TRANSPORTATION RESEARCH RECORD 1013

Improving Utilization of Transit Resources

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1985

Transportation Research Record 1013
Price $11.40
Editor: Elizabeth W. Kaplan
Compositor: Harlow A. Bickford
Layout: Harlow A. Bickford

mode
2 public transit

subject areas
12 planning
13 forecasting
54 operations and traffic control
55 traffic flow, capacity, and measurements

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, N.W., Washington, D.C. 20418.

Printed in the United States of America

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Improving utilization of transit resources.

(Transportation research record; 1013)
1. Local transit–United States–Cost of operation–Congresses.
2. Local transit–United States–Labor productivity–Congresses.
I. National Research Council (U.S.). Transportation Research Board. II. Series.

TE7.H5	no. 1013	380.5 s	85-18791
[HE4441]		[388.4'042]	
ISBN 0-309-03902-9		ISSN 0361-1981	

Sponsorship of Transportation Research Record 1013

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
William A. Bulley, H. W. Lochner, Inc., chairman

Urban Public Transportation Section
John J. Fruin, PED Associates, chairman

Committee on Transit Management and Performance
Nigel H. M. Wilson, Massachusetts Institute of Technology, chairman
Mark Abkowitz, Colin H. Alter, John P. Attanucci, A. Jeff Becker, Avishai Ceder, John Dockendorf, John T. Doolittle, Jr., Gordon J. Fielding, Ronald J. Fisher, Ronald I. Hollis, Joseph J. Kern, William R. Loudon, Joel E. Markowitz, Brian E. McCollom, James H. Miller, Joseph Misner, Subhash R. Mundle, Richard L. Oram, Marian T. Ott, Eve M. Wyatt, Robert J. Zerrillo

W. Campbell Graeub, Transportation Research Board staff

The organizational units, officers, and members are as of December 31, 1984.

Contents

Addresses of Authors

Attanucci, John P., Multisystems, Inc., 1050 Massachusetts Avenue, Cambridge Mass. 02138

Bladikas, Athanassios K., Polytechnic Institute of New York, 333 Jay Street, Brooklyn, N.Y. 11201

Carter, Douglas W., Price Waterhouse, 555 California Street, San Francisco, Calif. 94104

Chomitz, Kenneth, Institute of Transportation Studies and Department of Economics, University of California, Irvine, Calif. 92717

Dooley, Thomas, Transportation Systems Center, DTS-62, Kendall Square, Cambridge, Mass. 02142

Englisher, Larry S., Multisystems, Inc., 1050 Massachusetts Avenue, Cambridge, Mass. 02138

Giuliano, Genevieve, Institute of Transportation Studies and Department of Economics, University of California, Irvine, Calif. 92717

Lall, Upmanu, Department of Civil Engineering, University of Utah, Salt Lake City, Utah 84112

Lave, Charles, Institute of Transportation Studies and Department of Economics, University of California, Irvine, Calif. 92717

Markowitz, Joel E., Metropolitan Transportation Commission, Metrocenter, 101 8th Street, Oakland, Calif. 94607

McCollom, Brian, Urban Mass Transportation Administration, URT-41, 400 7th Street, S.W., Washington, D.C. 20590

Morgenbesser, Martin J., Consultant, 3204 Ramona Street, Palo Alto, Calif. 94306

Papadimitriou, Charles, Polytechnic Institute of New York, 333 Jay Street, Brooklyn, N.Y. 11201

Purdy, Jeffrey E., Transportation Consulting Division, Booz-Allen & Hamilton, Inc., 400 Market Street, Philadelphia, Pa. 19106

Smith, Robert L., Jr., Department of Civil and Environmental Engineering, University of Wisconsin, Madison, Wisc. 53706

Steinmetz, William R., Transportation Consulting Division, Booz-Allen & Hamilton, Inc., 30 Charles II Street, St. James' Square, London SWY1Y 4AE, England

Stephanedes, Yorgos J., Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis, Minn. 55455

Wells, Martin J., Gorove/Slade Associates, Inc., 1140 Connecticut Avenue, N.W., Washington, D.C. 20036

Yu, Jason C., Department of Civil Engineering, University of Utah, Salt Lake City, Utah 84112

SOLON: Interactive Method for Evaluating and Improving Transit Route Performance

YORGOS J. STEPHANEDES

ABSTRACT

SOLON is an interactive microcomputer graphic method that addresses the problems of transit scheduling and route efficiency and productivity improvement. It has three unique characteristics: First, it provides solutions at any specified time or continuously through the life of a transport service by tracing the interactions among scheduling changes, route ridership, and service cost-effectiveness. Second, it explicitly treats time delays that hamper route performance. Third, its modules operate interactively or independently and can be modified by the user. Because it addresses transit problems at the route level, SOLON can facilitate management and funding decision making, by determining the viability of individual routes, and aid in selecting the best performance policies. Further, because it assumes no previous computer knowledge, it can substantially increase personnel training productivity. SOLON inputs include conventional socioeconomic and transportation data. Output performance indicators are provided by mode and for every time instant, subject to selected policies and exogenous fluctuations such as changes in energy cost, inflation, and unemployment.

Declining subsidies, low ridership, and rising deficits are increasingly affecting management decisions in public transit operations. Although revenues from the gas tax bill may alleviate some transit fiscal problems temporarily, structural inefficiencies remain and act to discount the added benefits. Curtailing services, a simple short-term solution, may spell the end of urban public transit in the long term. Service changes, applied selectively to individual routes, are emerging as one of the most feasible policy alternatives for improving the cost-effectiveness of public transportation operations.

Transit managers can accomplish such improvements by selecting the best options from an extensive set of policy variables (e.g., fare, service frequency, number of stops, and route length). However, traditional planning methods cannot deal effectively with this complex problem, especially because it is subject to continuously changing constraints and often requires different solutions at different stages of development. To be sure, in real-world applications, strict optimality with respect to performance measures is seldom required. On the other hand, there is a need for decision-making procedures that are systematic and that reflect trade-offs and simplifications in the policy selection process. To be effective, such methods should be interactive and time sensitive and should be able to quickly sort through a great number of policy alternatives and advise the decision maker of the most desirable choice or choices given the stated evaluation criteria. Such are the capabilities of SOLON (selection of policy options interactively), the interactive method presented here.

Although several quick-response planning methods have been developed, SOLON is the only interactive system that begins to address the time interactions between the major elements that give rise to changes in transit route performance. The method is rather straightforward and is distinguished by three unique characteristics: First, it provides solutions at any specified time or continuously through the life of a transport service by tracing the interactions between scheduling changes, route ridership, and service cost-effectiveness. Second, it explicitly treats time delays (e.g., capital procurement and ridership) that hamper route performance. Third, its modules (demand, supply, and performance) operate interactively or independently and can be modified by the user. For instance, trip purpose (work or shopping) and market segments (by automobile ownership, transit availability, and so forth) may be selected; demand specifications can be updated; and management plans on fare, frequency, and other policies may be proposed by the user or adopted from those suggested. SOLON addresses these problems at the route level, where management decisions are usually made. It, therefore, can facilitate management and funding decision making by assessing the viability of individual routes (or route sets) and aiding in the selection of the best performance policies.

Route performance evaluation and policy selection could, of course, be performed without use of an interactive graphics system. A dynamic methodology was initially developed by this author for evaluating transport systems without the benefit of interactive graphics (1-3). However, inclusion of the interactive and graphic capabilities creates substantial savings in the time required to select a better set of policies. In addition, SOLON has been implemented in a way that allows easy access by decision makers with little or no computer experience. As a result, it enables experienced policy analysts to examine expected performance improvements in greater depth by experimenting with a wide range of plans and to fine tune selected policies before implementation. Further, it can substantially improve personnel training productivity as suggested by Twin Cities transit specialists who had hands-on experience with one version of this method. Recommendations for improvements, made by these specialists, have been incorporated in the current version, which is being continually updated with data from Metropolitan Transit Commission (MTC) routes in the Twin Cities.

OVERALL MODEL STRUCTURE

The structure of SOLON can be analyzed at several levels of detail. At the most general level, it may be pictured as a simple demand-supply model (Figure 1) with the transit route performance sector acting as a link between supply and demand. For instance, a transit frequency increase in the service supply sector results in waiting time reduction in the performance sector; as level of service improves, so does travel demand. In turn, as demand for and use of transit grow along a route, cost-effectiveness measures (e.g., load factor, operating ratio) improve and call for service adjustments. Following such adjustments (e.g., frequency, fare) in the supply sector, performance measures are further modified and the interactions continue full circle.

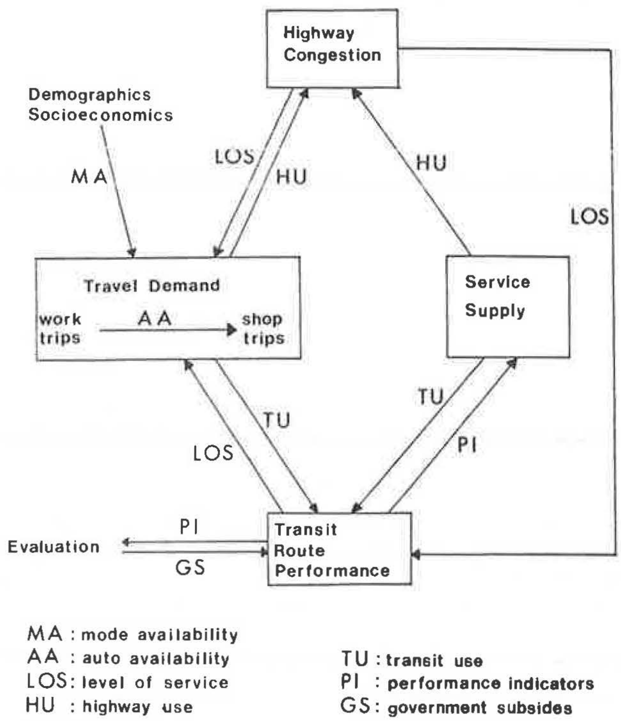


FIGURE 1 SOLON structure.

However, service as well as pricing adjustments are limited by government regulations. For example, it may be impossible to increase transit fares in order to clear the market at given levels of transit service supply. Further, many of the changes in transit route service (e.g., route length and frequency modifications), supply resources (e.g., equipment and funding acquisition), and travel patterns can be accomplished only over relatively long periods of time or at infrequent time intervals (1-3). Such substantial physical and information delays to transit supply and demand changes, together with regulatory restrictions, imply that a realistic model of transport supply-demand interactions should be able to treat the time dynamics of response to policy changes. SOLON achieves this by tracing the interactions shown in Figure 1 continuously through time.

As demand and supply continue to interact, each pass through the performance sector determines a new value for the set of efficiency, cost-effectiveness, and productivity indicators of that sector. The tables and plots of these indicators through the study period are the major products of SOLON and will be discussed in a later section. These performance records can facilitate the route evaluation process by assisting managers in assessing candidate operating policies and by advising funding agencies in making subsidy decisions on the basis of performance.

MODEL COMPONENTS

Tracing the demand-supply interactions through time is done on the basis of component equations housed within each of the four major SOLON modules shown in Figure 1. A summary of the contents of the modules (i.e., logic, empirical base, and data sources) is presented in this section. Although this exposition is necessarily constrained by space, it should nevertheless expose the strengths and limitations of the methodology. The summary should further aid the reader to identify the assumptions and, therefore, the applicability of SOLON to specific decision-making and policy evaluation situations.

Travel Demand Sector

Estimating the expected demand for a planned transit service is essential to system managers before implementation of specific service policies along a route. In SOLON the demand model should be capable of estimating ridership as a function of time and level-of-service measures based on initial route conditions only (i.e., without the need for extensive time-series data for the entire policy evaluation period). In addition, it should take into account the information delay between the time a service policy is implemented and the time residents of the service area become aware of the new service. However, most existing demand models are static (i.e., they estimate demand when the transport system is at equilibrium). Further, the static approach overlooks delayed interactions between the demand and the other transportation sectors and cannot treat transient policy impacts. Such impacts are of particular importance to decision makers during periods in which conditions exogenous to the transportation system, having created a trend, suddenly undergo major changes and force the system away from equilibrium.

Because of the limitations of the existing models, a dynamic demand equation was developed that fulfills the requirements of a time-sensitive technique such as SOLON. Assuming, for the purposes of this discussion, that the trip generation and trip distribution stages are completed, the dynamic demand equation states that P_{t+1}, the estimated probability of selecting a given mode of travel, is a function of the current probability (P_t), the estimated equilibrium probability (P_e), and an information time constant (T_p):

$$P_{t+1} = P_t + (P_e - P_t)/T_p \quad (1)$$

The information time constant can be determined experimentally and is generally longer in rural areas (2). The equilibrium probability is estimated by a disaggregate logit specification calibrated with data from the area of application. When SOLON is implemented, the suggested parameter values, which have been validated at several rural and urban routes (2-4), can be updated with a small data sample (5). Alternatively, any steady-state demand model, such as ULOGIT (6), or a simple regression equation could replace the existing specification. It should be noted that Equation 1 is based on the assumption that transportation demand, unless otherwise driven, behaves as a first-order system. Although this assumption is based on data from several systems (4), transportation operations may exist for which higher order effects are substantial.

All or some elements of the demand sector can be activated by the user, depending on the nature of the application. For instance, a modal split specification for work trips should be employed for rush-hour applications. Other equations available in this sector include a joint generation/distribution/modal-split shopping specification based on concepts presented in Stephanedes (2) and Adler and Ben-Akiva (7) and specifications developed particularly for area residents without a car, with one car, with two or more cars, and with or without transit access. Although the choice of specification or specifications is made by the user, each equation takes into account the results from other sector elements as appropriate; for example, the shopping trips equation is a function of automobile availability, which is estimated by the work trip equation.

The data needs of the demand specifications include conventional demographic, socioeconomic, and trip information traditionally required by logit models (e.g., household income, automobiles per household, workers per household, trip distance, travel time, waiting time, and cost). Most of the data are entered when the method is initialized. At that time, the user has the option of entering additional data or elementary formulas generating such data for the simulation period; however, such information is not necessary. For example, forecasts of the price of gasoline and inflation and unemployment rates can be entered for each month of the study period if these are expected to change; further, formulas that forecast such changes for the evaluation period could be directly incorporated in SOLON. The second kind of data needed by the demand sector is created internally in the highway and transit route performance sectors at every time instant of the simulation.

Transit Route Performance Sector

Efficiency, effectiveness, productivity, and quality indicators describing the route performance are needed by the supply and demand sectors for estimating service and ridership decisions, respectively, during simulation. The time records of these indicators for the study period are also needed for policy evaluation and funding decisions as mentioned earlier. Although there exist a large number of indicators that could be calculated with information from the demand and supply sectors, only a set of reasonable size is initially available to the user. The indicators included in this set have been selected on the basis of earlier work by Fielding and others (2,8-11) who sought to identify the measures that can best be employed to evaluate transit performance. To be sure, the user can easily create additional indicators using the components generated by SOLON (for a set selected by a typical urban user in the Twin Cities, see Table 1). Similarly, the user can work with a smaller set that is generated automatically at the end of an evaluation and never turn to the SOLON library for additional information; such a small set could, for instance, be adequate for performing a preliminary evaluation of a rural transit route.

Performance indicators are determined on the basis of identities, assumptions, and simple algebraic equations calibrated in the area of application. For instance, headway in minutes is estimated from frequency in vehicles per hour on the basis of the identity:

HEAD = 60/FREQ (2)

As another example, bus highway time along a route is determined by adding the stop time per transit trip (STIME) to the automobile highway time. In turn, STIME is a function of DAT, deceleration and acceleration time; STOP, the average number of stops experienced by a typical rider; DIST, the average trip length of a typical rider; STPR, stop time per rider getting on or off; and RPBM, rider per bus mile:

STIME = (DAT) (STOP) + (DIST) (2•STPR) (RPBM) (3)

The values for the parameters of Equation 3 are recommended by SOLON from previous urban and rural measurements and are supplemented with information entered by the user. "Riders per bus mile" is determined internally at every time instant of the evalu-

TABLE 1 Selected SOLON Library Performance Indicators

Condition To Be Measured	Indicator	Type
Absolute Measurements		
Demand		
Absolute ridership	Ridership	Output
Absolute demand	Demand per time period	Impact
Survivability of system	Years in operation	Effectiveness
Quality of services		
Frequency	Average headway and on-time arrival	Quality
Travel time	Average speed	Quality
Supply of services (quantity of service)	Vehicle-miles and route-miles	Output
Comparative Measurements		
Demand (relative)	Passenger trip modal split	Effectiveness
	Passenger trips/capita/time period	Effectiveness
	Passenger trips/capita/time period after service improvement compared to before	Impact
Survivability of system	Operating ratio	Cost-effectiveness
Quality of services		
Level of service	Trip time compared to automobile	Quality
Cost of service	Fare compared to automobile out-of-pocket cost	Quality
	Fare/mile	Quality
Comfort	Seat-miles/passenger-mile	Quality
Supply of services	Frequency compared to desired frequency	Output
Vehicle utilization	Vehicle-miles/vehicle	Efficiency
	Vehicle-hours/vehicle	Efficiency
	Vehicle-hours/vehicle/service-hour	Efficiency
Cost per produced unit	Cost/vehicle-mile	Efficiency
	Cost/vehicle-hour	Efficiency
Utilization of service	Passengers/vehicle-hour	Cost-effectiveness
	Passengers/vehicle-mile	Cost-effectiveness
	Passengers/seat-mile	Cost-effectiveness
	Passenger-miles/vehicle-mile	Cost-effectiveness
Cost per consumed output unit	Cost/passenger-trip	Cost-effectiveness
	Cost/passenger-mile	Cost-effictiveness
Subsidy	Subsidy	
	Per passenger trip	Cost-effectiveness
	Per vehicle-mile	Cost-effectiveness
	Per vehicle-hour	Cost-effectiveness
Diversion from other modes	Percentage of trips diverted from other modes	Impact

ation period from values generated by the demand and supply sectors.

