TRANSPORTATION RESEARCH RECORD 961

Transit Performance Evaluation and Auditing

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1984

Transportation Research Record 961
Price $7.00
Editor: Naomi C. Kassabian
Compositor: Harlow A. Bickford
Layout: Betty L. Hawkins

mode
2 public transit

subject areas
11 administration
12 planning
13 forecasting
54 operations and traffic control

Transportation Research Board publications are available by ordering directly from TRB. They may also be obtained on a regular basis through organizational or individual affiliation with TRB; affiliates or library subscribers are eligible for substantial discounts. For further information, write to the Transportation Research Board, National Research Council, 2101 Constitution Avenue, Washington, D.C. 20418.

Printed in the United States of America

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Transit performance evaluation and auditing.

(Transportation research record; 961)
1. Local transit–Evaluation–Congresses. I. National Research Council (U.S.) Transportation Research Board. II. Series.
TE7.H5 no. 961 380.5 s 84-22683
[HE4202] [388.4'068]
ISBN 0-309-03717-4 ISSN 0361-1981

Sponsorship of Transportation Research Record 961

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Kenneth W. Heathington, University of Tennessee, chairman

Public Transportation Section
John J. Fruin, Port Authority of New York and New Jersey, chairman

Committee on Transit Management and Performance
Gordon J. (Pete) Fielding, University of California-Irvine, chairman
Mark Abkowitz, William G. Allen, Jr., John P. Attanucci, A. Jeff Becker, Donald M. Chapman, John Dockendorf, John T. Doolittle, Jr., F. Norman Hill, Ronald I. Hollis, Joseph J. Kern, Arthur Thair Leahy, William R. Loudon, Brian E. McCollom, James H. Miller, Subhash R. Mundle, Richard L. Oram, Howard Permut, Robert E. Prangley, Nigel H. M. Wilson, Robert J. Zerrillo

W. Campbell Graeub, Transportation Research Board staff

The organizational units, officers, and members are as of December 31, 1983.

Contents

Addresses of Authors

Abkowitz, Mark, Department of Civil Engineering, Rensselaer Polytechnic Institute, Troy, N.Y. 12181
Attanucci, John, Multiplications, Inc., 1050 Massachusetts Avenue, Cambridge, Mass. 02138
Brenner, Mary E., School of Social Sciences, University of California, Irvine, Calif. 92717
de la Rocha, Olivia, School of Social Sciences, University of California, Irvine, Calif. 92717
Deacon, John A., Department of Civil Engineering, University of Kentucky, Lexington, Ky. 40506
Engelstein, Israel, Department of Civil Engineering, Rensselaer Polytechnic Institute, Troy, N.Y. 12181
Fielding, Gordon J., School of Social Sciences, Institute of Transportation Studies, University of California, Irvine, Calif. 92717
Hobeika, A. G., Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061
Kanok-kantapong, C., Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061
Sindzinski, John, Crain and Associates, Inc., 343 2nd Street, Los Altos, Calif.; formerly with Metropolitan Transportation Commission, Berkeley, Calif.
Tran, T. K., Department of Civil Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061
Vaziri, Manouchehr, Department of Civil Engineering, University of Kentucky, Lexington, Ky. 40506
Vozzolo, David, Multiplications, Inc., 1050 Massachusetts Avenue, Cambridge, Mass. 02138
Wilson, Nigel H. M., Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 12139

Methods for Maintaining Transit Service Regularity

MARK ABKOWITZ and ISRAEL ENGELSTEIN

ABSTRACT

Maintaining regular transit service has been a chronic operational problem that affects both travelers and operators. Although many researchers have studied aspects of this problem, major limitations of previous work have resulted in the recommendation of procedures that are neither representative of nor operational in the transit management environment. The development of a method for maintaining service regularity through improved scheduling and real-time control based on models developed and validated from empirical data is described. The method is simple to employ and does not require extensive data from the transit agency considering its implementation. Also included is a case study describing the application of the method to three routes in Los Angeles. The results indicate that the procedure can produce reasonable solutions and demonstrate its potential value to the transit community.

It is generally agreed that maintaining regular transit service intervals is an important operational problem that affects both travelers and operators. Operators rely on minimizing run-time uncertainty in specifying timetables and allocating vehicles to routes. Travelers are affected by the reliability of service, which stems directly from the predictability of vehicle run times. The proper amount of run time and slack to build into a schedule and how to control real-time reliability problems can have a profound effect on service regularity and thus on the productivity and efficiency of transit operations.

Many researchers have studied factors affecting bus running time and what corrective actions should be taken when reliability deteriorates (1-4). In the absence of empirical data because of the high costs associated with direct observation, most studies have been restricted to models that are analytically based. A major limitation of this work has been the assumptions made in order to derive closed-form solutions, which often result in recommendation of procedures that are neither representative of nor operational in the transit management environment.

The primary objective of this research was to make available to transit managers methods for maintaining service regularity through improved scheduling and real-time control based on models developed and validated from empirical data. An interest in developing transferable models and methods that are simple in nature and do not require extensive data from the transit agency was an important motivating factor in the research. The availability of empirically based methods of this kind would alleviate the need to analyze individual problems by manually collecting extensive amounts of data at each agency yet permit the identification of planned and real-time schedule modifications that can be implemented to improve service efficiency and productivity.

RESEARCH METHODOLOGY

The research design consisted of six sequential steps:

1. Determination of mean running time,
2. Determination of running-time variation,
3. Determination of headway variation,
4. Determination of passenger wait time,
5. Identification of optimal control strategy, and
6. Establishment of operator compatibility with the developed methodology.

Steps 1-4 are interdependent problems that, once resolved, serve as inputs to the fifth step. The last step concerns transferring the research results into an environment with which the operator is compatible.

Determining mean running time is the initial step in this process because the schedule and timetable are based on the mean running time. The research emphasis was on the temporal and spatial factors affecting mean running time.

Running-time variation is an important measure in defining unreliable service. The degree to which running-time variation propagates as the vehicle proceeds down the route is of particular interest when real-time control is studied. A priori, one would expect that running-time variation is correlated with mean running time and that delays tend to accumulate once a vehicle falls behind schedule.

It has been proven theoretically and demonstrated empirically that the waiting time of passengers at stops is related to the headway variation. To be able to reduce the headway variation effectively, it is important to know what influences it and how the headway variation propagates along the route. It is also important to understand the relationship between the headway variation before and after the control stop and to what extent a control strategy causes reductions in headway variation.

The effectiveness of both headway-based and scheduled-based strategies was considered in this study. A headway-based strategy is defined here as holding the bus for a certain amount of time (x_0). If the coming headway is less than x_0, the bus is held up to x_0. If the coming headway is greater than x_0, the bus is not held. Headway-based holding is most suitable for routes operating with short, uniform headways. When headways are short and uniform, it is assumed that passengers arrive more randomly at stops and that they are primarily concerned with the headway and not the schedule. Similarly, operators are concerned about keeping vehicles evenly spaced so that vehicle availability remains stable.

Schedule-based holding is considered suitable on routes that have long headways, which means that the schedule is not so tight and the procedure is simple to administer. It may also be appropriate for cases in which headways are uneven and the schedule is designed to meet certain demand requirements. In both cases the passengers' concern is not to miss a certain bus, so the buses should adhere to the schedule. To implement a schedule-based policy, there is a need to construct a reasonable schedule and enforce adherence to it by using proper incentives for drivers and a mechanism for accurate monitoring of their performance.

With either headway- or schedule-based holding, the choice of where to locate a holding point is extremely important. This problem is often solved by determination of minimized passenger wait time. Accordingly, the relationship among scheduled headway, headway variation, and wait time was examined. In this study empirical wait-time models covering a range of headways from 3 to 12 min were estimated and the results were compared with those of theoretical wait-time models and other research findings.

After the first four phases of the research had been completed, an optimization routine was developed to determine (a) the appropriate holding strategy to implement, given the schedule characteristics; (b) the effectiveness of holding on the route; and (c) the location of the control stop and the optimal corresponding holding time, given route and schedule characteristics. For the headway-based strategy the objective was to minimize the total waiting time of passengers, including those delayed on board the vehicle at the holding point. For the schedule-based strategy the objective was to maximize the effectiveness of control; effectiveness is defined in the subsequent discussion.

An important issue to consider is the eventual implementation of the models and methods by the transit operator. Computer software was developed so that the decision methodology could be utilized. The software is designed for a microcomputer system, because many transit operators are now using or considering the use of microcomputers in managing their operations and the program could be used by them without additional cost being incurred.

Mean Running Time

The data used in this analysis were collected in 1978 from Queen City Metro in Cincinnati, Ohio, by General Motors; automated-vehicle-monitoring (AVM) equipment was used. The data consisted of observations on two bus routes, each roughly 10 miles long, that traverse city streets and extend radially from the central business district (CBD) along a traffic corridor. The routes extend into the CBD and return to the suburban origin point. Except for layovers (time spent between the end of the previous run and beginning of the next run) at the CBD and suburban terminals, no other holding points are used on these routes. Peak-period headways are 12 min, increasing to 15 to 20 min during the off-peak period. Additional information on the routes included physical characteristics (length between observation points, number of traffic signals, parking restrictions, stop signs, yields, and unsignalized intersections) as well as dynamic characteristics (average boardings and alightings, average number of stops made, time of travel, and direction of travel). These data were segmented by observation point and operating period (5).

The analysis focused on determining the physical and dynamic factors affecting mean running time. This was accomplished by using linear regression with mean running time as the dependent variable and the route characteristics as the independent variables. Model specification in all phases of this research was guided by a criterion that included consideration of variables that could be justified a priori as explanatory variables of the dependent variable, had the expected coefficient signs, and had statistically significant coefficient estimates (t-statistics). The overall statistical fit of the model (corrected R^2) and potential dependencies among the independent variables were also considered.

The final model for mean running time is as follows:

$$\text{Mean running time (sec)} = \beta_1 + (b_2 * \text{link length}) + (\beta_3 * \text{passengers boarding}) + (\beta_4 * \text{passengers alighting}) + (\beta_5 * \text{percentage off-street parking}) + (\beta_6 * \text{signalized intersections}) + (\beta_7 * \text{daytime off peak}) + (\beta_8 * \text{afternoon peak}) + (\beta_9 * \text{outbound travel}) \quad (1)$$

It was found (Table 1) that mean running time is highly influenced by trip distance, boardings and alightings, and signalized intersections and to a lesser degree by parking restrictions on the route, time of day, and direction of travel. The model results tend to confirm earlier views. The order of importance of the explanatory variables also seems reasonable. The finding that running time is positively related to the number of signalized intersections is consistent with observations made by Welding (6).

TABLE 1 Mean-Running-Time Model

Variable	Coefficient Value	t-Statistic	Variable Mean	Avg Contribution[a]
Constant	-122.04		1.0	-122.04
Link length	216.54	10.89	2.05	443.91
Passengers boarding	6.03	5.74	9.37	56.5
Passengers alighting	3.83	3.83	11.57	44.3
Percentage of on-street parking	114.59	3.15	0.09	10.31
Signalized intersections	8.16	5.13	10.64	86.82
Daytime off-peak period	30.43	2.16	0.25	7.61
Afternoon peak period	41.73	2.78	0.25	10.43
Outbound travel	25.80	1.82	0.5	12.9

Note: Number of observations = 56, F (8, 46) = 76.5, corrected $R^2 = 0.92$, standard error = 41.9, Durbin-Watson statistic = 1.92.

[a] Coefficient value times variable mean.

It is interesting to note that the value of the constant implies a maximum average running speed of 21 mph. Adding the average number of boardings, alightings, signals, and typical parking restrictions, the average running speed decreases to 14 mph. These values are quite reasonable for bus movement in an urban corridor.

Data from Route 44 in Los Angeles were used to validate the mean-running-time model developed from Cincinnati data. Route 44 has the same characteristics as do the routes in Cincinnati. It is roughly 15 miles long and runs into the CBD and returns to a suburb. Peak-period headways are 5 to 7 min; they increase to 10 to 12 min during the off-peak period. The route was divided into 15 links that ranged from 0.5 to 2 miles long. The day was divided into four time periods as in Cincinnati. For each combination of link, period, and direction, the mean running time was calculated. The mean running time obtained from the model (predicted) and the mean running time observed for each link on Route 44 are plotted in Figure 1. The implication for transferability is encouraging, because most observations fall around the 45 degree line. In fact, a χ^2 goodness-of-fit test suggests that the hypothesis (95 percent confidence level) that the observations are from the same population cannot be rejected.

Running-Time Variation

Separate models for each time period were initially

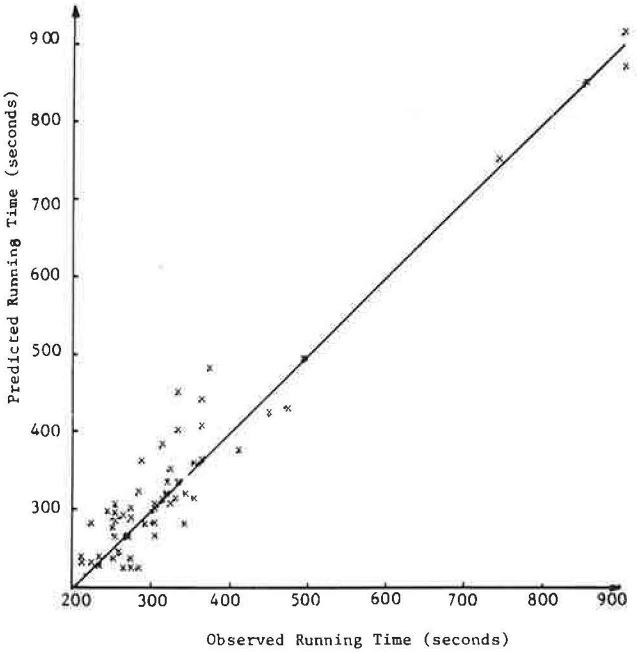


FIGURE 1 Validation of running-time model.

developed by using Cincinnati data. Trip origins and destinations were selected randomly, which resulted in a range of mean running time of only 10 to 47 min. The process of random selection also reduced the degrees of freedom significantly. Data collected in Los Angeles at a later stage provided a basis for both model validation and modification. Separate regression models were subsequently estimated based solely on the Los Angeles data.

For each Cincinnati and Los Angeles regression model, the error sum of squares was calculated. These values were then used to perform an F-test to check whether the Cincinnati and Los Angeles regression lines were significantly different. If the regression lines are not different, they can be pooled to estimate an improved model with more degrees of freedom. The confidence level (α) used in the F-test was 95 percent.

For both the morning and afternoon peak periods, the results suggested that one cannot reject the hypothesis that the two data sets produce the same regression line and therefore that the pooled regression line can be used as the final model. However, for the daytime off-peak period, the two regression lines were significantly different. Because the model estimated in Los Angeles had more degrees of freedom, it was selected as the day off-peak model.

The estimation results are as follows:

A.M. peak (6:00 a.m. to 9:00 a.m.), pooled data

$$\sigma_r = 1.399 + 0.0454U_r \qquad R^2 = 0.82,\ N = 45 \qquad (2)$$

Day off-peak (9:00 a.m. to 3:00 p.m.), Los Angeles

$$\sigma_r = 0.977 + 0.0530U_r \qquad R^2 = 0.90,\ N = 85 \qquad (3)$$

P.M. peak (3:00 p.m. to 6:00 p.m.), pooled data

$$\sigma_r = 0.707 + 0.08197U_r \qquad R^2 = 0.90,\ N = 49 \qquad (4)$$

where

σ_r = running-time deviation (min),
U_r = mean running time (min), and
N = sample size.

The final models cover a range of running time between 10 and 85 min. The highest deviation occurs during the p.m. peak, thus suggesting that service reliability is worse during this time period. This implies that perhaps priority should be given to controlling reliability during the p.m. peak.

Implicit in the model results is the suggestion that running-time deviation at early points on the route propagates as the vehicle proceeds further downstream. This is consistent with observations made by Doras (7) on bus routes in Paris and by Loo (8) in a study of a Minneapolis bus route.

The running-time models can also be used to improve schedules by allowing for the appropriate amount of slack time, so that succeeding runs are less likely to be affected by delays on earlier runs. Assuming a distribution for the running time, the appropriate slack time can be determined for a given confidence level. For example, for a normal distribution of running time, if mean running time from terminus to terminus is 30 min and the standard deviation of running time is 3 min, the operator can be 95 percent confident of having buses begin the next run on time by allowing just under 6 min of slack time in the schedule (union work rules are a separate consideration). This analysis can be extended rather easily to determine the vehicle requirements to operate a route given the desired headway, mean-running-time deviation, and confidence level.

Headway Variation

Headway-variation analyses focused on two issues: (a) deriving a headway-variation model based on scheduled headway and running-time variation and (b) assessing the impact of control on headway variation beyond the control point. The discussion in this and the following section applies only to the headway-based strategy, because the schedule-based strategy does not address regulating headways or the impact of headway variation on system wait time.

The data used to derive a headway-variation model were generated by using Monte Carlo simulation. The inputs to the simulation program included scheduled headway, average running time to each stop, and variation of running time. The scheduled headway was set at 3, 6, and 9 min, and running times were assumed to come from a beta distribution. Stop locations ranged from average running times of 5 to 90 min from the route origin and the coefficient of variation ranged from 0.05 to 0.17. The output of the simulation consisted of headway variation for each combination of scheduled headway, running time, and running-time variation. These results were used as inputs to model estimation by using headway variation as the dependent variable.

The simulation results indicated that the headway variation increases rather quickly near the beginning of the route and then reaches an upper bound. The time it takes to reach the upper bound depends on the scheduled headway and variation in running time. This was borne out by the following model estimation result:

$$v_h = (-12.2 + 6.94\bar{h})[1 - \exp(-0.0447v_t)] \qquad N = 554 \qquad (5)$$

where

v_h = headway variation (min),
$\bar{h}$ = scheduled headway (min), and
v_t = running-time variation (min).

The residual mean square for this model is 4.08, indicating a good statistical fit.

Data from Route 44 in Los Angeles were also available to evaluate the headway-variation model. This validation was performed by comparing the observed headway variation with the predicted headway variation. The predicted headway variation was computed by using the model in Equation 5, where the inputs to the model were derived from the data collected.

Route 44 was divided into links demarcated by the AVM location. For each AVM location, the running-time variation and the headway variation were calculated. Three sequences of scheduled headways were available for calculating the headway variation (6-, 8-, and 11-min headways), with each headway representing a different time period.

Plots of the predicted and observed headway variation appear in Figure 2. The results are generally encouraging, particularly for the 6- and 8-min headways, which are within the 10-min headway range under consideration for headway-based control (see discussion of passenger wait time).

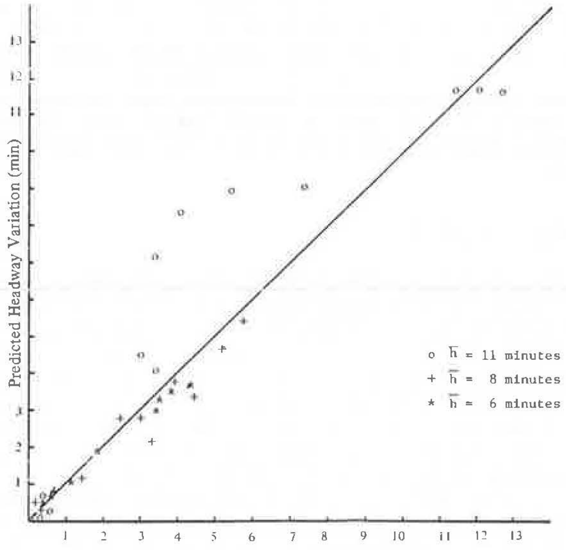


FIGURE 2 Headway model validation results.

The impact of headway-based control on headway variation downstream of the control point was also examined by using simulation, because no transit data on control are now available for model development. Recall that the headway-based approach is to hold to a threshold value (x_0) if the coming headway is less than x_0. The simulation design was to introduce control at stops located 10, 20, 30, and 40 min from the route origin, varying the threshold value in 0.5-min increments from zero to the scheduled headway. Headways of 3, 6, and 9 min were considered. The output measures included headway variation before and after the control stop.

The simulation output provided data for model estimation, which yielded the following:

$$v_a = 0.5448 v_b^{0.713} (\bar{h} - x_0)^{0.734} \qquad \text{corrected } R^2 = 0.94 \qquad (6)$$

where

v_a = departure headway variation at control stop (min),
v_b = incoming headway variation at control stop (min), and
x_0 = threshold value (min).

An interesting implication of this model is that the headway variation reduces to nearly zero when x_0 is equal to the scheduled headway for an extended operating period, independent of the level of variation before control. This does not suggest that it is always better to hold according to a threshold equal to the scheduled headway, because the optimal strategy also depends on the number of passengers on board at the control stop and those waiting downstream.

The model, when combined with Equation 5, implies that the benefits of control are not distributed uniformly to all stops after the control point. Instead it appears that the maximum benefits are felt near the control point; the headway variation begins to increase again downstream until it reaches an upper bound.

Passenger Wait Time

Two types of wait-time models were examined: (a) passenger wait time at stops along the route and (b) delay to on-board passengers when the bus is being held at the control stop.

The analysis of passenger wait time at stops was conducted by using data collected in Los Angeles as part of the evaluation of AVM equipment implemented at the Southern California Rapid Transit District (SCRTD). The data were collected on four routes with headways varying from 3 to 12 min. Checkers were located at specific stops on the routes and noted passenger and vehicle arrival times at stops and the weather conditions at the time of observation. Three days of data were collected on each route in both directions, with the exception of one route for which only one direction was observed. Separate analyses were performed on the 3-min headway route and the other routes (8- to 12-min headways), because there was reason to expect that passenger arrival patterns might be related to the scheduled headway.

The regression estimate for the 3-min route was as follows:

$$\bar{w} = 77.34 + 0.0028v_h \qquad \text{corrected } R^2 = 0.66 \qquad (7)$$

where $\bar{w}$ is the average passenger waiting time in seconds and v_h is the headway variation in seconds.

It is interesting to note that the wait times were 7 percent lower than would be predicted by using a theoretical model, which assumes random passenger arrivals. Late arrivals running to catch the bus at the last minute might account for this, because they incur no wait time.

The wait-time model for the routes with 8 to 12 min headway was as follows:

$$w = -47.02 + 0.497\bar{h} + 0.00121v_h \qquad \text{corrected } R^2 = 0.69 \qquad (8)$$

where $\bar{h}$ is the scheduled headway in seconds.

The negative constant and the coefficient for the mean headway, which is less than 0.5, result in lower wait time than that predicted by the theoretical model.

The wait-time analysis results are not unusual and are consistent with findings reported by Holroyd and Scraggs (9), O'Flaherty and Mangan (10), Seddon and Day (11), and Joliffe and Hutchinson (12). If anything, they suggest that the accepted assumption of random passenger arrivals for headways of 12 min or less should be modified to 10 min or less.

The passenger delay at the control stop depends on the headway variation at the control stop as well as on the threshold headway. Simulation was again used to obtain data for the model estimation. Control stops were introduced in the simulation and bus delays were calculated for different threshold values. These data were then used to estimate the following model:

$$d_j(x_0) = \{3.9245 + 0.0755\ \mathrm{var}_j(H)\,[x_0/E(H)]^4\} \qquad R^2 = 0.94 \tag{9}$$

where

$d_j(x_0)$ = the average delay (min) at control point j,
x_0 = the threshold headway (min), and
$E(H)$ = expected headway (min).

The model is sensitive to x_0 such that when x_0 approaches $E(H)$, the delay increases quickly. This occurs because large values of x_0 are causing many buses to be delayed.

Optimal Control Strategy

The results of the steps described in the previous discussion were used as inputs to the decision process in resolving the following questions:

1. Which kind of control is appropriate?
2. Should the strategy be implemented?
3. Where should the control point be located?
4. For headway-based control, what is the optimal threshold value?

Question 1 is determined outside of the decision algorithm and depends on the length and uniformity of scheduled headways for reasons described previously. The remainder of the questions are addressed within the decision algorithm.

The algorithm developed for headway-based control was to minimize the following objective function:

$$TW = \sum_{i=1}^{j-1} (n_i \times \bar{w}_i) + [b_j \times d_j(x_0)] + \sum_{i=j}^{N} (n_i \times \bar{w}_i) \tag{10}$$

where

TW = expected total wait time on route,
j = the control stop,
x_0 = threshold value,
n_i = number of passengers boarding at stop i,
b_j = number of passengers on board at stop j,
$\bar{w}_i$ = average wait time at stop i,
N = total number of stops on route, and
$d_j(x_0)$ = expected delay at the control stop for the threshold of x_0.

The first term represents the wait time of passengers upstream of the control point. The second term represents the delay caused to passengers on board the bus at the control stop. The final term represents the passenger wait time at stops downstream of the control stop.

The minimum expected total wait time will occur at a specific j and x_0, which will result in the identification of the optimal control point and threshold value. The minimum expected total wait time is then compared with the expected total wait time without control to determine whether control represents an improvement and the magnitude of the benefit provided.

Preliminary evaluation of this algorithm was conducted by using a 30-stop route with five different boarding and alighting profiles. In each scenario, the optimal threshold value and the control point were found and the percentage of reduction in the total wait time was computed (see Figure 3 for a sample of the scenario results).

It was found that the location of the control stop is quite sensitive to the distribution of passengers boarding at stops. Generally the control point occurs just before a group of stops at which many passengers are boarding. Thus, more passengers enjoy a reduction in the wait time, because the headway variation is mainly reduced at stops that are close to the control point. If the number on board is small, it is more likely that the threshold value will be larger. One must remember that the threshold value and the location of the control point are interrelated and that they are dependent on all the input parameters in the algorithm.

The objective for schedule-based control is to find the most effective location for returning service to the original schedule without causing excessive delay to passengers on board the bus at the control point. For this reason desirable locations for enforcing adherence to the schedule should be those points at which the schedule deviation is high and at which few passengers are expected to be on board the bus at the control point.

The algorithm for schedule-based control is to identify the stop that maximizes the following:

$$ER = SD_j(V_r)/(b_j/\sum_{i=j}^{N} n_i) \tag{11}$$

where ER is the effective ratio and $SD_j(V_r)$ is the standard deviation of running time at stop j. Thus, the best control stops would be those with high values of ER.

