways), construction activity, crowding, and passenger behavior. These factors differ from route to route and complicate comparisons of one route with another. Is a 2 percentage point improvement in OTP for a long route with numerous merges equal to the same improvement for a short route with no merges? How can OTP goals be set for different routes?

The Office of the Inspector General (OIG) of the Metropolitan Transportation Authority (MTA) made a statistical causal model of New York City's rush hour subway service to account for differing infrastructures, route designs, and other variables. The model allows performance analysis to control for these factors. Are some routes with low OTPs really doing well given the operating environment? Should some routes with high OTPs be expected to do better? Quantifying the effects on subway performance also provides a planning tool that can be used to anticipate the impacts on timeliness caused by capital projects and the redesign of route structure. Although practitioners have long considered such factors in predicting service impacts, quantification of these effects is needed when change involves multiple variables. When some variables have opposite effects, an evaluation of the trade-offs requires some precision. For example MTA's proposal to extend the 63rd Street Tunnel requires evaluation of numerous route design options. The results of this model show that increased throughput achieved under some options would be offset by degraded reliability caused by the addition of merges.

FACTORS AFFECTING OTP

Two data bases covering morning rush hour subway performance in 1988 and 1990 were used (*1,2*). Each was produced from an analysis by OIG staff of New York City Transit Authority (NYCTA) train movement records. Together they include times for over 54,000 trains from terminals to central business district (CBD) stations, including all trains arriving between 6:00 and 10:00 a.m. From these data and other sources, we quantified hypothetically important variables. The variables found to have a significant impact on OTP were (a) number of route merges, (b) whether public schools are in session, (c) scheduled headway, (d) distance traveled, (e) stops, (f) crowding, that is, an index of ridership compared with scheduled capacity, (g) whether construction occurred the night before, and (h) mechanical reliability of subway cars, measured by mean distance between failure (MDBF). Other variables lacked useful data, were statistically insignificant, or were accidentally correlated with OTP.

Coefficients for the estimated effects of each variable on OTP (Table 1) were produced by logistic regression (*3–5*). OTP, the dependent variable, is dichotomous; each train is categorized as success or failure, as on time or late. Logistic regression estimates how variables affect the probability of being on time. It is a nonlinear model; the magnitude of the effect changes depending on the starting level of OTP, with the largest effects occurring when OTP is near 50 percent. This is necessary mathematically, because OTP cannot exceed 100 percent. The nonlinear behavior of OTP also provides an interesting perspective on how to measure performance, which will be discussed later. The coefficients produced by logistic regression are expressed in terms of logits—the natural logarithm of the ratio of successes to failures (the odds ratio). Table 1 gives three coefficients for each variable, using the 1988 and 1990 data both separately and combined. In Table 1 a negative sign, as for *merges*, indicates that an additional merge can be expected to hurt OTP, whereas a positive coefficient, as for *headway*, suggests that routes with longer headways are more likely to be on time.

A variable's effect on OTP depends on the starting level of OTP. It requires a conversion from logits back into OTP, and this is done in Table 2. The coefficient of each variable is given just

---

Office of the Inspector General, Metropolitan Transportation Authority, 100 Park Avenue, 14th Floor, New York, N.Y. 10017-5516.

TABLE 1 Estimates of Change for Causal Variables

VARIABLE	UNIT OF CHANGE	BOTH YEARS	1990	1988
MERGES	One	-0.292	-0.237	-0.310
CROWDING INDEX	1.0	-0.881	-1.077	-0.511
SCHOOL DAY	True	-0.132	-0.071	-0.156
MDBF Starting Level				
ANY	10,000 Miles	n.a.	0.029	0.098
10,000 MILES	10,000 Miles	0.305	n.a.	n.a.
30,000 MILES	10,000 Miles	0.113	n.a.	n.a.
60,000 MILES	10,000 Miles	0.066	n.a.	n.a.
TRIP LENGTH	1 Mile	-0.013	-0.010	-0.034
HEADWAY	1 Minute	0.107	0.041	0.172
NIGHTWORK	True	-0.127	-0.017*	-0.198
INTERCEPT		1.780	2.540	1.344

Notes:

The coefficients show the expected change in the natural logarithm of the odds of being on-time, i.e., the ratio of on-time probability to late probability. To see how this translates into OTP itself, refer to Table 2.