As a third example, the operating ratio is defined as fare revenue, determined from ridership estimates in the demand sector, divided by operating cost. Previous experience with SOLON and its predecessor, TRANSIT (12), has indicated that most systems break the operating cost down to a variable and a fixed unit cost component, b_1 and b_2, respectively:

$$OCOST = b_1 \text{ (FREQ)} + b_2 \quad (4)$$

where b_1 is in dollars per round trip times hours per week and b_2 is in dollars per week; a third variable component, based on the number of route buses, can be added in Equation 4 if necessary. Route values for the b_1 and b_2 parameters are usually available from transit operations or can be determined from routinely recorded information on gas, oil, garaging, maintenance, wages, insurance, and other cost components.

Service Supply Sector

Using information from the route performance sector regarding key performance indicators such as load factor, operating ratio, and net revenue, the service supply sector determines the transit frequency that is necessary to keep the route within certain performance guidelines. These guidelines are declared by the SOLON user during initialization and are essential to the policy evaluation and decision-making process.

The guidelines reflect management's objectives and constraints and are usually expressed in terms of performance measures. For instance, most transit managers would not make a frequency change while the load factor remained within a range they consider acceptable. Further, frequency changes cannot be implemented unless the route performance fulfills certain constraints. A set of basic guidelines that depend on the load factor, operating ratio, and available capital has been designed with data from conversations with experienced managers (2-4) and is included in SOLON. The user can then indicate the specific threshold values of the performance constraints as well as the desired frequency modification at those values. Alternatively, the user can override the proposed guidelines and enter others that are more appropriate to the application at hand.

The service sector employs the information on the current state of the route available from the performance sector to determine the estimated frequency ($FREQ_{t+1}$) as a function of current frequency ($FREQ_t$), a management time (T_f), and the frequency control coefficient U:

$$FREQ_{t+1} = FREQ_t + (U)\ (FREQ_t)/T_F \quad (5)$$

where U reflects the performance guidelines discussed previously and is usually a function of the load factor, the operating ratio, and the available capital; namely, revenue, employees, and rolling stock (see Figure 2 for an illustration of U from a suburban route in the Twin Cities). The new frequency value is then fed back to the route performance sector where it is employed in the estimation of the performance measures and level of service, which, in turn, determine the demand estimate at the next time instant.

Highway Congestion Sector

Transit travel time and demand for transit are influenced by the level of service supplied by the

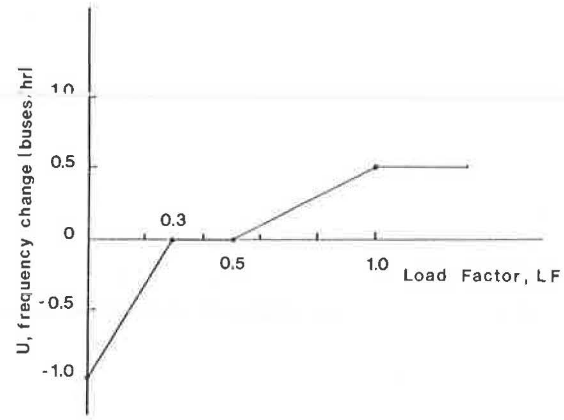


FIGURE 2 Sample SOLON policy input.

highway sector. This influence is substantially stronger in urban areas, where congestion effects are more likely. In addition, the differential effect of congestion on mode choice may be considerable where bus preferential treatment allows buses to avoid congestion areas.

In SOLON the effects of congestion are treated by estimating the highway travel time for varying highway characteristics via simple delay curves that can be updated by the user. In large urban areas, delay curves can be prepared off-line using simulation [e.g., the freeway priority entry control model FREQ6 (13)]. In simple corridor applications the same information can be prepared with limited field traffic measurements. In rural areas, where congestion is often not a problem, constant highway times can be used at different times of day and, therefore, this sector is not needed.

APPLICATION: POLICY EVALUATION AND SELECTION

The basic process for evaluating and selecting route performance policies using SOLON is now illustrated through application of the dynamic method to peak-period work trips on bus Route 35, Burnsville-Minneapolis, a radial route along a corridor south of the Minneapolis central business district. This route is characterized by local stops at the two ends of the trip and virtually no stops in the middle portion, which makes use of freeway I-35W. Census data on Burnsville and traffic information on I-35W were readily available to the project team in the course of the study.

Data initialization is the first step in applying the policy selection method. The data requirements include conventional demographic, socioeconomic, and trip information, readily available to transport planners. Most of these data are required by the logit specifications that are built in SOLON. [The statistical properties and data needs of logit are well documented (5,14-17) and not restated here.] Additional data, associated with the dynamic nature of the method, include T_p and T_F, the information and management time constants, respectively, and the initial values of ridership and frequency. A summary of data initialization, with values from the Burnsville application, is shown in Figure 3.

Performance guidelines, indicating how transit service should be modified in response to performance changes, must next be declared by the user. In this application, for instance, the Metropolitan Transit Commission (MTC) of the Twin Cities follows a policy that states that there should be no frequency change for round-trip load factors in the 0.3

```
SOLON here!

For some of the input values that follow hit RETURN
if [DEFAULT] value is desired, else enter the desired value.

U1...Round-trip transit fare, cents - - - - - - - - - - - - - - [200]
U2...Round-trip drive cost , cents  - - - - - - - - - - - - - - [225]
A1...Management implementation delay, wks - - - - - - - - - - - [3.0]
A2...Ridership information delay, wks - - - - - - - - - - - - - [3.0]

Are you happy with this input - - - - - - - - - - - - - - - - - (y/n)
n
Please enter the correct values now -

U1...Round-trip transit fare, cents - - - - - - - - - - - - - - [200]
U2...Round-trip drive cost, cents - - - - - - - - - - - - - - - [225] 425
A1...Management implementation delay, wks - - - - - - - - - - - [3.0]
A2...Ridership information delay, wks - - - - - - - - - - - - - [3.0] 4.0

Are you happy with this input - - - - - - - - - - - - - - - - - (y/n)
y

X1INIT...Initial value of ridership,  rtrips/wk - - - - - - - - [1000]
X2INIT...Initial value of frequency, bus/h- - - - - - - - - - - [1.0]
MINF.....Minimum value of frequency allowed, bus/h- - - - - - - [0]
DT1......Management conference interval, wks- - - - - - - - - - [5.0]
DT.......Value of time step for computing, wks- - - - - - - - - [0.1]
DTP......Value of time step for printing, wks - - - - - - - - - [5.0]
LEN......Time length of simulation period, wks- - - - - - - - - [60]

Are you happy with this input - - - - - - - - - - - - - - - - - (y/n)
y

You have an option to enter the input to the logit equation
or else default values will be assumed.

WPOP...Total origin work population - - - - - - - - - - - - - - [13600]
SERM...Fraction of origin served by transit - - - - - - - - - - [0.16]
PPH....Origin people per household- - - - - - - - - - - - - - - [4.12]
AALD...Autos available per licensed driver- - - - - - - - - - - [0.64]

DAYS...Days/wk system operates- - - - - - - - - - - - - - - - - [5.0]
DIST...One-way trip length, mi- - - - - - - - - - - - - - - - - [15]
NSTOP..# Stops experiencedin aver. 1-way trip - - - - - - - - - [15]
TWLKT..Transit round-trip walk time, min- - - - - - - - - - - - [12]
ADVTT..Auto out-of-veh rtrip travel time, min - - - - - - - - - [10]

Are you happy with this input - - - - - - - - - - - - - - - - - (y/n)
y
```

FIGURE 3 Sample SOLON data initialization.

to 0.5 range; below this range, frequency should decrease whereas above it (i.e., when there are standees), it should increase. To be sure, there exist an infinite number of candidate policies to deal with similar situations; see Stephanedes et al. (4) for a more detailed discussion. A visual display and hard copies of the policy are produced at this stage and can be used to facilitate communication among decision makers. The SOLON user can update this policy, if desired, at a later stage after the performance evaluation results have been reviewed. The management policy from the Burnsville application is shown in Figure 4. An illustration of this policy, which corresponds to control U of Equation 5, is shown in Figure 2.

Following initialization, SOLON begins operation by computing the initial values of the route performance indicators. Load factor, operating ratio, and net revenue are key indicators in this application because they determine the service frequency modifications. If the load factor (LF) exceeds the maximum value (LF_{max}) allowed by the size of the transit vehicles, then $LF = LF_{max}$ and SOLON signals the need for more service. If LF is below LF_{min}, the minimum value allowed by the operating ratio constraints, SOLON signals that present service conditions should be reevaluated. Fulfillment of additional performance criteria, critical to service change decisions, may also be similarly determined.

```
**** MANAGEMENT POLICY                                          ****


***  This is the transit management policy which                ***
***  determines frequency changes in terms of changes           ***
***  in the load factor, assuming maximum net                   ***
***  cost per passenger trip = $1.50.  The policy               ***
***  is summarized as a graph of                                ***
***  frequency change vs. load factor                           ***


***  You must determine four points on this graph.              ***


1.  When buses are empty ...
    Frequency should change by (buses/hr) - - - - - - [-1.0]

2.  When buses are (% full) - - - - - - - - - - - - - [30]
    Frequency should change by (buses/hr) - - - - - - [0]

3.  When buses are (% full) - - - - - - - - - - - - - [50]
    Frequency should change by (buses/hr) - - - - - - [0]

4.  When buses are (% full) - - - - - - - - - - - - - [100]
    Frequency should change by (buses/hr) - - - - - - [0.5]

Are you happy with this input - - - - - - - - - - - - (y/n)
y
A plot of your management policy is being prepared.
```

FIGURE 4 Transit route service policy input.

SOLON performance evaluation is here!

Please select up to 5 indicators from SOLON library.
?X1
?X2
?HEAD
?LF
?OR

time (weeks)	ridership (rtrips/wk)	frequency (buses/hour)	headway (minutes)	load factor	operating ratio
0	1000	1.0	60	1.0	0.30
5	1180	1.5	40	0.80	0.35
10	1410	1.8	33	0.80	0.41
15	1650	2.1	29	0.80	0.48
20	1880	2.4	25	0.80	0.54
25	2110	2.7	22	0.80	0.60
30	2250	3.0	20	0.77	0.63
35	2360	3.3	18	0.74	0.66
40	2410	3.5	17	0.70	0.66
45	2440	3.7	16	0.67	0.67
50	2470	3.9	15	0.65	0.67
55	2480	4.0	15	0.63	0.67
60	2490	4.1	15	0.63	0.67

*** press RETURN to continue ***

FIGURE 5 Sample route performance printout.

The next step is activated only at regular time intervals (DT1), the length of which is controlled by the user (Figure 3). DT1 indicates the points in time when transit management makes service change decisions. If current resources are adequate, frequency change is initiated on the basis of the state of the performance indicators determined previously. Owing to implementation time delays, the frequency change initiated in this step can only be implemented after a period of time. This is necessary as a result of regulations or because of needed adjustments in the rolling stock and the number of available drivers. In extreme cases, when the available transit vehicles are not sufficient for satisfying the needed service improvements, additional capital (R^d) is desired; in this case vehicles may be ordered or drivers may be hired, or both. In such an event, order and acquisition delays intervene and service change decisions may be implemented more than a year after the process is initiated. At the same time, depreciating vehicles are retired and create the need for additional equipment orders.

If additional capital is needed, orders (R^o) are placed according to the following plan:

$$R^o_{t+1} = R^d_t - R_t - \sum_{m=1}^{p} R^o_{t-m} \qquad (6)$$

As Equation 6 suggests, the orders are determined by desired capital (R^d) but take into account existing capital (R); that is, existing number of buses and all previous orders that have not yet arrived because of the capital arrival delay (p). Meanwhile,

$$R_{t+1} = R^o_{t-p} + R_t\ (1 - 1/L) \qquad (7)$$

An account of existing capital is kept, where R is estimated as the sum of R_t, current capital, and R^o_{t-p}, the capital ordered p weeks earlier and arriving now. Further, assuming a capital lifetime of L weeks, the appropriate amount of capital (R_t/L) is depreciated from this sum.

Following the change in transit service, the route performance measures (and highway levels of service if needed) are determined and the demand is estimated. A record of the ridership, frequency, and

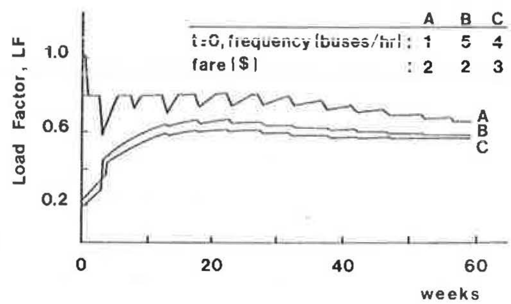


FIGURE 6 Load factor versus time for three different policies.

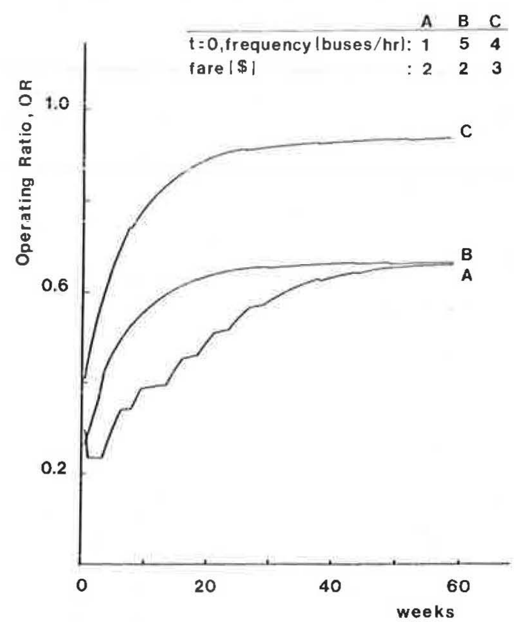


FIGURE 7 Operating ratio versus time for three different policies.

other performance indicators is kept in the form of tables and plots. These are then available to the user interactively beginning with time zero (i.e., the time SOLON is initialized) and ending with the end of the time horizon or any other instant specified by the user. In this application, the route performance is summarized in the printout sample (Figure 5) and in the plots shown by Line A in Figures 6 and 7, which illustrate the progress of load factor and operating ratio, respectively. The printout and the plots labeled "base case" represent the route performance under the policies initially proposed by the user.

Having inspected the selected printouts and plots of route performance indicators for the base policy, the user can indicate policy modifications and compare the resulting route performance against the base case. To illustrate this option, two such policy modifications have been included in Figures 6 and 7 along with the base case results. In this example, policy C succeeds in increasing the operating ratio so that it enters the desired region (i.e., becomes greater than 0.9) while the load factor remains within acceptable levels. The overplotting can continue and the user can evaluate additional policies until he is confident that one or more of the policies are superior and should be selected. A typical SOLON session in the Twin Cities application lasted 10 to 20 min before the user could make such a selection. This time length may, of course, vary depending on the complexity of the scenarios to be evaluated.

As a final step of the SOLON process, time contours are presented that indicate the time required for the route being evaluated to reach the desired performance conditions, for any combination of initial conditions. For instance, the user can determine how long it will take for a specific route to reach and remain within a set of performance constraints. This evaluation option is available for any policy selected at the end of overplotting in the previous stage. An illustration of this option with respect to the "base case" policy and policies B and C is shown in Figure 8. In that figure, the initial values of the base policy (A) are, load factor (at t = 0) = 1.0 and operating ratio (at t = 0) = 0.30 as previously computed and shown in Figures 5; 6, Line A; and 7, Line A. Similarly, the initial values of load factor and operating ratio for the other two policies were shown in Figures 6, Lines B and C, and 7, Lines B and C, respectively. As the data in Figure 8 indicate, the base policy (A) cannot guide the route to the desired conditions (i.e., to load factor $\geq$0.4 and operating ratio $\geq$0.9). Policy B, selected by the user after gaining some insights with SOLON, would also fail although it is substantially better than the initial "naive" policy. The third policy, selected after the user had some additional practice, is effective and is expected to lead the system to the desired region in approximately 25 weeks, in agreement with Figure 7, Line C.

From this application it can be seen that the number of policies that could be evaluated at once with respect to a pair of indicators, such as load factor and operating ratio, using Figure 8 covers the whole two-dimensional space and is, therefore, infinite. Thus, this figure can result in substantial time savings if the user only seeks to ascertain the degree of success (or failure) expected from a wide range of policies. Together, Figures 5-8 provide information that can substantially facilitate the policy selection process.

SOLON USES AND LIMITATIONS

SOLON has been applied to several transit routes in the Twin Cities (a validation example with Burnsville data is shown in Figure 9) and has confirmed the benefits expected to be derived from the use of systematic interactive planning methods. Further, users of the software experimental versions have made a number of helpful recommendations for improvements; incorporation of these in the current system has made it more responsive to the complex problems of the transportation practitioner.