In testing whether ER produces reasonable results, 10 different profiles of passengers boarding and alighting were used. In each case schedule-based and headway-based strategies were compared to see whether the two strategies selected similar con-

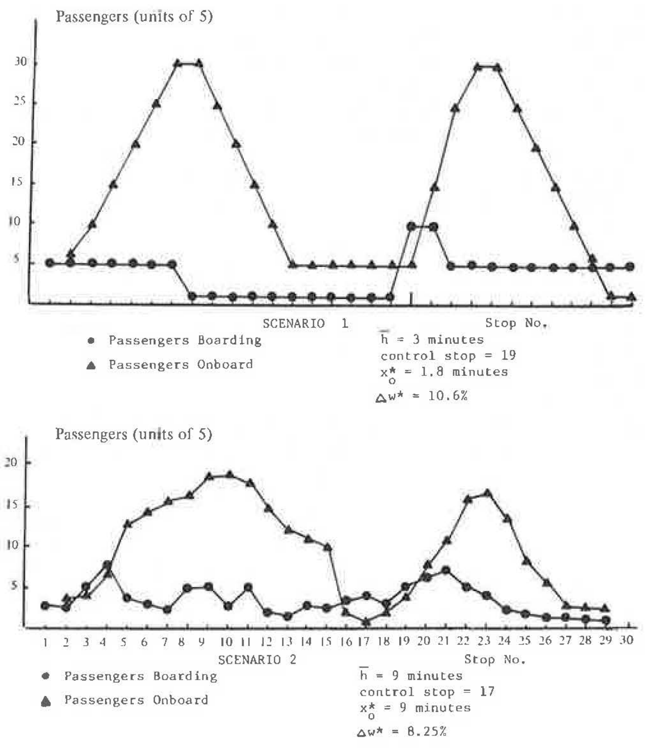


FIGURE 3 Sample evaluation scenarios.

trol stops. In most cases the same or nearly the same stops were selected, which implied that ER is a reasonable selection criterion for the schedule-based holding methodology.

Evaluating the Methodology

Although most of the models developed for the decision methodology were validated individually, there remains a need to test whether the entire methodology produces satisfactory results. A simulation written in the general purpose simulation system (GPSS) was formulated to (a) evaluate whether the control strategy and decision methodology are effective in reducing total passenger wait time on the route and (b) compare the optimal control parameters selected by the decision methodology with those identified by the simulation results.

Introducing the headway-based holding strategy into the simulation reduced passenger wait time on the route by a similar amount as predicted by the decision methodology. This pattern was consistent across different passenger boarding and alighting profiles.

When the optimal control parameters selected by the decision methodology were compared with those of the simulation, the stops chosen by the decision methodology corresponded to those selected by the simulation. In all cases the same stop and the same threshold value were identified or the second-best stop selected by the decision methodology was the first stop selected by the simulation; there were slight differences in the threshold values (1 min and 0.5 min, respectively). Because there is virtually no difference between the decision methodology and the simulation, one can conclude that the decision methodology is feasible for obtaining optimal control parameters.

Another important finding is that the average delay time of buses is small, on the order of 1 to 1.8 min for threshold values of 4 min. This implies that few buses will actually be held for long periods of time and further that the probability of holding more than one bus at the same time is small. This suggests that physical space restrictions are not likely to be a constraining factor in implementing the headway-based strategy.

The previous data used to evaluate the decision methodology were generated by simulation. Data collected in Los Angeles provided an additional opportunity to develop a case study for evaluating whether the decision methodology selects reasonable control parameters.

Routes 16, 30, and 44 were selected to perform the case study. On Route 30, headways are even and short (3 min), and a headway-based strategy was considered. The route follows a westward direction, beginning in the suburbs and passing through the downtown region to another suburb. On Route 16, headways are uneven, so a schedule-based strategy was evaluated. The route begins downtown and moves to the suburbs in a westerly direction. Route 44 represents a route with even but longer headways (5-8 min), which makes it a candidate for the headway-based strategy. The route is U-shaped and passes through the downtown region at its midpoint. Each route of the three routes has approximately 60 stops and data were available for each stop for Routes 16 and 30.

The drawback of using Route 44 in the case study is that the data were only available for AVM locations and not by stop. After consultation with the manager of the SCRTD Planning Department, it was agreed to use the AVM locations as proxies for stops. Thus, the decision methodology could only identify the best AVM location for control, which resulted in potential biases in the results. Data for this route were available for both directions in the morning and afternoon peak periods.

In applying the methodology to Route 30, two adjacent stops were identified as the best stops to control: Broadway and Third Street and Broadway and Fourth Street. Both stops have a threshold value of 1.2 min and are located at the beginning of the downtown region. At these stops there are few passengers on board and many passengers are boarding at stops immediately downstream of the control point. The estimated percentage reduction achieved in passenger wait time was about 3 percent.

For Route 16 a schedule-based strategy was evaluated. In this case the methodology predicted that control would not be effective. An examination of the route substantiates this conclusion, because the route origin is downtown and many people board near that point. In this case the best strategy is to ensure that vehicles depart on time from the route origin rather than detain them en route.

For both the morning and afternoon periods on Route 44, the same AVM location was selected for control; the threshold value in both cases was equal to the scheduled headway. The estimated reduction in passenger wait time was between 11 and 15 percent. The stops chosen are reasonable and are located just before the entrance to the downtown region where many passengers are boarding.

The manager of the SCRTD Planning Department was informed of the results obtained by using the decision methodology and agreed that the recommended strategies for each route were reasonable. It should be noted that the entire analysis was conducted based on data furnished by SCRTD supplemented by additional information that required observation of only a single run on each route.

Operator Compatibility with the Methodology

The decision algorithm has been coded in PASCAL for the Apple II microcomputer. For each stop the user defines the number of boardings and alightings, distance and number of intersections from the previous stop, direction and time period of travel, and, if available, the percentage of on-street parking allowed from the previous stop. Most of this information is available or can be easily collected by the transit agency. This data file serves as an input to the decision algorithm.

The user is prompted to describe the scheduled headway, which determines whether headway-based or schedule-based control will be considered. The input file of stop information is combined with the models previously described to form the inputs to the objective function.

The model output includes a statement of whether control is effective, a priority listing of the most effective control stops, and, for headway-based control, corresponding threshold values and absolute and relative benefits of control over the no-control case. The priority listing is useful in situations where it is impractical to implement control at a particular stop (e.g., traffic conditions) and near-optimal alternatives are worthy of consideration. The absolute and relative benefits provide for a comparison across routes, which is useful when there are constraints on the number of available street supervisors.

Sample output for headway-based control appears in Table 2. To assist transit operators in utilizing the methodology, a user's manual has been written that accompanies the software.

TABLE 2 Sample Output for Headway-Based Control

```
LIST OF EFFECTIVE CONTROL STOPS BY ORDER

STOP 13, THRESHOLD  3.75 MIN, REDUCTION  49.21 MIN,  %REDUCTION  4.98%
STOP 21, THRESHOLD  3.00 MIN, REDUCTION  47.15 MIN,  %REDUCTION  4.77%
STOP 20, THRESHOLD  3.00 MIN, REDUCTION  46.81 MIN,  %REDUCTION  4.74%
STOP 22, THRESHOLD  2.75 MIN, REDUCTION  46.70 MIN,  %REDUCTION  4.73%
STOP 12, THRESHOLD  4.00 MIN, REDUCTION  46.36 MIN,  %REDUCTION  4.69%
STOP 11, THRESHOLD  4.25 MIN, REDUCTION  42.04 MIN,  %REDUCTION  4.26%
```

VALUE OF MODELS

The models reported in this paper represent an attempt to use empirical data to establish factors that affect transit route performance and passenger level of service. They should not be interpreted entirely as cause-and-effect models because the collinearity between variables and lack of information on other potentially significant explanatory variables make it difficult to understand the individual contributions of each factor. Other assumptions that were made in conducting this research include independence of routes within the network and on-time vehicle departures from the route origin. Thus, the models should be considered primarily for their value in providing reasonable estimates of performance and service given the availability of information on route characteristics.

CONCLUSIONS

Several findings can be reported from this research activity. Mean running time is strongly influenced by trip distance, passengers boarding and alighting, and signalized intersections; other route characteristics have a lesser effect on this measure. Running-time deviation magnifies and propagates as vehicles proceed downstream. Headway variation is highly correlated with running-time variation; scheduled headway also affects this measure. Headway-based control decreases headway variation, and the magnitude of the change is dependent on the threshold level. Finally models of passenger waiting time that assume random passenger arrivals overestimate observed waiting times, even for short-headway routes.

Beyond the individual model implications, many general contributions can be attributed to this research effort. The research has addressed individually and collectively the issues that affect service regularity, which has resulted in pertinent information on setting timetables and allocating vehicles to routes. These effects are then represented mathematically and utilized in the development of a decision process that can be used to improve service regularity through the implementation of real-time holding strategies. Finally a mechanism is provided by which the operator can apply the methodology directly to address current reliability problems. The research product is based heavily on empirical analysis, which appears to be representative of actual operations.

The research results have direct practical application in metropolitan regions in which conventional transit service is operated. The decision algorithm is economical and does not require special data-collection activities to implement. It has the potential to bring about both cost reductions and increased productivity for public services, which are particularly important in these times of fiscal conservation.

ACKNOWLEDGMENT

This research was sponsored by the Office of Service and Management Demonstrations of UMTA. The support provided by Joseph Goodman of UMTA is particularly appreciated, as is the cooperation and advice of Amir Eiger at Rensselaer Polytechnic Institute and Joel Woodhull at SCRTD.

REFERENCES

1. M.A. Turnquist and L.A. Bowman. Control of Service Reliability in Transit Networks. Report DOT/RSPA/DPB-50/79/5. U.S. Department of Transportation, 1979.
2. A. Barnett. On Controlling Randomness in Transit Operations. Transportation Science, Vol. 93, May 1974, pp. 102-116.
3. R.L. Jackson and D. Stone. Experiments in Bus Service Control Using an Interactive Model of a Typical Urban Bus Route. University of Newcastle Upon Tyne, England, 1976.
4. D. Koffman. A Simulation Study of Alternative Real-Time Bus Headway Control Strategies. In Transportation Research Record 663, TRB, National Research Council, Washington, D.C., 1978, pp. 41-46.
5. M.D. Abkowitz and I. Engelstein. Factors Affecting Running Time on Transit Routes. Transportation Research, Vol. 17A, No. 2, March 1983, pp. 107-113.
6. P.I. Welding. The Instability of Close-Interval Service. Operational Research Quarterly, Vol. 8, 1957, pp. 133-142.
7. J.L. Doras. Irregularité des Autobus and Temps d'Attente des Voyageurs. T.E.C. (Transport Environment Circulation), 1979, pp. 13-21.
8. D.F. Loo. Evaluation of Schedule-Based Holding for Transit Vehicles: A Case Study of Bus Route 5 in Minneapolis, Minnesota. Report SS-24-U.3-201. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., 1981.
9. E.M. Holroyd and D.A. Scraggs. Waiting Times for Buses in Central London. Traffic Engineering and Control, Vol. 8, No. 3, 1966, pp. 158-160.

10. C.A. O'Flaherty and D.O. Mangan. Bus Passenger Waiting Times in Central Areas. Traffic Engineering and Control, Vol. 11, No. 9, 1970, pp. 419-421.
11. P.A. Seddon and M.P. Day. Bus Passenger Waiting Times in Greater Manchester. Traffic Engineering and Control, Vol. 15, No. 9, 1974, pp. 442-445.
12. J.K. Joliffe and T.P. Hutchinson. A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop. Transportation Science, Vol. 94, 1975, pp. 248-282.

Do Performance Audits Audit Performance?

JOHN SINDZINSKI

ABSTRACT

The requirements of the state of California for performance audits of publicly funded transit systems are examined. These performance audits are conducted by agencies that distribute state funds to support the operating and capital needs of public transportation systems. The objective of this examination is to discuss the intent of the audit requirement and how audits are conducted in order to determine the purpose of performance audits. The enabling legislation and its implementation are traced into practice and the processes used to conduct audits are critically examined. It is argued that performance audits focus on the management of transit systems at the expense of examining whether they are delivering the service required of them. In conclusion it is argued that performance audits that only evaluate how well systems perform do not fully evaluate transit performance. It is recommended that performance audits first determine whether transit systems are in fact meeting the demand for service. It is argued that performance audits that review the quality of service delivered are more helpful than those that focus solely on the management of the system.

In recent years federal, state, and local governments have become concerned about the rapid escalation in transit operating costs. Although operating costs have risen at rates equal to or greater than the overall rate of inflation, fare revenues have generally been unable to keep pace. Further large deficits in federal as well as state and local budgets have reduced the amount of funding available to support transit. As a consequence, agencies responsible for funding transit have begun to focus attention on evaluating the performance of transit systems. Such evaluations are considered useful in determining whether transit systems can become more efficient and maintain desired levels of service.

Recently several states, including New York, Pennsylvania, and Michigan, either began systematic performance audits of transit systems or were in the process of developing programs to do so. Furthermore, there has long been interest in using the annual reporting system of Section 15 of the Urban Mass Transportation Act of 1964 to conduct performance reviews. However, to date there has been no concerted government effort to do so. This is in large part because of many problems with the reliability of the data base.

Although the states just mentioned are now at the initial stages of their performance audit programs, California has been undertaking such audits for more than 6 years. In 1978, the California legislature passed a law that required all transit systems that receive state sales-tax assistance to have a performance audit conducted triennially. To date, all transit systems that have been in existence since 1979 have undergone at least two such performance audits. What the California performance audit requirement is and how audits have been conducted are examined in this paper. Although performance audits conducted throughout the state are considered, the focus is on those audits conducted for transit systems in the San Francisco Bay Area, which are within the jurisdiction of the Metropolitan Transportation Commission (MTC).

LEGISLATIVE MANDATE

In 1978 California Senate Bill 620 was passed, which amended certain provisions in the Transportation Development Act (TDA) (California Public Utilities Code, Sec. 99200, 1978). The TDA program was started in 1971 and provides sales-tax revenues to support transit systems. It is a multimillion-dollar-a-year program that serves as a major source of both operating and capital assistance to most California public transit systems. The requirement for performance audits is as follows (California Public Utilities Code, Sec. 99246):

> (a) The transportation planning agency shall designate entities other than itself, a county transportation commission, a transit development board, or an operator to make a performance audit of its activities, and those of county transportation commissions and transit development boards located in the area under its jurisdiction, with respect to these funds. The transportation

> planning agency shall consult with the entity to be audited prior to designating the entity to make the performance audit.
>
> Where a transit development board was created pursuant to Division 11 (commencing with Section 120000) or a county transportation commission exists, the board or commission, as the case may be, shall designate entities other than itself, a transportation planning agency, or an operator to make a performance audit of its activities and those of operators located in the area under its jurisdiction to whom it directs the allocation of funds. The board or commission shall consult with the entity to be audited prior to designating the entity to make the performance audit.
>
> (b) The performance audit shall evaluate the efficiency, effectiveness, and economy of the operation of the entity being audited and shall be conducted in accordance with the efficiency, economy, and program results portions of the Comptroller General's "Standards for Audit of Governmental Organizations, Programs, Activities, and Functions". A performance audit shall be submitted by July 1, 1980, and triennially thereafter.
>
> (c) With respect to an operator providing public transportation services by motor vehicles, the performance audit shall include, but not be limited to, a verification of the operator's operating cost per passenger, operating cost per vehicle service hour, passengers per vehicle service hour, passengers per vehicle service mile, and vehicle service hours per employee, as defined in Section 992457. The performance audit may include consideration of the needs and types of the passengers being served.

Thus the broad objectives of these performance audits are set forth. The principal purpose of these audits is to evaluate the efficiency and effectiveness of the transit system. Secondarily the audits are to verify a set of performance indicators that are to measure both efficiency and effectiveness. It is not specified how these audits are to be conducted except that they are to be done by independent auditors and that the Comptroller General's standards are to be adhered to. Beyond that the legislature has delegated full responsibility and authority for performance audits to the various planning agencies located throughout the state that administer the TDA program.

GENESIS OF THE REQUIREMENT

In 1977 the Auditor General for the state of California issued a report in which it was found, among other things, that there was no system for routinely measuring the efficiency and effectiveness of California transit systems (1). It was recommended that legislation be enacted that would require that each system undergo an evaluation. The Auditor General recommended an evaluation program that includes both performance indicators and specific comments on the efficiency and effectiveness of the operational components that affect these indicators.

The reason for this dual approach was that the Auditor General believed that indicators alone cannot be used to evaluate transit performance. This is because each transit system operates in different environments with different characteristics and audits must be sensitive to these differences. The Auditor General concluded that interoperator comparisons are difficult to make and, because of the unique circumstances under which each system must operate, are of doubtful utility. This reluctance to use interoperator comparisons has been a concern of the industry for many years. It is a point that has been stated often to those involved in performance audits.

TYPE OF APPROACH

The focus of performance audits is on evaluating the efficiency, effectiveness, and economy of the transit system's operation. The legislature clearly intended that the mandated performance measures be used to this end. However, the law was not specific in defining these measures, so the U.S. General Accounting Office (GAO) has defined two of them (2):

> Economy and Efficiency--determines whether the entity is managing or utilizing its resources (personnel, property, space, and so forth) in an economical and efficient manner and the causes of any inefficiencies or uneconomical practices, including inadequacies in management information systems, administrative procedures, or organizational structure.

In addition, the California Auditor General defines effectiveness specifically for transit as (1) "the measure of how well the system meets the needs of the residents of the area it serves."

It is important to note that the law focuses on an approach based on efficiency and effectiveness as opposed to a program-results approach to evaluation. The program-results approach is defined by GAO as follows (2):

> Program Results--determine whether the desired results are being achieved, whether the objectives established by the legislature or other authorizing body are being met, and whether the agency has considered alternatives which might yield desired results at a lower cost.

What is noteworthy about this focus is that the approach based on efficiency and effectiveness is more concerned with how the job gets done than the program-results approach, which focuses on whether the job gets done according to plan.

WHAT ARE PERFORMANCE AUDITS SUPPOSED TO DO?

The foregoing review of the enabling legislation and its genesis as well as the conceptual methodology give some indication as to what performance audits are supposed to do. The state legislature emphasized using outside independent and objective auditors. This suggests that the audits are to be candid and complete honest reviews of how well systems are being managed. This is tempered to some degree by the requirement that the systems being audited be consulted before the audit. Although the purpose and scope of this consultation are not elaborated in the law, it is clearly meant that the operators are to have a role in their audit before it takes place.

The audit requires that five performance measures be verified that are intended to provide a minimum level of information by which the systems can be audited. Although the ability and appropriateness of the measures may be questioned, it is important to realize that they require information (such as number of passengers carried and hours and miles of service provided) that is absolutely essential to the good management of a system. However, these measures are, at best, systemwide and provide only broad trends in performance. As such they are of limited use as a diagnostic tool in that they can only suggest where there may be problems. Therefore, the audits must go beyond a simple review of performance measures and determine why the measures have performed as they did.

WHAT PERFORMANCE AUDITS DO NOT DO

Although the task of the auditor as detailed earlier is quite broad and complex, there are certain things a performance audit is not intended to do. First, the performance audits do not evaluate the policies of transit systems, which are set by the board of directors. Second, these audits do not question whether a particular transit system is a worthy public endeavor and whether it should continue to provide service. Finally the audits are not meant to determine whether transit systems are accurately and correctly accounting for their funds. This is the objective of a financial audit. The verification of the five performance measures is limited to a review that checks whether the system is correctly defining the data and whether there are systems in place to calculate the measures in accordance with standards.

FROM LAW TO PRACTICE

As noted earlier, the legislature left it to each planning agency to implement the performance audit requirement. In 1978 the California Department of Transportation adopted regulations that provided some additional guidance to planning agencies as to what the audits were to accomplish (California Administrative Code, Sec. 66645.5):

> Performance Audits and Operators. A performance audit shall be made and submitted by July 1, 1980 and by July 1 triennially thereafter for each operator. Beginning with allocations for the 1981-82 fiscal year, no operator shall be eligible to receive an allocation under Article 4 of the Act until the transmittal of reports of the performance audit that is to be submitted by July 1 of the fiscal year prior to fiscal year of the allocation.
>
> The performance audit shall be made pursuant to Public Utilities Code Section 99246. The evaluation of the performance of the operator shall include but not be limited to:
>
> (a) The degree to which the management has established overall system goals and objectives and the degree to which these goals and objectives are being accomplished.
>
> (b) The manner and extent to which management seeks to improve the effectiveness and efficiency of its transit services by developing transit plans that are responsive to user needs.
>
> (c) The manner and extent to which management addresses the effectiveness and efficiency of the transit system's operations by developing, analyzing and acting upon information about specific performance measures.
>
> (d) The manner and extent to which management addresses effectiveness and efficiency of vehicle maintenance, complies with vehicle safety regulations and evaluates general maintenance activities and progress against established objectives and standards.
>
> (e) The manner and extent to which management addresses the effective and efficient conduct of marketing and public relations activities.
>
> (f) The manner and extent to which the budgeting and financial planning process reflects the goals and objectives for the effectiveness and efficiency of the transit system's operations.

Although these regulations provide specific aspects (such as operations, maintenance, safety) that audits are to evaluate, they provide little insight into how the audits are to be accomplished.

Not surprisingly, the first round of audit reports, due in 1980, showed little consistency in approach, style, or results. Some audits were complex reports that reviewed nearly every function of a large system. Others simply reported the five measures and included a broad statement attesting to the efficiency and effectiveness of the system audited. As to approach, some audits included interoperator comparisons of performance and included many more indicators than the five legally mandated. Others were limited to a trend analysis of the five mandated measures.

Only a few planning agencies had developed systematic audit programs for the 1980 audits. For example, the San Diego Association of Governments and MTC (for the San Francisco Bay Area) developed extensive guidebooks that specify the purpose of performance audits and lay out several approaches. The MTC guidebook, for example, is a two-volume set that identifies three approaches by which the audits may be conducted:

1. Key issues: Issues of importance to the effective and efficient operation of the system are identified by MTC for audit review.
2. Goals and objectives: It is recommended that actual performance be related to planned performance and that planned performance objectives and standards be reviewed for reasonableness.
3. Functions: Specific activities or functions (such as maintenance, purchasing, scheduling, and so on) typically performed by an operator should be reviewed.

In the MTC region (which includes nine San Francisco Bay Area counties) there are more than 20 transit systems for which performance audits must be conducted. The systems range in size from rural dial-a-ride services with one bus or van to the San Francisco Bay Area Rapid Transit District and Municipal Railway. With such a large number of operators and a wide range of complexity, MTC has had the opportunity to use all three approaches and various combinations thereof in conducting these audits.

Currently MTC uses the functional approach, for the most part, in its performance audits. This approach has been refined somewhat from the original one as set forth in the audit guide. Specifically, performance audits are broken down into two distinct phases. The first, or preaudit, phase, is a high-level examination of all functions or activities of a transit system. In addition the preaudit phase includes an examination of the five mandated performance measures. The goals of the preaudit phase are twofold: One is to collect basic information about the system and to identify issues or problems that may warrant in-depth analysis. Secondly, the preaudit phase is also intended to provide the auditor with sufficient information to verify the performance measures.

The preaudit phase culminates in a report that sets forth the auditors' preliminary findings and suggestions as to what aspects and specific issues warrant analysis in the second, or performance audit, phase. This report is reviewed by both MTC and the transit operator. On completion of this review process, which includes significant dialogue among MTC, the operator, and the auditor, the final audit work plan is developed. This work plan sets forth what issues are to be evaluated and what questions are to be answered. The purpose of this preaudit report and its review is to provide the means by which the audit can be focused and concentrated into specific and tangible aspects.

MTC has used this approach in all the major operator audits conducted over the last 3 years. The reasons for focusing the in-depth analyses are important. First, because of financial, time, and personnel constraints, the auditor cannot thoroughly and adequately review a large and complex transit system such as Alameda-Contra Costa Transit, which has an annual budget of $95 million and operates some 600 buses during the peak hour. Furthermore, it is doubtful whether a comprehensive audit would be worthwhile except to verify that some functions are operating well and with few problems. By focusing the audit, MTC hopes that those issues and problems that are most critical will be examined.

HOW IS PERFORMANCE ANALYZED?

Earlier in this paper, reference was made to interoperator comparisons and the problems associated with that diagnostic tool. In audits done several years ago for MTC, interoperator comparisons were used to evaluate performance. Since then transit operators have strenuously objected to this technique and today such comparisons are not allowed by MTC. This prohibition is quite common in California. However, in some regions comparisons have been made against the industry as a whole. Again MTC does not allow such comparisons because the yardstick--the industry average--is more amorphous than a comparison with a particular system. The industry average is only a statistic, and a comparison to an average is dubious at best. This prohibition obviously constrains the auditor. However, MTC has another practice that further constrains the auditor.

The prohibition is against using any measure to evaluate performance that is not embraced by the operator. MTC believes that audits should be evaluations of actual performance against standards set by the operator. Any other performance measurement may result in criticisms that have little bearing on what the operator recognizes as a problem or goal. Further, MTC actively encourages and supports transit systems in preparing transit development plans that specify goals, objectives, and standards.

Performance then is to be judged in this process by examining trends in the five mandated measures and any other measures the operator may use. Although the basis for the evaluation is goals and objectives, the analysis is not intended to be only a report as to whether the goals and objectives were attained. Rather the auditor is to determine why certain goals and objectives were attained and others were not. The focus on goals and objectives is meant to ensure that operators not be audited against standards or measures that are inappropriate or misleading. The use of performance measures by themselves can be especially misleading in that there are never enough measures (or base data from which they may be derived) to really explain what may be occurring in a particular function or subfunction.

The MTC approach to audits closely follows the original idea that the state Auditor General had for transit system audits. Specifically the Auditor General recommended an audit program that looked at performance measures as a tool to uncover problems and then a commentary on the efficiency and effectiveness of those components that affect the measures.