* The estimate is not statistically significant at the 95% confidence level. All coefficients except for Nightwork (1990) and Trip Length (1990) are significant at 99.9% confidence.

n.a. For the MDBF analysis of the combined 1988 and 1990 data, we used the natural log of MDBF to show how the strength of the effect differs for different starting levels of mechanical reliability. This non-linearity demonstrates diminishing returns (in terms of higher OTP) on investment. Such a logarithmic method applied to 1988 and 1990 data individually was not statistically significant.

Source: Analysis by OIG of NYCTA train movement records, General Orders, ridership counts, schedules, and other variables from 1988 and 1990. Analysis used logistic regression on the SAS system.

below the variable's name. In the column below that, the change in OTP expected for each additional merge, minute of headway, mile traveled, and so on, is provided for each level of OTP. For example an additional merge when OTP is only 5 percent would lower OTP to 3.8 percent. (To avoid confusion, we will use the term *percent* when we refer to OTP itself and the term *percentage point* when we refer to the change in OTP caused by the variable. For example, if OTP is 80 percent and a change in some variable causes a 10 percentage point decline, the resulting OTP would be 70 percent.) When OTP is 50 percent one more merge would lower OTP to 42.8 percent. Figures 1 and 2 illustrate the trends in Table 2.

Merges gives the number of times a given route converges on the same track with another route. For example the D route merges with the Q and B routes before West 4th Street (two merges). We did not include divergences: They should not delay trains or across-the-platform transfers, because rush hour trains are not supposed to be held for connections. Using data from both years, we estimated that each time a route merges OTP may drop as much as 7.2 percentage points (3 percentage points lower than when OTP is at 90 percent.) The effect of *merges* was greater in 1988 than in 1990, reflecting that OTP in general was lower in 1988 when mechanical reliability was so poor and schedule adherence so much worse than that in 1990 that the mistiming of trains at merge points was more acutely felt. Because the negative impact of merges results from the mistiming of train arrivals at the merge point, better schedule adherence can reduce the impact of this variable. However merges may always be a strong negative factor because of some inevitable lateness. For example schedule adjustments to manage service evenness are typically made by terminal dispatchers for one of the merging routes without conferring with the other terminal. A centralized, modernized control center planned by NYCTA may reduce the number of delays resulting from uncoordinated actions of decentralized, local decision makers, but we doubt that New York's system of merging routes can ever be completely rationalized.

Crowding index measures the ratio of the number of passengers to the capacity on that route (measured at the most crowded point) for each half-hour period. For example if the trains currently scheduled can carry 14,500 riders and 12,000 riders pass through