Several potential applications of the interactive method have been suggested. For the transit route manager, being able to anticipate ridership and other performance indicators in response to selected service policies before policy implementation is a most valuable asset. This was especially appreciated in cases involving long implementation delays such

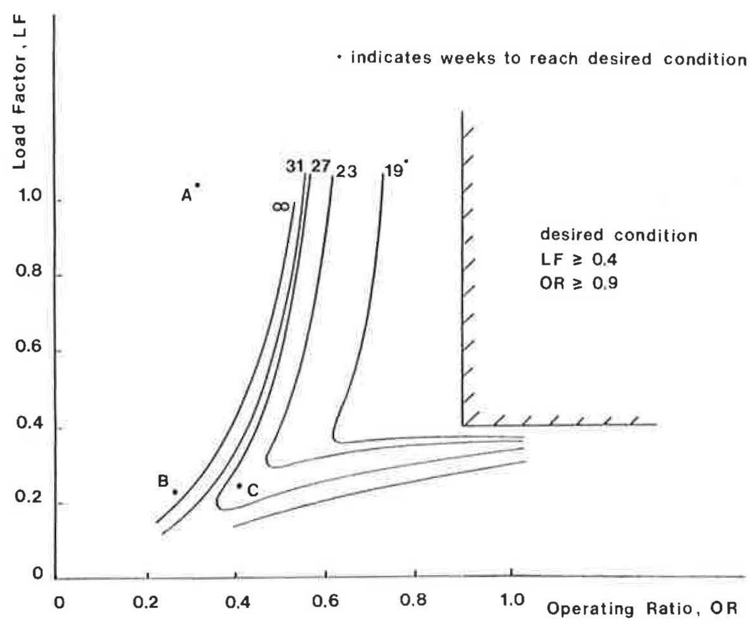


FIGURE 8 Performance time contours.

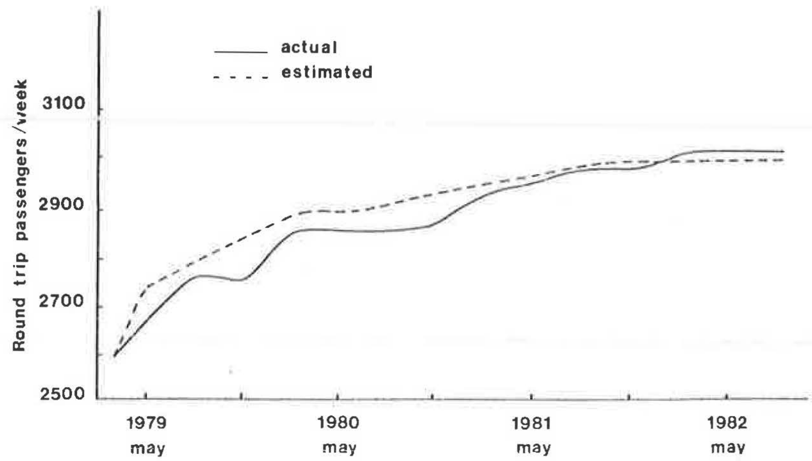


FIGURE 9 Weekly ridership in urban system.

as discontinuing service and laying off personnel. A recent movement of city suburbs toward opting out of the metropolitan transit system and beginning their own services has presented a different kind of application in which decentralized route service replaces the old provider. In this case SOLON can be employed to determine the benefits, if any, of the contemplated change and to estimate the viability and expected performance of the proposed service. Meanwhile, all questions dealing with financing transit deficits are of special interest to the transit administrators of the local department of transportation. Such issues are especially relevant in light of the recent discussions on the possibility of reducing, restructuring, or discontinuing service on selected routes on the basis of performance. Answers to such problems can be provided through use of SOLON as a performance monitoring and evaluation tool.

When implementing the performance evaluation strategies and the method developed in this work, it should be clear that the dynamic analysis described here has addressed a particular set of questions on a particular type of transport service. The questions addressed relate to ongoing policy making, and the type of service implied is one that may change in response to the strategies being implemented. The chosen application was commuter service in a metropolitan area that sought to improve the route operating ratio under a load factor constraint; fare and frequency were the major controls available to management. However, other SOLON versions can address different types of service, objectives, constraints, and controls, in urban or rural areas, in a similar fashion. For instance, a simple version that is designed to assist transit plans for rural routes is now being implemented in a small town in northwestern Minnesota. New versions, which will also relax certain simplifying assumptions currently made by the method (e.g., assumptions on carpool occupancy and route transfers), are under development. All current versions are being implemented in microcomputer PASCAL, APPLE-II, or IBM-PC and operate with 8-64K RAM.

CONCLUSIONS

SOLON is an interactive microcomputer graphic method that addresses the problems of transit scheduling and route efficiency and productivity improvement. It has three unique characteristics: First, it provides solutions at any specified time or continuously through the life of a transport service by tracing the interactions between scheduling changes, route ridership, and service cost-effectiveness. Second, it explicitly treats time delays that hamper route performance. Finally, its modules operate interactively or independently and can be modified by the user.

By addressing the transit problems at the route level, SOLON can facilitate management and funding decision making by determining the viability of individual routes and aid in selecting the best performance policies. Further, because it assumes no previous computer knowledge, it can substantially increase personnel training productivity.

ACKNOWLEDGMENT

The theoretical part of this work was partly supported by the U.S. Department of Transportation. Certain computer tasks on the microcomputer were performed by B. Padmanabhan.

REFERENCES

1. Y.J. Stephanedes. A Feedback Dynamic Model for Transportation Systems. Proc., 2nd International Symposium on Large Engineering Systems, IEEE and University of Waterloo, Waterloo, Ontario, Canada, Sanford Educational Press, 1978.
2. Y.J. Stephanedes. Techniques for Analyzing the Performance of Rural Transit Systems. Final Report II, DOT-RSPA-DPB-50/80/23. Research and Special Programs Administration, Office of University Research, U.S. Department of Transportation, 1980.
3. Y.J. Stephanedes, P.G. Michalopoulos, D. Gabriel, H. Hanna, and R. Plum. Development and Implementation of Dynamic Methodologies for Evaluating Energy Conservation Strategies. Final Report DOT-MN-11-0004. UMTA, U.S. Department of Transportation, 1981.
4. Y.J. Stephanedes, P.G. Michalopoulos, and T. Bountis. Dynamic Transit Scheduling under Efficiency Constraints. Transportation Research, forthcoming 1985.
5. T. Atherton and M. Ben-Akiva. Transferability and Updating of Disaggregate Travel Demand Models. Report 76-2. Center for Transportation Studies, Massachusetts Institute of Technology, Cambridge, 1976.
6. Urban Transportation Planning System. UMTA, U.S. Department of Transportation, 1976.

7. T. Adler and M. Ben-Akiva. Joint-Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips. In Transportation Research Record 569, TRB, National Research Council, Washington, D.C., 1976, pp. 136-150.
8. G.J. Fielding, R.E. Glauthier, and C.A. Lave. Development of Performance Indicators for Transit. CA-11-0014-3. Institute of Transportation Studies, University of California, Irvine, 1977.
9. J.J. Kern and G.A. Weiss. Transit Performance Evaluation Process. Presented at 63rd Annual Meeting of the Transportation Research Board, Washington, D.C., 1984.
10. J. Attanucci, I. Burns, and N. Wilson. Bus Transit Monitoring Manual. Final Report UMTA-IT-09-9008-81-1. UMTA, U.S. Department of Transportation, 1981.
11. Carter-Goble Associates. Rural Public Transportation Performance Evaluation Guide. Final Report DOT-I-83-31. Pennsylvania Department of Transportation, Harrisburg, 1982.
12. Y.J. Stephanedes. Modeling Dynamic Operating and Financial Strategies in Transportation. In Systems Dynamics and Analysis of Change, B.E. Paulre, Ed., North-Holland, Paris, France 1981, pp. 19-32.
13. P.P. Jovanis, W.K. Tip, and A.D. May. FREQ6PE--A Freeway Priority Entry Control Simulation Model. Research Report UCB-ITS-RR-78-9. Institute of Transportation Studies, University of California, Irvine, 1978.
14. T.A. Domencich and D. McFadden. Urban Travel Demand. North-Holland Publishing Co., Amsterdam, The Netherlands, 1975.
15. D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In Frontiers in Econometrics, P. Zarembka, Ed., Academic Press, New York, 1973, pp. 105-142.
16. D. McFadden. The Revealed Preferences of a Government Bureaucracy: Theory. The Bell Jar of Economics, Vol. 6, No. 2, 1975, pp. 401-416.
17. Y.J. Stephanedes. Improved Demand Estimation for Rural Work Trips. In Transportation Research Record 842, TRB, National Research Council, Washington, D.C., 1982, pp. 10-16.

System and Route Optimization Model for Minimizing Urban Transit Operating Deficits

JASON C. YU and UPMANU LALL

ABSTRACT

U.S. transit operators are faced with escalating operating deficits along with growing opposition to the increase in taxes required to offset them. This financial situation has created an immediate need to restructure inefficient and underproductive transit operations. In response to this need, this study developed an analytical framework to help control transit operating deficits. A bilevel optimization model based on nonlinear programming was developed at system and route levels of detail. The model postulates that transit operators could reach a feasible solution for minimizing operating deficits through modifications of current fare and service policies. The model has an economic framework (through the specification of appropriate cost and revenue functions) and solves for optimality through system supply-demand equilibrium. Solutions of the optimization model will provide transit operators with specific operating guidelines for minimizing deficits subject to resource and policy constraints. The nonlinear optimization model is solved using a large-scale (sparse matrix) successive linear programming algorithm. The model was implemented on a microcomputer and was tested with a real-world application to establish its practicality and usefulness.

The financial status of most urban transit properties in the United States is at best bleak. During the past decade, total operating deficits rose more than $4.5 billion, and the problem is likely to get worse. The underlying causes of operating deficits are escalation in transit operating costs, rapid service expansion, and operators' decisions to reduce fare levels. In the past, deficits have been met primarily by government subsidies, with a significant share coming from federal sources. However, as part of the Reagan administration's Program for Economic Recovery, federal operating assistance to local transit properties may be scheduled to be phased out. With the expected subsidy reduction and escalating operating costs, transit operators will obviously be faced with serious financial difficulties.

If the transit system is to be a viable element in urban transportation, it is imperative that the operators establish and maintain a self-reliant operating budget. In considering ways to operate without federal assistance, a first response of most transit operators is to alter fare or service policies, or both. Results from the American Public Transit Association survey indicate that 89 percent of the nation's operators will raise fares and 67 percent will reduce service (1). This trend has been substantiated by the results of another national survey (2). Although an increase in fare may be instrumental in improving transit financing, it must be made without unduly suppressing ridership, because this could actually result in decreasing the overall fare-box revenue. Along with an increase in fares, reductions in operating costs through service cuts may also be necessary to decrease deficits. However, the public demand for transit travel is usually more sensitive to the quality of service than to the level of fare (3).

The net effect is that a reduction in transit service will have a greater negative impact on ridership and hence on revenues than will an increase in fare. Further, transit fare and service variations leading to adjustments in operating costs and revenues should be considered as an interactive process. Although fare and service structures lead to a level of fare-box revenue, a targeted level of revenue can also dictate fare and service policies. Revenue increases resulting from fare hikes, for example, could lead to a demand for a commensurate improvement in the quality of service. This may result in an increase in costs and hence deficits, necessitating a further increase in fares to keep the deficit at the same level. Thus, fare, service costs, revenues, and deficits interact dynamically and sequentially.

Many transit operators have also indicated that the federal subsidy loss will be partly recovered by increased operational efficiency (4). The transit industry is being encouraged to become more productive, not only because of diminishing federal subsidies but also because of the overall economic conditions of the operators. However, this is a difficult task because transit policy making has many components and the problem of reducing operating deficits has many dimensions. Economic, social, and political factors all bear on fare and service decisions. The complexity of the problem stems not only from its magnitude but from the diversity of its parts. At present, there remains a scarcity of comprehensive, yet easy-to-use, procedures that model the realities of transit pricing and operation to aid operators in reducing deficits through fare and service modifications. The literature review on transit planning and optimization aids included in Kour (5) and Yu and Lall (6) bears out this statement.

STUDY OBJECTIVE

The principal objective of this study was to develop an optimization model to minimize or control transit operating deficits by manipulating current fare and service policies, instead of by capital-intensive system changes. The model, which is fully responsive to the typical environment of the urban transit system, can be used as an effective management and planning tool. Emphasis was placed on the practicality of the approach, so the data required are either readily available from or easily assembled by transit operators. The model can be simply applied by using a variety of microcomputers with an interactive preprogrammed package. The model is able to assess the impacts of and develop a strategy for the implementation of various fare and service policies, subject to resource limitations and policy constraints for different situations in which transit services are provided.

The optimization model was developed with the following specific objectives:

1. Generality of application (independent of route configuration and temporal period of application);
2. Focus on minor system modifications (i.e., fare, service frequency, stop spacing) of an existing system;
3. Satisfaction of transit goals specified at the system level through modifications implemented at the route level (i.e., treatment of each route individually and simultaneously in a systemwide context); and
4. Accurate representation of costs, service options, relationship of demand to fare and service, interactions between transit operation components and between supply and demand, and physical and social constraints on system operation.

MODEL FORMULATION

The complexities of transit operation and the need to provide solutions at a relevant level of detail necessitate the consideration of two hierarchical levels of analysis: the system as a whole and the individual route. The reason for this system and route structure is the need to provide solutions to the overall systemwide problem that can be implemented at the route level. The model structure is shown in Figure 1. The optimization model attempts to minimize systemwide operating deficit using route fare and service characteristics as decision variables. Transit operating deficits are total operating costs minus total fare-box revenues. Fare-box management and service cost control are inextricably intertwined. Optimal solutions to these two problems are not independent.

Objective Function

The objective function represents system operating deficits and is expressed in terms of a set of model variables and input parameters. Model variables are partitioned into two subsets: decision variables and relational variables. The latter are quantities defined as functions of the former and are used for a concise model presentation. Decision variables are defined for fare and service options over which transit operators have control. The optimal solution that leads to a minimum operating deficit is obtained by iteratively selecting values for these variables. Interaction between the costs of providing services and the revenue generated by these services determines the optimal state in a supply-demand equilibrium framework.

The objective function is formulated as the difference between system operating cost, defined as the sum of individual route operating costs, and system revenue, defined as the sum of individual route revenues. The objective function is stated as

$$\text{Minimize } D = \sum_{i=1}^{I} (C_i - R_i) \qquad (1)$$

where

D = total system operating deficit,
C_i = operating cost of route i,
R_i = fare-box revenue of route i, and
I = total number of routes in the system.

This function can be applied to any independent time-of-day (i.e., peak, off-peak, weekend) operation and for any length of the total planning period (i.e., month, season, year). In computing operating deficit, the planning period usually refers to 1 year because the operator's budget outlay is typically on an annual basis.

In the following sections the formulation of the operating cost function and of the operating revenue function as components of the objective function will be briefly presented.

Operating Cost Function

The operating cost function is formulated using a cost allocation procedure designed during this study. This procedure assigns all relevant variable costs to four resources: vehicle-hours, vehicle-miles, peak vehicles, and stops. The cost of vehicle-hours relates to labor costs, whereas the cost of vehicle-miles reflects vehicle operation costs. Vehicle and stop costs are confined to the local share of capital depreciation because both are subsidized through federal capital grants. This assignment procedure is a significant departure from the traditional one. Only those variable costs that vary directly with minor system modifications are taken into consideration. All fixed costs and some of the operating costs that do not vary with minor service changes are not included because they are not optimizable. A detailed discussion of the assignment procedure can be found elsewhere (6).

The operating cost function at the route level can be expressed as

$$C = c_h H + c_m M + c_v V + c_y Y \tag{2}$$

where

C = total operating cost of a route,
c_h = unit cost of vehicle-hours,
c_m = unit cost of vehicle-miles,
c_v = unit cost of vehicles,
c_y = unit cost of stops,
H = vehicle-hours operated for the route,
M = vehicle-miles operated for the route,
V = peak vehicles needed for the route, and
Y = total number of stops on the route.

All route costs are summed to obtain the total system cost. The unit costs are assumed to be constant for the range of system modifications and the duration of planning period considered. The unit cost of each resource is derived by dividing the total system cost allocated to a resource by the total use of that resource.

The four resources for each route are then presented in terms of decision variables (frequency of service and stop spacing), relational variable (vehicle operating speed), and other input parameters as follows:

$$H = \ell a(1 + L)n/u \tag{3}$$

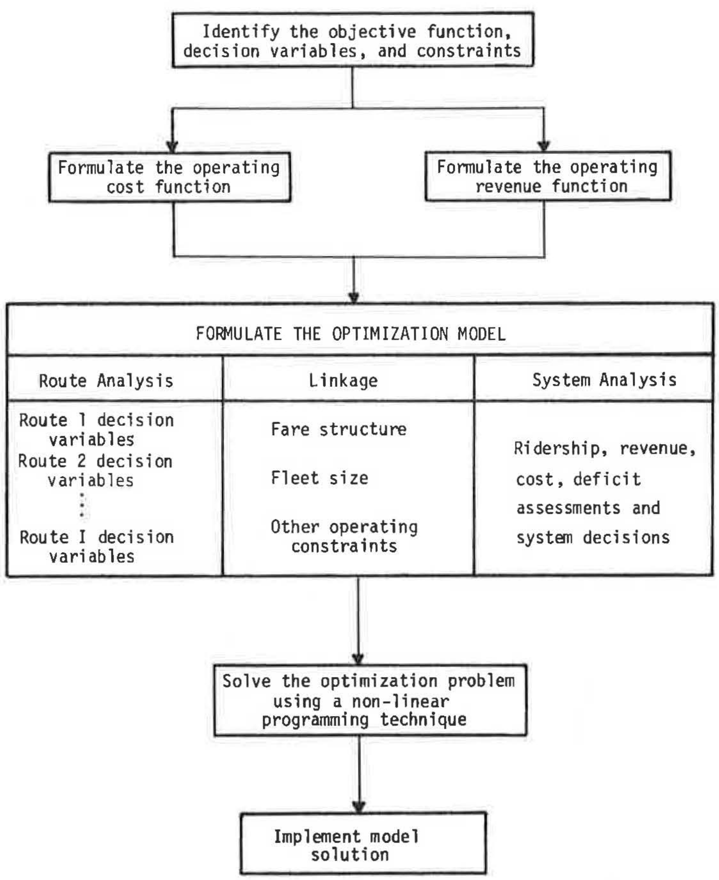


FIGURE 1 Developmental framework of the optimization model.

$$M = \ell an \quad (4)$$

$$V = \ell(1 + P_1)n/u \quad (5)$$

$$Y = \ell y \quad (6)$$

where

ℓ = round-trip route length in miles,
a = service hours per operating period,
L = layover time factor as a fraction of round-trip travel time,
n = frequency of service per hour,
u = average vehicle operating speed in miles per hour (mph),
p_1 = additional vehicle factor as a fraction of vehicles operated on the route, and
y = number of stops per route mile.