In contrast to MTC's approach is that of the state performance audit guide, which is a nonbinding audit program for small and medium-sized (up to 500 buses) operations. This guide was developed in 1982 after the first round of audits had been finished. Several small planning agencies requested that it be developed to serve as a source document for them to use in their regions to develop uniform audit reports. The approach set forth in this document is almost totally a performance-measure approach. For small systems (1 to 20 vehicles), the audit is predicated on a review of a set of measures followed by interviews about goals and objectives with management. The resulting report is a high-level one with little detail and few recommendations. For larger systems the state guide uses a preaudit and an audit phase as does MTC. However, whereas MTC's preaudit phase is based on interviews, document reviews, and site visits, the approach of the state guide relies on the use of measures to analyze performance. Likewise, the approach of the state guide uses many measures in the audit phase, whereas the MTC approach continues to rely more on interviews with management and other operator personnel.

ANALYSIS OF THE METHODOLOGIES

It can be argued that the MTC approach with its many prohibitions is a conservative one in that it is quite process oriented. The MTC approach places a great burden on the auditor to determine how well a system is performing and what its problems are within the context of what the operator has defined as his goals, objectives, and standards. Without using intersystem comparisons and measures not used by the operator, the auditor must quickly and accurately identify issues and develop justifications as to why they are significant enough to evaluate.

On the other hand, the approach of the state guide provides quick and easily verifiable information as to how a particular component of a system is working. Once there is agreement on the accuracy of the data used to derive the measures, the auditor can make objective judgments as to efficiency and effectiveness.

However, the performance-measure approach has its limitations. For example, it cannot tell the auditor why the system and its components are performing as indicated. Nor can it truly tell the auditor whether performance is good or bad. Measures, per se, do not judge performance. Rather they provide some basic information that must be put into a context that assesses whether the system is operating efficiently and effectively. There is an easy trap to fall into when audits rely exclusively on measures to judge performance. The trap is quite simply that numbers are merely information and not answers.

The MTC approach avoids this trap by making the auditor go behind the numbers to determine how well things are or are not working. Although the operator may be able to hide poor performance by judicious use of objectives and standards, the MTC auditor does not rely on standards alone. Instead, the auditor interviews staff, reviews documents, and visits facilities to develop conclusions. This is a conservative approach, but it tends to be more honest in that it recognizes the certain fallacies of an audit approach predicated on performance-measure analysis.

The MTC approach does not allow either the operator or the auditor distant reality by selected use of measures. Therefore, it should be more adept at identifying the causes of problems, assuming that the operator cooperates and that the auditor is competent.

AUDITOR AS EXPERT

Perhaps the most critical issue regarding the question of whether performance audits do in fact audit performance is this: "Who is the performance auditor?" The obvious assumption in the enabling legislation and in the auditing profession is that the auditor is an expert with, perhaps, special knowledge of how things are supposed to work. The auditor is seen as someone who can use an audit program to identify problems, determine their causes, and develop solutions.

If in fact it is believed implicitly or otherwise that the auditor is an expert, the next question is, "What sort of expert is this auditor?" Earlier in this paper it was argued that performance audits focus on how well things are done rather than on program results. Therefore, it is logical to conclude that the performance auditor is an expert on how organizations are supposed to perform. Or, more explicitly, auditors are experts on how well transit systems are managed and whether management is directing the system and its components to function efficiently and effectively.

If the argument just stated is sound, the inescapable conclusion is that what performance audits do is not to audit performance in the broadest possible interpretation of the term, but rather to audit the management of transit systems. It has been argued in this paper that performance cannot be judged solely by numbers, but that a context must exist within which performance indicators can be used to provide a diagnosis of what may be happening in a transit system. It has also been argued that a critical part of this is the human element--the auditor. To accomplish this, it is required that the auditor be an expert, a person with special knowledge who can judge whether things are operating well or not.

Obviously the level of expertise needed from the auditor can vary considerably. For instance, it takes little expertise to know that an operations division is having difficulties in providing service when the auditor has experienced waiting for a bus that did not show up a few times. It is clear that the schedule is not being met but not why. Is it an operator problem, a routing or scheduling problem, or a maintenance problem or are there simply circumstances beyond anyone's control? These are the types of questions performance auditors need to answer. Therefore, it appears that the auditor must be many experts--schedule, maintenance, operations, planning, and financing--to determine why a system cannot do its job well.

Given this conclusion, what then do audits reveal? Current thought and statutes have created an expectation that auditors are experts who can determine why things are not working well (if they are not) and how to fix them. But as currently constructed, performance audits evaluate process and not results. All operators have goals relating to the essential purpose of transit--to carry people from one place to another. However, the current practice of performance auditing is almost exclusively restricted to evaluating whether service is being provided in an economical manner.

Performance audits rarely determine whether transit systems are providing the type and amount of service needed at the right time. Furthermore, audits do not tell anything about the untapped market of those who are not using transit and why they are not using it.

The fundamental issues facing the transit industry are to maintain the current ridership and to find new riders. Performance audits oriented toward management do not wrestle with this issue. As currently practiced, audits can tell how well a system is fulfilling its prime objective. For example, a system that carries few people but maintains its schedule and operates in an efficient manner might be seen as a good system by the auditor. As such, the audit has not revealed whether anyone is using the system and if not, why not. Until these fundamental questions are considered, performance audits are not fully auditing performance. The first issue an audit might address is whether the system is being used and then to what extent. After these issues have been resolved, the audits might then examine the reasons behind underuse. At that time the auditor should focus on how well the system is managed to determine whether that is the cause of the low use. Critical components that would be addressed at that stage are the marketing and planning functions. It is conceivable that the operator does not know what and where the market is and how it can be reached.

For those systems that suffer from overuse the primary focus of the auditor would have to be somewhat different. In such cases the auditor should examine how capacity can be increased in the most economical manner. It can also be argued that from a program-results standpoint there is little need for the audit. The danger with this arrangement is that a shift too far toward a results-only approach would also be narrow. It would ignore other performance problems. The ideal compromise then should be performance audits that look at both process and results with equal attention.

REFERENCES

1. Financing and Evaluating Public Transit Systems in California. Office of the Auditor General, California Legislature, Sacramento, 1977.
2. Standards for Audit of Governmental Organizations, Programs, Activities, and Functions. U.S. General Accounting Office, 1972.

Peer Comparisons in Transit Performance Evaluation

MANOUCHEHR VAZIRI and JOHN A. DEACON

ABSTRACT

A methodology by which to group urbanized areas for the purpose of peer comparisons in transit performance evaluation is presented. A suitable basis for grouping was found to be those market and environmental variables that effectively constrain attainable performance levels. By using U.S. Bureau of the Census data for 1980, homogenous clusters of urbanized areas were formed and the key market and environmental variables were reduced by means of factor analysis to one size index. Reporting-system data as outlined in Section 15 of the Urban Mass Transportation Act of 1964 were used both to evaluate adequacy of the grouping scheme and to establish attainable target performance levels. General relationships were observed between the mean transit performance of the peer groups and their mean size indices. It was concluded that regression models were the most effective way to eliminate the effects of market and environmental dissimilarities in establishing target performance levels. Models relating individual performance measures to significant market and environmental variables were calibrated for each peer group.

Comparative studies form an indispensible component of transit performance evaluation. Such studies include comparisons of performance changes over time, comparisons of actual performance with preestablished target levels, and comparisons of the performance of a subject system with that of other similar systems. This last type of study, sometimes termed peer comparison, is frequently used in the establishment of feasible performance targets. However, it is also used for other purposes such as evaluation of the effectiveness of management and has even been suggested as a basis for distribution or allocation of financial aid (1-4).

Regardless of purpose, peer comparisons offer great promise in the quest for improved transit performance. In the past, however, their application has been hampered by two restraints. First, detailed performance data were not consistently and uniformly reported by transit agencies. Second, procedures for the formation of reasonable peer groups were not well understood. Implementation of a uniform public mass transportation reporting system, mandated by Section 15 of the Urban Mass Transportation Act of 1964, as amended, has been instrumental in eliminating the first difficulty. The second is the primary subject addressed here.

The objectives of this research were primarily to

1. Develop a methodology for the classification of transit systems that would be useful for enlightened peer comparisons and
2. Apply that methodology to systems throughout the United States.

This research was part of a more comprehensive study of transit performance that has been reported elsewhere (5).

STUDY APPROACH

Central to the peer-comparison concept are two important notions, namely, that

1. All transit systems cannot be expected to achieve the same high level of performance and
2. The subject system can, with proper action by management, achieve performance levels demonstrated by the best within its peer group.

The first of these notions establishes one of the necessities for the formation of homogenous groups in peer comparisons. The second suggests a basis for selection of systems to make up each of the groups, namely, that systems within a given group should have the same attainable or potential levels of performance. Because potential performance levels cannot be measured directly, the formation of peer groups is not at all a straightforward process.

It is helpful, however, to understand that actual performance levels are dictated or determined by both controllable and uncontrollable variables; the distinction is made on the basis of whether the determining variables are within or beyond the influence of the transit provider. In this sense the controllable or endogenous variables are those manipulated by the provider to influence performance: They include such examples as fare, routing, maintenance, and vehicle replacement strategies. These are also sometimes termed policy variables. Uncontrollable or exogenous variables have no lesser effect on transit performance, but at the same time they cannot be reasonably manipulated by the transit provider. Uncontrollable variables can be exemplified by such diverse characteristics as size of population served, development density, automobile ownership, and extent of freeway development.

It is hypothesized that the level of performance attainable by a subject system is theoretically constrained by the uncontrollable variables, which reflects primarily the nature of the market served and the environment within which the service is provided. When these conditions have been identified, the controllable variables can be set at levels that will enable performance to reach its potential limits. Because potential performance cannot be directly measured, the formation of peer groups should be based solely on those uncontrollable market and environmental variables that significantly influence transit performance. This finding had a major impact on the structure of the data base used here and largely dictated the approach taken in the grouping or clustering procedure.

The first phase of this study was to create a data base from which performance and market and environmental data could be extracted for transit systems throughout the United States. Then, by means of factor analysis, these data were reduced to a simpler, nonredundant dimension. The reduced market and environmental data were next used with another multivariate statistical procedure, cluster analysis, to form homogenous peer groups. By means of analysis of variance, the resultant peer groups were analyzed to determine their similarities and dissimilarities. Finally, target performance levels were developed for each peer group.

DATA BASE

The nationwide scope and limited resources of this study required that centralized data sources be used. The only reasonable source of transit performance data was the Section 15 reporting system. Annual data, as reported for fiscal years ending between July 1, 1979, and June 30, 1980, were used (6). The most reasonable source of market and environmental data was the U.S. Census. Original plans were to use only 1980 data. However, reporting delays required some 1960 and 1970 data to be forecast to 1980.

The merger of Section 15 and census data required a common unit of analysis, that is, a comparable level to which data could be aggregated. Census data are reported both for various levels of governmental jurisdiction and for various geographical levels. Section 15 data, on the other hand, are reported only by transit operator. Limits of the transit-served region are not accurately defined in Section 15 reporting and there is often more than one operator in a particular geographic region. Given this situation, the most reasonable unit of analysis was judged to be the urbanized area, and where necessary transit data were aggregated to this level. The urbanized area is indicative of the entire transit-serviceable region and should serve as a better unit for comparison than that used in most peer comparisons, that is, the region often ill defined, served by the individual transit operator. Only 188 of the 366 urbanized areas were included in this study. The remainder were excluded either because they were not served by transit or because transit data had not been adequately reported.

Development of transit performance measures was an important task, which is described in detail elsewhere (5). The 25 measures that were ultimately used are identified, together with their means and standard deviations, in Table 1. Suffice it to say here that this list reflects the major performance dimensions available from the Section 15-census data base. It also reflects many of the major performance variables that have been used by other researchers and practitioners.

A composite measure of system performance was also developed to expedite the analysis. Termed overall sum of the Z scores (OSZ), this normalized variable reflects an equal weighting of those six dimensions of performance identified in Table 1, namely, output and input, consumption and input, and so on. In constructing OSZ, accidents per revenue vehicle mile (SCSO14) and the input-market variable set (SIMEs) were treated negatively. Increases in the levels of all other variables were taken to be indicative of improved performance.

Because the peer groups of urbanized areas were to be formed in consideration only of the exogenous market and environmental (ME) variables, selection of these variables was critical. The ME variables had to be significantly related to potential transit

TABLE 1 Performance Variables

Symbol	Variable Name	Mean	Standard Deviation	Unit
Output and input				
SOSI3	Rev Veh Hr/Opr Exp	0.047	0.024	Vehicle hours per dollar
SOSI5	Rev Veh Mi/Rev Veh	27,648.965	7,779.455	Vehicle miles per vehicle
SOSI8	Rev Veh Hr/Transit Empl	1,057.533	265.061	Vehicle hours per employee
SOSI12	Rev Veh Hr/Equiv Gal Gas	0.313	0.196	Vehicle hours per gallon
SOSI13	Pass Cap Mi/Rev Veh	1,660.531	1,242.439	1,000 passenger miles per vehicle
Consumption and input				
SCSI4	Rev/Rev Veh	54,203.054	24,669.277	Dollars per vehicle
SCSI6	Pass/Opr Exp	1.397	0.846	Passengers per dollar
SCSI8	Pass Mi/Opr Exp	4.840	3.684	Passenger miles per dollar
SCSI16	Rev/Veh Opr Cost	1.939	1.306	Dollars per dollar
SCSI18	Pass Rev/Opr Exp	0.311	0.124	Dollars per dollar
Consumption and output				
SCSO1	Pass Mi/Pass Cap Mi	152.633	99.561	Passenger miles per 1,000 capacity miles
SCSO3	Pass/Rev Veh Mi	2.662	2.492	Passengers per vehicle mile
SCSO9	Pass/Transitway Length	22,334.092	27,205.245	Passengers per mile
SCSO14	Accidents/Rev Veh Mi	77.357	93.707	Accidents per million vehicle miles
SCSO15	Rev Veh Mi/Tot Road Calls	4.132	9.840	1,000 vehicle miles per call
Input and market				
SIME1	Opr Exp/Pop	17.607	16.182	Dollars per person
SIME2	Opr Assistance/Pop	10.563	9.489	Dollars per person
SIME5	Rev Veh/Pop	0.314	0.174	Vehicles per 1,000 persons
SIME6	Transit Empl/Pop	0.646	0.427	Employees per 1,000
Output and market				
SOME1	Pass Cap Mi/Pop	540.113	508.906	Passenger miles per person
SOME2	Rev Veh Mi/Pop	8.443	4.974	Vehicle miles per person
SOME8	Transitway Length/Area	3.048	3.210	Miles per square mile
Consumption and market				
SCME1	Pass/Pop	23.341	25.290	Passengers per person
SCME6	Pass Mi/Area	215.253	343.107	1,000 passenger miles per square mile
SCME12	Avg Trip Dist/Equiv UA Radius	0.637	0.507	Miles per mile

performance, otherwise proper interpretation of the results of peer comparisons would be difficult or impossible. Some assistance in the selection was found in the literature (4,7,8), but to a great extent selection of the 41 ME variables of Table 2 was based on the authors' judgment. The rather large number of ME variables assures, it is hoped, that the critical market and environmental dimensions, as they can be extracted from census data, have been included.

FACTOR ANALYSIS OF ME VARIABLES

High correlations between pairs of ME variables could potentially cause a significant bias in groupings based thereon. For example, because housing

TABLE 2 ME Variables

Symbol	Variable Name	Mean	Standard Deviation	Unit
Automobile Ownership				
ME1	Avg Persons/Auto	2.088	0.379	Persons per automobile
ME30	Workers/Auto	0.998	0.246	Workers per automobile
ME31	Autos/Hsld	1.376	0.213	Automobiles per household
Urbanized Area Size				
ME2	1980 Pop	495.379	894.967	1,000 persons
ME7	Area	202.498	261.069	Square miles
ME37	Housing Units	193.505	344.298	1,000 units
Income of Residents				
ME3	Families < $5,000	6.680	1.865	Percent
ME4	Families > $10,000	75.847	12.973	Percent
ME13	Median Family Income	21,017.187	2,793.683	Dollars per year
ME32	Families < Low Income	7.605	2.745	Percent
Age of Residents				
ME5	Pop < 18 Yr	27.039	2.941	Percent
ME6	Pop > 65 Yr	11.264	3.123	Percent
ME15	Pop < 5 Yr	7.092	1.142	Percent
ME35	Median Age	29.855	3.027	Years
Occupation of Residents				
ME18	Pop in School	21.498	2.072	Percent
ME22	Pop Employed	40.029	3.655	Percent
ME23	Pop in College	5.535	3.112	Percent
ME24	Empl in Manufacturing	21.859	8.857	Percent
ME25	Empl in Sales	22.603	2.890	Percent
ME26	Empl in Construction	5.628	1.789	Percent
ME27	Empl in Government	15.575	6.287	Percent
Education of Residents				
ME20	Median School Yr Completed	12.544	0.349	Years
ME21	Pop Completed College	14.547	4.879	Percent
Gender and Race of Residents				
ME28	Civ Labor Female	19.267	2.625	Percent
ME29	Pop Female	51.964	1.134	Percent
ME34	Pop Nonwhite	16.701	11.407	Percent
Housing				
ME14	HU Renter-Occupied	36.014	5.914	Percent
ME36	Hsld Size	2.662	0.156	Persons per household
ME38	HU Single Unit	71.759	8.151	Percent
Land Use Distribution				
ME8	Area in Central City	43.988	25.192	Percent
ME9	Pop Density	2,271.002	766.456	Persons per square mile
ME10	Land Area	95.227	8.802	Percent of total area
ME16	Housing Density	881.812	291.949	Housing units per square mile
ME33	1900 Pop	61,521.074	99,014.069	Persons
ME39	SMSA Pop in UA	78.868	11.471	Percent
ME40	UA Pop in Central City	59.826	23.387	Percent
ME41	Pop Density Central City	3,851.457	2,356.825	Persons per square mile
Growth				
ME17	Pop Growth 70-80	14.169	26.863	Percent
ME19	Housing Growth 70-80	38.090	23.809	Percent
Climate				
ME11	Avg Jan. Temp	34.162	12.317	Degrees Fahrenheit
ME12	Annual Rainfall	35.907	12.265	Inches

TABLE 3 Factor Dimensions of ME Variables

Symbol	Variables with High Positive Loading	Variables with High Negative Loading	Interpretive Meaning of Factor Dimension	Percentage of Variance Explained by Factor
F_1	Families < \$5,000 (low income) Families < Low Income Avg Jan. Temp Housing Growth 70-80 Empl in Construction	Families > \$10,000 (high income)	Poverty	16.9
F_2	1980 Pop Area 1900 Pop HU	UA Pop in Central City	Size	13.5
F_3	Pop < 18 Yr Pop < 5 Yr	Pop > 65 Yr Median Age	Youthfulness	10.5
F_4	Pop in College Pop Completed College HU Renter-Occupied	Empl in Manufacturing	Education	9.3
F_5	Autos/Hsld	Avg Persons/Auto	Automobile availability	6.3
F_6	Pop Density Housing Density Land Area		Density	4.7

units and population are highly correlated, including both in the cluster analysis is equivalent to counting the effect of size twice. To eliminate such possible problems yet still retain all of the important ME dimensions, factor analysis was used. The Factor Analysis Program of the Statistical Package for the Social Sciences (9) with varimax rotation and a minimum eigenvalue of 2 was found effective in reducing the 41 ME variables into 6 factor dimensions that all together explained 61 percent of the total variance. In Table 3 the dimensions of these factors are identified and interpreted. The urbanized areas are thus characterized by the six dimensions of poverty, size, youthfulness, education, automobile availability, and density of development. These are clearly distinguishing characteristics that intuitively seem to be the most important determinants of potential transit performance.

PEER GROUPS

All clusters or peer groups were formed by using the Biomedical Computer Program cluster analysis with K-mean clustering (10). In this algorithm the Euclidean distance was used as a measure of the deviation of an individual case from the cluster mean. Initially all cases were considered in one cluster. With each succeeding iteration, a new cluster was formed until the requested number of clusters had been reached.

In each cluster analysis, the 188 urbanized areas were divided into two sets, one of 150 and the other of 38. The smaller set was considered to be a homogeneous peer group and was not included in the clustering. Such a procedure was required because some of the ME data were missing. Because each of the 38 areas had become newly classified as urbanized in 1970 or 1980, reasonable estimates of the 1980 forecast variables were not available and hence these variables were treated as missing. A complete set of data was available for the larger group.

Also in each cluster analysis, the number of clusters for the remaining 150 areas was preselected at 10. This was intuitively judged to be sufficiently large to assure the necessary intragroup homogeneity while retaining, on the average, a sufficient group size to permit intragroup statistical analyses. Actually, in preliminary analyses, 6, 8, and 10 groups were investigated. Both the maximum group size and the proportion of sparse groups were judged excessive when the number of groups was less than 10.

FORMATION OF GROUPS

Four different schemes for clustering the urbanized areas were subjected to detailed analysis. Two were based on the use of ME data, one was based on transit performance data, and one was based on a combination of ME and performance data.

In the first scheme, formation of peer groups was based on the six ME factors, as identified in Table 3. The 10 groups are described in Table 4.

The clusters that had been thus formed were intuitively appealing. However, there were inconsistencies within the groups when measures of transit activity and performance were examined. It was reasoned that perhaps the six independent factor dimensions were not of equivalent importance in their effects on transit and that a second set of peer groups should be formed on the basis of the singularly most significant factor.

In order to determine which of the six factors was most significant, correlation coefficients were computed between each of the six factor dimensions and summed, normalized scores (SZ scores) for each of the six sets of performance measurements as identified by Table 1. In Table 5, which gives a portion of the resulting correlation matrix, the second factor, that relating to size, is most significantly correlated with performance.

Accordingly, the size factor (F_2) was then used as the basis for a second clustering of the urbanized areas. The 10 groups were clearly distinguish-

TABLE 4 Peer Groups Based on ME Factors

Group	Distinguishing Characteristic
1	Large cities located throughout the United States
2	Florida cities with large retired populations
3	Southern cities with youthful, low-income, low-automobile-owning residents
4	Northeastern cities with older residents
5	Low-density cities predominantly in Midwest
6	Low-density and low-automobile-owning cities
7	Automobile-dominated cities of West
8	Average cities with younger residents predominantly in Midwest
9	Low-income cities of the South and West
10	Small university cities with highly educated young residents

TABLE 5 Correlation Matrix of Six Factor Dimensions of ME Variables with Six Dimensions of Transit Performance

Factor Dimension	SZ_1	SZ_2	SZ_3	SZ_4	SZ_5	SZ_6
F_1 (poverty)	NC	NC	NC	−0.290	−0.262	NC
F_2 (size)	NC	0.401	0.324	0.583	0.476	0.185
F_3 (youthfulness)	NC	NC	NC	−0.142	NC	NC
F_4 (education)	NC	0.152	0.237	0.329	0.206	NC
F_5 (automobile availability)	NC	−0.213	−0.206	NC	NC	NC
F_6 (density)	NC	NC	NC	NC	0.260	0.174

Note: NC = not correlated at level of significance of 0.05. Performance variables: SZ_1 = sum of normalized output-input variables; SZ_2 = sum of normalized consumption-input variables; SZ_3 = sum of normalized consumption-output variables; SZ_4 = sum of normalized input-market variables; SZ_5 = sum of normalized output-market variables; SZ_6 = sum of normalized consumption-market variables.

able from each other in terms of size variables such as population and area. Subjective analysis of the adequacy of this clustering scheme was favorable.

Clustering of urbanized areas based on transit performance has little utility in most peer comparisons. The motivation for comparing a subject urbanized area with others of similar performance is not compelling. At the same time, a comparison of performance clusters with market clusters offered potential for revealing new insights. Therefore, performance clusters were formed by using the overall sum of Z scores (OSZ) as a basis. The groups that were so formed were clearly different in their overall level of transit performance.

To complete the analysis, a final grouping was developed on the combined basis of overall performance (OSZ) and market (F_2). These 10 groups can be described as shown in Table 6.

TABLE 6 Peer Groups Based on Size Factor

Group	Size	Performance
1	Below average	Above average
2	Below average	Below average
3	Below average	Well below average
4	Below average	Average
5	Below average	Well above average
6	Above average	Below average
7	Above average	Above average
8	Above average	Well above average
9	Well above average	Average
10	Well above average	Above average

ANALYSIS OF GROUPS

Four complete sets of peer groups had been formed by using the four different bases for clustering, namely, six factors (market), F_2 (size), OSZ (overall transit performance), and the combination of F_2 and OSZ. The four sets were markedly different, and two additional procedures remained to be carried out. First was a determination, for each of the four sets, of the transit characteristics that were different among the groups. Second was the selection of one of the four sets for more-detailed analyses.

Previously described transit characteristics that were available for testing included the 25 performance variables. In addition there were 18 system-input (SI) variables, reflecting the resources dedicated to transit and generally including labor, capital, and operating monies; 15 system-output (SO) variables, reflecting the level of transit service that is produced and generally including vehicle miles, vehicle hours, and capacity miles; and 12 system-consumption (SC) variables, reflecting utilization of the service and including passengers, passenger miles, and revenue of various types. Analysis of variance was used to determine which of these variables were significantly different among the groups for each clustering scheme. Results of the analysis are summarized in Table 7. The larger entries in Table 7 are statistically preferred because they indicate a larger percentage of variables that are significantly different among the groups and hence a more discriminating clustering scheme.

TABLE 7 Percentage of Difference Among Groups of Variables at 0.05 Level of Significance

	Percentage by Clustering Basis			
Category of Variables	Six Factors	F_2	OSZ	F_2 and OSZ
SI (system input)	77	83	39	67
SO (system output)	86	87	67	87
SC (system consumption)	74	50	75	75
Performance	60	60	68	84
All	72	70	61	78

In addition to the summary statistics of Table 7, five consumption variables, including passengers, total revenue, passenger revenue, passengers per capita, and passenger miles per square mile, were judged to have special significance. Each of the four clustering schemes was successful in forming groups that differed with respect to these five variables. Because the number of groups had been held constant at 10, the degree of success is indicated by the η^2-statistic. This statistic assumes a minimum value of 0.0 if the grouping has been completely unsuccessful in reducing variability in the chosen measure and reaches a maximum value of 1.0 when all intragroup variability has been eliminated. More effective clustering techniques thus yield larger values of η^2. In Table 8 the η^2-statistic is summarized for the five selected consumption variables.