TABLE 2 Effects of Variables on Morning Rush Hour Subway OTP

							10,000 MILE IMPROVEMENT IN MDBF WITH STARTING MDBF OF		
ON-TIME PERCENT (OTP)	ONE MORE MERGE	SCHOOL IN SESSION	SCHEDULED HEADWAY (+1 MIN)	ONE MORE MILE	CROWDING INDEX (+0.1)	NIGHT WORK	10,000	30,000	60,000
	-0.292	-0.132	0.107	-0.013	-0.881	-0.127	0.305	0.113	0.066
1	-0.3	-0.1	0.1	0.0	-0.1	-0.1	0.4	0.1	0.1
5	-1.2	-0.6	0.5	-0.1	-0.4	-0.6	1.7	0.6	0.3
10	-2.3	-1.1	1.0	-0.1	-0.8	-1.1	3.1	1.1	0.6
20	-4.3	-2.0	1.8	-0.2	-1.4	-2.0	5.3	1.9	1.1
30	-5.8	-2.7	2.3	-0.3	-1.8	-2.6	6.8	2.4	1.4
40	-6.8	-3.1	2.6	-0.3	-2.1	-3.0	7.5	2.7	1.6
50	-7.2	-3.3	2.7	-0.3	-2.2	-3.2	7.6	2.8	1.6
60	-7.2	-3.2	2.5	-0.3	-2.1	-3.1	7.1	2.7	1.6
70	-6.5	-2.8	2.2	-0.3	-1.9	-2.7	6.0	2.3	1.4
80	-5.1	-2.2	1.7	-0.2	-1.4	-2.1	4.4	1.7	1.0
90	-3.0	-1.3	0.9	-0.1	-0.8	-1.2	2.4	1.0	0.6
95	-1.6	-0.7	0.5	-0.1	-0.4	-0.6	1.3	0.5	0.3
99	-0.3	-0.1	0.1	0.0	-0.1	-0.1	0.3	0.1	0.1

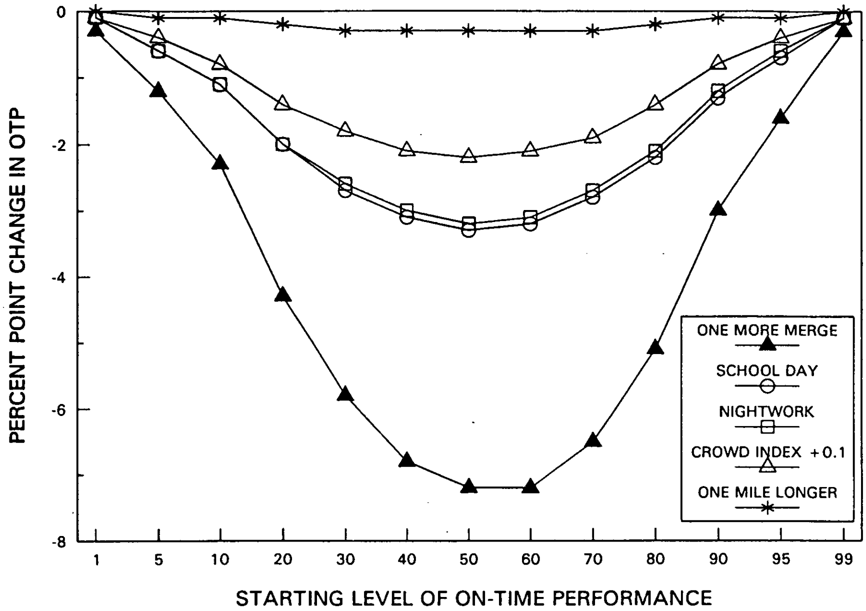
the maximum load station, then *crowding index* would be 0.83. The number of passengers came from NYCTA's annual counts of riders entering the CBD. Capacity is based on the service that NYCTA actually provided on the days of these counts. The index was calculated for each route for every half hour. In the 1990 data it ranged from the sparsely used downtown Q-route service (0.009 at West 4th Street between 6:00 and 6:30 a.m.) and the heavily used downtown E and F routes (1.193 at Fifth Avenue between 8:30 and 9:00 a.m.). The coefficient for *crowding index* (Table 1) gives the change in the logit for a change in the index from 0 to 1.0 (i.e., from no passengers to total capacity), but in the discussion of the measure—and in Table 2 and Figures 1 and 2—the expected change for an increase in the index of 0.1, for example, an increase from 0.5 to 0.6 was shown. It was found that an increase in the index of 0.1 would lower OTP by 2.2 percentage points when OTP is 50 percent and 1.4 percentage points when OTP is 80 percent.