Equation 3 states that the annual vehicle-hours operated on a route are given as the product of round-trip travel time (round-trip length divided by operating speed), annual operating hours, and frequency of service per hour. The amount of time spent on each round trip includes the layover time required. Equation 4 states that the number of annual vehicle-miles of a route is given by the round-trip route length multiplied by the number of round trips during the operating period. Equation 5 states that the number of peak vehicles used on a route is given by the product of round-trip time and frequency of service, divided by the average operating speed, plus additional vehicles expected to be on the maintenance schedule and needed to meet other requirements. Equation 6 states that the total number of stops on a route is equal to the product of round-trip route length and number of stops per mile on that route.

The average operating speed (relative variable) for a route can be expressed in terms of decision variables and input parameters as follows:

$$u = \ell/\{[(\ell - d\ell y)/r] + [(d\ell y)/V_a] + [(c\ell y)/3600] + (bQ/3600a)\} \quad (7)$$

where

r = vehicle peak running speed between stops on the route (mph),
V_a = average vehicle speed during acceleration and deceleration (mph),
d = average distance traveled during acceleration and deceleration per stop (miles),
c = vehicle clearance time per stop (sec),
b = boarding and alighting time per rider at a stop (sec), and
Q = ridership of an operating period for the route.

It is widely recognized in the transit industry that the costs associated with providing service during peak, off-peak, and weekend periods might differ substantially due to the quantity and quality of service required. To account for this temporal variation, a procedure based on Cherwony's and Mundle's peak-base model (7) was employed to derive unit cost adjustment factors for each of the three periods of service. In support of using such a relatively simple costing procedure, a comparative study recently performed by Carter, Mundle, and McCollom on various costing procedures (8) found that the increased sensitivity and complexity of the more detailed procedures did not increase relative model accuracy for minor service modifications. In addition, information required by the peak-base model was found to be more easily obtainable from the transit operator.

The following equation, based on the peak-base model, was used to calculate the unit cost adjustment factors for temporal variations:

$$A_p = (H_p^d/H_p^v)(\sum_p H_p^v/\sum_p H_p^d) \quad (8)$$

where

p = peak, off-peak, weekend, or night period index;
A_p = temporal variation adjustment factor for period p;
H_p^d = driver pay-hours for period p; and
H_p^v = revenue vehicle-hours for period p.

The adjustment factors are multiplied by the basic vehicle-hour unit cost to achieve separate unit costs for different service periods. The factors are applied only to the vehicle-hour unit cost because the major percentage of temporal variation in cost results from variability in driver pay-hours and corresponding benefits, which are the main input in calculating the basic vehicle-hour unit cost, for various periods of operation.

In calculating the total system operating cost using the cost functions for different operating periods, the vehicle cost and stop cost would be repeatedly counted. To avoid this cost repetition, unit cost weighting factors are devised to rationally distribute the peak vehicle cost and the stop cost among the different operating periods. The peak vehicle cost is basically the capital depreciation of the vehicles, which in turn is a function of vehicle mileage. Thus the weighting factors are determined by

$$W_p^v = M_p/\sum_p M_p \quad (9)$$

where W_p^v is weighting factor of vehicle cost for period p and M_p is vehicle-miles generated during period p. The vehicle unit cost weighting factors are then normalized so that the sum of the factors equals one (i.e., $\sum W_p^v = 1$).

The unit cost weighting factors (W_p^y) for the stop costs are determined using a similar procedure. Instead of using revenue vehicle-miles, the relative number of revenue vehicle-hours for different operating periods of stop utilized are used. The stop cost weighting factor is therefore obtained by

$$W_p^y = H_p/\sum H_p \quad (10)$$

where W_p^y is weighting factor of stop cost for period p and H_p is time-sharing stop utilization during period p. Again, the factors W_p^y are normalized and applied to the unit stop cost in each time-of-day cost equation so that $\sum W_p^y = 1$.

Incorporating Equations 3 through 9 and 8 through 10, the final total cost function for a route is obtained as follows:

$$C = c_h A_t \ell a(1 + L)n/u + c_m \ell an + c_v W_v^p \ell(1 + P_1)n/u + c_y W_y^p \ell y \quad (11)$$

Operating Revenue Function

The revenue function is formulated by examining the effects of service and fare structure modifications on ridership and hence on fare-box revenue. The fare-box revenue on a particular route is computed as the product of the weighted average fare and the total ridership. The demand function is specified in terms of model variables through elasticity considerations.

The average fare for a route is determined by

$$F = F_c \sum_j w_j e_j \tag{12}$$

where

- F = average fare for a route;
- F_c = base cash fare;
- w_j = fraction of riders on the route using method of payment j; and
- e_j = discount rate (i.e., the ratio of fare paid by the jth method to base cash fare).

In addition to the level of fare, the average trip time of riders is selected as a crucial measure of the quality of transit service and its impact on rider demand responsiveness. Although other measures of service quality (e.g., reliability) were considered, total trip time was believed to be of greatest concern to riders, and it lends itself most easily to the quantitative treatment required for inclusion in the model. Travel time is divided into in-vehicle time and out-of-vehicle time.

Ridership response is modeled with a "shrinkage-ratio" elasticity formulation. Although other formulations may have somewhat greater theoretical validity, the paucity of data precludes their use in practice. It was thought that a shrinkage-ratio formulation would provide adequate accuracy over the constrained range of response modeled. The route revenue function for each route is represented as

$$R = FQ^{o}\{1 + \sum_j \alpha_j w_j[(F_c - F_c^{o})/F_c^{o}] + \beta_1[(t_1 - t_1^{o})/t_1^{o}] + \beta_2[(t_2 - t_2^{o})/t_2^{o}]\} \tag{13}$$

where

- R = route fare-box revenue,
- Q^o = existing route ridership,
- α_j = fare elasticity of rider group j,
- β_1 = in-vehicle time elasticity,
- β_2 = out-of-vehicle time elasticity,
- t_1 = average in-vehicle time, and
- t_2 = average out-of-vehicle time.

The superscript o represents a value for existing conditions (i.e., input).

The average rider in-vehicle time on a travel route is defined by

$$t_1 = L_a/u \tag{14}$$

where L_a is the average round-trip length for a rider on the route.

In the framework of the optimization model it is assumed that the average trip lengths (L_a) are not affected by minor system modifications during the course of optimization. This implies that origin-destination characteristics of riders on a given route are relatively stable.

Out-of-vehicle time has two components: walking time and waiting time. The former is the time spent from the rider's origin or destination to a stop and the latter is the time spent waiting for a vehicle after arriving at a stop.

Average walking time is obtained by dividing the average walking distance by the average walking speed of the riders (normally 3 mph). Determination of average walking distance is based on the assumption that potential riders are uniformly distributed in the neighborhood of a stop and that there is a maximum walking distance (W_m) beyond which no riders are attracted. The average walking time derived by this study is given as

$$t_w = \{1 + 2W_m y(n_s + m_s + 1) + 2[W_1 n_s(n_s + 1) + W_2 m_s(m_s + 1)]\}/[4y(1 + n_s + m_s)V_s] \tag{15}$$

where

- n_s = average number of blocks walked parallel to a route = 2 x int (W_m/W_1),
- m_s = average number of blocks walked perpendicular to a route = 2 x int ($1/2yW_2$),
- W_1 = average block length along the route,
- W_2 = average block length perpendicular to the route,
- V_s = average walking speed, and
- int = integer operator.

Estimation of average waiting time is based on the assumption that the rider arrival rate is uniform during a final waiting time interval (t) and is a mixture of an exponential and triangular distributions during the early waiting time (from the time of departure of the previous vehicle to the start of the final arrival period) if the headway (h) is greater than t. An exponential arrival rate distribution implies the response of well-informed and knowledgeable riders served by a reliable transit system; a triangular distribution implies riders who are misinformed or not well aware of service schedules. Figure 2 shows the concept of waiting times as defined. If no information on riders' awareness is available, the average waiting time is

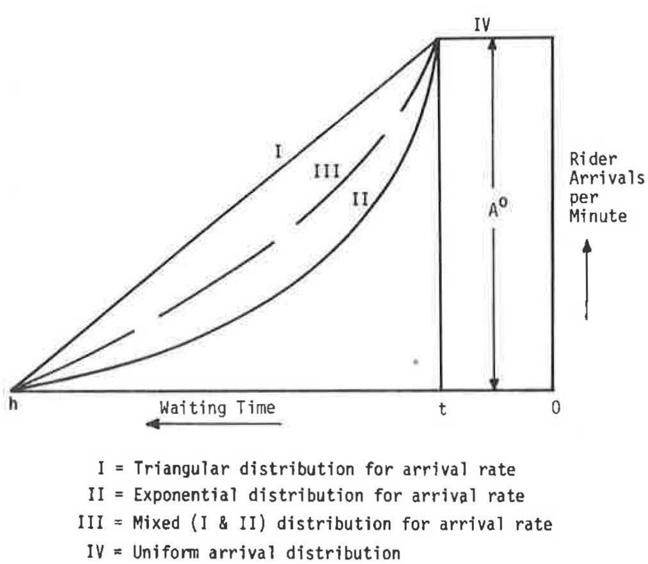


FIGURE 2 Function of waiting time distribution.

$$t_s = 1/2(t/3) + [1200/(tn^2 + 60n)] + \{[n(t^2 + 2t + 2)e^{-t} - (1200n + 2n^2 + 3600)e^{-60/n}] \div [2n^2(t + 1)e^{-t} - 2(n^2 + 60n)e^{-60/n}]\} \tag{16}$$

The total out-of-vehicle time (t_2) is then given by $t_w + t_s$. The detailed derivation of t_w and t_s can be found in the final research report (6).

The complete specification of the revenue function for a route can be expressed as follows:

$$R = (F_c \sum_j W_j e_j) Q^o \Big(1 + \sum_j \alpha_j w_j [(F_c - F_c^o)/F_c^o] + \beta_1 (u^o/u - 1) + \beta_2 \{[(t_w + t_s)/(t_w^o + t_s^o)] - 1\}\Big) \quad (17)$$

Constraints and Bounds

The objective function is minimized subject to a set of existing resource and policy constraints. Explicit constraints are specified to (a) limit system peak vehicle use to a ratio of the existing fleet size, (b) limit peak ridership per vehicle to maximum vehicle loading capacity for each route, and (c) constrain total system ridership to a desired ratio of total existing ridership.

Constraints

System Fleet Size

The peak number of vehicles used for the system should not exceed a specified fleet size (N_f^u) and should have a value of at least N_f^ℓ.

$$N_f^\ell \leq (1 + P_1) \sum_i \ell_i \, n_i/u_i \leq N_f^u \quad (18)$$

Superscripts u and ℓ represent upper and lower limits and subscript i refers to the ith route.

Permissible Vehicle Loading

The number of riders on any vehicle should not exceed the capacity of the vehicle (L_p). The constraint applies to each route.

$$P_2 P_3 (Q_i/a_i u_i) \leq L_p \quad (19)$$

where P_2 is ratio of peak load in major flow direction to average round-trip loading and P_3 is average vehicle occupancy factor between major loading points in major flow direction.

Total System Ridership

Total system ridership should remain above some fraction (q) of total existing ridership (Q_t^o):

$$\sum_i Q_i \geq q \, Q_t^o \quad (20)$$

Bounds

Upper and lower bounds (superscripts u and ℓ) are placed on each decision variable (x) where x includes F_c, n, and y for all routes:

$$x^\ell \leq x \leq x^u \quad (21)$$

SOLUTION TECHNIQUE

As indicated previously, the optimization model is based on a nonlinear programming technique. The nonlinear programming problem formulated is solved using successive linear programming. The algorithm used is one developed by Palacios-Gomez et al. (9) and Lasdon and Kim (10).

The general statement of the nonlinear programming problem is

$$\text{Minimize } f(x)$$
$$\text{subject to}$$
$$g^\ell \leq g(x) \leq g^u$$
$$x^\ell \leq x \leq x^u$$

where

x = a vector of decision variables;
f(x) = the objective function;
g(x) = a vector of constraints, with g^ℓ and g^u vectors of constraint for lower and upper bounds, respectively; and
x^ℓ and x^u = vectors of lower and upper bounds on the variables, respectively.

The constraint set is assumed to be composed of a mix of purely linear and nonlinear constraints. The vector of variables is also partitioned into two subsets: linear and nonlinear. The nonlinear constraints are then transferred to the objective function using penalty weights specified by the user.

The optimization problem can then be stated as

$$\text{Minimize } f(x) + W\, g(x)$$
$$\text{subject to}$$
$$b_1 < \sum_j a_{ij} x_{j\ell} < b_2$$
$$x^\ell < x < x^u$$

where

W = penalty weights,
a_{ij} = coefficients of the jth linear variable in the ith linear constraint,
$x_{j\ell}$ = linear variables, and
b_1 and b_2 = bounds on the linear constraints.

The problem is then linearized by evaluating the nonlinear objective function using a Taylor series approximation at a current solution. The resulting linear problem is solved using a standard linear programming (LP) algorithm designed for large sparse matrices. The LP solution is then used to compute the feasibility of the nonlinear constraints. If the solution to the nonlinear constraints is infeasible, Newton's method is used to find the closest feasible point. A new LP iteration is initiated at this stage and the process is repeated. For a purely nonlinear problem, the algorithm behaves in the same manner as the well-known gradient projection algorithm. A number of criteria are used for termination of the iterative scheme. These include (a) satisfaction of the Kuhn-Tucker conditions, (b) cycling between iterations, (c) slow rate of improvement of objective function, and (d) slow rate of change in penalty functions at an infeasible point.

The implementation of the standard LP package is transparent to transit operators; no mathematical sophistication on their part is required. Input can be provided from data files or interactively with a matrix-generating program that provides data prompts and input instructions.

MODEL COMPUTERIZATION

An interactive, user-friendly, machine-independent, modular structure was adopted for the computer implementation of the optimization model. All the computer programs are developed as a modular package with three basic components--a preprocessor, an optimization process, and a postprocessor. Each of these packages is basically independent allowing for ease of modification. The programs are written in ANSI standard FORTRAN 77 for portability and are fully interactive with the user with respect to data input, result output, and help displays. Backup files for all data entered into the programs are automatically provided, with a label for each piece of data. Provisions have also been made for data entry using data files. The data saved in the backup files may be edited and submitted as a data file.

MODEL APPLICATION--CASE STUDY

The optimization model was applied to a medium-sized transit system operated by the Utah Transit Authority (UTA) to demonstrate its real-world usefulness and practicality. Fiscal year 1983 data were used for this case study. UTA serves an area covering approximately 200 square miles and encompassing two main urbanized areas, Salt Lake City and Ogden. During the study year the population of the service area was estimated to be about 910,000. UTA employment was a total of 745, of which 410 were hired as bus drivers for a fleet of about 400 vehicles. There are 89, 69, and 60 routes for regular peak, off-peak, and weekend services, respectively. UTA received total operating subsidies amounting to more than $21 million in 1983 with more than $4.6 million coming from federal sources and earned fare-box revenues covering only 21 percent of total operating expenses.

In applying the model the bulk of the data was obtained from UTA, mostly from their 1983 Section 15 annual report (11). However, because the Section 15 report provides systemwide data only, route-level data were obtained or derived from UTA's surveys, schedules, monthly passenger count summaries, and technical study reports, as well as personal interviews and special studies conducted by the authors with the help of UTA.

The total expenses incurred and resources provided, along with the unit costs calculated for four resources, are given in Table 1. These values are operating statistics during peak, off-peak, and weekend service periods. The basic unit costs were then subjected to the temporal and weighting factor adjustments discussed previously. All adjustment factors and the final unit costs for each of the three service periods are given in Table 2.

The revenue function for UTA was estimated by using the values of those parameters shown in Equation 17. A summary of these parameter values is given in Table 3. All parameter values vary by route and are not presented in this paper.