As indicated by the data of Tables 7 and 8, the clusters of F_2 and those of the combined F_2 and OSZ are superior to those based on the other two

TABLE 8 η^2 of Selected Consumption Variables

		η^2 by Clustering Basis			
Symbol	Variable Name	Six Factors	F_2	OSZ	F_2 and OSZ
SC1	Pass	0.517	0.844	0.137	0.655
SC2	Rev	0.618	0.916	0.111	0.694
SC12	Pass Rev	0.522	0.884	0.122	0.700
SCME1	Pass/Pop	0.295	0.522	0.224	0.532
SCME6	Pass Mi/Area	0.234	0.371	0.412	0.601
Avg		0.437	0.707	0.201	0.636

Note: η^2 = between-group sum of squares divided by total sum of squares.

schemes. Objectively, the choice between clusters based on F_2 and those based on F_2 and OSZ is a mixed one. At the same time, clustering based on F_2 is simpler and more comprehensible. Furthermore, it supports the critical notion that ME variables in themselves largely dictate transit potential. For these reasons, clustering based on the size factor (F_2) was chosen as the preferred basis for further investigation. The resulting clusters are identified in Table 9. Distributional characteristics of the F_2 scores for these groups are summarized in Table 10.

COMPARATIVE TRANSIT PERFORMANCE

One objective in peer comparisons of transit performance is establishment of feasible performance targets. The assumption is that the subject system could achieve performance levels demonstrated by others of its peer group if the proper policy decisions were made. If the peer group were truly homogeneous, the subject system could even reach performance levels at least as good as the best demonstrated within the peer group. Most peer comparisons use average performance as the target, however, and this convention seems reasonable given the uncertainty in identifying truly homogenous peer groups. A three-level structure for comparative analysis is developed here, each level of which requires the use of averages.

The first level is an uncontrolled comparison. The peer group simply represents the set of all 188 urbanized areas, and the average performance levels are used in establishing first-cut targets. For some systems, however, such averages will be unrealistic targets because of ME constraints. For others, ME considerations will be so favorable that the averages will represent unacceptably low targets. Therefore, the use of uncontrolled comparisons is only recommended as a supplement to those of a more controlled nature and even then their findings must be cautiously interpreted and applied. Data useful for uncontrolled comparisons within the 1980 time frame are summarized in Table 1.

TABLE 9 Recommended Peer Groups

Cluster	City
1	Amarillo, Bay City, Boise, Champaign, Dubuque, Eugene, Kenosha, Lexington, Lubbock, Madison, Pittsfield, Pueblo, San Angelo, Sioux City, Sioux Falls, Springfield (Mo.), Terre Haute, Topeka, Tuscaloosa, Utica-Rome, Wichita Falls
2	Altoona, Asheville, Beaumont, Billings, Cedar Rapids, Colorado Springs, Decatur, Duluth, Evansville, Fort Wayne, Fresno, Green Bay, Jackson (Miss.), Johnstown, Kalamazoo, Knoxville, Lewiston, Lynchburg, Manchester, Muskegon, Racine, Rockford, Spokane, Springfield (Ohio), Stockton, Waco, Wichita
3	Albany (Ga.), Albuquerque, Allentown, Austin, Bakersfield, Binghampton, Brockton, Canton, Charlotte, Chattanooga, Davenport, Des Moines, Erie, Flint, Grand Rapids, Jackson (Miss.), Lake Charles, Lancaster, Little Rock, Monroe, Montgomery, Nashville, New Bedford, Oklahoma City, Peoria, Portland (Maine), Reading, Reno, Roanoke, Saginaw, Santa Barbara, Shreveport, South Bend, Syracuse, Tucson, Tulsa, Worcester, York, Youngstown
4	Akron, Albany (N.Y.), Augusta, Aurora-Elgin, Baton Rouge, Birmingham, Dayton, Harrisburg, Indianapolis, Jacksonville, Lowell, Memphis, Mobile, Omaha, Salt Lake City, Savannah, Springfield (Ill.), Tacoma, Toledo, Trenton, West Palm Beach, Wilkes-Barre
5	Galveston, Hartford, Louisville, New Haven, Norwalk, Pensacola, Phoenix, Portland (Oreg.), Richmond, Rochester, Sacramento, San Antonio, Stamford, Tampa, Wilmington (Del.)
6	Buffalo, Cincinnati, Denver, Honolulu, Kansas City, Milwaukee, New Orleans, Providence, St. Petersburg, San Diego, San Jose, Seattle
7	Atlanta, Cleveland, Miami, Minneapolis, Pittsburgh
8	Baltimore, Dallas-Fort Worth, Houston, St. Louis
9	Boston, Detroit, San Francisco, Washington
10	Los Angeles
11	Alexandria, Anchorage, Anderson (Ind.), Appleton, Battle Creek, Biloxi, Bloomington, Bristol, Brownsville, Charleston (W. Va.), Daytona Beach, Fayetteville, Fort Myers, Gainesville, High Point, Lafayette (Ind.), Lafayette (La.), Mansfield, Melbourne, Modesto, New London, Orlando, Oshkosh, Oxnard-Ventura, Petersburg, Pine Bluff, Poughkeepsie, Raleigh, St. Cloud, Salinas, Santa Cruz, Santa Rosa, Seaside-Monterey, Tallahassee, Waterbury, Williamsport, Wilmington (N.C.), Winston-Salem

TABLE 10 Distributional Characteristics of Groups Developed by Clustering Based on Size Factor

Group	Min F_2	Center F_2	Max F_2	Standard Deviation of F_2	No. of Cases
1	−1.787	−1.089	−0.899	0.222	21
2	−0.884	−0.701	−0.518	0.114	27
3	−0.485	−0.299	−0.119	0.108	39
4	−0.075	0.089	0.270	0.099	22
5	0.327	0.474	0.665	0.099	15
6	0.778	0.983	1.169	0.137	12
7	1.468	1.661	1.791	0.144	5
8	1.876	1.932	1.993	0.050	4
9	2.622	2.771	2.929	0.165	4
10	5.499	5.499	5.499	0.0	1
11	N/A	N/A	N/A	N/A	38

Note: N/A = not applicable.

The second and third levels represent controlled comparisons: Performance of the subject system is compared with that of a more selective peer group. To the extent that the F_2 (size) clusters represent homogenous groups of urbanized areas having equivalent potential for transit, they are considered an appropriate basis for controlled comparisons.

In the second-level analysis, the performance target is the average performance of the peer group. Table 11 shows such averages for 10 of the peer groups. Since the original group 10 has only one member, Los Angeles, it is not useful for controlled comparisons and hence is not included in Table 11. In Table 12 it is demonstrated that rather distinct differences result from use of peer-group averages as target values rather than overall U.S. averages. For simplicity, only 6 of the 25 performance variables are included in Table 12, and tabulated values have a normalized value of 1 at the overall mean. It is apparent that there are distinct differences in the group means for each variable and that a general relationship seems to exist between the group mean and the size of the average urbanized area (with size increasing from top to bottom of the table). Because the second-level comparison significantly reduces the effects of dissimilarity among the urbanized areas, it is judged to be more reliable and useful than the uncontrolled comparison.

However, within each group, there remain inherent market differences that influence transit performance. To further control for these differences, a third-level comparison is sometimes useful in which the target becomes the expected performance computed from regression models of peer-member statistics. As in all controlled comparisons, the intent is to eliminate, insofar as possible, the effect of ME dissimilarities.

The stepwise multiple regression analysis of SPSS (9) was used. Two forms of regression equations were screened, the linear and the multiplicative. Because the linear is simpler and seemed to be of comparable or superior accuracy, it was chosen for the detailed analysis.

The independent variables were chosen from the set of 41 ME variables. In order to reduce collinearity, the following selection procedure was employed. The ME variables were first rank ordered with respect

TABLE 11 Average Peer-Group Performance

Symbol	Performance Variable Name	U.S. Avg	Performance by Peer Group 11	1	2	3	4	5	6	7	8	9
SOSI3	Rev Veh Hr/Opr Exp	0.047	0.059	0.054	0.052	0.044	0.042	0.039	0.034	0.031	0.025	0.019
SOSI5	Rev Veh Mi/Rev Veh	27,600	29,900	25,400	28,000	27,100	26,400	27,200	29,900	24,300	26,500	26,700
SOSI8	Rev Veh Hr/Transit Empl	1,060	1,120	1,100	1,150	1,040	1,030	1,040	928	1,070	856	743
SOSI12	Rev Veh Hr/Equiv Gal Gas	0.313	0.338	0.321	0.312	0.292	0.269	0.253	0.279	0.659	0.374	0.345
SOSI13	Pass Cap Mi/Rev Veh	1,660	1,410	1,170	1,420	1,710	2,340	1,650	1,970	1,890	1,720	2,500
SCSI4	Rev/Rev Veh	54,200	45,300	39,700	48,600	52,200	52,700	53,100	80,100	75,700	100,000	124,000
SCSI6	Pass/Opr Exp	1.40	1.37	1.24	1.38	1.36	1.30	1.36	1.56	1.38	1.19	1.30
SCSI8	Pass Mi/Opr Exp	4.84	4.58	4.27	4.83	5.23	3.60	6.19	5.37	5.68	5.50	4.68
SCSI16	Rev/Veh Opr Cost	1.94	2.10	1.78	1.79	1.89	2.23	1.50	1.88	1.79	2.70	2.32
SCSI18	Pass Rev/Opr Exp	0.311	0.288	0.282	0.290	0.303	0.349	0.359	0.345	0.316	0.354	0.314
SCS01	Pass Mi/Pass Cap Mi	153	133	166	151	141	112	197	189	215	222	200
SCS03	Pass/Rev Veh Mi	2.66	1.81	1.78	2.16	2.35	2.99	2.77	3.43	8.14	3.29	5.33
SCS09	Pass/Transitway Length	22,300	10,200	11,900	13,500	18,100	20,900	27,000	52,300	38,100	30,100	121,000
SCS014	Accidents/Rev Veh Mi	77.4	45.9	52.5	72.2	72.4	91.0	73.2	92.8	326	149	86.1
SCS015	Rev Veh Mi/Tot Road Calls	4.13	4.00	2.89	4.31	7.66	2.86	2.51	2.05	1.65	1.34	1.79
SIME1	Opr Exp/Pop	17.6	10.7	12.4	13.2	13.1	15.2	24.9	31.7	43.8	30.0	74.2
SIME2	Opr Assistance/Pop	10.6	6.51	7.90	9.03	8.99	9.33	11.1	15.8	20.7	19.8	53.6
SIME5	Rev Veh/Pop	0.314	0.239	0.299	0.285	0.269	0.307	0.379	0.448	0.594	0.391	0.688
SIME6	Transit Empl/Pop	0.646	0.472	0.531	0.542	0.531	0.624	0.746	1.10	1.27	0.861	2.015
SOME1	Pass Cap Mi/Pop	540	350	347	399	487	655	622	925	1,200	662	1,770
SOME2	Rev Veh Mi/Pop	8.44	7.12	7.62	7.53	6.98	7.95	9.86	13.4	13.7	10.3	18.7
SOME8	Transitway Length/Area	3.05	3.43	2.85	3.39	2.38	2.40	2.44	3.21	5.09	5.28	5.28
SCME1	Pass/Pop	23.3	12.3	13.7	16.5	17.5	19.5	28.5	48.5	58.9	36.7	103
SCME6	Pass Mi/Area	215	85.2	110	130	186	126	280	589	626	408	1,140
SCME12	Avg Trip Dist/Equiv UA Radius	0.637	0.832	1.02	0.755	1.04	0.362	0.546	0.343	0.289	0.276	0.203

TABLE 12 Comparison of Target Performance from Peer-Group Average Versus Overall U.S. Average

Group	Ratio of Group Average to Overall Average by Variable SOSI3	SCSI4	SCS01	SIME1	SOME1	SCME1
11	1.255	0.835	0.872	0.607	0.649	0.527
1	1.149	0.733	1.090	0.706	0.642	0.586
2	1.106	0.898	0.990	0.751	0.739	0.705
3	0.936	0.963	0.925	0.744	0.901	0.748
4	0.894	0.972	0.731	0.866	1.213	0.837
5	0.830	0.979	1.290	1.417	1.151	1.220
6	0.723	1.478	1.240	1.801	1.713	2.078
7	0.660	1.397	1.407	2.488	2.225	2.525
8	0.532	1.855	1.451	1.704	1.226	1.574
9	0.404	2.293	1.308	4.217	3.281	4.401

TABLE 13 Independent Variables of Regression Models

Symbol	Variable Name
ME33	1900 Pop
ME4[a]	Families > $10,000
ME9	Pop Density
ME14	HU Renter-Occupied
ME21[a]	Pop Completed College
ME22[a]	Pop Employed
ME25[a]	Empl in Sales
ME34	Pop Nonwhite
ME19[a]	Housing Growth 70-80
ME31[a]	Autos/Hsld
ME29	Pop Female
ME30[a]	Workers/Auto
ME10	Land Area
ME37	HU
ME40	UA Pop in Central City
ME39	SMSA Pop in UA

[a] Variables that could not be included for regression models of group 11.

to the magnitude and frequency of their correlations with the 25 performance variables. The top-ranked ME variable was selected, and all remaining ME variables with which it was correlated (correlation coefficient of 0.4 or more) were discarded. The highest-ranked of the remaining variables was next selected, and again correlated variables were discarded. The process was repeated until the 16 variables of Table 13 remained. These variables made up the set of independent variables considered as candidates for inclusion in the models: Each variable is significantly related to transit performance but no pair is highly correlated.

The number of independent variables in each regression equation was somewhat arbitrarily limited to 5. This number seemed to be sufficient with regard to accuracy but not so large that the relationships became completely meaningless and the computations laborious. The best 5 of the 16 candidate variables were selected by the stepwise routine for each model developed.

An example of the regression models is that which relates passengers per capita to five ME variables for the group 5 urbanized areas. The calibrated model is

Passengers per capita = - 70 + 0.0085 x population density + 0.94 x HU renter-occupied - 0.28 x housing growth, 1970-1980 - 0.96 x population completed college + 1.8 x population employed.

Intuitively, these five independent variables appear to have been appropriately selected and the signs of the coefficients appear reasonable. The coefficient of determination (R^2) was 0.79 and the adjusted coefficient, an indicator of the accuracy of the simulation for the entire population, was 0.67.

Other results of the modeling effort are too extensive to include here, but the complete set is available elsewhere (5). A total of 175 equations were calibrated, one for each combination of the 25 performance variables and 7 peer groups. Four of the original 11 peer groups (groups 7-10) were eliminated from this calibration because of small group size.

The adjusted coefficients of determination were judged to be quite acceptable. Table 14 presents, in summary form, the range in the adjusted R^2. To illustrate the meaning of the tabulated entries, consider those for the SOSI variables. After R^2 ranking, the 75th percentile of the 35 regression

TABLE 14 Range of Adjusted R^2 for Regression Models

	R^2 by Percentile of Models		
Variable Type	75th	50th	25th
SOSI (output and input)	0.72	0.43	0.23
SCSI (consumption and input)	0.74	0.51	0.24
SCSO (consumption and output)	0.65	0.47	0.19
SIME (input and market)	0.81	0.60	0.46
SOME (output and market)	0.76	0.56	0.33
SCME (consumption and market)	0.63	0.46	0.33

equations (5 performance variables and 7 peer groups) had an R^2 of 0.72; the 50th percentile (median), 0.43; and the 25th percentile, 0.23. In forming judgments regarding the acceptability of the adjusted R^2, it must be recalled that performance was being related only to the exogenous, ME variables: Policy variables related to the provision of transit, which also affect performance, were properly excluded.

To illustrate application of the foregoing procedures in establishing performance targets, consider a case in which a target is being established for passengers per capita in a hypothetical city within group 5. Alternative target measures are as follows:

Measure	Amount (passengers/ capita)
U.S. avg	23.3
Peer-group avg	28.5
Peer-group regression	30.6

These numbers suggest that a ridership of at least 30.6 passengers per capita is achievable if appropriate transit decisions are made. It is imperative to note, at the same time, that numerous factors must be incorporated into the development of performance targets, especially the importance or weight given by the community to various transit objectives. Although a ridership of 30.6 passengers per capita may be achievable, the community might appropriately decide that the commitments necessary to reach this level are not justified.

In summary, of the three levels presented, the regression models best eliminate the effects of market dissimilarities and hence best represent attainable performance levels. Analysts uncomfortable with their use should turn to the peer-group averages as a reasonable alternative. Availability of these two alternatives frees the analyst from reliance on nationwide averages and their attendant inaccuracies. However, it should be remembered that in each case the targets are scaled to average rather than to exceptional performance. Prudent transit decisions may well yield performance superior to the target averages.

CONCLUSIONS

Peer comparisons are an invaluable component in transit performance evaluations. A former impediment to such comparisons, the paucity of uniform statistical data, has been largely overcome by implementation of the Section 15 reporting system. Attention can now be turned to refined techniques for forming reasonable peer groups.

Comparable transit systems--that is, those within each peer group--should be homogenous with respect to their potential performance levels. Because it is practically impossible to quantify potential performance directly, attempts to form peer groups on this basis are not currently feasible. However, potential performance is a direct function of many exogenous ME variables. Properly selected, these variables offer great potential for identifying comparable transit sytems.

Data from the U.S. Bureau of the Census form a sufficient set for adequately characterizing ME conditions. Key dimensions can be identified by means of factor analysis, and cluster analysis using these key dimensions is an effective tool for formation of peer groups. The constituency of the various peer groups, however, is sensitive to the basis and method of clustering.

Target performance levels can be established by using either uncontrolled or controlled techniques. In uncontrolled comparisons the target levels represent means for systems throughout the United States without regard to homogeneity. Such targets must be augmented by others representing peer-group means if necessary recognition is to be given to key ME dissimilarities. A second level, controlled comparison using regression-based targets, is recommended to further account for ME constraints on attainable performance levels.

ACKNOWLEDGMENT

The research on which this paper is based was sponsored by the Office of Policy Research, University Research and Training Program, UMTA, U.S. Department of Transportation.

REFERENCES

1. W.B. Allen. The Productivity and Efficiency of Inputs in the Provision of Transportation Services of the Southeastern Pennsylvania Transportation Authority. Report UMTA-IT-09-0073-79-1. UMTA, U.S. Department of Transportation, May 1979.
2. J.M. Holec, Jr., D.S. Schwager, and A. Fandialan. Use of Federal Section 15 Data in Transit Performance Evaluation: Michigan Program. In Transportation Research Record 746, TRB, National Research Council, Washington, D.C., 1980, pp. 36-38.
3. S.R. Mundle and W. Cherwony. Diagnostic Tools in Transit Management. In Transportation Research Record 746, TRB, National Research Council, Washington, D.C., 1980, pp. 13-18.
4. S.C. Anderson and G.J. Fielding. Comparative Analysis of Transit Performance. Report UMTA-CA-11-0020-82-1. Institute of Transportation Studies and School of Social Sciences, University of California, Irvine, Jan. 1982.
5. M. Vaziri and J.A. Deacon. Application of Section 15 and Census Data to Transit Decision Making. Report UMTA-KY-11-0002-83. Department of Civil Engineering, University of Kentucky, Lexington, Aug. 1983.
6. National Urban Mass Transportation Statistics, Second Annual Report: Section 15 Reporting System. Report UMTA-MA-006-0107-82-1. UMTA, U.S. Department of Transportation, June 1982.
7. T.F. Golob, E.T. Canty, and R.L. Gustafson. Classification of Metropolitan Areas for the

Study of New Systems of Arterial Transportation. Res. Publ. GMR-1225. General Motors Research Laboratories, Detroit, Mich., 1972.

8. T.F. Golob, E.T. Canty, and R.L. Gustafson. National Studies of Urban Arterial Transportation: A Research Framework. Res. Publ. GMR-1274. General Motors Research Laboratories, Detroit, Mich., 1973.
9. N.H. Nie et al. Statistical Package for the Social Sciences (SPSS). McGraw-Hill, N.Y., 1975.
10. W.J. Dixon, ed. Biomedical Computer Programs (BMDP). University of California Press, Berkeley, 1983.

An Assessment of the Use of Part-Time Operators at the Massachusetts Bay Transportation Authority

JOHN ATTANUCCI, NIGEL H.M. WILSON, and DAVID VOZZOLO

ABSTRACT

The impact of introducing part-time bus operators at the Massachusetts Bay Transportation Authority (MBTA) in Boston is evaluated and the likely impact of various future scenarios regarding the size and utilization of a part-time labor force at MBTA is analyzed. In January 1982 MBTA had no part-time operators; there are now 280, representing almost 19 percent of the surface-operator classification. Introduction of this number of part-time operators to provide the current level of service has resulted in an annual saving of more than $5 million through reduction in unproductive paid hours, spread penalties, and fringe benefits. However, three factors mitigate this financial benefit: higher accident rates, absenteeism, and turnover among the part-time operators compared with that among the full-time operators. Although there are clear opportunities to obtain further financial benefits from the introduction of more part-time operators, the high accident rate to date suggests that caution is appropriate in expanding their role. Strategies to improve productivity by using the existing complement of part-time operators are also discussed.

In January 1982 the Massachusetts Bay Transportation Authority (MBTA) introduced part-time operators (PTOs) on surface bus lines with the assignment of 20 PTOs to the Quincy bus garage. This was the result of enactment by the Massachusetts Legislature in 1980 of a bill that gave MBTA management, among other things, the right to hire and assign part-time employees as they thought appropriate, notwithstanding previous collective bargaining agreements and past labor practices. This right to use part-time employees, when applied to the typical bus-scheduling requirements of MBTA, provided an opportunity to make substantial savings by reductions in 8-hr work day guarantees and long working hours (called spread penalties).

The first 1 1/2 years of MBTA experience with PTOs is assessed and alternative uses of part-time employees in the Transportation Department are examined. An attempt has been made to evaluate all impacts of the use of PTOs, although the effort was limited by the relatively short period of experience to date and, in some instances, a lack of primary data on the particular issue at hand. Where possible, the impacts of the current and projected use of PTOs have been quantified.

BACKGROUND

The introduction of PTOs at MBTA has clearly been accomplished in an accelerated manner over the past year and a half. The initial 20 PTOs who were assigned to the Quincy garage in January 1982 were primarily from the ranks of former full-time operators (FTOs) who had been laid off in April 1981. In each quarterly timetable through March 1983, an increasing number of PTOs were trained and assigned a daily run of up to 6 working hr per day. Today, 280 PTOs are assigned throughout the bus system.

Throughout late 1981 and 1982, MBTA negotiated with the Boston Carmen's Union (Local 589 of the Amalgamated Transit Union) to set conditions for hiring and utilizing PTOs. These discussions did not result in an agreement, and so, while MBTA management pressed ahead with the hiring and assignment of part-time drivers, the Carmen's Union brought the matter to interest arbitration. Although MBTA maintained that the right to hire PTOs was not subject to collective bargaining or arbitration under Chapter 581 of the 1980 Acts and Resolves of Massachusetts, it presented a proposal that called for unrestricted use of PTOs under the following conditions:

1. A maximum of 30 hr of work per week;
2. A guarantee of 2 hr pay for each scheduled work day;
3. A schedule of work on a 7-day basis;
4. A 6-month probationary period after instruction;

5. A pay rate of 58 percent of the top operator's hourly rate;
6. Eligibility for the following benefits: standard uniform allotment, free transportation, other benefits mandated by federal or state law, and pension plan as amended for part-time employees;
7. No requirement to join the union but requirement of agency fee equal to the dues paid by members of Division 589, less the international portion, as a condition of employment; and
8. The same qualifying standards as those for regular operators.

The Carmen's Union maintained that part-time employees were different from full-time employees only in guaranteed hours of work per day and should otherwise be accorded the same wages, rights, and benefits. The Carmen's Union also contended that there should be an agreed limit to the number of part-time employees working in any MBTA job classification.

On January 15, 1983, the arbitrator made an award covering salary and working conditions for both FTOs and PTOs. Concerning PTOs, the arbitrator ruled that they should indeed be represented by the Carmen's Union and that they should be paid at the same rate as FTOs, subject to the new-hire progression. Benefits for PTOs were set as previously stipulated by MBTA. Although MBTA management immediately put into place the pay and benefits aspects of the award, it objects to, and has refused to comply with, several elements on the basis that they are in violation of Chapter 581. Specifically, MBTA refused to

1. Place an upper limit on the number of part-time employees at 15 percent of the number of full-time employees in each classification (e.g., surface operator, rapid transit motorman),
2. Preclude the layoff of full-time employees while part-time employees remain on the payroll in the same classification,
3. Prohibit PTOs from working on Saturday or Sunday, or
4. Prohibit PTOs from substituting for absent FTOs.

MBTA is in violation of the first of these provisions because PTOs make up almost 19 percent of the full-time surface operators. Because MBTA did not reduce the number of PTOs but rather maintained their number at 280, the Carmen's Union sought injunctive relief from the courts. A preliminary injunction was denied and the case (for a permanent injunction) is still pending. Thus far, MBTA has elected not to increase the number of PTOs further, pending disposition of the court case. Thirty-five new PTOs were hired in April and May 1983 to replace the same number, who were promoted to full-time status when the summer timetable began in June.

CURRENT USE OF PTOs

The difficulty of scheduling transit service to meet the demands for morning and afternoon school and commuter travel is well known in the industry. MBTA must schedule in the peak travel hours, on average, 2.5 times the number of vehicles scheduled during midday. To meet such uneven daily demands for service with only full-time employees, MBTA has historically scheduled many operators to work split shifts to provide service in both peaks (e.g., 6:00 to 9:00 a.m. and 2:00 to 7:00 p.m.). Because this results in long work days with an unpaid break in the middle, labor has successfully bargained over the years for compensating work rules that provide higher pay for split runs, guarantee a minimum of 8 hr of pay, and restrict the number and manner in which split runs are designed.