In measuring crowding characteristics of passenger behavior was inevitably included. To the extent that the social norms of New York subway riders differ from those of riders in other cities and of other cultures, the estimates may not be applicable elsewhere. The policy implications may differ as well. As trains grow more crowded OTP tends to be lower because dwell times can increase dramatically when trains are crowded. In New York passengers on the platform appear more inclined to crowd aggressively around doors when trains are very crowded, perhaps because it often happens that the meek sometimes cannot board the train at all. This increases dwell time by slowing down deboarding passengers. Moreover crowded platforms require train operators to drive more slowly into stations to maintain safety.

The effects of all variables in the model were stronger in 1988 than in 1990 with the exception of *crowding index.* The effect of crowding was almost three times stronger in 1990 than in 1988. For example in 1988 the addition of 0.1 on the index, when OTP was 80 percent, would decrease OTP to 79.2 percent, but would decrease OTP to 78.2 percent in 1990. NYCTA reported that systemwide rush hour OTP improved from 89.6 percent in 1990 to 91.7 percent in the first 10 months of 1992, when ridership was lower. It is estimated that ridership loss accounts for one-fourth of the total improvement.

School indicates that New York City public schools were in session. Its values are the same for all routes. When school is in session the likelihood of a morning rush hour trip being on time declines by a maximum of 3.3 percentage points, probably because of the higher rate of pulled emergency cords, held train doors, and so on.

Nightwork indicates construction activity on the previous night. All work should have ended before the rush hour began, although

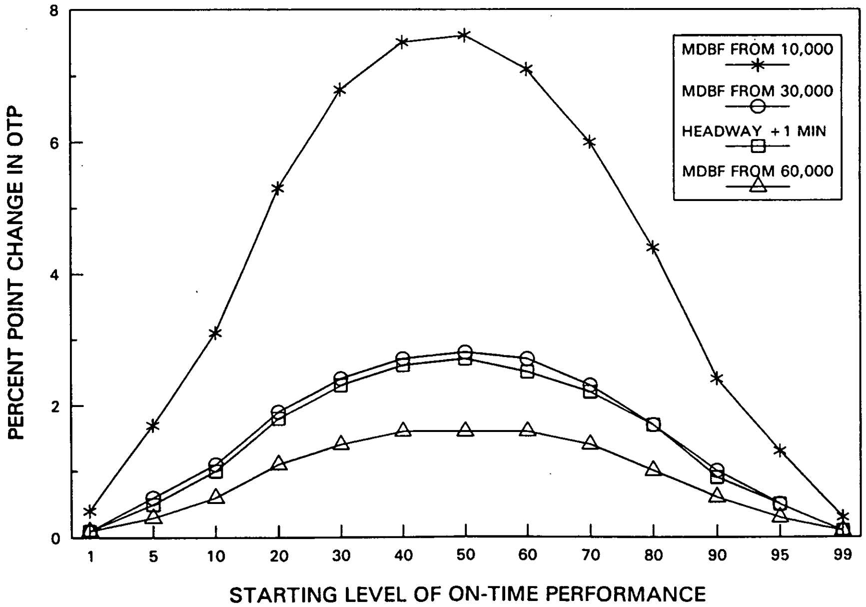


Source: Analysis of NYCTA morning rush hour subway data, 1988-90 (See Table 2)

FIGURE 1 Variables that decrease OTP.

track conditions are sometimes not perfect. If nightwork was scheduled the value of *nightwork* for every train on that route until 6:30 a.m. would be "true". All others are counted as "untrue." For passengers arriving in the CBD before 6:30 a.m. nightwork decreased the chance of being on time by a maximum of 3.2 percentage points. (In 1988 the effect could be as great as 4.9 percentage points.) This is caused when work cannot be finished on time or by slow work trains returning to the yard after construction duties. The coefficient for nightwork in 1990 was much weaker than that in 1988 and was not statistically significant. This suggests that NYCTA has been successful in scheduling and planning capital construction.