TABLE 1 System Cost and Resource Totals and Unit Costs for UTA Regular Services

Resource	Expenses Assigned ($) (A)	Resource Provided (B)	Unit Cost ($) (A ÷ B)
Vehicle-hours	7,742,483	533,564	c_h = 14.51
Vehicle-miles	7,014,497	8,461,880	c_m = 0.83
Peak vehicles	231,815	361	c_v = 642.15
Stops	55,750	11,150	c_y = 5.00

TABLE 3 Selected Input Parameters for UTA[a]

Parameter	Peak	Off-Peak	Weekend
Cash fare, F_c ($)	0.50	0.40	0.40
Methods of payment, j	5	5	5
Discount rate, e_j (percentage of cash fare)			
Student	0.6337	0.6337	0.6337
Adult pass	0.8770	0.8770	0.8770
Special group pass	0.4385	0.4385	0.4385
Commuter pass	1.0719	1.0719	1.0719
Fare elasticity, α_j			
Cash	−0.33	−0.43	−0.43
Student pass	−0.44	−0.54	−0.54
Adult pass	−0.32	−0.42	−0.42
Special group pass	−0.35	−0.45	−0.45
Commuter pass	−0.11	−0.21	−0.21
In-vehicle time elasticity, β_1	−0.52	−0.12	−0.12
Out-of-vehicle time elasticity, β_2	−0.59	−0.51	−0.51
Final waiting time interval, t (min)	10	10	10
Average street block length, W_1 and W_2 (mile)	0.1	0.1	0.1
Rider walking speed, W_s (mph)	3	3	3
Maximum walking distance, W_m (mile)	0.25	0.25	0.25
Rider awareness factor, w_t	0.5	0.5	0.5

[a]From UTA and the literature.

The model was independently applied to peak, off-peak, and weekend periods of operation. The overall deficits for optimal conditions could be compared with the actual deficits incurred for the study year 1983. Comparisons of actual and estimated quantities for existing conditions revealed insignificant differences (less than 3 percent error for all service periods), indicating that the model performs well as a forecasting tool.

The computerized model produces an extensive amount of information at the system and route levels, such as the amounts of revenue, cost, deficits, and resources used for existing and optimal conditions. The model also produces recommended service and fare policy changes to achieve the goal of minimizing operating deficits. In addition, values of system and route performance indicators are provided. Performance indicators are formulated to represent a variety of perspectives on transit system performance. A sample computer output for system-level and route-level results is shown in Figures 3 and 4, respectively.

TABLE 2 Adjustment Factors and Final Unit Costs for UTA

Period	Vehicle-Hour Unit Cost Adjustment Factor (A_p)	Peak Vehicle Cost Weighting Factor (W_p^v)	Stop Cost Weighting Factor (W_p^y)	Vehicle-Hour Unit Cost (c_h)	Vehicle-Mile Unit Cost (c_m)	Peak Vehicle Unit Cost (c_v)	Stop Unit Cost (c_y)
Peak	1.067	0.353	0.450	15.48	0.83	226.04	2.25
Off-peak	0.962	0.491	0.419	13.96	0.83	315.30	2.10
Weekend	0.955	0.157	0.313	13.86	0.83	100.82	0.66

_DCO/TRANS.PEAK/OUTRPT..2   11-Sep-1984 11:21:48

TEST SLC AND OGDEN FOR APRIL 12 1984 PEAK HOUR

SYSTEM LEVEL RESULTS

CATEGORY	EXISTING	OPTIMAL	%CHANGE
TOTAL OPER. DEFICIT ($)	3181855.85	2501739.40	-21.37
TOTAL OPER. REVENUE ($)	2382920.49	2561575.36	7.50
CASH FARE	.50	.54	8.00
AVERAGE FARE	.35	.38	8.57
ANNUAL RIDERSHIP	6833147.70	6805441.91	-.41
TOTAL OPER. COST ($)	5564776.34	5063314.76	-9.01
TOTAL REV.VEH. HOURS	191417.00	171084.41	-10.62
COST OF REV.VEH. HOURS($)	3035181.25	2712779.97	-10.62
TOTAL REV.VEH. MILES	2942509.00	2735465.27	-7.04
COST OF REV.VEH. MILES($	2439200.38	2267570.95	-7.04
FLEET UTILIZED	158.99	141.69	-10.88
TOTAL FLEET SIZE	361.00	321.72	-10.88
COST OF REV.VEH. ($)	60619.23	54023.46	-10.88
TOTAL NUMBER OF STOPS	11217.97	10903.34	-2.80
AV. STOP SPACING/MILE	5.48	4.95	-9.67
TOTAL COST OF STOPS ($)	29775.48	28940.37	-2.80
AVERAGE FREQUENCY/HR.	1.53	1.40	-8.50
PERFORMANCE INDICATORS			
REV.VEH. MILES/REV.VEH.	8150.99	8502.60	4.31
REV.VEH. HOURS/OPER.COST	.03	.03	.00
REV.VEH.MI./REV.VEH.HRS	15.37	15.99	4.03
RIDERS/REV.VEH. HOUR	35.70	39.78	11.43
REVENUE/REV.VEH. HOUR	12.45	14.97	20.24
RIDERS/OPER.COST	1.23	1.34	8.94
OPER.REVENUE/OPER.COST	.43	.51	18.60
IN VEH TIME (MIN/MI)	3.08	3.08	.00
AV. WALKING TIME (MIN)	4.65	4.60	-1.08
AV. WAITING TIME (MIN)	9.81	9.71	-1.02
AV OUT OF VEH TIME(MIN)	14.46	14.31	-1.04
AV OPER SPEED(MPH)	19.50	19.46	-.21
OPER.COST/RIDER	.81	.74	-8.64
OPER.DEFICIT/RIDER	.47	.37	-21.28
REV.VEH. HOURS/REV.VEH.	530.24	531.78	.29
REVENUE/REV.VEH. MILE	.81	.94	16.05
RIDERS/PEAK REV.VEH.	44.06	48.08	9.12
RIDERS/REV.VEH. MILE	2.32	2.49	7.33
REV.VEH. MILES/OPER.COST	.53	.54	1.89

FIGURE 3 Sample computer printout of system-level results.

System-Level Study Results

Using the cost and revenue input data, the optimization model was independently applied to peak, off-peak, and weekend periods of service. The system cost, revenue, and deficit totals are obtained by aggregating the statistics for each of the three periods of operation.

Because ridership is a major indicator of social benefit and is a crucial performance measure with respect to system productivity, it was meaningful to conduct a parametric sensitivity analysis to see how ridership levels influenced deficit totals. For this case study, optimal solutions were obtained for five levels (80, 90, 100, 110, and 120 percent) of present ridership to illustrate the interactive effects between cost, revenue, deficits, and ridership. Thus the relative change in ridership can be estimated for various expected deficit levels, and, conversely, future deficits can be estimated using targeted ridership levels. In addition, deficits, costs, and revenues, the key factors of interest to all transit operators, were examined together with the underlying causes affecting any forecast changes in ridership, service, and fare levels.

Operating Deficits

The relationship between the optimal UTA deficits for the five levels of ridership and existing UTA deficits is shown in Figure 5. Point A in the figure represents the optimal level of deficit corresponding to no federal operating subsidies required. Achieving this indicated only a 5 percent decrease in ridership. However, most transit operators would like to reduce deficits while maintaining or increasing ridership. It can be seen from the figure that it is possible to keep ridership between existing and approximately 112 percent of existing levels for UTA while keeping the deficit at or below existing levels. Also, a reduction of 34 percent (about

_DCO/TRANS.PEAK/OUTRPT..2   11-Sep-1984 11:21:48

STATISTICS FOR ROUTE 11

CATEGORY	EXISTING	OPTIMAL	%CHANGE
OPER. DEFICIT ($)	71584.09	49479.40	-30.88
REVENUE ($)	36665.56	38554.99	5.15
AVERAGE FARE ($)	.35	.37	5.71
ANNUAL RIDERSHIP	105964.00	103170.80	-2.64
OPERATING COST ($)	108249.65	88034.39	-18.67
REV.VEH. HOURS	3602.31	2818.78	-21.75
REV.VEH. MILES	59552.68	50602.96	-15.03
STOPS/MILE	6.90	5.53	-19.86
ROUTE FLEET SIZE	6.99	5.47	-21.75
FLEET UTILIZED	3.08	2.41	-21.75
FREQUENCY PER HOUR	1.77	1.50	-15.25
PERFORMANCE INDICATORS			
REV.VEH.MILES/REV.VEH	19340.99	21002.59	8.59
REV.VEH. HOURS/OPER.COST	.03	.03	.00
REV.VEH.MI./REV.VEH.HRS.	16.53	17.95	8.59
RIDERS/REV.VEH. HOUR	29.42	36.60	24.41
REVENUE/REV.VEH. HOUR	10.18	13.68	34.38
RIDERS/OPER.COST	.98	1.17	19.39
OPER.REVENUE/OPER.COST	.34	.44	29.41
IN VEH TIME (MINS/MILE)	3.24	2.99	-7.72
AV. WALKING TIME (MIN)	4.64	4.68	.86
AV. WAITING TIME (MIN)	8.80	9.76	10.91
AV. OUT OF VEH TIME(MIN)	13.45	14.44	9.09
OPERATING SPEED (MPH)	18.51	20.10	8.59
OPER.COST/RIDER	1.02	.85	-16.67
OPER.DEFICIT/RIDER	.68	.48	-29.41
REVENUE/REV.VEH. MILE	.62	.76	22.58
RIDERS/PEAK REV.VEH.	52.58	60.25	14.59
RIDERS/REV.VEH. MILE	1.78	2.04	14.61
REV.VEH. MILES/OPER.COST	.55	.57	3.64

FIGURE 4 Sample computer printout of route-level results.

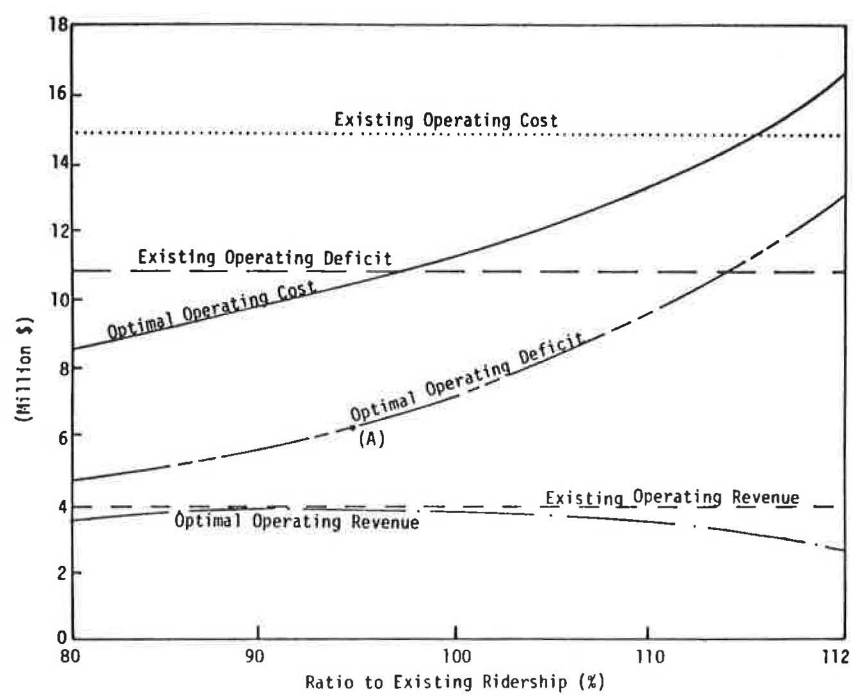


FIGURE 5 System costs, revenue, and deficits as a function of ridership.

$3.7 million) of the total annual operating deficits could be achieved without decreasing ridership. The transit operator may focus on this operational range of increased ridership in optimizing the overall system. Also shown in the figure is that optimal operating deficits rise faster with ridership levels in excess of current levels than with lower ridership levels.

Figure 6 shows the variation in UTA optimal deficits as a function of time period of operation (peak, off-peak, weekend) and as a function of ridership level. It is interesting to note that, for all three operation periods, the percentage reduction in the optimal deficit is almost identical when a reduction in ridership from 80 to 90 percent of existing levels is considered. When the level of ridership is increased, it is observed that the most improvement in deficits occurs for weekend operation, followed by off-peak, and then peak. Increasing the level of ridership beyond existing levels leads to a much higher rate of deficit increase for peak than for off-peak and weekend operation. This observation is consistent with intuition and actual system observations. The justification is that the ridership level is the highest for peak, followed by off-peak and weekend operation, implying higher incremental costs for providing additional service. These in turn imply higher fares (marginal revenue) for system equilibrium and consequently reduced ridership increases resulting in increased operating deficits.

Operating Costs

Operating costs are reduced substantially from existing levels to produce operating deficit reduction. The main goal of the optimization model is to reduce deficits by increasing efficiency and productivity; thus the lowered operating cost resulting from increased efficiency and productivity is a key element in lowering total deficits.

The optimal costs with respect to various ridership levels for UTA are shown in Figure 5. As can be seen in the figure, cost reduction provides the main contribution to the overall deficit reduction. Costs can be reduced by approximately $4 million while maintaining the present ridership level. Ridership can be increased up to approximately 115 percent of existing ridership without increasing costs beyond the present level.

Optimal costs are achieved by modifying service policies to increase vehicle use, thereby reducing vehicle-hours, vehicle-miles, and number of vehicles required for service. Frequencies are reduced slightly leading to a decrease in the number of peak vehicles required. The number of stops per mile is also decreased leading to an increase in vehicle operating speed. These changes result in an overall increase in productivity and efficiency, which leads to reduced operating cost.

Operating Revenue

In examining the total revenue produced for the various levels of ridership as shown in Figure 5, it is seen that the total UTA operating revenue remains relatively stable with respect to ridership change. As indicated previously, cost reduction provides the main contribution to deficit reduction for the UTA application.

UTA revenue remains close to existing levels from approximately 88 percent ridership through 100 percent ridership. Revenue decreases from the 100 percent ridership level to approximately 75 percent of existing revenue at 120 percent of existing ridership. This downward trend is brought about because, in order to attract more riders, not only must the cost associated with providing better service increase but fare levels should simultaneously decrease. Because of the fare elasticities given earlier, the fare levels decrease at a higher rate than ridership increases; thus total operating revenue is decreased. Drastic service cuts (cost reduction) lead to reduced ridership. To satisfy the ridership constraints, fares have to be cut substantially, leading to reduced fare-box revenue.

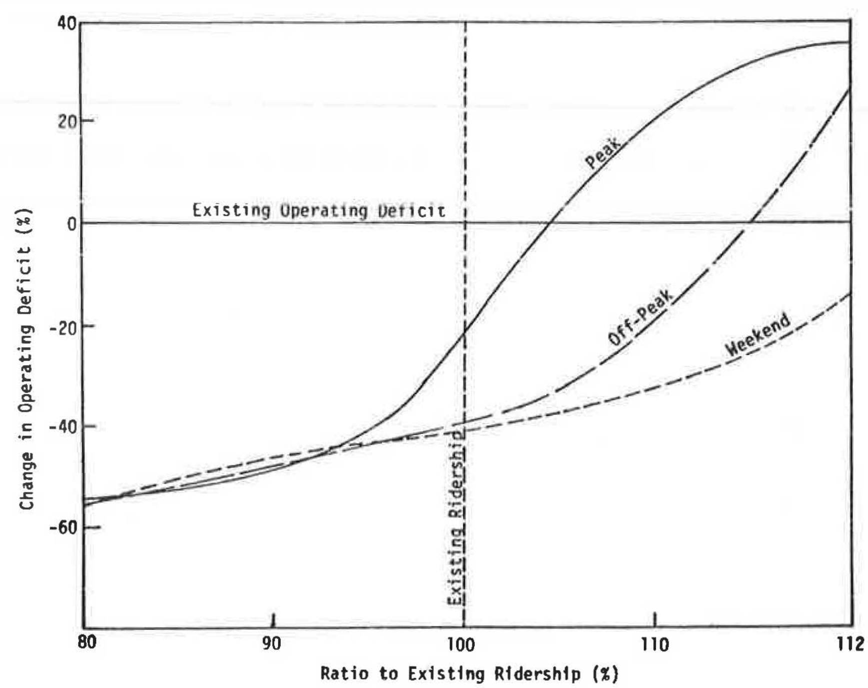


FIGURE 6 Percentage changes in UTA operating deficit versus ridership by time of day.

Route-Level Study Results

The results presented in the preceding sections are based on a system-level optimization; therefore no individual route-level statistics are presented. As indicated previously, the optimization model was developed within a system and route context. Therefore, it is capable of optimizing the entire system at the route level of detail. Any fare and service policy changes recommended to bring about overall system deficit reductions must be implemented at the route level. The system may, however, be defined as comprising all routes, or as few as one route, of a given transit system. In this way the model can be employed to optimize route operations within the system context.

For the sake of brevity, not all details of the route-level analysis are presented. To illustrate model application at the route level, the results obtained for an example route, UTA Route 11, are presented in Figure 4. The economic and operating statistics representing peak-period service correspond to the case where existing ridership is maintained. It is seen that operating cost can be reduced by almost 19 percent and revenue increased more than 5 percent, leading to an overall deficit reduction of approximately 31 percent. The two main modifications responsible for the deficit reduction are the decrease in frequency per hour and the decrease in stops per mile. Frequency per hour, representing the average frequency for both directions, can be decreased 15 percent, and stops per mile can be reduced almost 20 percent. As a result of changes in frequency and in stops per mile, operating speed increased from 18.5 mph to 20.1 mph. On the basis of the modification of cash fare from $0.50 to $0.55 at the systemwide level, the average fare for Route 11 during the peak period should change from the existing $0.35 to about $0.37 per rider.

It should be noted that the results are theoretical. Slight modifications would have to be made in implementing the suggested modifications. For example, the optimal headway may be 28.7 min. In actual practice, a headway of 30 min would be used. The slight modifications necessary for application to the real-world will, however, change the deficit reduction only slightly.

CONCLUSIONS

The model formulated by this study optimizes system deficits, subject to systemwide constraints, through minor individual-route service changes. Model results at the system and route level of detail represented by the decision variables (average fare, frequency of service, and number of stops per mile) are suitable for direct implementation by a transit operator. The optimal condition results in an overall increase in system and route efficiency and productivity, which leads to a reduction in transit operating deficits.