Three of the most restrictive and costly rules that MBTA faced before hiring PTOs included

1. A requirement that all scheduled pieces of work be included in 8-hr driver runs (which forced some unnecessary and unproductive cover time to be scheduled);
2. A requirement that 70 percent of all runs in a rating station (an organizational unit that corresponds in general to a garage) be less than 11 hr in total length, including breaks (which also forced the scheduling of unproductive cover time in so-called additional runs); and
3. The requirements for spread penalties for all runs that exceed a total elapsed time of 10 hr (time worked in the 11th hour is paid at the rate of time and a half; time worked in the 12th and 13th hours is paid at the rate of double time).

Just before the introduction of PTOs, MBTA's distribution of weekday surface operator runs was as follows:

Type of Run	No.	Percent
Weekday ≤ 10 hr	427	40
Weekday between 10 and 11 hr	309	29
Weekday between 11 and 13 hr	335	31
Additional that included an average of >1.5 hr of cover time (included in the previous figures)	92	9

Extra pay for the spread penalties and unnecessary cover time added about 900 daily pay hours to the schedule or about 10.5 percent to the total cost of the schedule.

In scheduling PTOs, MBTA has been, thus far, successful in eliminating the longer spread penalties and unnecessary cover time. It is impressive that so many scheduling changes have been accomplished in so short a time through a completely manual process. Nearly every PTO used thus far has replaced an FTO who had unnecessary paid cover time or who had received large spread-pay penalties. This has been accomplished, however, through scheduling the majority of the PTOs over a 12- to 13-hr work day during which they have, on average, a 6- to 7-hr unpaid break in the middle of the day. Thus far the peak-period work has not been split into two pieces to be assigned to two different PTOs because of the difficulty of recruiting and training new operators and the perception that the overall objective was to maximize cost savings for a given number of PTOs.

Each PTO is assigned and trained for a single run from a particular garage on Monday through Friday each week. The assignment of a specific number of PTOs to a garage is based on an informal analysis of existing spread penalties and additional runs. However, the allocation by the Plans and Schedules Department has at times been altered based on the Transportation Department's limited ability to reassign and retrain PTOs and FTOs. Thus, the current allocation of PTOs among garages does not maximize spread-penalty savings. No PTOs are currently used on the weekends because the spread-penalty savings are much greater during the week. As mentioned earlier, 23 of the 280 PTOs are not assigned to scheduled runs but fill in as substitute personnel for absent PTOs. Finally, no PTOs have yet been assigned to any rail operations because of the special training required.

FINANCIAL IMPACT OF PTOs

There is no doubt that the use of 280 PTOs is saving

MBTA substantial operating resources. Assuming that the spring (March 1983) timetable is carried forward during the next year and the level of part-time use (280) remains constant, MBTA will save approximately $5.6 million over the next year. These savings come primarily from three sources:

1. Wage savings because of reduction of unnecessary scheduled driver time (i.e., 6 productive hr of a PTO have been substituted for a guaranteed 8-hr day, which included extensive unnecessary cover) and because a quicker turnover of PTOs keeps their wage rate somewhat below the average rate of a comparable number of FTOs;
2. The elimination of costly FTO spread penalties for runs that were scheduled more than 10 hr to cover both the morning and evening peak travel requirements; and
3. Fringe benefit savings, because PTOs receive only the benefits required by statute (social security, worker's compensation), a scaled-down retirement fund contribution, free transportation, and a uniform allowance, whereas FTOs receive holiday, vacation, and sick pay and health, dental, life, and accident insurance as well.

The annual financial impact of using 280 PTOs is summarized as follows:

	Annual Financial Impact ($000,000)		
	Operating Cost		PTO
Type of Impact	FTO	PTO	Savings
Wages	6.1	4.3	1.8
Spread penalties	1.4	-	1.4
Fringe benefits	3.4	1.0	2.4
Total	10.9	5.3	5.6

The wage impacts reflect an average PTO rate of $9.76/hr (81 percent of top scale) and an average FTO rate of $10.45/hr (87 percent of top scale), which reflects the different turnover rates among PTOs and FTOs and the 8 months that it takes for a PTO to earn a 5 percent progression step increase working 6 hr a day as compared with 6 months for an FTO working 8 hr a day. The wage difference was computed by adding 2 hr per day to each part-time run and applying the respective wage rates to the total hours worked. The spread-penalty savings were computed directly based on the spread hours currently worked by part-time operators and applying the top wage rate ($12.065/hr) because more senior FTOs would generally pick these lucrative runs. The savings in benefits were computed directly on a per-operator basis by using the respective average wage rates for PTOs and FTOs and other unit cost data supplied by the MBTA Budget Office.

There are also some second-order financial benefits (which are not accounted for here), including up to 15 min of overtime and make-up time, which would have to be paid to FTOs and for which PTOs do not qualify. (For example, although the goal for scheduling PTO runs is 6 hr a day, the average PTO run is only 5 hr and 48 min. The additional 12 min a day for each PTO is saved, whereas all FTOs must be paid for a full 8-hr day.)

The average total annual savings is about $20,000 for each PTO currently assigned. As shown in the foregoing tabulation, the PTO savings are split relatively evenly among the three types of expenses. Although the fringe-benefit category is approximately linear with the number of PTOs, it should be noted that the other two categories will vary significantly at different levels of PTO use. After a certain point, the assumption of substituting one PTO run directly for one FTO run (and thus automatically saving 2 hr of pay per day) will not apply; rather, four PTO runs will be needed to replace three FTO runs and the only difference will be the difference in wage rates between the two classifications. The threshold point at which the dramatic wage savings are eliminated is about 310 PTOs for the bus system and 335 PTOs for the entire bus and light rail system. Similarly, the spread-penalty savings per operator are reduced as the number of PTOs increases, because the largest spread penalties are eliminated first as PTO runs are developed. For weekday service, it has been estimated that each PTO hired for the surface system beyond a total of 335 will save MBTA about $6,300 annually, that being the difference in fringe benefits for a PTO and an FTO.

ACCIDENT RATES

A critical concern is the effect on the accident rate of introducing PTOs. In this section, the accident rates for PTOs and FTOs are compared and the role that the difference in experience and working hours between these two groups plays in accident rates is investigated.

Table 1 shows the numbers of operators employed and accidents by month for PTOs and FTOs. It is clear that the accident rates for part-time employees are significantly higher than those for full-time employees. For example, in July 1982 the accident rate for part-time employees was fully three times that for FTOs. Although the accident rate for part-time employees had decreased considerably by the first quarter of 1983, it was still 75 percent higher than that for full-time employees. Clearly, as PTOs acquire experience their accident rate is declining, but it remains to be seen how far and how quickly this decline will proceed.

Furthermore, if the accident rate is computed on the basis of hours worked rather than number of op-

TABLE 1 Accident Rates for PTOs and FTOs

	FTO			PTO		
Date	No. of Operators	No. of Accidents	Annual Rate[a]	No. of Operators	No. of Accidents	Annual Rate[a]
July 1982	1,367	151	1.33	67	24	4.30
Aug. 1982	1,367	133	1.17	101	38	4.51
Sept. 1982	1,367	140	1.23	177	44	2.98
Oct. 1982	1,367	165	1.45	203	49	2.90
Nov. 1982	1,379	150	1.31	227	61	3.22
Dec. 1982[b]	1,379	107	1.44	281	44	2.91
Jan.-March 1983	1,370	449	1.33	280	161	2.33
Total		1,295	1.31		421	2.83

[a] Accidents per operator per year.
[b] Up to December 20th only.

erators, the discrepancy between PTOs and FTOs is much more marked; the accident rate for PTOs in the first quarter is 133 percent above the accident rate for FTOs. This is, in fact, a fairer way of looking at the accident rate because the exposure increases with hours worked.

In June 1981 the MBTA Safety and Training Department analyzed the relationship between accidents and driver experience. For 9 months of accident records the accident rate was computed by the number of years of MBTA service by the driver. The results summarized in Table 2 show that accident rates indeed decline with experience, being twice as high in the first 2 years as rates for those with more than 5 years' experience.

TABLE 2 Accident Rate as a Function of Experience Level

Accidents per Operator per Year	Years of Service
2.1	0-1
2.2	1-2
2.0	2-3
1.7	3-4
1.8	4-5
1.1	5-10
0.8	10-20
0.9	20-30
0.8	30-40+

A comparison of the PTO accident rate with the FTO accident rate for operators with comparable experience (0 to 2 years of service) indicates that PTOs have a 30 percent higher gross annual accident rate, or 73 percent higher hourly accident rate. These findings indicate that the higher PTO accident rate is related to their lower level of driving experience.

It has also been suggested that the higher PTO accident rate may be related to the use of PTOs during peak periods. This issue was addressed by examining the relationship between accidents and time of day. March 1983 accident data for FTOs were studied to see whether accident rates varied between time periods. This analysis showed that there was no statistically significant difference between peak-period and off-peak accident rates; in fact, the peak-period rate was marginally lower.

In addition to the impact on public safety, the higher accident rate exhibited by PTOs will clearly increase MBTA costs for settling accident claims and repairing damaged vehicles. A comparative analysis with 1983 data on the use of PTOs and FTOs indicates that the use of PTOs incurs an additional $0.9 million, representing a 20 percent increase, in annual costs for accident claims. (Note that this analysis did not include any increase in the cost to MBTA for repairing vehicles damaged in accidents.) This analysis compared current 1983 operations using 1,370 FTOs and 280 PTOs with a scenario of operations using no PTOs and 1,650 FTOs. For each scenario the number of accidents per year was estimated by applying the annual accident rate per driver (FTO or PTO), based on data from July 1982 through March 1983. Annual costs for accident claims and suits were estimated by assuming that each accident cost MBTA $2,000, based on data from recent years. The scenario with no PTOs resulted in an annual total of 2,162 accidents at a cost of $4.3 million. On the other hand, current operations (with 280 PTOs) are estimated to yield 2,587 accidents at an annual cost of $5.2 million. Applying the more recent lower 1983 PTO accident rate (annualized to 2.33 accidents per year) to the scenario of 1,320 FTOs and 280 PTOs results in an annual total of 2,447 accidents and an annual cost of $4.9 million. This represents a 13 percent increase in annual accident cost over similar operations without the use of PTOs.

It should be noted that this analysis uses average (inflated) claims and suit settlement data from the past several years, which excluded several large settlements. It is impossible to predict how the higher PTO accident rate will translate into a probability of encountering a number of extremely costly claims that could quickly erase the savings realized from introducing PTOs. (In 1981 one suit, now on appeal, was decided at a cost of $1.5 million, and in 1982 another suit was decided at a cost of $1.8 million.) Already one PTO has been involved in a fatal bus-pedestrian accident.

The high PTO accident rates at the MBTA are clearly disturbing. Individual safety is of utmost importance in a public transit operation and MBTA should continue to monitor this situation closely in the coming months. Perhaps it should be required that stricter standards and disciplinary actions accompany any PTO vehicle violations or that any PTO involved in an accident undergo remedial training during the midday breaks. Data from three other agencies using PTOs suggest that the PTO accident rates should continue to decline to the level of FTO rates; both Los Angeles and Baltimore report about the same accident rate for PTOs and FTOs, and Seattle reports a slightly lower PTO accident rate.

One other possible explanation for the difference between PTO and FTO rates should be recognized. It may be that PTOs are reporting a larger number of minor accidents that would go unreported by FTOs. With the data available, it has not been possible to test this hypothesis, but it is a possibility and if it is true, this would eliminate safety as a major factor in the analysis.

ABSENTEEISM

Another potential effect of the introduction of part-time employees is a lower rate of absenteeism. If this is the case, service performance should be improved or the cover list can be reduced, resulting in cost savings.

In Table 3 absence hours are given as a percentage of total scheduled hours for full-time surface operators, rapid transit operators, and PTOs for each year. The data indicate that the absence rates for PTOs are in fact significantly higher than those for FTOs, primarily because of more sick time, terminations, suspensions, and unauthorized leave. On the other hand, the absences of full-time surface and rapid transit operators are attributable primarily to sick time and industrial accidents.

Figure 1 shows the annual absenteeism data by employee class over time. The graph of total hours absent clearly indicates the higher rate for PTOs and also that the absence rates for all employee classes declined in 1983. (Because the 1983 data represent only the first 4 months of the year, the lower absenteeism rate may be a result of seasonal fluctuation.) Note that FTO absenteeism increased significantly for a period in 1982, which is related to the introduction of PTOs. The graph of absences due to sick leave and industrial accidents shows the higher rates for FTOs and rapid transit motormen. The PTO rate increased substantially from 1982 to 1983, almost entirely because of an increase in hours of sick leave. The 1982 increase for FTOs was entirely attributable to a dramatic rise in industrial accidents. The graph of terminations, suspen-

TABLE 3 Absenteeism

	Hours Absent[a] (%)									
	FTO				Rapid Transit Operator				PTO	
Type of Absence	1980	1981	1982	1983	1980	1981	1982	1983	1982	1983
Sickness	3.4	3.7	3.7	3.1	4.5	4.3	4.2	3.5	2.4	4.0
Industrial accident	1.2	1.9	3.3	2.8	1.5	1.8	1.8	2.0	0.2	0.5
Excused	0.3	0.3	0.2	0.2	0.3	0.3	0.4	0.3	0.6	0.4
Absent without leave	0.1	0.1	0.3	0.1	0.1	0.1	0.3	0.0	0.3	0.1
DIF	0.1	0.1	0.1	0.1	0.1	0.1	0.0	0.0	0.1	0.0
Jury duty	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.0	0.0	0.0
Termination	0.1	0.1	0.1	0.0	0.0	0.2	0.1	0.0	2.5	1.6
Military duty	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.2	0.1	0.3
Suspension	0.3	0.5	0.6	0.6	0.3	0.5	0.5	0.3	2.6	2.1
Union	0.0	0.2	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Unauthorized	0.1	0.0	0.0	0.2	0.0	0.0	0.1	0.1	1.9	0.9
Total	5.9	7.1	8.9	7.3	7.0	7.5	7.7	6.4	10.7	9.8

[a]Hours absent expressed as a percentage of total scheduled hours.

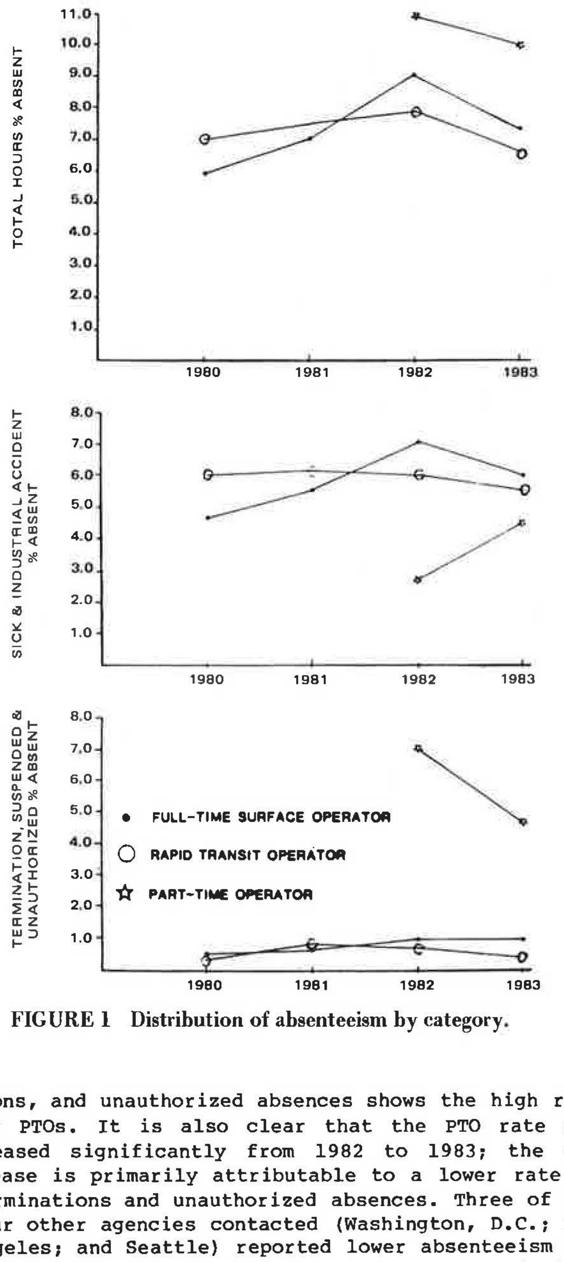


FIGURE 1 Distribution of absenteeism by category.

sions, and unauthorized absences shows the high rate for PTOs. It is also clear that the PTO rate decreased significantly from 1982 to 1983; the decrease is primarily attributable to a lower rate of terminations and unauthorized absences. Three of the four other agencies contacted (Washington, D.C.; Los Angeles; and Seattle) reported lower absenteeism for PTOs; the fourth (Baltimore) reported a significantly higher PTO absence rate.

TURNOVER RATE

The turnover rate is important because it affects the amount of hiring and training required and also reveals the amount of experience that PTOs acquire before leaving. There is no reason a priori to expect similar turnover rates for PTOs and FTOs because the working conditions, pay, and benefits are quite different.

Annual FTO turnover rates were computed based on the number of annual terminations and total number of FTO employees, shown as follows for 1980 through 1982:

Year	Total No. of FTOs	No. of Terminations	Turnover Rate (%)
1980	1,930	103	5.3
1981	1,785	225	12.6
1982	1,468	114	7.3

Note that in 1981 a substantial number of drivers were laid off. As a result, the analysis is based on the 1982 data; that is, an annual rate of 7.3 percent, or a little more than 100 FTOs laid off each year.

Two major components of PTO turnover are examined: the promotion of PTOs to full-time status, and PTO terminations (discharge or resignation).

Promotion to FTO Status

The estimated annual rate of PTO promotions over the next few years is 36 percent (100/280). Estimation of this rate assumes that a PTO staff size of 280 employees is maintained and that approximately 100 FTOs terminate and must be replaced annually.

PTO Terminations

Experience to date has indicated a 21 percent turnover rate because of PTO discharges or resignations based on 13 discharges and 30 resignations. Most of these terminations involved PTOs hired in the summer and fall of 1982, relatively early in the MBTA experience with PTOs. Evidence suggests that many of these early terminations were a result of confusion regarding the implementation of new procedures and a new work force. In fact, the early 1983 termination rate is 16 percent, lower than the 21 percent overall rate to date. Therefore, it appears likely that the part-time operation is stabilizing and that the number of discharges and resignations can be reduced over time. However, the current termination rate for PTOs is still twice as high as that for FTOs.

Experience from other agencies (Seattle and Los Angeles) suggests that although turnover rates are still significantly higher for PTOs than FTOs, the turnover rates did decrease as more experience was gained in screening, hiring, and monitoring PTOs.

One implication of the high PTO turnover rate concerns the impact on costs of training operators. By incorporating the estimated 57 percent annual PTO turnover and any additional training requirements for PTOs, it is estimated that the introduction of PTOs has increased annual costs for training by almost $206,000, approximately 113 percent.

ALTERNATIVE SCENARIOS

An analysis was conducted of various MBTA options under three future scenarios:

1. Keeping the number of PTOs constant;
2. Complying with the recent arbitrator's ruling of a 15 percent limit (per classification) for part-time employees, and
3. Increasing the use of PTOs without restrictions.

Options for the future were analyzed separately for the three scenarios related to the overall number of PTOs that might be available. In the first scenario, the potential for improving the utilization of PTOs was explored under the assumption that the total number remains at 280. This scenario is strictly a short-term one assuming a continuation of the MBTA policy of neither increasing nor decreasing the number of PTOs until the litigation resulting from the arbitration award is resolved. The second scenario focuses on the implications of the arbitration award's being upheld in the courts, requiring the number of PTOs to be reduced to 15 percent of FTOs. Finally, the third scenario is that the courts remove the arbitration award restrictions on the number of PTOs and MBTA continues with its initial plan to expand to 350 to 400 PTOs.

Keeping the Number of PTOs Constant

Under the first scenario, a detailed analysis of the current FTO and PTO runs for three bus rating stations suggested two ways to improve the utilization and productivity of PTO assignments:

1. Schedule PTOs to cover all pullouts (beginning of run) before 5:00 a.m., thereby eliminating all paid meal breaks, and
2. Adjust most of the FTO runs to bring the spread times to just under 10 hr in all (i.e., increase those now about 9 hr).

These two improvements together would eliminate most spread-time penalties.

Currently there are about 80 weekday straight runs on the MBTA surface bus system, each of which by contractual agreement includes a paid break of at least 20 min. An analysis of these straight runs shows that the average weekday paid break is approximately 30 min, yielding 40 paid hr per day without work. Most of these straight runs have pullouts before 5:00 a.m. because by contractual agreement all such runs must be straight if assigned to an FTO.

Currently all pullouts before 5:00 a.m. are served by FTOs who cannot participate in the afternoon peak because of the restriction that they must work a straight 8 hr (they will all be out of service by 1:00 p.m.). By structuring all pullouts before 5:00 a.m. as PTO runs, two benefits are obtained:

1. The paid breaks are eliminated, and
2. By scheduling runs with spreads of approximately 13 hr, these drivers can also participate in the afternoon peak.

It is estimated that the annual savings from the elimination of all weekday straight shifts and the associated paid breaks would be approximately $145,000. This is based on the observation that one driver run could be eliminated when the runs were recut in each of the Charlestown, Cabot, Somerville, and Arborway rating stations.

By eliminating all straight shifts and by lengthening the spread time for many evening runs from the current level of about 9 hr to close to 10 hr, many of the runs that currently have spreads requiring premium pay can be reduced to about 10 hr in total spread. In general, this will involve increasing the average spread time for PTOs slightly, but few (if any) would be required to work spreads of more than 13 hr, which is the policy maximum currently used by MBTA schedulers. In most garages, PTOs would be divided into two groups. One group would serve the earliest pullouts (including all of those before 5:00 a.m.) and work 3.5 to 4 hr in the morning and 2 to 2.5 hr in the afternoon peak. The other group would serve the 6:30-7:00 a.m. pullouts and work 2 to 2.5 hr in the morning and 3.5 to 4 hr in the afternoon peak. This pattern of assigning PTO runs in two groups could be consistently applied systemwide, because our analysis showed it to significantly reduce spread penalties in each of the Cabot, Quincy, and Somerville garages. An additional annual savings of approximately $350,000 is projected from this restructuring of existing part-time runs to reduce spread penalties. If implemented along with the elimination of FTO runs with paid meal breaks, a total of about $0.5 million in transportation costs would be saved annually.

Limiting PTOs to 15 Percent

One possible outcome of the court's review of the arbitration award is that the award will be upheld with respect to the terms of MBTA use of PTOs. The particular implications of such a finding would arise from one clause that MBTA is not now honoring: The maximum number of part-time employees should be 15 percent of the number of full-time employees in the same classification.

The 15 percent limit would imply a reduction from about 280 PTOs to 224 PTOs in the surface-operator classification. In order to estimate the increased cost of service, again assuming that the amount of service provided does not change, it is assumed that the following process is used:

1. Select those part-time runs with the minimum spread time;
2. Convert each run to a full-time run by adding sufficient cover time to bring the run up to 8 hr paid time; and
3. Compute the additional cost of the full-time runs by costing the additional time worked, the spread-time penalties, and the increased benefits accruing to FTOs.

The total cost of the additional 56 FTOs rather than PTOs is estimated to be about $870,000 annually. Some additional economies might be realized by using fewer FTOs to replace the 56 PTOs and increasing the number of FTO runs with large spread penalties, but these additional savings (mainly from lower FTO fringe benefits) would probably not amount to more than 10 percent of the total estimated additional cost.

Although the $0.9 million annual cost is significant, it appears that it can be at least partly offset by the use of PTOs in other MBTA classifications. There are two other classifications in which PTOs could be valuable in reducing premium pay components: rapid transit motorman and rapid transit doorman (guard). In each of these classifications, approximately 20 part-time employees could be hired within the 15 percent limit ruling. After allowing for part-time cover, this implies that the 18 most expensive full-time runs in each classification may be converted to part-time runs, eliminating spread-time penalties and unnecessary cover time for these runs. The estimated savings for these new part-time employees total $0.75 million, as shown in the following:

Saving	Amount ($000s)
Reduced cover hours (20 hr/day)	63
Spread penalties (45 hr/day)	142
Reduced benefits, 20 PTOs	165
Total	370

If this use of part-time employees is adopted, it would require careful prior review of required training to ensure that accidents are prevented. Certainly the increase in accident rates observed for part-time surface operators must be avoided in the rail system. For this reason, it would be much easier to start the part-time employees in the doorman classification than the motorman classification. Nonetheless, in all likelihood, the length of training and consequently its cost would increase, erasing some of the hypothesized savings.

A final element in increasing the savings under the 15 percent limit is the use of some of the 280 part-time surface operators on the streetcar system. Specifically, if use of part-time employees for streetcar operation is sanctioned, with the same safety proviso given earlier for rail transit, a total of 20 PTOs could be shifted to streetcar operation to eliminate all spreads of more than 11 hr and 30 min. This would produce a net additional annual savings of about $63,000.

In sum, it appears that if the arbitration award is upheld, the immediate net impact on annual MBTA operating cost would be almost $0.9 million. However, these costs could be partly recovered by introducing part-time employees into the classifications of rapid transit motorman and doorman (an annual savings of about $0.43 million can be achieved by using 20 part-time doormen and shifting 20 part-time bus operators to streetcars) and by the better use of existing part-time bus operators as discussed under the first scenario.

Increasing Use of PTOs

The final scenario is based on the overturn of the arbitration award as it affects part-time employees, which would allow MBTA to increase the number of part-time employees in any classification without limit. MBTA would have the greatest number of options available under this scenario and the following additional PTO uses were identified with their projected transportation cost savings:

1. Thirty additional PTOs to eliminate all remaining bus system spread penalties, $650,000;
2. Twenty-five more PTOs if midday service adjustments on heavy bus routes are made according to the MBTA Service Policy, $465,000;
3. Twenty-five PTOs to eliminate spread penalties and unproductive cover time on the Green line, $537,000;
4. Thirty PTOs to eliminate spread penalties and cover time for the rapid transit doorman classification, $500,000;
5. Thirty PTOs to eliminate spread penalties and cover time for the rapid transit motorman classification, $500,000;
6. Fifty PTOs to provide all Sunday bus service and eliminate paid meal breaks on Sundays, $750,000; and
7. Any additional PTOs beyond those just shown, $6,300 per PTO.