MDBF is the average monthly number of miles that trains travel before a mechanical failure causes a cancellation or a delay of



Source: Analysis of NYCTA morning rush hour subway data, 1988-90 (See Table 2)

FIGURE 2 Variables that improve OTP.

more than 5 min. The car class of every train was known and the monthly MDBF for that class was assigned to each record. MDBF is a crucial factor and perhaps the one most amenable to managerial intervention. The significantly improved MDBF of the subway fleet, primarily as a result of new car purchases and car overhauls, helped produce the improvement in New York City's subway reliability that occurred between 1988 and 1990. The coefficients in Table 1 give the effect of raising MDBF by 10,000 mi.

MDBF was treated to test whether improvement would have a greater effect when MDBF is low than when it is high. This approach was statistically significant when considering both years together. Therefore when combining data from both years, a different coefficient is given for cars with MDBFs of 10,000 mi than for cars with MDBFs of 30,000 or 60,000 mi. Table 2 gives three MDBFs to show that the effect depends on the magnitude of MDBF itself and on the starting level of OTP. However when 1988 and 1990 were looked at separately, this ''nonlinear'' approach was not statistically significant, so only one coefficient is listed for the individual years. When MDBF is 10,000 mi before a service disruption, improving it to 20,000 mi could improve OTP by as much as 7.6 percentage points. A 10,000-mi improvement is modest in light of the accomplishments of the new car and car overhaul programs. Now no cars have MDBFs as low as 10,000 mi, but in 1988 an MDBF of less than 6,000 mi was not unusual for certain car classes.

For most routes further improvement in mechanical reliability will not improve OTP by many percentage points. With an MDBF of 30,000 mi and OTP of, say, 80 percent, the expected improvement for raising MDBF to 40,000 mi is 1.7 percentage points. However with MDBF at 60,000 mi and OTP at 90 percent an increase in MDBF to 70,000 would raise OTP only to 90.6 percent. This conclusion was borne out by recent data. The 12-month rolling average MDBF for the subway fleet after August 1993 was 50,048 mi, a healthy improvement over the average in August 1992 of 41,452 mi. However this 8,500-mi improvement (more than 20 percent) had no measurable effect on OTP, which actually went down slightly during the same period (*6*). Similarly Table 1 shows that the effect for 1988 (0.098 for every mile) was three times greater than the effect for 1990 (0.029 for every mile). MDBF may have played itself out as a means of improving OTP, but not reliability generally, as will be seen.

Trip length gives the effect of adding a mile to the train's run. This variable was used in place of scheduled travel time because the latter varies during the rush hour; for example, the scheduled travel time of the Lexington Avenue Express between 125th Street and Grand Central is 7 min longer at the peak than early in the rush hour. The longer the trip, the more time there is for something to go wrong. Moreover the OTP standard is 5 min, not a percentage of the running time.

Headway is the scheduled time between trains at the most congested point. For example an A train from Lefferts Boulevard merges with the A train from Far Rockaway and the C train before Canal Street. The combined headway at Canal Street was used; for example, the time between an A train and a following C train is the C train's headway. The chance of being on time improves as headway grows larger. When more trains are scheduled and headway consequently decreases, OTP will decline by some amount. Adding 1 min to the headway of a route with 80 percent OTP raises OTP to 81.7 percent.

LESSONS FOR OTP MEASUREMENT THEORY

Logistic regression is the appropriate method for modeling a probability measure like subway OTP, and this provides an important insight into the uses and drawbacks of an OTP measure. Figures 1 and 2 illustrate that a given variable affects OTP most when OTP is about 50 percent—an extremely poor level of service—and least when it is closer to zero or 100 percent. As service improves it becomes more difficult to increase the OTP statistic; more resources are needed to produce the same percentage point change in OTP. However passengers may detect significant improvement in reliability even though the OTP statistic changes little (*7*).