The model was developed in conjunction with continual input from a typical transit operator (UTA) and comprehensively incorporates most of the modeling consideration relevant to transit operators. It has been implemented as a portable, efficient, user-friendly, interactive computer routine.

The solution algorithm (standard LP) used for the nonlinear optimization program performed successfully and satisfactorily. The exploitation of efficient, commonly available, large, sparse-matrix-oriented linear programming solution algorithms makes the choice of standard LP particularly attractive. Convergence to optimal solutions was fairly rapid for the size of the problems solved.

The experience with model applications for UTA indicated that the data needed for the model can be readily assembled by a transit agency. The model results were meaningful, implementable, and intuitively consistent. For all applications of the model using UTA data, significant deficit reductions were achieved without major system modifications, loss of ridership, or undue fare increases.

In summary, the performance of the developed optimization model was judged to be good and representative of the type of model presented.

ACKNOWLEDGMENT

This study was supported by the U.S. Department of Transportation, Urban Mass Transportation Administration University Research and Training Program Grant No. UT-11-0003. The grant was awarded to the Department of Civil Engineering, University of Utah, Salt Lake City, Utah. The authors are especially thankful for the help of the Utah Transit Authority in supplying necessary data for the case study. The study benefited from the assistance of J.Z. Meade and R. Jensen-Fisher of UMTA. The valuable assistance of L.I. Choi is also appreciated.

REFERENCES

1. Cuts Could Force End of 27 Systems. Passenger Transport, Vol. 39, NO. 51, Dec. 18, 1981, pp. 1,9.
2. A Survey of Local Mechanisms for Financing Transit Operating Costs. American Public Transit Association, Washington, D.C., July 1982.
3. A.M. Lago, P.D. Mayworm, and J.M. McEnroe. Ridership Response to Changes in Transit Services. In Transportation Research Record 818, TRB, National Research Council, Washington, D.C., 1981, pp. 13-19.
4. R. Cerverao and G. Black. Probable Effects of Eliminating Federal Transit Operating Subsidies. In Transportation Research Record 936, TRB, National Research Council, Washington, D.C., 1983, pp. 25-31.
5. G.A. Kocur. A Unified Approach to Performance Standards and Fare Policies for Urban Transit Systems, Vols. 1-4. Final Report. U.S. Department of Transportation, Dec. 1982.
6. J.C. Yu and U. Lall. A Bi-Level Optimization Model for Integrating Fare and Service Structures to Minimize Urban Transit Operating Deficits. Final Report. UMTA, U.S. Department of Transportation, Jan. 1985.
7. W. Cherwony and S.R. Mundle. Peak-Base Cost Allocation Model. In Transportation Research Record 663, TRB, National Research Council, Washington, D.C., 1978, pp. 52-56.
8. D.W. Carter, S.R. Mundle, and B.E. McCollom. Comparative Evaluation of Bus Route Costing Procedures. In Transportation Research Record 994, TRB, National Research Council, Washington, D.C., 1984, pp. 1-8.
9. F. Palacios-Gomez, L. Lasdon, and M. Engquist. Nonlinear Optimization by Successive Linear Programming. Management Science, Vol. 28, No. 10, Oct. 1982.
10. L.S. Lasdon and N.H. Kim. SLP User's Guide. University of Texas at Austin, July 1983.
11. Utah Transit Authority. Section 15 Report FY 12-31-83. Salt Lake City, Utah, 1984.

Guided Tour Through the Section 15 Maze

ATHANASSIOS K. BLADIKAS and CHARLES PAPADIMITRIOU

ABSTRACT

Before the first year's Section 15 data were released hopes were high among academics, researchers, and policy makers. It was believed that this wealth of new, detailed, consistent, and accurate information would help answer conclusively questions about transit productivity and performance and about whether subsidies contribute to better performance or are simply wasted in inefficiencies and wage increases. In addition, the data base was eagerly awaited as a tool that would assist in making "peer" comparisons among transit systems, determine the reasons for performance variations, and possibly help in shaping future federal and state subsidy allocation formulas. After 4 consecutive years of data collection, however, Section 15 proved to be far from what was originally envisioned. Although the quality of data has been improving, for all practical purposes, a uniform reporting system that includes all transit properties receiving federal assistance does not yet exist because of numerous problems that may be classified broadly into four categories: (a) access and structural problems, (b) erroneous and missing data, (c) inconsistencies and definitional ambiguities, and (d) exclusion of important data elements. The problems that were encountered in these areas when using the first 4 years' Section 15 data are presented. Suggestions are also made about how users may solve some of these problems and how future editions of Section 15 data may be improved.

Section 15 of the Urban Mass Transportation Act of 1974 as amended requires transit operators receiving federal operating assistance to annually file auditable reports on their system's operations and finances, create an internal system of accounts and records that can provide accurate and detailed information, and improve overall budgeting and operating management. To minimize data collection burdens a "Uniform System of Accounts and Records and Reporting System" was established with a minimum set of mandatory reporting requirements as well as three possible levels (C, B, and A) of voluntary, more detailed reporting. The Transportation Systems Center (TSC) is responsible for editing, tabulating, storing, and releasing to the public data that transit operators supply annually. So far (January 1985) 4 consecutive years of Section 15 reports are available (FYs ending June 30, 1979, 1980, 1981, and 1982, which are also referred to as Years 1, 2, 3, and 4). All information supplied by transit system operators is available on magnetic tape. Most information supplied at the required level is also available in hard copy form in Annual Section 15 Reports (1).

Before the first year's Section 15 data were released, hopes were high among academics, researchers, and policy makers. It was believed that the wealth of this new, detailed, consistent, and accurate information that was about to become available would be used to answer conclusively questions about transit performance and thus enable researchers to determine the impact of subsidies on performance, formulate or restructure federal or state subsidy allocation policies, make transit system "peer" comparisons, and use the results to train future transportation professionals better. Unfortunately, Section 15 proved to be a disappointment. Missing, inaccurate, and badly structured data made the use of Section 15 information a frustrating experience. For all practical purposes, a uniform reporting system that includes all transit properties receiving federal assistance does not yet exist because of a variety of problems that are discussed in this paper.

ACCESS AND STRUCTURAL PROBLEMS

File WDSPSC, which contains the Weekday Service Period Schedule, is a good, but certainly not the only, example of the structural problems of Section 15 data. In all 4 years of existing data, the first 14 columns of the file provide system identification, fiscal year, and mode and day codes. The remaining columns provide information on 11 more variables, as the TSC documentation given in Table 1 indicates. However, this detailed file description is misleading, because variables MYSRB, PMPSB, NTSRB, and NTSRE are actually five and not the four columns wide that the column description from the Year 2 documentation on the table's left side indicates. TSC attempted to correct the documentation

TABLE 1 File WDSPSC Variable Length Specifications for Year 2 (left) and Years 3 and 4 (right)[a]

Column	Name	Type	Description	
15-18	AMSRB	Real	A.M. service begins	15-18
19-22	AMPSB	Real	A.M. peak service begins	19-22
23-26	MYSRB	Real	Midday service begins	23-27
27-30	PMPSB	Real	P.M. peak service begins	28-32
31-34	NTSRB	Real	Night service begins	33-37
35-38	NTSRE	Real	Night service ends	38-42
39-42	AMPRD	Real	A.M. peak period	43-46
43-46	MDYPD	Real	Midday period	47-50
47-50	PMPRD	Real	P.M. peak period	51-54
51-54	NGTRD	Real	Night period	55-58
55-58	TOTHR	Real	Total hours	59-62

[a]File = 39: (DSNAME=WDSPSC.XMI); from area: OPERA.

for Years 3 and 4 and came up with the column widths on the right side of Table 1, but, again, this is wrong because the record is 60 and not 62 characters long. TOTHR is two not four columns wide as indicated. In addition to the incorrectly specified variable lengths, there are considerable discrepancies between the documentation describing the file and its actual structure, and, to make things worse, file structures vary from year to year in terms of record lengths and block sizes. Fortunately, after some changes in the first 3 years, the third and fourth year structure is identical. Furthermore, structural problems are not major, provided that the researcher remembers that block lengths produced by TSC's computers are multiples of four and, from the second year on, record lengths are also multiples of four to avoid the first year practice of adding blanks in the last record of each block. These problems made trial-and-error the only appropriate technique for accessing some of the files, at least during the first 3 years, and occasionally the discrepancies forced users to abandon a file entirely (2).

ERRONEOUS AND MISSING DATA

One major problem that Section 15 data users have to deal with is the distinction between missing data and legitimate zero entries. For some variables this distinction is trivial. Zero vehicle-miles (VM) for a time period a system operates obviously means that the system failed to report. However, a zero for a minor expense item may mean either that the system spent nothing for that item or that it failed to report. Fielding et al. (3) have covered extensively the problem of missing data, which are fortunately progressively decreasing in consecutive years of Section 15 releases, and it is hoped that this problem can be completely corrected in the future. But, even when nonzero entries are provided, on many occasions they are erroneous.

Often erroneous data cannot be detected by inspection. A bus system may report 3 million vehicle-miles accumulated on its vehicles during the period, and that it used 100,000 gallons of fuel. These figures appear perfectly legitimate individually. However, their ratio indicates an average fuel consumption of 30 MPG, which is impossible for buses. Ratios of other pairs of variables from all 4 years of data produce many more surprises. Average operator salaries exceed $100,000, and vehicle-hours per operator exceed the total hours available in a year. These errors do not exist only in the machine readable data but in the raw data and ratios contained in the annual reports as well. According to the second year annual report, taking the ratio of revenue-vehicle-miles to revenue-vehicle-hours, which can be interpreted as average speed of revenue service, the trains in Boston provided service at 153.9 MPH.

This is the most serious problem with the Section 15 data, because users may unsuspectingly use erroneous data. For example, a person who wants to investigate what determines the observed variations in expenses per vehicle-hour may read only operating expenses, vehicle-hours, fleet size, unlinked trips, and hours of system operations and attempt various regressions. Because errors in the data become apparent only when ratios are formed, this person will not suspect the validity of the data and will include in the regression all nonzero values. However, if even a few of the provided entries are twice or three times (and occasionally 1,000 times) larger or smaller than they should have been, the regression fits will become completely meaningless.

All 4 years of Section 15 data contain .XXXXXXEYY entries (Xs and Ys are integers), mainly because the columns allocated to a variable are not enough to accommodate its digits. In Year 1, for example, the four files that contain information on service supplied, service consumed, and service personnel for rail (Form 407) and nonrail (Form 406) systems by hour of day and day of week (files NRSDWK, RSDWK, NRSTDY, and RSTDY) contain a number of entries (56 to be exact) of the type .XXXXXXE69. In later years, the number of .XXXXXXE69 entries was reduced and was replaced by entries of the form .XXXXXXE11. This could be interpreted as 10 to the 11th and it is correct in some instances. But, on other occasions, the power should be 9 or 10. However, whether the power is correct or not, the letter E thrown in unexpectedly will stop execution by causing conversion from decimal to character errors no matter what the access language or package is.

INCONSISTENCIES AND DEFINITIONAL AMBIGUITIES

For the purposes of this paper, a Section 15 data element is inconsistent if it can be read or derived from more than one data file and the alternative derivations produce values that differ by more than round-off errors would warrant. Inconsistent data exist in all 4 years of Section 15 data. However, unless otherwise indicated, Year 3 examples involving single-mode, motor bus systems will be presented here due to lack of space. There is a total of 187 such systems in the third year of data.

Data on Service Supplied and Consumed

Annual vehicle-miles (VM) is an important variable because it is often used to produce a number of performance indicators. It can be computed by multiplying average weekday, Saturday, and Sunday VM by 253, 53, and 59, respectively (from file NRSDWK containing Form 406 information), or by summing the miles accumulated over the reporting period on all classes of vehicles that a system operates (file RVINV containing information on Form 408, Revenue Vehicle Subsidiary Schedule). Taking the ratio of VM from Form 406 to VM from Form 408 should produce numbers probably in the range of 0.98 to 1.02, because all systems may not operate on holiday schedule the same number of days and the multipliers used (253, 53, and 59) may be inappropriate. However, the ratios are distributed as given in Table 2. Ratios that are close to 1,000 are easy to explain. Forms often indicate that figures should be reported in thousands. The line where the number has to be entered may read "Total vehicle miles (000)," although UMTA reversed itself and advised in the first year to report on Form 406 in whole numbers because the small systems could not report meaningfully in thousands. Some op-

TABLE 2 Distribution of Form 406 to Form 408 Vehicle-Mile Ratio

No. of Systems	Ratio Range
5	More than 1,000
5	31.13-889.19
26	1.11-2.65
50	1.01-1.09
14	Exactly 1.00
62	0.99-0.90
17	0.89-0.56
8	0.00

erators obviously became confused in the beginning and remained confused in subsequent years. But if the ratios that are close to 1,000 can be explained, a ratio of 2.65 or 0.6 is an unexplainable and inexcusable variation. Similar discrepancies, although not as large, exist if the ratio of vehicle-hours (VH) from Form 406 to platform-hours (time drivers spend operating a vehicle) from Form 321, Operators Wages Subsidiary Schedule, are calculated.

Although VM are accumulated every time a vehicle moves, revenue-vehicle-miles (RVM) are accrued only when a vehicle is serving the public (i.e., RVM is MV minus deadheading miles). Therefore, RVM should always be less than VM, and, similarly, revenue-vehicle-hours (RVH) should always be less than VH. However, using only Form 406 information, the ratio of VM to RVM is identically 1.00 for 42 systems, or 22 percent of those investigated and VH to RVH is 1.00 for 55 systems or 29 percent of the total. On the other hand, some systems appear to have enormous deadheading miles. There are 12 that have a VM-to-RVM ratio greater than 1.2, and 42 for which the ratio exceeds 1.1.

The definition of RVM and RVH is rather confusing and it is certainly a factor that contributes to the inconsistencies. In defining Service Supplied, the original Volume III of the Uniform System of Accounts and Records that contains the required level forms provides the following instruction for Form 406:

> Revenue miles (line 04) and Revenue Hours (line 05) should exclude charter and school bus miles and hours respectively.

A later version of Volume III (Revised July 1982) changes the instruction to read:

> Revenue miles (line 04) and Revenue Hours (line 05) should exclude charter and deadhead miles and hours.

And finally, Volume II of the Uniform System of Accounts and Records that contains definitions provides on page 8.7-1 yet a third version stating that what should be excluded is miles and hours traveled ". . . to and from storage facilities and other deadhead." In summary, deadhead travel is excluded by the second and third definition but not by the first. Charter service is excluded by the first and second definitions but not by the third, and school bus service is excluded only by the first definition. This is certainly sufficient to confuse and discourage even the most conscientious and well-meaning form preparer.

The existence of two alternative derivations for VM and VH implies that speed (an explanatory variable for many performance measures) can be computed four different ways as follows:

(a) VM/VH (Form 406)
(b) VM (Form 406)/platform hours (Form 431)
(c) VM (Form 408)/VH (Form 406)
(d) VM (Form 408)/platform hours (Form 431)

Using these four alternative derivations to compute system speeds, values that differ among themselves by up to a factor of 2 can be obtained as the small sample given in Table 3 indicates.

Data on service consumed are derived from samples (4) and their accuracy depends on how rigorously the sampling instructions are followed. Some, and fortunately few, operators apparently fill the forms by copying the figures from the previous year's report. There are three systems that reported exactly the same unlinked trips in Years 3 and 4, one system

TABLE 3 Sample of System Speed Variations Depending on Calculation Method

	Speed Derivation Method			
Case	a	b	c	d
1	11.15	8.73	12.10	9.47
2	8.13	6.30	9.36	7.26
3	7.53	6.96	10.61	9.80
4	12.95	13.54	14.67	15.35
5	11.83	13.74	10.38	12.05
6	16.65	17.36	12.50	13.04
7	12.52	13.77	6.35	6.98
8	13.52	19.97	10.74	15.91
9	13.67	11.29	15.50	12.80
10	6.63	5.63	9.63	8.18

that gave the same figures for Years 2 and 3, and finally one that reported identical trips through the first 3 years. The accuracy of data on service supplied (e.g., vehicle-hours and vehicle-miles) depends on the systems' record keeping diligence. Judging from the identical entries for vehicle-hours that exist throughout the 4 years of data, operators are copying data on service supplied from one report to the next more frequently than they are copying on service consumed. There are 10 identical vehicle-hour entries between Years 1 and 2, 10 between Years 2 and 3, and 8 between Years 3 and 4. In addition, there are five systems that report the same vehicle-hours continuously from the first to the third year, and five more that give the same figures for Years 2 through 4. Finally, there is one system that reports exactly the same vehicle-hours for all 4 years.

Service-Profile Data

Variables that describe a transit system's peaking characteristics provide valuable information for the analysis of efficiency indicators. Information on the peaking characteristics of supplied and consumed service can be obtained from Form 406 (file NRSTDY) and Form 401 (Transit System Service Period Schedule contained in files WDSPSC and WESPSC). A number of indicators can be derived from this information such as peak-to-base ratio in terms of vehicles and VH, shoulder-to-shoulder time (start of a.m. peak to end of p.m. peak), and interpeak hours. This wealth of information should satisfy even the most demanding researcher. However, a detailed examination of these indicators and the raw Section 15 data that produce them causes serious doubts about their validity. First of all, the beginning and ending of each period is not defined clearly. Volume II instructions simply say that the a.m. or p.m. peak begins when "additional service is provided to handle higher passenger volumes . . . when scheduled headways are reduced" and ends when "headways return to normal." The problem with this definition is that it is not concise and it can be interpreted differently by the reporting transit systems. A typical service profile may look like the one shown in Figure 1.