In considering this third scenario, however, the safety issue becomes paramount. Before there is any expansion of PTO participation in the MBTA work force, the exact causes of the high accident rate must be investigated and identified, and strategies to combat it must be developed and implemented. One possible exception to this generalization, however, is the potential to introduce part-time rapid transit doormen, which should not have any significant safety impact and can further reduce operating costs by about $0.5 million annually.

It is important to recognize the potential long-range implications of a policy of aggressively increasing the proportion of part-time employees at MBTA. Because PTOs are now represented by the Carmen's Union, it must be anticipated that as the number of PTOs increases, their impact on the contract negotiation and bargaining process will also increase. In the next round of bargaining, negotiation might focus on possible limitations and pay penalties on spread time for PTOs as well as FTOs and the incorporation of more fringe benefits into the part-time employee package.

Although it is impossible to predict what the outcome of such negotiations might be, it is important to recognize the potential for a narrowing of the cost differential between part-time and full-time MBTA employees. This again suggests that a policy of slow expansion of part-time employee participation is the most appropriate policy.

SUMMARY

MBTA has made significant progress in the last year and a half in improving productivity through the use of PTOs, and further productivity improvements appear possible. The high rate of accidents, however, suggests caution in expansion of the PTO labor force (except perhaps for the rapid transit doorman classification). MBTA management should develop careful monitoring and remedial training strategies to deal with the PTO accident problem.

Several other possibilities exist to increase the productivity of PTOs:

1. MBTA should experiment (especially under scenario 3) with hiring and assigning PTOs to work only 2 to 5 hr a day for one peak period. The performance of these one-piece PTOs should be compared with that of the existing PTOs.
2. MBTA should consider making the selection

process for promotion (to FTO) more formally structured and weighted more toward a merit rather than seniority basis.

3. A programmed hiring approach should be put into place to ensure that all newly hired and reassigned PTOs have adequate training time and reassignment of both PTOs and FTOs can be made more often to correspond to productivity changes identified by the scheduling department.

4. A range of short-term improvements should be made in planning and scheduling to more easily and quickly respond to changing work-force requirements and to allow a fine-tuning of MBTA service to better meet the region's travel demands.

Using Section 15 Data: Adapting and Evaluating the Magnetic Tape Version for Statistical Analysis

GORDON J. FIELDING, MARY E. BRENNER, and OLIVIA de la ROCHA

ABSTRACT

Data reported as required by Section 15 of the Urban Mass Transportation Act of 1964 have already proved useful in transit decision making. Yet wider use of these data has been inhibited by the difficulty of access to it electronically. A set of strategies for extracting, reorganizing, and evaluating data originating in the electronic data files disseminated by Transportation Systems Center on magnetic tape is described. The current organization of information within the files is unsuitable for most statistical software packages. Therefore, it is necessary to extract information from the Section 15 files and rearrange it in a form suitable for analysis. Different classes of missing data are also defined and remedies for the problem are addressed. In addition the cross-validation of values and the computation of basic transit variables are considered. Many statistical models make assumptions about the distributional characteristics of variables. Differences of scale among transit systems on such measures as size of fleet often result in variables the distributions of which violate these assumptions. Transformations that remedy the problem are recommended.

Since its first release for FY 1979, the reporting system outlined in Section 15 of the Urban Mass Transportation Act of 1964 has proved itself a powerful tool in transit decision making. It has provided standardized definitions of transit activities and recording procedures (1); replaced burdensome and nonuniform data-collection efforts by local operators (2); allowed local, regional, and nationwide comparison of transit performance (3); and facilitated management, performance evaluation, and the allocation of financial assistance at all jurisdictional levels (4-7). In short, analysis of Section 15 data offers greater leverage for understanding transit performance than has hitherto been possible.

The most complete version of Section 15 available for analysis is distributed by the Transportation Systems Center (TSC), Cambridge, Massachusetts, in the form of 62 electronic data files stored on magnetic tape. Although this version promises to be the most useful in the long run, current use of the tape is inhibited by the difficulty associated with reading it and adapting the information to a form suitable for statistical analysis. Considerable time and effort must be allocated to the development of a system for accomplishing the adaptation.

As TSC adopts a new operating system and develops new software for Section 15 data, a wider variety of data tape formats may become available. However, the first 4 years of Section 15 data (FY 1979-1982) share the same organization described in this paper.

An alternative to the magnetic tape is the National Urban Mass Transportation Statistics, UMTA's annual report (8), which provides tabular summaries of Section 15 data. But there are two drawbacks to substituting the printed annual report for the tape version. First, the tape is a comprehensive set of data including far more information than the printed annual report. All levels of reporting are included in the tape, whereas only the required level of information is given in the printed annual report. Entire classes of data such as operating schedules and peak loads are available only on the tape. This additional information permits the cross-validation of values, a critical step in assessing the accuracy of these data. Second, for users who wish to analyze transit systems on a nationwide level or use many variables, the cost of making the printed annual report machine readable could rival or exceed that involved in adapting the tape. For example, the data to be used require keypunching. Then a number of preliminary computational steps, such as converting percentages back to raw values, must be carried out before actual analysis commences. Therefore, it would be useful if a set of strategies could be outlined that would facilitate the use of Section 15 data as it originates on magnetic tape.

This paper describes such a set of strategies. A conceptual scheme underlying the conversion of the magnetic tape data to a conventional statistical format is first described. This is followed by a discussion of data preparation steps that precede statistical analysis and include the treatment of missing data. In conclusion there is a brief evaluation of the distributional characteristics of basic transit variables for FY 1980.

DATA REORGANIZATION

In this section a discussion is presented of why the tape data must be reorganized to make them acceptable to a statistical software package like the Biomedical Computer Programs (BMDP) and the Statistical Package for the Social Sciences (SPSS). The objective is to explain why a software package can "read" the data but cannot, without reorganization, perform a statistical analysis on them. Why reorganization is needed and what steps are required to reorganize are the focus points.

The data reorganization process revolves around four questions concerning data files:

1. What are the basic organizational features common to all numerical data files?
2. What are the distinguishing features of a data file organized for statistical analysis?
3. How are the data files on TSC's magnetic tape different from the statistical convention?
4. What steps are required to reorganize them?

Basic Organizational Features

All data files are organized in rows and columns. A sample file is shown in Table 1. However, the meaning of the data is not inherent in this simple physical organization but must be conveyed to the computer by the programmer. The system or scheme used by the programmer to give meaning to the array of numbers is called the logical organization.

The specification of the logical organization is laid out in a document called a codebook. In a codebook the meaning of data is defined by the way the numbers are organized into sets of columns. A large number like $4,000,000 takes up seven columns, for example. The assigned sets of columns are called fields.

Table 2 shows a codebook from TSC's documentation. According to the codebook, columns 1-4 of the number array have been reserved for transit system numerical identification. Columns 5-12 are reserved for the fiscal year end date for the system identified in columns 1-4. Column 13 is assigned to the mode code. With the help of this scheme the computer can be informed about the meaning of the data by the way fields in the block of numbers are assigned. This process is called formatting.

TABLE 1 Sample Data Set

1001063019801011467500000
1001063019801022138100000
1001063019801031761400000
1002123119801014287100000
1002123119801025891600000
1002123119801033892500000
1002123119802015411600000
1002123119802027382700000
1002123119802039188400000
1004063019801014816500000
1004063019801021810200000
1004063019801031718400000

TABLE 2 Sample Codebook

Column	Name	Type	Description
1-4	TRSID	Integer	Transit system identification
5-12	FY	Date	Fiscal year
13	MODE	Integer	Mode code
14-15	EMCOD	Integer	Employee class code
16-21	OLABR	Real	Operating labor
22-27	CLABR	Real	Capital labor

In formatting, space is set aside in the array and named so that any number found in that space by the computer can be presumed to have the assigned meaning. Any number found in columns 1-4 of the sample block of data (Table 1) will mean Transit system identification number to the computer as long as it is formatted in that way.

It is important to realize that there is some flexibility in the way that data may be formatted. That is to say, there may be more than one meaningful logical organization for the same physical file.

Finally, an actual line of data like 10041114676 that can be formatted is called a record. A record may take up one or more rows and there may be more than one type of record in a data file.

Statistical Files

Statistical procedures operate by making systematic comparisons among objects. The objects are compared on those attributes that have been measured in some way. For example, in Section 15 analyses transit systems are compared on such attributes as size of fleet and speed.

In statistical data files the most important organizational units are cases (objects) and variables (attributes). A case may be thought of as the full collection of information items defined in the codebook for a single transit agency. If some defined item is missing, the statistical case is incomplete, and a place-holding code must be inserted to fill it out.

A variable, like a case, is a statistical concept. When all cases have been measured on a given attribute, the resulting collection of values is organized in a list called a variable. Statistical procedures compare these lists and depend on the fact that the cases always appear in the same order. Once again, if no place-holder resides in the position of a missing item, the order is disturbed and statistical results are rendered meaningless.

One danger to be avoided in comparing all numerical data files to a smaller subset of them, i.e., statistical files, is that the distinctions between their separate terminologies will blur. It is important to keep in mind the differences between the horizontal concepts such as the row, the record, and the case on the one hand and the vertical concepts such as the column, the field, and the variable on the other. In general use, the members of these trios are often used interchangeably. Because understanding the data reorganization process may hinge on the distinctions among them, a glossary is provided at the end of this paper.

Organization of TSC Tape Files

The organization of the TSC tape files is closely linked to that of the reporting forms. Form 404, Transit System Employee Count Schedule, provides an example. Figure 1 shows Form 404 (top) and the information for one transit system as it appears in a data file on the tape (bottom). Spaces have been inserted between the fields for ease of reading. In the actual file there are no spaces.

A comparison of the form and the data shows that the first three fields, transit system identification, fiscal year ended, and mode, come from the top of Form 404, and are repeated on every record in the data. The next two fields, employee code and operating labor, are taken from the Employee Classification and Operating Labor sections of the form. Figure 1 shows a one-to-one correspondence between the numbers assigned to employee categories on the form (11, 12, 13, etc.) and the values under EC in the data file. However, the one-to-one correspondence is not quite complete. If Form 404 were used to construct a codebook that acted as the logical organization for the data appearing in Figure 1, there would be a discrepancy between what the logical organization predicts and what actually appears in the physical file. There is no record appearing for category 22, Maintenance Support Personnel, in the data.

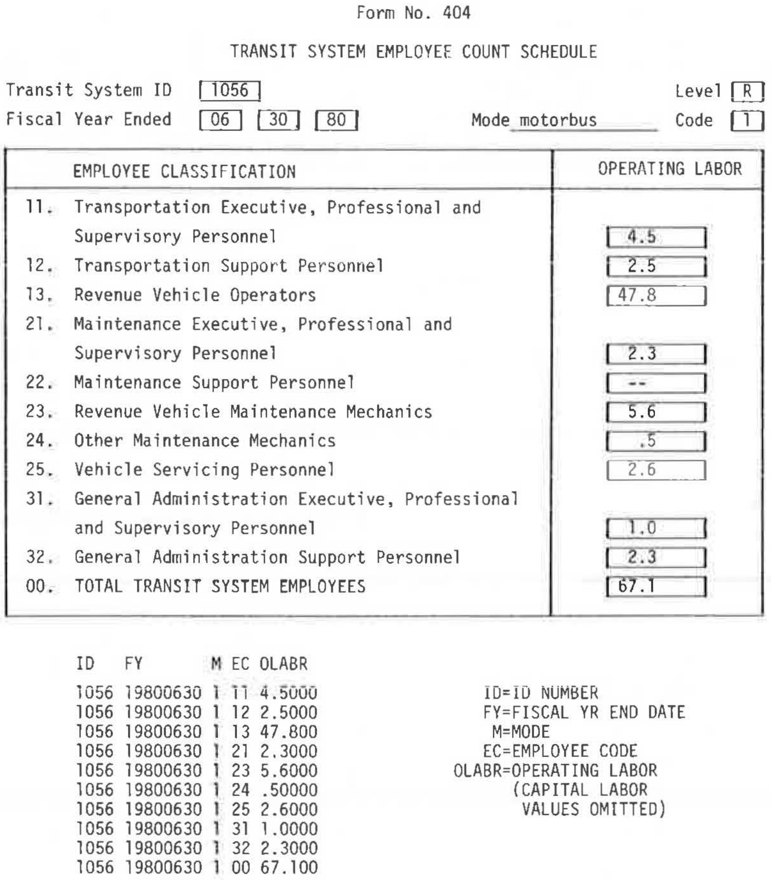
To reiterate, most statistical software packages require some entry to stand in for the missing category 22. Until a stand-in value is substituted, the information cannot be said to form a complete case. Therefore, all such instances of missing records must be remedied before statistical analysis can proceed. Only two widely available software packages, SAS and SPSS-X, are known to have methods for dealing with this problem.

Another important consequence of the correspondence between the data and the forms is the way in which values are being compared, that is, which values are making up the variables. Consider once again Figure 1. In a statistical routine, the OLABR value of 4.5000 cannot be compared with the OLABR value of 2.5000 beneath it. The 4.5000 must be compared with another value, not shown here, which also has an EC of 11. OLABR, therefore, is not one variable, but 11 variables (the number of employee classifications) collected together in one field.

Informing the computer of this relationship between the values in the OLABR field requires devising a new logical organization to replace that found in the TSC codebook. For statistical purposes, OLABR is too general a category to qualify as a variable. It would not be useful to compare the number of revenue vehicle operators in one system with the vehicle servicing personnel in another system. Instead, revenue vehicle operators must be compared with revenue vehicle operators. A variable, then, would be all instances of OLABR for category 11 or all instances of OLABR for category 00, Total Transit System Employees.

The TSC data file organization is common, economical, and often used as input to management information systems using customized software. In computer science it is referred to as hierarchical ordering.

To summarize, two major differences needing reconciliation between statistical files and TSC files are the omission of stand-ins for missing records and hierarchical organization. The concept of missing data is an important issue in its own right and is discussed more fully in a later section.



Form No. 404

TRANSIT SYSTEM EMPLOYEE COUNT SCHEDULE

Transit System ID 1056 Level R
Fiscal Year Ended 06 30 80 Mode motorbus Code 1

EMPLOYEE CLASSIFICATION	OPERATING LABOR
11. Transportation Executive, Professional and Supervisory Personnel	4.5
12. Transportation Support Personnel	2.5
13. Revenue Vehicle Operators	47.8
21. Maintenance Executive, Professional and Supervisory Personnel	2.3
22. Maintenance Support Personnel	--
23. Revenue Vehicle Maintenance Mechanics	5.6
24. Other Maintenance Mechanics	.5
25. Vehicle Servicing Personnel	2.6
31. General Administration Executive, Professional and Supervisory Personnel	1.0
32. General Administration Support Personnel	2.3
00. TOTAL TRANSIT SYSTEM EMPLOYEES	67.1

```
ID   FY       M EC OLABR
1056 19800630 1 11 4.5000             ID=ID NUMBER
1056 19800630 1 12 2.5000             FY=FISCAL YR END DATE
1056 19800630 1 13 47.800              M=MODE
1056 19800630 1 21 2.3000             EC=EMPLOYEE CODE
1056 19800630 1 23 5.6000          OLABR=OPERATING LABOR
1056 19800630 1 24 .50000                (CAPITAL LABOR
1056 19800630 1 25 2.6000                 VALUES OMITTED)
1056 19800630 1 31 1.0000
1056 19800630 1 32 2.3000
1056 19800630 1 00 67.100
```

FIGURE 1 Correspondence between reporting system forms and logical and physical organization of TSC data files.

Implementing Reorganization

The main goals of reorganization are to supply stand-in values for missing records and to reformat instances in which several variables have been grouped together in one field. A hypothetical example of the results of reorganizing is shown in Figure 2.

There are several noteworthy features in Figure 2. First, in File 1 under the field System Identification, there is no information for system number 1003, and systems 1002 and 1004 appear to have only half the information they need.

Also in File 1, the field Wages can be seen to contain six different variables. The values in the fields Mode and Employee Category must be used to find these variables. For example, the first wage value, 500, has a mode of 1 and an employee category of 0. These values indicate that the first 500 is for motor bus drivers' wages. Hence, the only other value it can be compared with is wages of 650, six lines down in case 1002, which also has a mode of 1 and an employee category of 0. There are six wage variables possible because in addition to the mode and employee category combination of 1 and 0 there are also the combinations of 1 and 1 or 1 and 2, and so on. Because there are two values of mode and three values of employee category, it takes two times three, or six, combinations to exhaust all pairs possible.

The six variables each have their own separate field in File 2. The information in the mode and employee category fields from File 1 has been incorporated into the new logical organization of File 2. Therefore, they disappear from File 2. File 2 also has full sets of information (complete cases) for all transit system identifications, although missing value codes of 999 had to be inserted to make this possible. For example, even though system 1002 has no trolleybusses, stand-in values of 999 were inserted in the three trolleybus variables for this case.

In summary the basic reorganization steps can be reduced to four:

1. Using the logical organization in the TSC codebook, in which a case is not a transit system but a single record, read and write the data, eliminating unwanted information;
2. Locate the positions in the retained data needing stand-in values;
3. Insert the stand-in values; and
4. Format the data with a new logical organization that considers all the records belonging to a single transit system as a case. Once the stand-in values have been inserted, this number of records will be the same for all transit systems.

Working with the Tape

The objective of this section has been to explain the reasons for data reorganization and the steps that are necessary to accomplish this task. A technical manual has been prepared that explains some of these steps in more detail (7). The complexity of

DATA FILE 1. HIERARCHICAL ORGANIZATION

SYSTEM ID	MODE	EMPLOYEE CATEGORY	WAGES
1001	1	0	500
1001	1	1	600
1001	1	2	600
1001	2	0	400
1001	2	1	700
1001	2	2	700
1002	1	0	650
1002	1	1	600
1002	1	2	700
1004	2	0	700
1004	2	1	000
1004	2	2	000

MODE
1 = MOTOR BUS
2 = TROLLEY BUS

EMPLOYEE CATEGORY
0 = DRIVER
1 = MAINTENANCE
2 = ADMINISTRATION

DATA FILE 2. STATISTICAL ORGANIZATION

SYSTEM ID	MTRBUS DRIVER WAGES	MTRBUS MAINT WAGES	MTRBUS ADMIN WAGES	TRBUS DRIVER WAGES	TRBUS MAINT WAGES	TRBUS ADMIN WAGES
1001	500	600	600	400	700	700
1002	650	600	700	999	999	999
1003	999	999	999	999	999	999
1004	999	999	999	700	000	000

999 = MISSING VALUE CODE

FIGURE 2 Hypothetical data file before and after reorganization.

the task lies not in the nature of the problems so much as in the large amounts of data that must be manipulated and the number of steps required to carry out the manipulations. Some statistics concerning the data files make this clear.

In FY 1980 there were 62 data files. Twenty were text files containing labels and 42 were numerical data files. The files ranged in size from approximately 300 records to 22,000 records, and all 62 files combined required 775,000 words or 3,800,000 characters. For comparison, the printed annual report is made up of approximately 2,100,000 characters.

The large number of steps required to reorganize a file is quite surprising. The most complicated expense file contained 22,000 records and required the use of more than 75 temporary data files during the process of inserting more than 2,000 needed stand-in values.

DATA PREPARATION

Once the data have been reorganized, additional data preparation is required before analysis can commence. There are three phases to preparing the data: calculating basic variables, identifying and flagging missing information, and validating existing data.

The Section 15 database contains a wealth of information too detailed for many purposes. The data to be used for statistical analysis must be customized. The purpose of this report was a comparative analysis of motor-bus performance in terms of general concepts such as labor efficiency and utilization of service. Thus aggregation of many small pieces of information into more comprehensive variables that contain only information about the motor-bus mode and that are applicable to an entire year's operation was necessary.

The information about transit employees is a clear example of too much information that must be summarized into broader categories. Ten employee categories are reported--three in vehicle operations (i.e., supervisors, revenue vehicle operators, support), five in maintenance, and two in general administration. These 10 categories are further subdivided into capital labor and operating labor. For the purpose of this report it was necessary to know the number of vehicle operators, the number of maintenance employees, and the number of administrative employees. The first step in creating these variables was to add together operating and capital employees because there was no interest in this distinction. At this point, the number of revenue vehicle operators was ready for use. The number of maintenance employees was calculated by adding together the five categories of maintenance employees. The number of administrators was calculated by adding together the supervisory personnel in vehicle operations and maintenance to the two categories of administrative personnel.

Other variables that must undergo this aggregation process include the total number of accidents, total amount of subsidies, and the miles of roadway used on bus routes.

Estimating Annual Data

The data on service supplied by a transit agency and the service consumed by passengers must undergo a different kind of calculation before they can be used in a general analysis. Although the Section 15 reporting system requires that all financial data be reported for a complete fiscal year, information on service variables such as unlinked passenger trips and revenue vehicle hours is collected by a sampling procedure and reported for an average weekday, average Saturday, and average Sunday. This information must be combined with a formula that annualizes it so that it is comparable with the financial data. A formula was used that allowed for 253 weekdays, 53 Saturdays, 52 Sundays, and 7 holidays (also calculated as Sundays); each of these numbers was multiplied by the given values for average weekdays, Saturdays, and Sundays.

A series of calculations was also needed to disaggregate data so that they applied only to the motor-bus mode. Revenue and subsidy information is reported in the Section 15 system for entire transit systems, not by mode. In addition multimodal systems have the option of reporting expenses as joint expenses between modes, and a few systems report most of their expenses in this way. A series of weighting formulas were designed that allow assignment of revenues or joint expenses to specific modes. For example, a proportion of passenger revenue is assigned to the motor-bus mode by multiplying the system's total passenger revenues by the ratio of motor-bus passengers to total passengers. Although the resulting values are only estimates, they are better than the distortions caused by using overly large figures or dropping the multimodal systems (32 percent of the systems reporting in 1980) from the analysis.

Missing Data

The second phase of preparing data for analysis is detecting those cases that have missing data and that therefore must be eliminated from further analysis. A database prepared for statistical analysis will usually have a special symbol such as -9 that indicates that information is missing. However, the Section 15 data tape has no such special symbol, and the analyst must therefore insert one during the process of calculating the variables. The analyst is able to detect missing-data problems by considering the logical properties of specific variables, by comparing a variable to other information in the database, and by comparing the Section 15 data with other sources of information (including the analyst's own knowledge of transit systems).

For some variables, detecting missing data is straightforward and quite logical. For instance, a transit system that has zero operating expenses can readily be assumed to have a missing-data problem. But most variables require more judgment on the part of the analyst. It is possible for a transit system to have zero accidents for a given fiscal year, but the larger a system, the less likely it is that it will have no accidents. The analyst must examine other transit systems of similar size to the one reporting zero accidents to see whether zero is a possible number. A cross-year comparison of reported accidents gives the analyst further evidence on which to base a decision. For this report it was decided that any system with more than 10 revenue vehicles could not have zero accidents, and a missing-data symbol was inserted for these systems. Other systems were then judged individually, taking into account their peak vehicle size (a better measurement of size than the revenue vehicle fleet), their safety record in other years as reported in annual reports or reports to the American Public Transit Association (APTA) (9), and the performance of like-sized systems.

Some judgments about missing data involve making decisions about whether a concept is adequately measured by a combination of several different variables. For instance, vehicle maintenance can be supplied by employees on the transit agency payroll or by contract with other organizations. Thus if a system reports zero maintenance employees, the analyst would expect to have zero maintenance wages reported but a substantial expenditure for services indicated under either the maintenance function or general administration. In the absence of wages and service expenses, a missing-data symbol would be used to indicate that maintenance expenses are missing.

For some other variables, the decision is more complex because a zero value can be a real value or it can be an indication of a problem. The example of total vehicle miles will make this clear. Total vehicle miles, as noted above, is constructed from three variables--average weekday miles, Saturday miles, and Sunday miles. If weekday miles are zero, it can be assumed that information is missing. However, many systems do not offer weekend service, so a zero for Saturday or Sunday miles might be real or might be an indication of a problem. Because this information is based on a time-consuming sampling procedure, there is a definite possibility that a transit system failed to collect this information and thus has a missing-data problem. The Section 15 data tape includes information about the service schedule of each system. Therefore, it is possible to determine whether a system offers Sunday service and whether it has a missing-data problem. This kind of cross-checking of variables is possible only with the data tape because the annual report does not carry information on service schedules.

The problem of missing data has received detailed attention because it is an inevitable problem with a database as complex as the one mandated by Section 15. More than 300 different systems must learn to interpret and fill out numerous forms, which range from 17 pages for a small, single-mode system to 90 pages for a large, multimodal system. Because 1980 was only the second year in which this information was reported, some systems were still in the process of instituting accounting systems compatible with Section 15 requirements. Although missing data will become less of a problem as transit systems become accustomed to the reporting requirements, there will always be new systems completing the forms for the first time. In 1980, 321 systems reported; in 1983, 414 systems are expected to report.

Missing data are not evenly distributed across variables or transit systems. In 1980 the most complete data available were for economic variables such as operating expenses and passenger revenue (see Table 3). The most incomplete information available was for passenger measures such as unlinked passenger trips and passenger miles.

The missing-data problem is particularly acute for small systems, those with fewer than 25 revenue vehicles. In 30 percent of these systems, information on passenger trips and in 6 percent information on expenses is missing. Although it is still possible to analyze the smaller systems, because more than one-third of all systems fall within this size category, generalizations to all small systems must be made cautiously.

The failure to identify missing values with a special symbol can greatly distort the results of a statistical analysis. If too many zeroes are allowed to remain in the data, the mean for a variable will be unrealistically low whereas the standard deviation will be too high or distorted. Unwarranted conclusions will also be drawn if care is not taken. For instance, it would look as though small systems carry many fewer passengers per peak vehicle because small systems are missing 30 percent of the data on this measure whereas the large systems are missing only 13 percent of the data.

TABLE 3 Distribution of Missing Data in Selected Transit Variables

Variable	Missing Values[a] (%)
Passenger revenue	0.7
Total operating expense	2.0
Total employees	2.6
Total vehicle miles	8.2
Unlinked passenger trips	18.1
Passenger miles	24.0

[a] Out of 304.