Figure 3 shows how the odds ratio—the ratio of on-time to late trips—responded to changes in four variables. The change from 95 to 96 percent OTP may appear small, but in terms of the odds it is substantial; instead of experiencing a delay once for every 19 on-time trips, riders have one delay for every 24 on-time trips. By contrast the change in OTP from 75 to 76 percent is imperceptible: with OTP at 75 percent, riders are late once for every 3 on-time trips; with OTP at 76 percent, riders are late once for every 3.2 on-time trips.

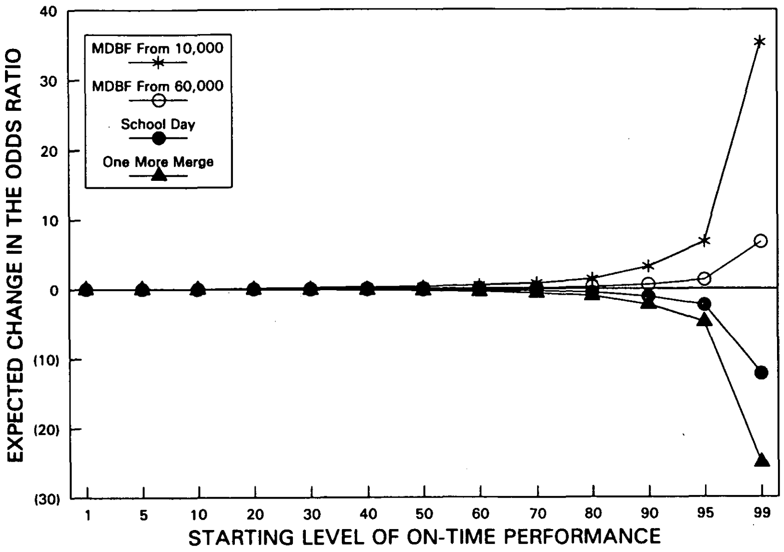
USING CAUSAL MODEL TO EVALUATE ROUTE PERFORMANCE

The results in Table 1 provide an equation for predicting OTP. These predictions, based on 1990 data and coefficients for that year, are given in Table 3. The intercept of 2.54 (a logit) translates into a base OTP of 92.7 percent. Each variable adds or subtracts from this intercept. For example the downtown no. 3 route has one merge before the CBD, which changes the logit to 2.303 (OTP 90.9 percent). Actual and predicted OTPs can be compared to see whether a route performed to expectations and to compare routes in terms of the variance between actual and predicted OTPs. Twenty-one of 33 routes or directions were predicted within 5 percentage points. The model had the hardest time predicting routes with high OTPs. This suggests that there are important quantitative factors—especially positive factors—that are not represented in the model.

Table 3 ranks each route or direction according to success against expectations. The J/Z route performed best. The downtown Q route was predicted to have the highest performance given its short length (to West 4th Street) and sparse ridership, but it finished 11th to 12th in terms of actual performance. Relative to its expected performance the downtown Q route was rated 31st out of 33. The downtown no. 5 and the uptown B routes were the most hopeful surprises. In terms of actual performance they were among the lowest 20th percentile, but given the merges and other obstacles experienced on those routes, they did better than expected.

CONCLUSIONS

Mechanical reliability has an obvious, positive impact on timeliness, and NYCTA's success in upgrading the subway car fleet had a significant effect on service throughout the period from 1988 to 1990. However further increases in MDBF will have diminishing



Source: Analysis of NYCTA morning rush hour subway service, 1988-90

FIGURE 3 Change in odds ratio.

returns in terms of OTP. This was predicted by the model, and recent NYCTA data support this conclusion.

A trade-off is involved in considering merges, because the merging structure provides beneficial routing alternatives when service disruptions occur. However the route merges that are so common to the New York City system hinder the delivery of timely service and constitute a major cause of delays according to the model. Each time a route merges, OTP may be as much as 7.2 percentage points lower (3 percentage points lower when OTP is at 90 percent). Because the negative impact of merges results from the mistiming of train arrivals at the merge point, better adherence to schedule can reduce the impact of this variable. The strong negative effect of merges on timeliness underscores the importance of careful scrutiny of route design plans that rely on additional merges. Even if significant increases in throughput and decreases in crowding can be achieved, the addition of merges can erode the benefits by impairing reliability.