When does a.m. peak service really begin in Figure 1? If the Section 15 definition is followed, just about 4 a.m. (the beginning of a.m. service). A few operators reported just that, but most of them reported some later time, which they felt was more appropriate. This leads to a variety of starts, ends, and durations of service periods that are superficial. The service periods of 37 of the 187 systems investigated cannot even be determined because they have zero entries. This may be a problem introduced by the instructions, because operators are provided with an example on page 6.3-4 of the

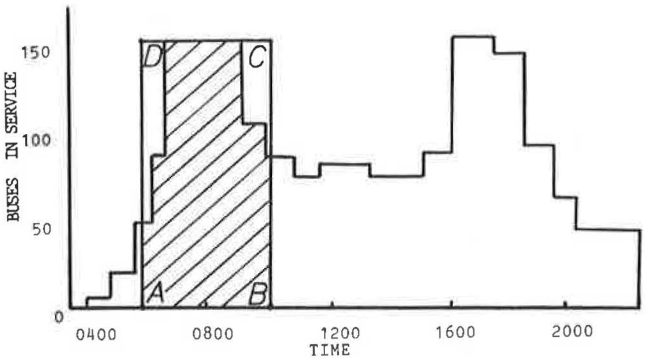


FIGURE 1 Typical transit service profile.

revised Volume III where it is shown that entries should be left blank if service does not fluctuate by time of day. It is therefore impossible to distinguish between systems that did not report and systems that should have zeros due to lack of peaks. Of the remaining 150 systems that report times for their various service periods, 40 report the same time as the start of a.m. service and start of a.m. peak service, and some have the peculiarities given in Table 4. Some of these entries may be legitimate. Systems may really start their a.m. peak before 6:00 a.m. But an a.m. peak duration of 13.8 hr or a p.m. peak longer than 7 hr makes the researcher wonder how valid the rest of the service-period data really are.

TABLE 4 Suspicious Service Period Figures

No. of Systems	Reported Variable	Value
29	Start of a.m. peak	Before 6:00 a.m.
3	Start of a.m. peak	Before 5:00 a.m.
7	End of a.m. peak	11:00 a.m. or later
21	Duration of a.m. peak	4 hr or more
1	Duration of a.m. peak	13.8 hr
40	Duration of p.m. peak	4 hr or more
10	Duration of p.m. peak	5 hr or more
2	Duration of p.m. peak	More than 7 hr

The problem is compounded if an examination is made of the VH reporting during each period. It would be expected that the ratio of (VH during period)/{(vehicles during period)•(duration of period)} would not exceed 1.0 for the peaks and would roughly be in the 0.80 to 0.99 range. This is obvious from Figure 1. Vehicles during the period are, according to the definitions, the maximum number of vehicles. Therefore, the ratio for the a.m. peak is graphically the rectangle ABCD over the cross-hatched area under the service profile. If the midday period is considered, the ratio should be expected to exceed 1.0 for the same reason. However, when this ratio is computed, it is over 1.0 for 39 systems in the a.m. peak and 34 systems in the p.m. peak.

Vehicle Data

Fleet size is critical because it is used in the denominator of many performance indicators. If the number of vehicles is incorrect, particularly if fleet sizes are small, the performance ratios can be over- or underestimated by rather wide margins. There are actually three files containing information on vehicles. File TRSYS gives the total number of revenue vehicles, file TRSVEH gives the total number of revenue vehicles by type, and file RVINV describes revenue vehicles in more detail and groups them according to make and model year. The distinction between active and revenue vehicles is also made in the last file, and, according to the definitions, the number of revenue vehicles should be larger than the number of active vehicles.

However, the ratio of revenue to active vehicles is less than 1.0 for 32 systems. In addition, the ratio of total revenue vehicles from file RVINV to revenue vehicles from file TRSYS is greater than 1.1 for 24 systems and greater than 1.5 for 8 systems, which are not small, and therefore the discrepancy does not involve just a few vehicles.

Examining the ratio of revenue vehicles to peak vehicles in order to find out the percentage of reserve buses a system has (spare ratio), often produces surprises as the sample of some of the worst cases given in Table 5 indicates. The revenue vehicles for this table were taken from file TRASVEH, and those are also the figures that TSC provides in the third year annual report. Peak vehicles are the vehicles serving the highest of the two peak periods. A revenue-to-peak vehicle ratio of up to 1.5 may be believable. However, it is extremely difficult to think of a reason why a system would have three or four times the maximum number of vehicles it needs. A closer examination of the data given in Table 5 together with the data in file RVINV reveals that the system of the third case does not have 102 revenue vehicles; it owns only 52 outright and leases 4 more. Similarly, the system of the fourth case owns only 59 and leases from related parties 10 vehicles, and the system in the fifth case owns just 34 vehicles.

TABLE 5 Revenue-to-Peak Vehicle Ratios for Some Systems

Case	Revenue Vehicles	Peak Vehicles	Revenue-to-Peak Vehicles
1	3,362	1,948	1.73
2	392	206	1.90
3	102	24	4.25
4	107	41	2.61
5	47	17	2.76
6	25	11	2.27

Even if a system's fleet size can be determined, calculating the fleet's capacity is not an easy task. There are two ways to compute average capacity. The first is to divide revenue capacity miles (RCM) by RVM from file NRSDWK. A second figure can be obtained directly by multiplying the number of vehicles times the sum of standing and seating capacity for every vehicle class in file RVINV and dividing by the total number of active vehicles. Table 6 gives 10 of the worst discrepancies between the two alternative derivations of average capacity. The figures differ between 12 percent and more than 100 percent. The inconsistencies in Table 6 are serious because the cases were taken from Year 4 data and from a reduced set of 87 systems that had reasonably clean data in Year 3. This implies that, even if a researcher works hard to select a set of systems that have reasonable data for a set of variables in a given year, data for the same systems cannot be extracted blindly from another year of data either for the same or for different variables. The same painstaking clean-up and cross-checking efforts have to be undertaken for each year and for every variable.

TABLE 6 Average Fleet Capacity Discrepancies

	Average Fleet Capacity	
Case	RCM/RVM (File NRSDWK)	Direct Computation (File RVINV)
1	64	73.5
2	54	63.2
3	45	58.3
4	60	53.8
5	62	51.4
6	50	59.1
7	80	90.2
8	41	61.0
9	45	99.9
10	42	53.4

Along with fleet age, the variable average cumulative miles per vehicle could be used for the prediction of maintenance efficiency variables such as road calls per vehicle-mile. Average cumulative miles per vehicle can be obtained from Form 408 (file RVINV). The data in Table 7 indicate that about one-third of the 187 systems in the sample have provided absolutely useless information by reporting unbelievably high or low figures. The figures that are in the millions are obviously operator reporting errors. Some saw the column heading "Average Accumulated Miles per Vehicle (000)" and instead of reporting in thousands, they multiplied the actual figure times one thousand, thus creating the 1-million-mile discrepancy.

TABLE 7 Erroneous Mileages on Vehicles

No. of Systems	Average Cumulative Miles per Active Vehicle (range in thousands)
6	157,892-630,000
2	18,903-32,751
6	1,487-3,122
11	503-955
6	Less than 18
27	0.0

Employee Data

Serious inconsistencies exist mainly in the files containing information on operating statistics (400 series Forms). Some of them have already been presented. Another common inconsistency exists between total labor-hours for inspection and maintenance (I/M) from file MNPENC (Form 402) and the sum of employees in the maintenance function contained in file EMPSCH (Form 404). These employees are grouped in five personnel classes as follows:

- Class 21 Maintenance executive, professional, and supervisory personnel;
- Class 22 Maintenance support personnel;
- Class 23 Revenue vehicle maintenance mechanics;
- Class 24 Other maintenance mechanics; and
- Class 25 Vehicle service personnel.

The ratio of total labor-hours for I/M to the sum of employees in these five classes should be 2,000, because this is the annual person-hour equivalent for one employee according to the instructions (although TSC appears to use 2,080 to produce the hard copy annual reports). However, the ratio ranges from the teens to the hundred thousands. This is another example of the confusing reporting instructions. The original Volume III contained the following Form 402 instruction on page 6-5 concerning total labor-hours for I/M:

> See definition Volume II Section 8.5, Page 8.5-1. NOTE: You should include all hours worked by employees in Employee Record hours to nearest hour.

However, on page 8.5-1 of Volume II labor-hours for I/M are defined as: "The labor-hours of transit system maintenance personnel working on revenue vehicles for the period." Examining the definitions of Employee Elements and Classification on pages 8.4-2 to 8.4-7 of Volume II, it seems that the only personnel working on revenue vehicles for the period are the revenue vehicle maintenance mechanics and the vehicle service personnel (Classes 23 and 24) and not all classes of maintenance employees as the Volume III instructions indicate. The revised Volume III changed the instruction to read:

> See definition Volume II Section 8.5, Page 8.5-1. NOTE: You should include all hours worked by employees whose labor expenses were charged to Function 061 INSPECTION AND MAINTENANCE OF REVENUE VEHICLES. (See definition on pages 7.4-29 to 7.4-31 of Volume II).

Some operators interpreted this as: "include only Class 23." In addition, Function 061 pertains only to the most detailed, voluntary level of reporting (A). Taking the ratio of employee-hours to employees using the latest interpretation, a range of about 10 to 100,000 is again obtained. Apparently operators became so confused that they simply threw in numbers, and 14 did not even bother to report.

A useful indicator for the examination of operator productivity is revenue vehicles per operator. The value of this ratio is determined by the number of additional operators a system wishes to have during a period either because of its standby policy or because of nonoperating duty assignments. Form 406 and file NRSTDY contain information on vehicles operated during each time period and the number of full- and part-time operators during the same period. Taking the ratio of vehicles over full- plus part-time operators, a researcher would expect values roughly in the 0.85 to 0.99 range. A ratio higher than 1.0 is impossible unless the vehicles are automatically controlled, and 15 percent or more extra operators would be an extreme waste of resources. However, the sample of cases given in Table 8 produces unreasonable ratios that are another example of sources of errors that produce useless Sec-

TABLE 8 Erroneous Data on Operators

Case	Period	Vehicles	Operators (full + part time)	Vehicles per Full-Time + Part-Time Operators
1	X	113	12.8	8.83
	Y	46	4.6	10.00
2	X	8	2	4.00
3	X	32	52	0.62
	Y	18	38	0.47
4	X	281	340	0.83
	Y	264	358	0.74
	Z	101	200	0.50
5	X	7	13	0.54
6	X	30	56	0.54
7	X	16	27	0.59
8	X	27	46	0.52

tion 15 data. The first and most illogical case can be corrected by moving the decimal one place to the right in the number of operators; this is obviously a transcription error. The up to 100 extra operators of the fourth case are probably the result of a confusing definition. Operators are defined to be those scheduled to operate vehicles. Thus it is conceivable that if during a time period, for example midday from 10:00 a.m. to 4:00 p.m., the number of vehicles in operation were 100, the system might be inclined to report the number of operators as 210 because 105 operators worked until 2:00 p.m., and at that time another 105 replaced them.

The conflicting instructions and misunderstandings presented so far are certainly not the only ones. Holec et al. (5) mention more in their discussion of Year 1 data. However, Section 15 data inconsistencies cannot always be blamed on conflicting or unclear instructions. The revised Volume III instructions for Form 321 state clearly that the figure to be placed in the dollar column for total operating and nonoperating time ". . . must balance to the dollar amount reported in Object Class 501.01--Operators' Salaries and Wages on the appropriate Section 15 expense reporting forms." The appropriate form suggested in the instructions is Form 301 (Expenses Classified by Function), and the total for Object Class 501.01 can be found in file X0. Although the instructions are perfectly clear, some operators report figures that differ between -3 and 22 percent.

Financial Data

Financial and expense-related data that do not come from samples or inconsistent collection procedures should be expected to be more accurate, and generally they are, although occasionally some expense figures are questionable. Average salaries by function and for the entire system can be computed by taking the ratio of salaries over the number of employees. Five such ratios are possible in the most aggregate function level (i.e., operations, vehicle maintenance, nonvehicle maintenance, general administration, and total), and they would not be expected to fall outside the approximate range of $10,000 to $30,000. However, some systems appear to pay large amounts for some employees, whereas other systems do not even pay the minimum wage. There are eight systems that report average salaries of more than $60,000 and as high as $800,000 and 16 systems that appear to be paying less than $7,000 and as low as $436 per year.

Major discrepancies and inconsistencies exist in the financial data reported in series 100, 200, and 300 Forms. Balance sheets and revenue summaries were reproduced for all systems and for each of the reporting years and checks were performed on the additions and to see whether assets were equal to liabilities and capital. If discrepancies existed, efforts were made to resolve and correct them.

Year 1 data contain nine resolvable discrepancies most of which arise from the omission of the minus sign in depreciation items ranging from $2,000 to $48 million. Year 2 contains 28 resolvable discrepancies most of which involve the reporting of items 10 times as large as they should have been, and the omission of the minus sign from accumulated losses, which appear as gains. Year 3 data contain 30 resolvable discrepancies about half of which involve the improper addition of revenues. Finally, Year 4 contains 12 resolvable discrepancies produced by all of the previously mentioned causes. Although the errors in the financial data are small in number, they are serious when they involve large systems. For example, the Washington Metropolitan Area Transit Authority's assets were overestimated by $21 billion in the second year. This figure should be compared with the approximately $15 billion in assets, $6 billion in liabilities, $9 billion in capital, and $7 billion of revenues for all Year 4 systems together. Thus the errors of one system can make the calculation of totals or averages completely useless. This is also true for variables that represent service supplied and consumed.

Although most major errors in the financial data can be resolved, there are a large number of unresolved discrepancies that are quite minor (less than $10 in most cases). Table 9 gives a summary of the unresolved discrepancies. Year 3 is the best of all. Sometimes the dollar amounts are smaller than the number of systems, because the real deviations between totals and the sum of their component parts were used. Thus, if a total is larger than its parts by $2 in one system and smaller by $2 in another, the discrepancies will cancel each other out. The large dollar amounts that appear in Table 9 are contributed mostly by a single system in each case.

The errors in the financial data provide an indication of what portion of the Section 15 inaccuracies can be blamed on transcription errors because there are no ambiguities involved in the preparation of a balance sheet. However, revenue inaccuracies may arise from the confusion of cash accounting, which most systems use, and accrual accounting that is required for Section 15 reporting (5,6). If there are about 30 major (more than $10 that can arise from round-off errors) balance sheet-related errors in each year of Section 15 data, it might be safe to assume that about 10 percent of the reporting systems have errors in nonfinancial data as well because of transcription errors and poor quality control.

EXCLUSION OF IMPORTANT DATA ELEMENTS

On the basis of the structure of the Section 15 system, it appears that those who originally conceived the breakdown of the data into nine areas and 62 files had intended to create a system that grouped the reported information better than the forms them-

TABLE 9 Unresolved Discrepancies Between Totals and Component Parts in Financial Variables

	Year 1		Year 2		Year 3		Year 4	
Discrepancies	Systems	$	Systems	$	Systems	$	Systems	$
Assets	41	329	37	2,479	9	4	12	13,859
Liabilities	6	3	10	896,664	6	8	9	1,006
Capital	27	172	20	17,975	12	6	8	2
Revenue	7	22,702	9	79,501	6	38	8	24
Total	81		76		33		37	

selves. This would allow the Section 15 data user to go easily to a file or set of files and obtain the pertinent information no matter what his interest was. Whether the data were to be used to evaluate issues pertaining to transit financing, performance evaluation, safety, or the cost-effective provision of service, the analyst would have a small number of files to access and could perform the intended task efficiently. The noble effort was also undertaken to provide users with even more data than the operators supplied. File UAREA is the produce of such an effort, and it is supposed to contain two variables, the square miles of area (USQMI) and the urban population (UPOP) that a system serves. First of all, variable USQMI is always zero, so half of the file is useless. The values of UPOP are for all practical purposes useless too. The New York City transit system with more than 10,000 vehicles is shown to serve 15.59 million people, and so do 35 more systems, including ones like the Resort Bus Lines of Yonkers, New York, with all of its 8 vehicles. Apparently, the system's address was the only determinant of population served. Area and population served are important elements in the analysis of a system's effectiveness, and for all practical purposes they are missing from the data. An attempt was made to include the variable UPOP in models that analyze bus systems' performance, and it failed to enter any of the equations with even a minimal degree of explanatory power (see paper by Bladikas and Papadimitriou in this Record.) Obtaining data for USQMI and UPOP is certainly not an easy task. UMTA could obtain this information with a research grant, or at production cost from the Bureau of the Census.

The usefulness of the Section 15 data would improve significantly if some additional information were collected. Most of this information could be collected by simply expanding Form 001 (Transit System Identification Schedule) to two pages or by replacing Form 332 (Pension Plan Questionnaire), required for systems with 25 or more vehicles and apparently ignored by everybody, because only a handful fill it out every year. The additional information could be useful not only because it would provide more explanatory variables for research in the area of performance evaluation, as others have already suggested (7), but because it would also make possible additional consistency checks. The minimum additional information should be

- The system's fare structure by mode;
- Entry, average, and maximum salary by function;
- The system's organizational and management structure (city agency, independent authority, whether it is managed by an independent company, and so forth);
- Union contract data (whether employees are unionized and some key contract provisions such as split shifts, spread times);
- Vehicle retirements and purchases during the reporting period; and
- Revenues (at least transportation revenues) broken down by mode.