Data Validation

The final phase of data preparation consists of cross-checking the data for validity. Errors can enter the database in many ways--misinterpretation by a transit system of what number should be reported, miscalculation of totals, and keypunching errors as data are prepared for the computer. Four major methods were used to validate the data: recomputation of totals, comparisons of redundant information, comparisons of related information, and comparison with feasible value ranges. An example of each of these methods with specific variables will be given.

The total number of employees reported for each system was compared with the sum of the separate categories. In about 10 cases, the totals differed by more than could be accounted for by rounding errors. In most cases the differences were apparently caused by keypunching errors (e.g., reversal of digits) or simple miscalculations. For these cases, reported totals were replaced by the recalculated totals and cross-checks were made with the annual reports.

Much of the financial data are reported in several different places. For instance, the Revenue Summary Schedule (Form 201) summarizes the information on the Revenue Subsidiary Schedule (Form 203). Total operating expenses are also reported in two different places on the magnetic tape. A simple comparison of these numbers reveals differences and the correct number can often be identified by the other validation methods.

Different variables in the database are sometimes different measures of the same thing. For instance, employee counts and employee wages are two different measures of labor utilization. If a transit system has a large number of vehicle operators, it must have a proportionately large amount of vehicle operator wages. However, caution must be used in some of these comparisons. Maintenance-employee counts and maintenance wages are subdivided into distinct, noncomparable subgroups.

The final method of identifying mistakes is to look for values that lie outside an expected range for that specific variable. This method works best for measures that are combinations of two variables such as miles per hour or cost per passenger. Miles per hour (speed) has an expected range of about 5 miles per hour (dense urban areas) to 30 miles per hour (commuter service). If the values for any system fall outside this range or are in the wrong part of the range for the kind of service offered, there has probably been a mistake in its measure of either miles or hours.

A variable such as cost per passenger is a little more difficult to work with because inflation and difference in fiscal years cause the feasible range to change over time and the boundaries of a feasible range are indefinite. In this instance all cases that lay more than three standard deviations from the mean as well as the largest and smallest cases were examined. Although some of these outliers had apparent, real causes, such as extremely long trip lengths, others were so different from the norm that they were obviously wrong. In these cases the correct values were sought in other parts of the database or in other sources. If a correction was impossible, incorrect values were designated as missing.

In the future many of these validation procedures will be incorporated into the preparation of the Section 15 data tape. Beginning with the 1981 data, totals and internal measures of validity were checked for each transit system by TSC. However, the last validation procedure outlined in the foregoing will remain a useful procedure for the next few years because it looks at a transit system in relation to other systems. TSC also compares a system to itself across years as another validity check. Although this was done for specific problems in this validation procedure, it was not done systematically.

UNIVARIATE PROPERTIES OF VARIABLES

To be useful for statistical analysis, a data set must meet the basic assumptions of the specific technique to be used. One common assumption is that a variable is normally distributed. Another common assumption is that the variances of two variables are equal. Tables 4 and 5 show a set of variables from the Section 15 database and the statistics that show whether they are normally distributed.

The most striking characteristic of the Section 15 data is the great variation of values for many variables. The major reason for this is the great range in size of transit systems. As the number of peak vehicles in Table 4 shows, transit systems can be small or large. Most other variables such as expenses or passenger trips will have correspondingly wide variation.

A normal distribution can be described in terms of a few characteristics. The mean is an average value for a variable. Most values will be quite close to the mean. In fact, 95 percent of the cases will be within 2 standard deviations larger or smaller than the mean. As the value of a variable gets farther from the mean, fewer cases will have that value. In addition there are just as many values larger than the mean than smaller in a normal distribution.

TABLE 4 Extreme Ranges of Typical Transit Variables

Variable	Mean	Standard Deviation	Range	
			Min	Max
Unlinked passenger trips per dollar cost (000s)	1.35	0.57	0.203	3.55
No. of passengers per peak vehicle (000s)	95.70	51.80	10.290	360.00
No. of peak vehicles	124.70	316.40	1.000	3,378.00
Operating expense ($000s)	12,462.00	41,560.00	10.000	441,060.00
Unlinked passenger trips (000s)	22,118.00	88,655.00	10.000	1,139,560.00

TABLE 5 Comparison of Normal and Nonnormal Variables

Variable	Variance	Skewness	Kurtosis
Unlinked passenger trips per dollar cost (000s)[a]	0.32	0.81	1.27
Unlinked passenger trips per peak vehicle (000s)	2683.24	1.57	4.47
No. of peak vehicles	100,109.00	5.98	46.40
Operating expense ($000s)	1,727,233,600.00	7.80	76.75
Unlinked passenger trips (000s)[b]	7,859,709,000.00	9.42	105.68

[a] Most normal.
[b] Least normal.

The skewness of the variable shows, relatively, how many of the cases are either larger or smaller than the mean. A normal distribution has a skewness of zero because there is no difference between the number of larger and smaller cases. The kurtosis shows, relatively, how many cases are closely bunched together. A normal kurtosis is also zero.

As Table 5 shows, the Section 15 variables vary greatly in terms of how normal they are. In fact, most common variables that describe aspects of a transit system--such as number of peak vehicles, operating expense, and unlinked passenger trips--deviate greatly from normality. The distributions of these variables are greatly skewed because many more systems fall below the mean than above it. The distributions have a high kurtosis because the small systems are quite similar to each other, whereas the large systems are more disparate.

Figure 3 shows this graphically. The solid line shows a normal distribution. The segmented line shows a typical transit-variable distribution. The high skewness of the distribution can be seen in the way most cases fall to the left of the mean. The high kurtosis can be seen in the way the transit variable's peak is higher than that of the normal distribution.

Although many statistical techniques can tolerate some departure from normality, this work has shown that the direct use of these variables produces meaningless results. For instance, a near-perfect regression correlation can be obtained between passenger revenue and unlinked passenger trips, but predictions are wrong by as much as 10,000 percent for small transit systems.

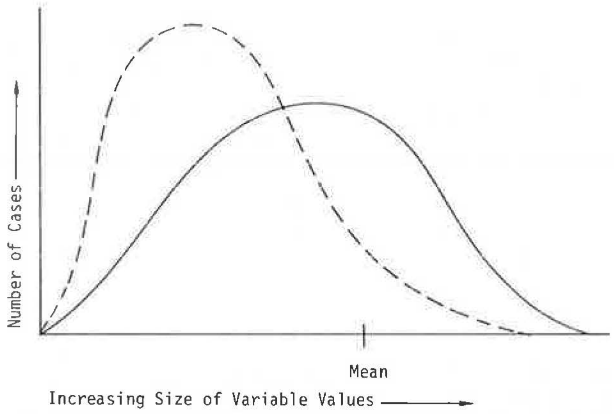


FIGURE 3 Comparison of normal distribution and distribution of nonnormal transit variable.

Table 5 also shows that the variances of different variables are quite different. Although transit systems show great variation in their size, this variation is exponentially increased in the variance measure. Thus great care must be used when variables are combined in statistical analyses. For some purposes, standardization will take care of variance problems. But for other purposes, the entire distribution must be transformed.

Because the departures from normality are a consequence of the great range in size of transit systems, any correction for size will make a more normal distribution. The first variable in Tables 4 and 5, unlinked passenger trips per dollar cost, shows this effect in action. Large numbers of passengers and high expenses tend to go together, so the ratio of the two corrects for the largeness or smallness of a transit system. This ratio variable is more normal than either of the variables that were used to compute it. However, some ratios such as passengers per peak vehicle are less normal because the original variables were not equally distorted by the effects of size, as shown in the greater differences in their variance, skewness, and kurtosis.

Another technique that corrects for nonnormality is a logarithmic transformation of a variable. This transformation causes the outlying, large systems to be more proportionately scaled to the rest of the transit systems. Other methods for coping with the nonnormality can be devised, including elimination of large outliers and analysis with smaller peer groups of transit properties that are relatively homogeneous with respect to size. However, these methods reduce the sample size and potentially eliminate important variance in the data. The method chosen should depend on the goals of the statistical analysis.

SUMMARY

The Section 15 reporting system has created a rich, new source of data for analyzing the performance of the transit industry for both researchers and transit managers. For those who want a limited amount of information on a few systems, the published annual reports provide easy access to basic information. However, for those who wish to use large samples, information in great detail, or information reported at the A, B, or C levels, the magnetic data tape provided by TSC is the better source.

In this paper methods have been outlined for using the magnetic tape, including the reorganization of data for use with statistical software, calculating basic variables, identifying missing information, and validating the data values reported. In addition, some cautions are given for using the data because the pattern of missing data makes the existing data not perfectly representative of the transit industry and many of the data variables are not normally distributed.

In coming years, access to valid, reliable data on the transit industry will become increasingly available. Missing-data problems will decrease as the transit industry becomes familiar with the Section 15 reporting requirements. Beginning with the FY 1981 data, TSC has begun extensive validation checks. In addition, the data are being distributed in new ways. The same information that is reported in the printed annual report for 1981 is now available on diskettes for minicomputers. A magnetic data tape in a sequential format is also available for the 1981 data. Although this data tape reduces the 62-file structure into two files, it has the same formatting problems delineated in this paper.

Beginning with the FY 1983 data, TSC will be using a new operating system and will begin to explore new ways of distributing the data for specific purposes such as statistical analysis. However, analysts who wish to work with the first 4 years of data will need to reorganize and clean the data before beginning further analyses.

GLOSSARY

- Case: A statistical concept; the full collection of information items defined in a codebook for a single transit company; to be distinguished from a row or a record.
- Codebook: The scheme by which data are organized in sets of columns; the logical organization of a data file. A sample codebook appears in Table 2.
- Field: A set of columns reserved for one kind of information, to be distinguished from a column or a variable.
- Format: The imposition of a logical organization on a physical file; the act of communicating to the computer how data are defined.
- Hierarchical ordering: A data file organization scheme in which a field may contain more than one variable and missing records are permissible.
- Logical organization: The scheme contained in a codebook by which data are broken up into fields. There may be more than one meaningful logical organization for a given data file.
- Record: A formattable string of numbers; not to be confused with a case or row.
- Variable: A statistical concept; when all cases have been measured on a given attribute, the resulting collection of values organized into an ordered list.

ACKNOWLEDGMENT

Research for this paper has been supported by UMTA, U.S. Department of Transportation.

REFERENCES

1. J. M. Holec, D.S. Schwager, and A. Fandialan. Use of Federal Section 15 Data in Transit Performance Evaluation: Michigan Program. In Transportation Research Record 746, TRB, National Research Council, Washington, D.C., 1980, pp. 36-38.
2. E. L. Owens. Utilization of Section 15 Data by State Governments. Presented at 59th Annual Meeting of the Transportation Research Board, Washington, D.C., 1980.
3. Uniform Data Management System: System Development and Testing. Report RPT-3715(001)-93-52, DOT-I-81-2. Iowa Department of Transportation, Ames, Oct. 1980.
4. F. W. Sherkow. Billing and Accounting by Use of a Computerized Data Reporting System: The Iowa Experience. In Transportation Research Record 823, TRB, National Research Council, Washington, D.C., 1981, pp. 12-18.
5. Florida Transit System Performance Measures and Standards. Division of Public Transportation Operations, Florida Department of Transportation, Tallahassee, 1979.
6. J. M. Holec, Jr., D. S. Schwager, and M. J. Gallagher. Improving Usefulness of Section 15 Data for Public Transit. In Transportation Research Record 835, TRB, National Research Council, Washington, D.C., 1981, pp. 9-15.
7. G. J. Fielding, M. E. Brenner, and O. de la Rocha. Using Section 15 Data for Transit Performance Analysis. Interim Report. UMTA, U.S. Department of Transportation, Jan. 1983.
8. National Urban Mass Transportation Statistics--Second Annual Report: Section 15 Reporting System. UMTA, U.S. Department of Transportation, July 1982.
9. Operating Statistics Report 1981: Transit System Operating Statistics for Calendar/Fiscal Year 1980. American Public Transit Association, Washington, D.C., Oct. 1981.

A Methodology for Comparative Transit Performance Evaluation with UMTA Section 15 Data

A. G. HOBEIKA, C. KANOK-KANTAPONG, and T. K. TRAN

ABSTRACT

The obstacles to the comparative evaluation of transit performance lie chiefly in the nonconformity and inaccuracy of the early data and also the inadequate coverage of the local operating characteristics. With the publication of the annual reports required by Section 15 of the Urban Mass Transportation Act of 1964 since May 1981, the first obstacle has been overcome. However, because of human error in compiling the data and the format of the report, there have been many shortcomings in the first two annual reports. These shortcomings together with their solutions were outlined. In an attempt to overcome the second obstacle, a set of indices related to the costs, demand, and revenues was developed for each bus system. Each index is defined as the ratio of the difference between the actual and the expected performance measures to the expected performance measure. The expected performance measure was derived from the regression models fitted on the second-year Section 15 data. With this approach a positive index means that the bus system performs better than its expected performance and also better than its peers. A negative value, in contrast, denotes an inferior performance. A zero value of the index is thus the average. Only results for systems with 25 to 99 vehicles are presented.

To measure the performance of a transit company, performance measures or indicators are generally used. In the past the development of transit performance measures was difficult because of the lack of a systematic, consistent, and accurate database (1). The early financial and operating statistics of transit systems were found to have many errors and limitations because of the structure of the reporting system, the lack of precise definitions of terms, and the lack of cooperation from transit companies (2). As a consequence, most early researchers had to rely on the data from local transit systems for performance studies, which presented the problem of nontransferability (from one system to another). Hence, any comparative evaluation involving more than one system was considered impossible by some practitioners (3) and dubious by the transit association (4).

An attempt to correct many of the early problems was made through the amendment of the Urban Mass Transportation Act in 1974. Section 15 of this act requires that all transit systems receiving federal aid must report common data to UMTA (2). The result of this requirement has been the publication of UMTA Section 15 annual reports on common transit data since May 1981 (5). The reporting system employed in this publication is now regarded as the industry standard. For example, the presentation format of the 1981 edition of the American Public Transit Association (APTA) Transit Fact Book (6) was radically altered from the previous one to be in line with this standard; the same is also true for state data sources (7).

Because the Section 15 reporting system is in its initial stage, an analysis of the reported data is needed. The availability of the data also allows for the development of a comparative performance evaluation methodology. This is the primary purpose here. The specific objectives of this paper are to identify the shortcomings of the UMTA Section 15 data and to formulate indices for evaluating the performance of transit systems through the use of multiple regression analysis.

UMTA SECTION 15 REPORTING SYSTEM

Contents

Pursuant to Section 15, transit systems are required to use one of four reporting levels: Required, A, B, or C (8). Although the Required level contains the least compulsory transit information, level A, the most comprehensive level, is recommended for systems with more than 500 revenue vehicles; level B for systems with 101 to 500 vehicles, and level C for systems with 1 to 100 vehicles. The forms (9-12) for reporting transit data for the four levels cover information related to capital resources, revenues, expenses, and system characteristics. The total number of data items to be reported varies depending on the reporting level required, the size of the system, and the number of modes operated. This number may be as low as 396 items for single-mode systems with fewer than 25 vehicles and as high as 2,385 items for multimode systems with more than 25 vehicles.

Shortcomings and Their Solutions

The inauguration-year data of the UMTA Section 15 reporting system contain many errors; attempts by Anderson and Fielding (13) to fit cost models were not significant, in spite of a rigorous checking method employed (14). As for the second year (15) (covering the period between July 1, 1979, and June 30, 1980), there are also some inherent and acquired shortcomings.

Inherent Shortcomings

These shortcomings are ingrained in the reporting system and thus cannot be corrected for the current use of the data. However, in order to improve the quality of future data, the reporting system itself must be modified. Following are some of the problems found in the data.

Revenues Not Reported by Mode

The reporting forms do not require multimode systems to separate revenues by mode, making it impossible to estimate the revenues generated by each mode. Hence, data items for multimode systems must be omitted from any analysis that requires the revenue of an individual mode.

Service Area and Population Not Well Defined

The reporting forms do not require data on population and region served, but these two items are crucial for the building of the demand and revenue models. The values provided in the tape and the report are estimated according to the UMTA area codes. As a result, confusion arises. For example, 10 systems were reported to serve urban area code no. 2 in California. All systems are listed as serving the same urban area population, even though their revenue vehicles number from 4 to 2,731. To remedy this problem, the latest census of population and housing was used in estimating the population. As for the land use, the County and City Data Book (16) was consulted. The use of the land area data for this study was based on the assumption that the urban land area remained unchanged. Although this assumption is not true, no other sources of data were available during the research.

Joint Expenses Not Reported by Mode

For multimodal systems, the direct and joint expenses by mode and object class are given in File 14 and File 17, respectively. However, there is no listing for joint expenses contributed by each mode. Therefore, the total expenses by mode and object class cannot be derived. In order to prevent the unnecessary omission of all multimodal systems in the study, a check was performed on File 17 to see whether any joint expense was zero. If the value was zero, the corresponding system was included.

Conflicting Demarcation Point in System Grouping

It is compulsory for systems with more than 25 revenue vehicles to complete Form 321 (Operator's Wages Subsidiary Schedule) and Form 331 (Fringe Benefits Subsidiary Schedule). However, in the annual reports, transit systems are stratified according to number of revenue vehicles, as follows: 1 to 24, 25 to 49, 50 to 99, 100 to 249, 250 to 499, 500 to 999, and 1,000 or more. Hence the compulsory transit data in Forms 321 and 331 are missing for the 25-revenue-vehicle systems (in the second group). This oversight has been made twice, and it is likely to be repeated again in the forthcoming reports.

Acquired Shortcomings

These shortcomings are acquired because of human error during the filing, keypunching, and checking of the transit data reports. They were corrected for the current use of the data based on the reasonableness of the values reported. The main errors found were as discussed in the following.

Missing Data

All missing data were replaced with zero values in both the tape and the report, which makes it difficult to know whether the zero value is real. To correct the situation, the reporting level of each transit system was examined. For the data items required to be reported, there are variables that cannot realistically assume a zero value, such as operating expenses and number of employees. Thus, if a zero value was reported, it must mean that data are missing. For data items that are optional it is harder to determine whether the zero is real or whether it indicates missing data. The nonzero variables identified at all reporting levels are given in Table 1 along with the number of bus systems having missing data items. It is quite obvious that the smaller the fleet size, the larger the number of missing data items. This may be because there were not enough personnel to handle the reporting task.

Data in Annual Report Do Not Match

Close examination of the report contents revealed

TABLE 1 Number of Bus Systems with Missing Data

	No. of Systems with Missing Data by System Size						
Nonzero Variable	1 to 24 (N=108)	25 to 49 (N=67)	50 to 99 (N=49)	100 to 249 (N=35)	250 to 499 (N=21)	500 to 999 (N=9)	1,000 or more (N=13)
Revenue vehicle operators (DR)	5	2	1	-	-	-	-
Operator's wages (DRWG)	10	-	-	-	-	-	-
Total employees (EMP)	3	2	1	-	-	-	-
Operating expenses (EXP)	1	-	-	-	-	-	-
Average fleet age (FLEETAGE)	9	-	-	-	-	-	-
Fuel consumption (FUEL)	6	1	1	1	-	-	-
Line miles (LM)	1	-	-	-	-	-	-
Material and supply expenses (MATX)	6	-	-	2	-	-	-
Unlinked passenger trips (PAS)	22	8	4	4	2	-	-
Unlinked passenger miles (PASM)	29	12	5	6	4	1	-
Passenger fare revenue (PASR)	2	-	-	-	-	-	-
Road calls—failures (RCAL)	12	-	3	1	-	-	-
Revenue capacity miles (RCM)	14	2	1	2	-	1	-
Time per linked passenger trip during weekdays (TRIPTIME)	28	10	5	6	3	2	-
Transportation revenue (TRR)	2	-	-	-	-	-	-
Vehicle hours (VH)	12	2	2	2	-	1	-
Vehicle miles (VM)	14	2	1	2	-	1	-
Vehicle operator expenses (VOX)	7	-	-	-	-	-	-
Vehicle revenue hours (VRH)	13	2	2	2	-	1	-
Vehicle revenue miles (VRM)	14	2	1	1	-	1	-
Employees' salaries and wages (WGX)	5	-	-	2	-	-	-

Note: Dash indicates no data.

that values representing the same variable in different tables were not the same. For example, total operating expense by function in Table 002.08.1 differed from the corresponding value by object class in Table 002.09.1, shown as follows (partial listing):

System Identification	Table 002.08.1	Table 002.09.1
4001	2,076,678	4,144,829
5029	1,028,048	720,393
8003	8,083	459,674
5030	1,071,857	1,192,799

These conflicts were solved by checking the same data in the tape.

Data in Tape and Annual Report Do Not Match

To illustrate the problem, total revenue variables with conflicting values in Table 002.01.1 and File 10 are listed in Table 2. In most cases, the data in the report are more reasonable. This is because there are fewer items in the report, which came out a few months after the tape. During this period, errors might have been found and corrected.

TABLE 2 Unmatched Data Between Tape and Report

System Identification	File 10 of Tape	Table 002.01.1 of Report (000s)
1047	536,107	351.2
2029	2,305,654	2,276.5
2044	6,280,910	6,503.9
3013	3,730,439	3,575.2
4025	2,513,384	2,566.3
6015	752,630	785.1
6005	4,138,832	4,530.3
7012	1,282,323	1,276.8

Vehicle Hours by Time of Day Excessive in Tape

This error was found when a need to calculate the values of new basic transit variables occurred (17). The error is in File 30. As an illustration, the average vehicle hours per vehicle during the nighttime period for system 1055 is $(VH)/V = 300/4 = 75$ hr, which is greater than the 24 hr allowed in one day.

After all errors had been eliminated, the database was established for the development of the proposed performance evaluation methodology as discussed in the following sections.

DEVELOPMENT OF A PERFORMANCE EVALUATION METHODOLOGY

To evaluate the performance of transit systems, measurements regarding their effectiveness and efficiency are usually employed (18,19). Effectiveness measures may cover a wide range of financial indicators of the transit systems as well as those related to the social well-being, economic development, and environmental quality of the community (3,20). The financial indicators, however, are greatly influenced by policy decisions, transit demand, management practices, and the local operating environment. Thus, for practical purposes, a performance evaluation methodology should provide information regarding the financial and operational characteristics of the system and enable a peer comparison for policy making. The proposed methodology presented here is an attempt to fulfill this purpose by using UMTA Section 15 data (with the adjustments discussed in the previous section).

Approach and Its Advantages

The cornerstones of the proposed methodology are the indices related to expenses, the number of unlinked passenger trips served, the passenger revenues, and the total transportation revenues (including fare revenues and such items as advertisement charges). The idea behind this choice of indicators is the opportunity to evaluate the cost effectiveness of transit systems from the points of view of management (expenses), service (passenger trips), and income (fare revenues or total transportation revenues).

All indices except the one for expenses were computed by using the following equation:

$$\text{Index } (i,j) = [\text{actual value } (i,j) - \text{expected value } (i,j)] \div \text{expected value } (i,j) \qquad (1)$$

where i is the parameter of interest (i.e., unlinked passenger trips, passenger revenues, and transportation revenues) for bus system j.

Equation 1 ensures that a positive index implies a superior situation (i.e., actual performance is greater than expected performance), whereas a negative index implies an inferior situation. For the case of expenditure, a superior situation would be one in which the actual expenses are less than the expected expenses. Thus, in order to denote a positive index as the superior one, the following equation must be used:

$$\text{Index (expense, } j) = [\text{expected expense}(j) - \text{actual expense}(j)] \div \text{expected expense } (j) \qquad (2)$$

Equations 1 and 2 show that an index of zero means that the transit system performs at the expected level compared with its peer. In this study, the expected performance values were estimated by using a set of multiple-linear regression equations the rationale and detailed description of which will be given later.

One of the advantages of using the indices is their simplicity and directness. By examining the sign (positive or negative) and the magnitude of the index, one can quickly gauge the performance of a transit system against its expected performance and those of its peers. This knowledge will facilitate the choice of corrective actions to improve the performance of the transit system. Another advantage is that they offer a straightforward basis for comparison, which may be illustrated as follows.

Transit performance comparisons may be grouped into three types: uncontrolled, controlled, and combined (21). In an uncontrolled comparison, an individual performance is pitted against a standard value or the average value of the peer group. The outcome of the comparison can be either better or worse than the peer group performance as shown in Figure 1A. In a controlled comparison, some form of equation, algorithm, or simulation is usually used to compute the expected performance for the whole group of systems. The expected performance is then used as a basis for comparing the performance of all individual systems. Hence, a company can perform better than, the same as, or worse than what is expected from it, as shown in Figure 1B. For the combined comparison, both controlled and uncontrolled concepts are used. There are four possible results, as shown in Figure 1C. They are

1. Better than the standard value (uncontrolled)

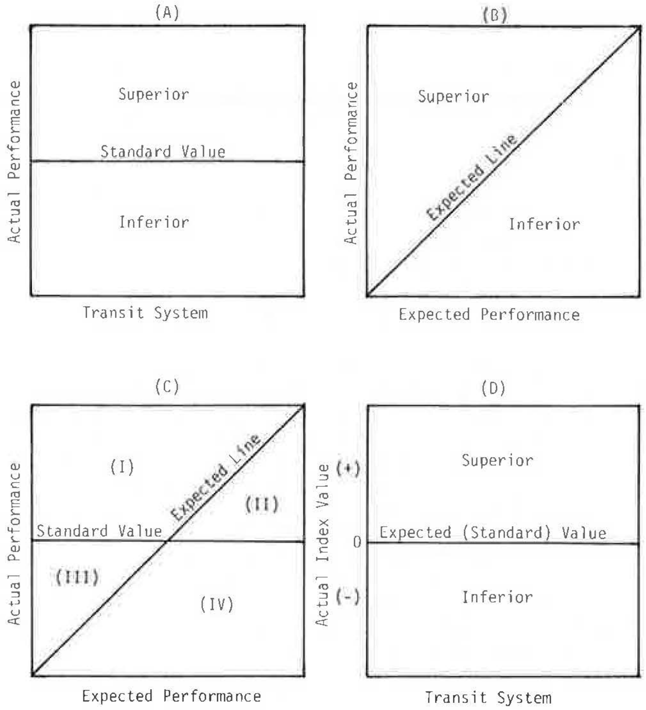


FIGURE 1 Comparison concepts in performance evaluation of transit systems.

and better than the expected value (controlled) (region I),

2. Better than the standard value but worse than the expected value (region II),
3. Worse than the standard value but better than the expected value (region III), and
4. Worse than the standard value and worse than the expected value (region IV).