Higher levels of ridership (holding the number of trains and all else constant) cause a decline in performance. As timeliness improved from 1988 to 1990, the negative effect of crowding on OTP became stronger. Although the effects of the other variables diminished, crowding emerged as a more serious problem. NYCTA measured a 2.1 percentage point improvement in rush hour OTP from 1990 to 1992, and the analysis suggests that one-fourth of this is because of lower levels of ridership.

Because the goal of NYCTA and transit advocates generally is to increase use of the subway, the adverse effect of increased crowding threatens to place constraints on service quality. Further improvement in subway service may be difficult as ridership levels increase from the recent slump, as they appear to be doing. Scheduling more trains can keep crowding levels constant on most routes. However on the most crowded routes more trains cannot be scheduled with current operating and safety rules and signalization. To make the matter more intractable, scheduling more trains makes headways smaller, and smaller headways are associated with lower OTPs. This analysis also suggests that reducing the number of trains (as is occasionally proposed in the interests of efficiency) may have unforeseen performance impacts, because the increase in headways from a service cut may be offset by an increase in crowding.

New subway cars being tested by NYCTA will carry more passengers and have design features that should allow passengers to board more quickly. If this helps reduce the times that trains spend at subway platforms, it could reduce crowding. Another way to increase throughput is to install new signal systems that permit more trains to operate in the peak interval.

A final possibility for increasing throughput to alleviate crowding is to change passenger behavior that produces unnecessarily long dwell times. Part of the strength of the crowding variable tested in the model results from passenger behavior. In a real sense passengers are members of the organization. Riders outside trains gather directly in front of doors, blocking exiting passengers, and riders inside cluster around doors, blocking entering passengers. Riders exiting from the middle of a car have tremendous difficulty making their way to the doors. Passengers hold doors for others. The clustering around doors by riders inside the train is being addressed by the design of the new cars; time will tell if this can help. Posters urge riders not to hold doors, but these focus on safety. NYCTA may need to consider public information campaigns that will inform riders how their actions delay trains.

Performance is likely to be slightly worse during official school days, another indication of the importance of passenger behavior. Nighttime construction lowered OTP in 1988 and 1990, but the 1990 result was not statistically significant, suggesting that improvements in operating procedures by NYCTA worked.

The research results show that relying on OTP to measure service quality may obscure a significant improvement once OTP has reached a high level. Odds ratios better reflect the improvement that passengers experience. Odds ratio show that the most significant improvement can be achieved, for a given amount of change in OTP, only after OTP has surpassed 90 percent. This is especially important in the context of state and local government fund-

TABLE 3 OTP Predictions and Comparisons for 1990 Morning Rush (7:00 to 9:00 a.m.)

RANK	ROUTE	DIRECTION	PREDICTION	ACTUAL	VARIANCE	ACTUAL RANK
1	J/Z	Downtown	87.3	98	10.7	1
2	N	Downtown	83.4	93	9.6	3-4
3	No.3	Downtown	83.5	92	8.5	5-7
4	No.7(loc)	Downtown	82.1	90	7.9	9
5	M	Downtown	88.2	96	7.8	2
6	No.3	Uptown	85.7	93	7.3	3-4
7	No.7(exp)	Downtown	82.0	89	7.0	10
8	No.5	Uptown	78.6	85	6.4	16-18
9	R	Uptown	85.5	91	5.5	8
10	C	Uptown	78.9	84	5.1	19-23
11	No.5	Downtown	77.1	82	4.9	26-28
12	B	Uptown	77.7	82	4.3	26-28
13	No.2	Uptown	84.2	88	3.8	11-12
14	Q	Uptown	82.3	86	3.7	13-15
15	B	Downtown	82.4	86	3.6	13-15
16	No.2	Downtown	76.4	80	3.6	29-30
17	No.1	Downtown	89.0	92	3.0	5-7
18	D	Downtown	73.2	76	2.8	32
19	L	Uptown	89.3	92	2.7	5-7
20	N	Uptown	83.4	86	2.6	13-15
21	No.4	Downtown	81.9	84	2.1	19-23
22	D	Uptown	82.9	85	2.1	16-18
23	R	Downtown	82.8	84	1.2	19-23
24	No.4	Uptown	82.8	84	1.2	19-23
25	A	Uptown	82.4	83	0.6	24-25
26	E	Downtown	76.6	77	0.4	31
27	A	Downtown	85.8	84	-1.8	19-23
28	F	Uptown	87.4	85	-2.4	16-18
29	No.6	Downtown	84.4	82	-2.4	26-28
30	M	Uptown	85.4	83	-2.4	24-25
31	Q	Downtown	90.9	88	-2.9	11-12
32	C	Downtown	85.7	80	-5.7	29-30
33	F	Downtown	76.0	70	-6.0	33

RANK shows the ranking of routes from best to worst, i.e., in the order given by "VARIANCE" (the difference between actual and predicted performance). ACTUAL RANK gives the ranking of routes according to their actual 1990 performance, as measured by the OIG (column called 'ACTUAL').

Source: ACTUAL OTP was calculated by the OIG; it is not an official NYCTA statistic. Predictions were made using the logistic regression coefficients listed in Table 1 for 1990.

ing decisions. Legislators may believe that a system with an OTP of 90 percent has already achieved an excellent level of service and that additional investment is not needed because it will not improve OTP by much.

OTP for each route was predicted by using the coefficients estimated by the model and compared the predictions with actual OTP. Assuming that a model that more accurately predicts OTP can be made, managers can use its predictions in several ways. (a) Routes ranked at the bottom—which perform worse than expected—can be targeted for managerial initiatives to improve performance. (b) Line superintendents can be given realistic performance improvement targets on the basis of the variances calculated in this way. NYCTA asked line superintendents to raise OTP by 1 percentage point in 1993. That may be unrealistic for some routes, and others may be expected to do better. (c) Employees on routes that have severe handicaps—for example, the oldest equipment, many merges, and crowding—may be assigned pay differentials in proportion to difficulty to attract the most experienced and capable workers and managers.

ACKNOWLEDGMENT

This research was funded in large part by the Federal Transit Administration (Section 8 Technical Studies grant NY-08-0185).

REFERENCES

1. *The Inspector General's Evaluation of Morning Rush Hour Subway Performance, 1987–1988.* New York State Inspector General for the Metropolitan Transportation Authority, Aug. 14, 1990.

2. *Commuting to Work by Subway: The Inspector General's Review of Morning Rush Hour Subway Performance.* New York State Inspector General for the Metropolitan Transportation Authority, Dec. 10, 1991.
3. Hosmer, D. W., and S. Lemeshow. *Applied Logistic Regression.* John Wiley & Sons, Inc., New York, 1989.
4. Hanushek, E. A. and J. E. Jackson. *Statistical Methods for Social Scientists.* Academic Press, Inc., New York, 1977.
5. *SAS/STAT User's Guide, Version 6*, 4th ed., Vol. 1. SAS Institute, Inc., 1989.
6. *Transit Authority Committee Agenda.* New York City Transit Authority. Oct. 1993.
7. Henderson, G., H. Adkins, and P. Kwong. Subway Reliability and the Odds of Getting There On-Time. Public Transit Research: Management and Planning 1991. In *Transportation Research Record 1297*, TRB, National Research Council, Washington, D.C., 1991.

The findings reflect the views of the authors and not necessarily the official views or policies of FTA or the New York Metropolitan Transportation Council (which administered the grant).

Publication of this paper sponsored by Committee on Rail Transit Systems.