It is obvious that the first type of result is superior and the fourth is inferior. For the second and third the performance of the company cannot easily be classified as good or bad.

If indices are used, ambiguous results will not occur in the comparison. The expected values obtained from the regression equations form the basis for a controlled comparison and the zero index value assumes the standard for an uncontrolled comparison. This way, there are three clear types of results: the company's performance is better than the expected and better than the peer-group performance, the company's performance is worse than both the expected and peer-group performance, and the company's performance is the same as the expected and the peer-group performance (see Figure 1D).

By having only distinctive results, the index can lend itself to the ranking of the systems according to their performance levels. The ranking will be based on the index values; i.e., a positive index will indicate a superior performance. The larger the magnitude of the positive index value, the better is the performance. Therefore the index not only shows the performance status of a company but also indicates how good or how bad its performance is.

Besides the above four individual indices, composite indices may also be developed to represent the overall performance of a system. Because the income of a transit system may come solely from the fare box or from the total transportation revenues, two composite indices may be suggested. They are the passenger-revenue-based composite index (CI_{pr}) and the transportation-revenue-based composite index (CI_{tr}):

$$CI_{pr} = W_e I_e + W_p I_p + W_f I_f$$

$$CI_{tr} = W_e I_e + W_p I_p + W_t I_t$$

where W_e, W_p, and W_f are the weighting factors for the expense, passenger, and fare-revenue indices, respectively. They are the subjective measures that represent the importance of each type of index (which may be obtained from an opinion poll, for example). I_e, I_p, and I_f are the index values for expense, passenger, and fare revenue, respectively. W_t is the weighting factor for transportation revenue and I_t is the index value for transportation revenue.

The composite indices can eliminate many inherent biases that generate from the operating policy of individual bus systems. For instance, if a system tries to attract more riders by reducing the fare, the index on passengers served would be high compared with that of the peer group. Moreover, if a system has a good revenue source from the local government (e.g., transit-dedicated tax), it generally does not stress revenue generation, but the system will always be viewed as effective if the passenger-based performance index is used. In addition, a passenger-based index will be biased against systems with a good charter service that generates a high level of revenue but not of passengers. On the other hand, a fare-revenue-based index alone may penalize the ingenuity of the transit manager to earn more income outside the traditional fare-box revenue.

In order to use Equations 1 and 2 discussed earlier, the expected value for each performance indicator must be estimated. A simple approach to estimating these values may be to use the mean value of each indicator. However, this approach would not provide a fair basis for comparison because of the nonhomogeneity of the transit systems' characteristics and environment, which are the primary determinants of the system's performance. Thus, to overcome this barrier, multiple-linear regression analysis was used in an attempt to explain the level of transit performance through causal indicators. [Nonlinear regression, however, was also employed; the results are presented elsewhere (17).] The main idea is that for each transit system with a certain number of characteristics (whether operational or environmental), there exists a reasonable expected value of performance relative to its peer. If this value is derived from the characteristics of all systems, biases may be greatly reduced.

Application

In UMTA Section 15 data, transit systems were stratified into seven groups according to the number of vehicles, as mentioned earlier. Ideally, the indices must be developed for each group to provide a complete basis for evaluating the transit industry. Unfortunately, because of the small number of observations available for the last three groups (i.e., 250 vehicles or more), regression analysis would not produce meaningful results. For the 100- to 249-vehicle systems, the propagation of missing data on basic variables prevents the formulation of meaningful regression equations. This leaves three groups that could feasibly be studied.

Two groups (25 to 49 and 50 to 99 vehicles) were merged into one because of the small differences between them in labor utilization, unit expense, and patronage. This also enlarged the database for regression analysis. The smallest group, 1 to 24 vehicles, was treated independently because of their distinct characteristics of costs, fuel efficiency, and patronage. For the purposes of illustration, only the results of the study on the systems with 25 to 99 buses will be presented here. The study on the systems with 1 to 24 buses may be found elsewhere (17).

To formulate the regression equations for estimating the expected performance of all systems, a set of potential causal variables was selected from UMTA Section 15 data. This set includes all the basic transit variables, three generic variables [vehicle hour miles (VHM), vehicle revenue hour miles (VRHM), and revenue capacity hour miles (RCHM)], and a number of other performance indicators, as follows:

1. Forty-six indicators officially listed in UMTA Section 15 reports (5,15);
2. The important indicators suggested by Sinha (22), Anderson and Fielding (13), and the Organization for Economic Cooperation and Development (OECD) (23) that do not duplicate those of Section 15; and
3. Additional indicators derived by substituting vehicle hours (VH), vehicle miles (VM), and revenue capacity miles (RCM) for the generic variables VHM, VRHM, and RCHM, respectively.

Initially 170 variables were selected for regression analysis. This list was later reduced to a manageable size by eliminating redundancy, omitting variables with low values of correlation, and so forth.

With the reduced set of data, a default level of significance of 10 percent in the System Analysis Study (SAS) stepwise procedure was used in model formulation. This value applied to all the estimated coefficients of causal variables and the overall significance of the models. The degree of explanation for the variation of the response variables (R^2) was set at a minimum of 50 percent for the purpose of screening. The explanation of the variation could be improved by discarding outlier points. This, however, was not done in this study because there was more interest in learning about the actual variation phenomenon. The results of the regression analysis led to the selection of the following models:

Operating expenses

$$\begin{aligned}\text{EXP} = {} & -1{,}361{,}530.89 + 18{,}975.71{*}\text{EMP} \\ & + 22{,}915.85{*}\text{FLEETAGE} + 243{,}062.60{*}\text{MNFAC} \\ & + 0.43022{*}\text{VMCM} - 479{,}199.63{*}\text{DRPT1PV} \\ & + 232{,}489.19{*}\text{DRWG1DRH} - 111{,}736.85{*}\text{VMC1FUEL} \\ & R^2 = 0.9375,\ n = 92 \qquad (3)\end{aligned}$$

where

EXP = total operating expenses,
EMP = total number of employees,
FLEETAGE = average fleet age,
MNFAC = light-maintenance facilities,
VMCM = vehicle miles plus charter miles,
DRPT1PV = part-time drivers per peak-period vehicle,
DRWG1DRH = driver wages per driver hour, and
VMC1FUEL = vehicle miles plus charter miles per gallon of fuel.

This equation shows good correlation between transit expense and the causal variables selected. The use of part-time drivers and fuel-efficient vehicles greatly reduces the operating costs of the system.

Unlinked passenger trips

$$\begin{aligned}\text{PAS} = {} & 77{,}540.35 + 206.48{*}\text{POPD} + 26{,}531.77{*}\text{PV} \\ & + 1.4208{*}\text{VRM} - 2{,}473{,}544.84{*}\text{PASR} \\ & R^2 = 0.5961,\ n = 73 \qquad (4)\end{aligned}$$

where

PAS = unlinked passenger trips,
POPD = population density,
PV = peak-period vehicles,
VRM = vehicle revenue miles, and
PASR = passenger fare revenue.

This equation does not explain the correlation with the independent variables well. The demand characteristics of the urban region served are represented only by the population density, which does not completely indicate the operating environment of the system. If data related to the vehicle ownership in the region were available, the significance of the model could be greatly improved. It must also be noted that PASR is an indicator of the fare charged. This variable, however, has some degree of collinearity with the dependent variable and therefore should be replaced in future studies with some other variable that is more representative of the average fare (e.g., fare per trip mile).

Passenger fare revenue

$$\begin{aligned}\text{PASR} = {} & 757{,}631.72 + 77.506{*}\text{POPD} + 20{,}258.79{*}\text{PV} \\ & + 0.0046613{*}\text{RCM} + 30{,}622.79{*}\text{TRIPTIME} \\ & + 624{,}889.26{*}\text{PASR1PAS} - 286{,}565.57{*}\text{PV1BV} \\ & R^2 = 0.8594,\ n = 69 \qquad (5)\end{aligned}$$

where

TRIPTIME = average time per unlinked passenger trip,
PASR1PAS = passenger revenue per unlinked passenger trip, and
PV1BV = peak-period vehicles per base-period vehicle.

The variable TRIPTIME in the equation reflects the average length of the trip. The sign of its coefficient should be consistent with this interpretation. Otherwise, travel time would have a negative impact on the demand.

The negative sign of PV1BV may be interpreted as the impact of relative services provided during peak and off-peak hours. If the off-peak service is poor, the overall efficiency of the system may be perceived as low, which leads to a decline in patronage.

Transportation revenue

$$\begin{aligned}\text{TRR} = {} & -1{,}486{,}294.38 + 75.126{*}\text{POPD} + 0.72520{*}\text{VMCM} \\ & + 1{,}029{,}784.37{*}\text{PASR1PAS} + 9{,}306.11{*}\text{RCM1VRM} \\ & + 30.924{*}\text{VM1LM} \qquad R^2 = 0.7899,\ n = 73 \qquad (6)\end{aligned}$$

where

TRR = total transportation revenue,
RCM1VRM = revenue capacity miles per vehicle revenue mile, and
VM1LM = vehicle miles per line mile.

As for the overall transportation revenues, vehicle miles of travel is the determining factor. Also, charter miles are important in revenue generation as shown in Equation 6.

RESULTS

The indices for each category (i.e., expense, passengers, passenger revenue, and transportation revenue) were computed by using the data for each bus system considered. The results for most cases were reasonable enough for comparative evaluation. However, some problems were encountered during the analysis that will be discussed later.

The number of indices developed for each category depends on the number of valid data items reported. There were 92 indices for the expense category, 73 for passengers, 69 for passenger revenue, and 73 for transportation revenue. Only 66 systems could meet the requirement that all four types of indices be considered for each bus system. Thus, for illustration the results for these 66 systems are presented as shown in Table 3. Each type of index is discussed in the following.

TABLE 3 Results of the Application Analysis for Systems with 25 to 99 Buses

	EXPENSE		PASSENGER		FARE REVENUE		TRANSP REV	
RANK	ID	INDEX	ID	INDEX	ID	INDEX	ID	INDEX
1	4041	0.24	7012	1.21	6016	435.04	4005	12.22
2	1043	0.23	9020	1.03	4005	54.71	6016	7.31
3	4025	0.17	3012	0.92	5059	2.77	1002	5.32
4	5084	0.17	2058	0.75	5001	1.67	1043	1.75
5	6024	0.14	5097	0.67	7011	1.58	5045	0.77
6	1002	0.14	2034	0.54	1002	1.45-	7011	0.74
7	5097	0.14	3014	0.48	1056	1.16	2029	0.69
8	4007	0.12	5059	0.41	4023	1.02	4009	0.64
9	5001	0.12	4044	0.40	4009	0.98	1056	0.61
10	6019	0.11	6024	0.39	6001	0.82	2060	0.57
11	6033	0.10	2029	0.37	2029	0.55	2034	0.50
12	4035	0.10	3025	0.31	2025	0.54	6017	0.44
13	6021	0.10	2025	0.29	1043	0.51	1016	0.41
14	4002	0.09	2060	0.29	4044	0.42	3008	0.41
15	4043	0.09	1043	0.26	6021	0.39	4036	0.41
16	4012	0.09	6016	0.24	2034	0.29	2025	0.38
17	3008	0.08	6021	0.19	1016	0.28	5097	0.35
18	4009	0.07	1016	0.16	6024	0.27	3002	0.34
19	5011	0.07	9018	0.15	4036	0.23	2044	0.34
20	3001	0.07	4025	0.14	4038	0.22	4017	0.34
21	5044	0.07	3008	0.13	6017	0.20	4007	0.31
22	5006	0.06	4005	0.13	2044	0.20	5001	0.24
23	4005	0.05	1056	0.12	4007	0.19	6024	0.23
24	5058	0.04	6019	0.09	3025	0.18	4044	0.20
25	4044	0.03	4007	0.09	3014	0.17	4012	0.17
26	1056	0.01	2003	0.09	9020	0.15	1005	0.17
27	2029	-0.00	9035	0.09	4017	0.15	3012	0.16
28	2067	-0.00	4043	0.06	9	0.13	3025	0.14
29	5059	-0.01	4023	0.06	3001	0.11	4025	0.14
30	1016	-0.01	5058	0.06	3008	0.10	5059	0.10
31	5003	-0.01	7011	0.04	4043	0.08	4002	0.10
32	9020	-0.01	1002	0.04	2060	0.07	3014	0.09
33	9	-0.02	9	0.01	1005	0.05	2003	0.08
34	9006	-0.02	5052	-0.01	9018	0.05	4038	0.04
35	12	-0.03	4036	-0.03	4035	0.04	4023	0.03
36	5045	-0.03	4017	-0.04	2058	0.01	5044	0.01
37	5052	-0.03	9006	-0.04	3011	-0.02	2058	0.00
38	3025	-0.03	4035	-0.06	4025	-0.02	9018	-0.05
39	3002	-0.04	4002	-0.07	2003	-0.03	4043	-0.06
40	4017	-0.04	5045	-0.08	6033	-0.05	4035	-0.06
41	4023	-0.04	6033	-0.09	5045	-0.05	6033	-0.07
42	6017	-0.05	3007	-0.10	5084	-0.06	9020	-0.08
43	3014	-0.06	5044	-0.13	5044	-0.12	5058	-0.10
44	5060	-0.06	5001	-0.13	3007	-0.12	3007	-0.11
45	2034	-0.06	2044	-0.17	5056	-0.12	9	-0.15
46	4036	-0.07	1005	-0.19	4012	-0.12	3001	-0.17
47	5056	-0.07	4009	-0.19	9035	-0.12	6021	-0.17
48	7012	-0.07	4012	-0.20	4002	-0.13	9035	-0.21
49	3013	-0.08	6017	-0.20	5003	-0.14	4041	-0.21
50	4038	-0.08	5060	-0.21	3012	-0.14	5056	-0.21
51	3007	-0.09	4038	-0.23	5058	-0.20	6019	-0.25
52	2060	-0.10	3001	-0.24	5052	-0.21	5052	-0.30
53	9018	-0.10	5057	-0.24	4041	-0.26	3011	-0.35
54	7011	-0.10	5056	-0.24	5060	-0.26	6010	-0.35
55	1005	-0.12	4041	-0.25	6019	-0.28	12	-0.37
56	2058	-0.13	12	-0.27	3013	-0.28	3013	-0.38
57	6010	-0.13	3002	-0.28	7012	-0.29	6001	-0.39
58	2025	-0.18	5006	-0.36	5006	-0.30	5003	-0.41
59	5057	-0.18	5084	-0.46	9006	-0.31	2067	-0.43
60	2044	-0.19	2067	-0.50	12	-0.33	5006	-0.45
61	2003	-0.27	5011	-0.55	5097	-0.38	5057	-0.47
62	9035	-0.29	3013	-0.56	2067	-0.41	5060	-0.49
63	6016	-0.30	6001	-0.57	6010	-0.42	5084	-0.52
64	3011	-0.37	6010	-0.65	5057	-0.51	9006	-0.54
65	3012	-0.54	3011	-0.70	3002	-0.61	7012	-0.57
66	6001	-1.38	5003	-0.70	5011	-0.62	5011	-0.65

Expense Index

Of the 92 expense indices obtained, 39 (or 42 percent) are positive and 53 (58 percent) are negative. The highest (positive) value is +0.24 and the lowest (negative) value is -1.38. This shows that the positive extreme is not far from the expected performance, whereas the negative extreme is much lower than the expected performance. A closer look at the magnitude of the indices reveals that 88 percent of all systems have an index value greater than -0.20 (which means that their actual expenditure is not greater than 120 percent of their expected expenditure).

To gain more insight into the reasonableness of the index values, an analysis was performed for a number of systems. The objective of the analysis was to compare the performance indicators of the selected systems to see whether they agreed with the performance implied by the indices. The indicators used were the ratios of expenses per driver hour, expenses per employee, expenses per passenger, expenses per passenger mile, expenses per peak-hour vehicle, expenses per vehicle hour plus charter hour, and expenses per vehicle mile plus charter mile. This analysis showed that the majority of the indicators agreed with the value of the index. That is, if a bus system has a higher index than that of another system, its ratios of expenses are likely to be smaller than those of the inferior system (see Table 4).

The analysis also reveals that bus systems with a high number of part-time drivers per peak vehicle are likely to have lower index values. This is because of the negative sign of the partial coefficient of this variable in Equation 3. The problem may be corrected and the regression equation would give more reasonable expected expenses if systems with part-time drivers were treated separately.

Passenger Index

The results show that approximately 45 percent of the bus systems studied have a positive passenger index (33 out of 73 systems). The range of the index is between -0.70 and +1.21 as shown in Table 3, and the majority of the systems (73 percent) have an index value greater than -0.20. The high value of the positive index in this case may be because some transit systems provide extensive services in return for financial assistance from other sources besides the fare box.

The analysis of the passenger indicators for individual systems shows a reasonable degree of consistency between the indicators and the index value. The indicators used were passenger miles per vehicle revenue mile, passengers per peak vehicle, passengers per line mile, passengers per vehicle revenue mile, passengers per service land area, passengers per employee, passengers per person in the population served, passengers per revenue capacity mile, passengers per vehicle hour, and passengers per vehicle mile. In comparing the values of the indicators for the selected systems, it was found that the population density of the region served (ranging from 500 to more than 19,000) has a great effect on the value of the index. For regions with lower population density, the expected number of passengers carried is quite low, causing an inflated value in the index. To avoid this problem, two approaches may be taken. First, as mentioned before, the population density does not accurately represent the demand for transit services and therefore should be replaced by some indicators of automobile ownership or the

TABLE 4 Analysis of Expense Indicators

	System Identification					
Indicator	4041	4025	3001	1056	5060	3011
Rank	1	3	20	26	44	64
Expense index	0.24	0.17	0.07	0.01	-0.06	-0.37
EXP/DRH	14.10	14.70	17.7	18.2	19.3	17.1
EXP/EMP	19,714	20,377	23,959	26,992	27,259	25,237
EXP/PASS	0.70	0.57	1.08	0.93	1.00	2.27
EXP/PASM	0.07	0.21	0.24	0.30	0.29	0.74
EXP/PV	45,806	62,622	73,138	72,445	89,474	50,473
EXP/RCM	0.02	0.03	0.03	0.04	0.04	0.05
EXP/VHCH	15.13	14.49	10.9	20.94	23.93	22.35
EXP/VMCM	1.04	1.36	1.42	1.94	1.89	2.18

availability of alternative modes. For the second approach, transit systems could be grouped according to the population density before regression equations are developed.

Passenger Revenue

The indices for this category are shown in Table 3. There are 36 positive indices out of a total of 69. The two highest-ranking indices with values of 435 and 54.7 are completely out of proportion. A check was made on their data and revenue indicators that revealed that the indices were inflated by the regression equation. These are the two points that are not well represented. The third-ranked system with an index of 2.8 was also misrepresented because of its reported average trip time of 1.1 min, which is not realistic. Thus, excluding these three systems, the reasonable range of the passenger revenue index is between -0.62 and +1.67.

An analysis of the individual indicators was also performed by using the ratios of passenger revenue per driver, passenger revenue per dollar of expenditure, passenger revenue per line mile, passenger revenue per passenger mile, passenger revenue per peak vehicle, passenger revenue per revenue capacity mile, passenger revenue per vehicle hour, and passenger revenue per vehicle mile. The indicators show a reasonable degree of consistency with the value of the index.

For bus systems with a value of passenger revenue per passenger greater than $1.00, the expected revenue estimated by the regression equation is high, which gives a low index value. However, their passenger-revenue-related indicators show that their performance is much superior to that implied by the index. The reason for the inaccurate estimates is that there are only three bus systems out of the 69 systems used to formulate the regression equation that have a value of passenger revenue per passenger greater than $1.00.

Because passenger revenue also depends on the demand for transit services, the same comments made on the use of population density hold here. In addition, if the fare rate per mile per passenger had been used, the estimates of the expected passenger revenues would have been more accurate.

Transportation Revenue

As shown in Table 3, transportation revenue indices vary from -0.65 to +12.2. The number of systems with a positive index is 40 of a total of 73 systems. The index values of the top four systems are significantly different from the other systems because of the biases created from the population density variable used in Equation 6. The problems encountered in the analysis of these indices are the same as those in the passenger-revenue category; that is, the regression equation causes the index values of those systems with a low population density to be inflated and those with a high value of passenger revenue per passenger to be deflated. Besides these special cases in which the index values are biased, other indices agree quite well with the transportation-revenue-related indicators of the individual systems.

CONCLUSION

Because of the uniform reporting of common transit data under the UMTA Section 15 reporting system, it is widely expected that these data sets will correct many previous problems on data adequacy for research and planning purposes. However, the inauguration-year data have so many errors that this study had to switch to the second-year data. Still, many errors remain, leading to the discussion of the shortcomings and the solutions to overcome them in this paper.

In order to improve the usefulness of the data, some modifications to the reporting forms should be made to give more specific information and to avoid ambiguity. In addition, the screening process should be improved to avoid missing data. From the point of view of data application, information related to the demand for transit services appears to be lacking. It was found in this study that although many transit systems have similar supply characteristics, their performance measures can be distinctively different because of the different demand environment. The data related to the population served, the service land area, and automobile ownership seem to be useful in representing the demand for transit services.

With respect to the performance indices, the main aim was to try to relate the performance of a system to its operating attributes and environment. Even with a limited set of data, the indices developed are indicative of the system's performance. They can be used to screen the weaknesses of a system so that improvements can be made. They are also useful in assessing the performance of the whole industry in any of the four categories considered; for example, if the majority of the systems do not perform at their expected level, corrective action needs to be sought in the attributes that the index represents.

The use of the indices can overcome the many differences in assessing transit performance. It is hoped that with the future improvements of the data, a more complete set of performance indices may be developed.

ACKNOWLEDGMENT

The authors wish to extend their special thanks to the following personnel of the Virginia Department of Highways and Transportation: David Berg for his support and interest in this study and Brandon Ford for her valuable comments and discussion. The authors are also indebted to many state and federal agencies for providing the much-needed and valuable data for this paper.

REFERENCES

1. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System, Volume 1: General Description. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
2. G. J. Fielding and R. E. Glauthier. Obstacles to Comparative Evaluation of Transit Performance. Report UCI-ITS-SP-77-1. Institute of Transportation Studies, University of California at Irvine, April 1977.
3. J. S. Dajani and G. Gilbert. Measuring the Performance of Transit Systems. Transportation Planning and Technology, Vol. 4, No. 2, 1978, pp. 97-103.
4. Revised Policy Statement on Transit Performance. Passenger Transport, Feb. 17, 1979.
5. National Urban Mass Transportation Statistics--First Annual Report: Section 15 Reporting System Transit Financial and Operating Data Reported for Fiscal Years Ending Between July 1, 1978 and June 30, 1979. Report UMTA-MA-06-0107-81-1. Office of Transportation Management, UMTA, U.S. Department of Transportation, Nov. 1981.
6. Transit Fact Book, 1978-1979. American Public Transit Association, Washington, D.C., Dec. 1979.
7. Public Transportation in Virginia: Service, Operations, Costs, Revenues During Fiscal Year 1980. Public Transportation Division, Virginia Department of Highways and Transportation, Richmond, Oct. 1981.
8. Urban Mass Transportation Industry Uniform System of Accounts and Records. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
9. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System: Required Reporting Manual and Sample Forms. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
10. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System: Level A Reporting Manual and Sample Forms. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
11. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System: Level B Reporting Manual and Sample Forms. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
12. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System: Level C Reporting Manual and Sample Forms. Report UMTA-IT-06-0094-77-1. UMTA, U.S. Department of Transportation, Jan. 1977.
13. S. C. Anderson and G. J. Fielding. Comparative Analysis of Transit Performance. Report UMTA-CA-11-0020-82-1. Office of Policy Research, UMTA, U.S. Department of Transportation, Jan. 1982. NTIS: PB 82-196478.
14. J. M. Holec, Jr., D. S. Schwager, and M. J. Gallagher. Improving Usefulness of Section 15 Data for Public Transit. In Transportation Research Record 835, TRB, National Research Council, Washington, D.C., 1981, pp. 9-15.
15. National Urban Mass Transportation Statistics--Second Annual Report: Section 15 Reporting System Transit Financial and Operating Data Reported for Fiscal Years Ending Between July 1, 1979 and June 30, 1980. Report UMTA-MA-06-0107-82-1. Office of Technical Assistance, UMTA, U.S. Department of Transportation, June 1982.
16. County and City Data Book, 1977. U.S. Bureau of the Census, 1978.
17. C. Kanok-kantapong. Comparative Transit Performance Evaluation: Cost, Demand and Revenue Index Approach with UMTA Section 15 Data. Ph.D. dissertation. Virginia Polytechnic Institute and State University, Blacksburg, Nov. 1983.
18. J. H. Miller. The Use of Performance-Based Methodologies for the Allocation of Transit Operating Funds. Traffic Quarterly, Vol. 34, No. 4, Oct. 1980, pp. 555-574.
19. R. G. Knighton and N. S. Erlbaum. Factors Influencing Transit Productivity. Preliminary Research Report 121. Planning Research Unit, New York Department of Transportation, Albany, Aug. 1977.
20. J. A. Caywood, J. W. Fuller, J. M. Colucci, and M. L. Downey. Factors Affecting Transportation Productivity. Traffic Quarterly, Vol. 34, No. 2, April 1980, pp. 159-196.
21. S. R. Mundle and W. Cherwony. Diagnostic Tools in Transit Management. In Transportation Research Record 746, TRB, National Research Council, Washington, D.C., 1980, pp. 13-19.
22. K. C. Sinha and D. P. Jukins. A Comprehensive Analysis of Urban Bus Transit Efficiency and Productivity: Part I. Definition and Measurement of Urban Transit Performance. Report UMTA-IN-11-0003-79-2. UMTA, U.S. Department of Transportation, Dec. 1978. NTIS: PB 295221.
23. Urban Public Transit: Evaluation of Performance. Organization for Economic Cooperation and Development, Paris, France, Oct. 1980.