SUGGESTIONS FOR IMPROVEMENT

There are three parties responsible for the introduction of errors in Section 15 and, therefore, the situation will be corrected only if all three make efforts to improve. UMTA has to work on the four volumes of the Uniform System of Accounts and Records and Reporting System to clear up the definitional ambiguities and conflicting instructions that the various volumes contain. In addition it should become stricter and reject operator reports that are obviously erroneous or inconsistent. Operators certainly see Section 15 as an additional burden imposed on them by UMTA and appear to be the major cause of errors. Unfortunately, the lack of clear instructions on the forms that they are filling out does not make that burden easier. TSC can improve its quality control and the accuracy of the published information, as it is capable of doing if judged on the basis of the excellent quality of other, more voluminous than the Section 15, data that it releases.

As far as the machine readable data are concerned, the field lengths for the variables that do not fit in them should increase so that the EYY entries can be eliminated. The detailed file descriptions should be changed to include the blanks (when they exist) at the end of every record, so that the record lengths match those described in the access JCL. The last set of files (LOOKUP) is practically redundant and useless because it contains the codes that exist on the reporting form anyway. The information contained in the 17 data files should be placed on just 7 typewritten pages to provide the user with an easy reference. Mileage-weighted fleet capacities, ages, and capacity miles would be more valuable indicators in the annual reports, because it is most likely that the largest, smallest, oldest, or newest vehicles are used only during certain periods of time and the entire fleet is not in service from the beginning to the end of the service period.

There is a set of about 40 flags in the machine readable Section 15 data. Only eight of the files have from one to twelve flags each. Apparently when the files were initially created these flags were supposed to act as checks for the values of every variable. However, soon this effort was abandoned, leaving only eight files with flags. Unfortunately, even the existing flags are of little value because they do not appear to be flagging anything. Flags should be placed on all files and put to use. It is possible to devise a flag system as follows:

Flag Value	Meaning that Item Is
0	Correct
1	Not reported (missing)
2	Out of a preset range
3	Inconsistent

Of course, if such a flag system is to be implemented, TSC has to develop software that will perform range checks and look for all possible inconsistencies in the reports. When a report is received, it should be analyzed and returned by UMTA to the operator for corrections. Data should be flagged if, even after the corrections, they are still erroneous or inconsistent.

Looking at Section 15 data for all 4 years suggests that there is some improvement in their quality with every subsequent edition. However, the improvements have to be speeded up by a factor greater than the normal learning process for all involved parties is bringing about. This is needed particularly because (a) the existence of a few years' data will soon allow the analysis of time series, (b) states are using Section 15 data to develop their own performance measures [e.g., Michigan (8)], and (c) future legislation may include in the allocation formulas even more performance measures than does the Surface Transportation Assistance Act of 1982.

REFERENCES

1. National Urban Mass Transportation Statistics. Report UMTA-MA-06-0107-81-1. UMTA, U.S. Department of Transportation, May 1981.

2. G.J. Fielding, M.E. Brenner, and O. de la Rocha. Using Section 15 Data for Transit Performance Analysis. Interim Report, UMTA-CA-11-0026-1. UMTA, U.S. Department of Transportation, Jan. 1983.
3. G.J. Fielding, M.E. Brenner, and O. de la Rocha. Using Section 15 Data: Adopting the Magnetic Tape Version for Statistical Analysis. In Transportation Research Record 961, TRB, National Research Council, Washington, D.C., 1984, pp. 28-35.
4. Circular UMTA C 2710.1. UMTA, U.S. Department of Transportation, Feb. 22, 1978.
5. J.M. Holec, Jr., D.S. Schwager, and M.J. Gallagher. Improving the Usefulness of Section 15 Data for Public Transit. In Transportation Research Record 835, TRB, National Research Council, Washington, D.C., 1981, pp. 9-15.
6. Circular UMTA C 2710.5A. UMTA, U.S. Department of Transportation, Oct. 1, 1981.
7. S.C. Anderson, G.J. Fielding, and D. Methe. Comparison of Supply, Demand and Cost Models Using UMTA Section 15 Data. UCI-ITS-WP-81-4. Institute of Transportation Studies, University of California, Irvine, Dec. 1981.
8. 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.

Research Implications of Proposed Changes in the UMTA Section 15 Reporting System

JOEL E. MARKOWITZ

ABSTRACT

An intensive effort started in 1983 to review the Urban Mass Transportation Act (UMTA) Section 15 reporting system for transit statistics. Although many transit industry professionals have been involved, few researchers are aware of the ways in which proposed changes would alter the national data base. A summary of the efforts to date is presented, and the implications of the proposed changes for those who have been routinely relying on Section 15 data for the conduct of research on U.S. transit systems are highlighted.

In 1974 Section 15 of the Urban Mass Transportation Act of 1964 was amended to require that transit agencies receiving federal formula grant funds submit a uniform report on their financial and operational characteristics each year (1). The requirement grew out of a large-scale study that examined transit industry accounting practices in detail. The result was a series of forms and manuals documenting accounting definitions that would be used for the required annual reports (2). The standards laid the framework for upgrading the management information systems in the industry as a whole. The nonfinancial data did not receive as much careful study in the early days and have continued to cause some problems, especially now that certain of those data have been incorporated into the new Section 9 transit block grant formula program (3).

Although transit industry representatives were actively involved in the work leading to the adoption of the Section 15 standards, some problems in the reporting system appeared only after the first few years of implementation (FY 1978-1979, 1979-1980 and 1980-1981). A massive amount of information is involved, from a few hundred data elements for the lowest level of required reporting to a few thousand elements for the larger multimodal systems. Inaccuracies in reporting, misunderstanding of definitions, inconsistencies within reports, and difficulties in quality control joined with some instances of outright refusal to cooperate. The result was a national data base with serious limitations. At the 1984 Transportation Research Board Annual Meeting, several presentations were made on transit performance analysis using Section 15 data. All illustrated the many problems inherent in the data that required either elaborate cleaning procedures or simply the exclusion of whole sets of agency reports. Fielding et al. (4,5) reported on the difficulties in organizing the magnetic tape version of the FY 1980 data for statistical analysis. Of 304 agencies that reported that year, 106 had missing data that prevented their being used in the performance analysis work. Vaziri and Deacon (6) similarly used the FY 1980 data base and had to work around problems caused by missing data. Hobeika et al. (7) found so much missing data on the items of interest that the analysis was done only for systems with fewer than 100 revenue vehicles. Patton (8) used only 17 systems from the FY 1981 data base for exploratory analysis. All of these researchers tried to look for regularities in a variety of performance measures within a single year's Section 15 data. All were at least partly stymied by data problems and none could use data from several years for time-series comparisons.

Beginning in 1983 several groups began actively examining the shortcomings in the current system of reporting. The TRB Committee on Transit Performance and Management formed a subcommittee to look at the analytical uses of the data. The American Institute of Certified Public Accountants (AICPA) formed a group to examine the new requirements under Section 9 to certify the nonfinancial as well as financial data. The American Association of State Highway and Transportation Officials (AASHTO) formed a committee to examine how state transportation departments use Section 15 and how changes might affect them. The American Public Transit Association (APTA) formed its committee, including several individuals who had participated in the development of the Section 15 system, to look broadly at the present problems and future prospects for national transit reporting. Finally, the Urban Mass Transportation Administration (UMTA) took the unusual step of appointing a special advisory committee, the first UMTA has ever had, to examine all aspects of Section 15 policy and practices. (46 C.F.R. 43352 established the committee for its first term, September 1, 1981 to September 1, 1983; 48 C.F.R. 41124 extended the committee's charter to September 1, 1985.)

The UMTA committee, which is staffed by UMTA and meets quarterly, receives input from all interested parties. Typically, the APTA committee or other parties have prepared recommendations for consideration by the UMTA committee, which then acts by resolution. Consultants to UMTA have also assisted the committee's deliberations by preparing background papers. One paper, for example, presented the results of asking the analysts responsible for checking the validity of the Section 15 data to assess its reliability. The assessment covered 233 data items in the 100, 200, and 400 series of reporting forms (excluding the 300 series detailed financial data). Of those items, 98 (42 percent) were rated as either inconsistently accurate or generally inaccurate or missing (9). Fortunately, there is some overlap in the membership of the groups and a direction for reforming Section 15 is beginning to appear. In the remainder of this paper the efforts to date, focusing on the APTA committee, and the potential benefits and problems such changes would present to researchers who now depend on the data are described.

USE OF SECTION 15 DATA IN RESEARCH

There appear to be several principal research uses of Section 15 data. Perhaps the most widespread use is the most difficult to document. That is the routine use of Section 15 data, especially in the form of the published annual report, as an encyclopedic reference, as if it were the transit equivalent of census data. When a researcher wants a national summary statistic on transit or a particular statistic on an individual agency, the book is there to provide the numbers. Although much of the data is straightforward, there are many underlying limitations that the casual user of the data could not be expected to know about. The most common situation is that circumstances affecting the data were not documented in the report. Unusual weather, service disruptions caused by labor disputes, major service changes, or fare changes could greatly affect an agency's data in comparison either with other agencies' or with the same agency's data from prior years. Unless the user of the data either knows enough about the data to tell if a number looks suspicious or takes the trouble to contact the agency to confirm the numbers, there are possibilities of misinterpretation. This has been one of the sorest points with transit agencies. The format of lengthy tables of raw data and ratios, listing each agency in fleet-size groups, encouraged such comparisons but had few cautions to the user.

A further and more subtle limitation on using the published data for routine reference needs is that some numbers are not what they appear to be because of the way individual agencies collect the data. Some discrepancies are matters of local policy, such as standing capacity on vehicles. Some are the results of estimates, such as scheduled versus actual vehicle-miles of revenue service provided, which depend on the local agency's ability and willingness to provide accurate information. Other data leave out key elements due to local institutional arrangements. For example, when a transit agency is part of a state, city, or county government, certain services, such as purchasing or personnel, may be provided to the transit agency. The full amount of these expenses may not be reflected in the transit agency's budget, with the result that the "true" operating expenses are underrepresented in comparison with an independent agency that must provide all of its own services. Again, only prior knowledge or checking with each agency would prevent misinterpretation or misuse of the data. It is likely that few casual users have the time or inclination to double check the published figures, and even less likely that they know they need to be concerned about the data at all.

In addition to using the published Section 15 data for simple reference work, researchers have worked with the entire set of published data or have gone to the much more detailed computer tapes to conduct analyses that might be categorized by the purpose of the research and the style or method of analysis used. The principal purposes, judging from papers such as those presented at the 1984 TRB Annual Meeting and others, are (a) the development and testing of statistical tools to aid transit managers in analyzing performance and (b) the analysis of the data to answer particular national policy questions. In both cases the emphasis is on cross-sectional comparisons of "similar" operators, with the bulk of the research effort devoted to defining the dimension of similarity. Researchers have expressed interest in longitudinal analysis as well, but they generally have found that the quality of the data over time has not been sufficient to the task. Thanks to increasingly sophisticated data validation by UMTA, the accuracy of the FY 1982 report was greatly improved, and recommendations by the UMTA advisory committee to further improve the data have been adopted or are being considered by UMTA. In time, these improvements should allow meaningful time-series analysis.

A particular focus of Section 15 research efforts has been on statistically determining a summary set of descriptive performance measures on the basis of which an operator may be compared to the "average" performance of a group of similar operators. These "peer groups" have been a controversial aspect of this type of research. Many transit managers readily admit that they compare their performance with that of other systems, but they also bring to such comparisons some direct knowledge of the operational, institutional, and managerial character of the selected agencies. Managers have some misgivings about surrendering the selection process to a statistical procedure less able to capture such qualitative distinctions and with which they may be technically unfamiliar. Proponents of such research efforts, however, point out that the purely qualitative selection of peers invites comparisons designed to be favorable to the agency making the selection. Another school of thought is that any such cross-sectional comparisons of operator performance are of little real use to managers and what is needed instead is analysis of an individual agency's performance indicators over time. Regardless of where one stands in this debate, more consistent, accurate, and reliable Section 15 data are essential for any meaningful research.

ISSUES IN REVIEWING SECTION 15

Both the APTA and UMTA committees set ambitious goals for their review. Among the issues identified for study were the following:

- Administrative and procedural issues
  - Certification and audit requirements,
  - Definition of reporting period,
  - Quality and availability of Section 15 instructions, and
  - Treatment of overdue reports.
- Policy questions
  - Should very small systems be exempt?
  - How should purchased transportation be treated?
  - How should private, noncontract service be reported?
- Specific areas for improvement
  - Section 9 formula factor definitions,
  - Urbanized area definitions,
  - Commuter rail,
  - Sampling for service-consumed data,
  - Fleet inventory data,
  - Safety and accident data, and
  - Maintenance data.
- Changes in the published report
  - Format and content of tables,
  - Addition of explanatory material to aid interpretation,
  - Performance indicators, if any, to be used,
  - Graphic summaries, and
  - Groupings, if any, by size or other "peer" categories.
- What shall be reported?
  - What should be reported at the national level?
  - Mandatory versus voluntary levels of reporting,
  - Amount of detail required versus need at national level,
  - Need for cross-classification of expenses by function and object by mode, and
  - Modal versus system data.

By the end of 1983 some of these issues had been discussed thoroughly, some superficially, and few conclusively. Both the APTA and UMTA committees decided that it would be most appropriate to focus their efforts on the cluster of issues under the rubric of "What shall be reported?" When the principles were established, it was thought, there would be a firmer basis for discussing all other issues. The remainder of 1984 was spent developing a framework for considering "What shall be reported?" Both the APTA and UMTA committees are scheduled to meet in early 1985 to consider these recommendations.

SECTION 15 REVISION OBJECTIVES AND PROPOSALS

Assumptions

The APTA committee began with several working assumptions. Each is subject to further discussion and revision, but together they establish a springboard for the debate. There were four major assumptions. First, certain problems in the data are the result of the entire reporting system's being too cumbersome. With so many data items required and so many different forms to complete, errors and inaccuracies are inevitable. Therefore, simplification through reduction in the sheer number of discrete data elements should be a goal. Second, part of the reason for noncomparability of data items across operators comes from basic problems in data definitions and difficulties in obtaining the data from typical operator information systems. Therefore, the feasibility of accurately and efficiently collecting the desired information should temper demands for data. Third, if the data are not useful to a transit manager, they probably are of little use to national policy makers or researchers. National reporting should be no more than summary reporting on an annual basis of much more detailed data the transit agency must itself keep on an ongoing basis to manage its own operations. (A notable exception may be data required for the formula grant program.) Fourth, multiple levels of mandatory and voluntary reporting contradict the need for comparability of data items across all operating systems. A single, required level of reporting is desirable, but a two-tiered system of more detailed reports for larger systems and less detailed reports for smaller systems should be considered.

Criteria

With this as a basis, the committee developed three types of recommendations: (a) restructuring existing data items, (b) reducing the existing data items by consolidating them, or (c) eliminating data items. For the financial reporting forms, the approach taken was to examine each of the functional, line item (object), and revenue categories, irrespective of the current reporting format. When the categories and grouping of categories had been set out in principle, issues involving specific definitions and forms design could be addressed. Each nonfinancial reporting form was examined on its own merits. The committee developed the following criteria for screening out unnecessary data items:

- Is the information required for the Section 9 formula?
- Is the information useful for the purposes defined by Section 15 (i.e., federal, state, and local policy decisions, information for the public and for transit agencies)?
- Is the information comparable from operator to operator? Similar data should be collected for each mode. Voluntary reporting of selected data should be discouraged. Contracted services and directly provided services should be clearly separated.
- Is it feasible to collect the information so that it is comparable, statistically valid, clear, simple, and easy to understand? Data should flow from the operator's regular reporting system, with no special collection procedures or excessive costs required. Data should only be reported if they will be suitable for annual publication by UMTA (i.e., if data will not be readily accessible to most users, they should not be required).

Application of these criteria resulted in recommendations for a substantial reduction in the number of data items required currently under the voluntary levels and some reduction or slight increase in the number of items for the required level. The specific recommendations for the financial data items are described next.

Recommendations

Revenue Classes

The revenue classes recommended for national reporting are in place of the 80 under the most detailed voluntary (A) level and the 15 under the current required (R) level:

1. Operating revenue
   - Passenger fares (currently item 401);
   - Other transportation revenue (currently items 402, 403, 404, 405, and 406); and
   - Nontransportation revenue (currently item 407).
2. Other revenue
   - Contributed services (currently item 430) and
   - Subsidy from other sectors of operation (currently item 440).

It is proposed that these revenue items be combined on the form that presents operating assistance by source and type. This is intended to reduce duplication and ensure internal consistency in reporting.

Functional Classes

The committee recommended that there be only 12 functional categories, compared to the current level A number of 44. The 12 collapse into the same four functions currently used at the lowest (required) level of reporting:

1. Transportation/operations
   - Administration (currently item 011);
   - Revenue-vehicle operation (currently item 031); and
   - All other transportation/operations (currently items 012, 021, 151, and 161).
2. Vehicle maintenance
   - Administration (currently item 041);
   - Revenue-vehicle maintenance (currently items 061, 062, and 071); and
   - Non-revenue-vehicle maintenance (currently items 051, 081, and 091).
3. Nonvehicle maintenance
   - Administration (currently item 042);
   - Track/roadway (currently item 121);
   - Other structures/grounds (currently items 122, 123, 124, and 125); and
   - All other nonvehicle maintenance (currently items 101, 111, 126, and 141).
4. General administration
   - General support functions (currently items 165, 166, 167, 168, 169, 170, 171, 172, 174, 175, 176, and 181); and
   - Planning and public information (currently items 145, 162, 163, 164, 173, and 177).

The committee has also recommended certain changes in the classification of expense items to improve the usefulness and comparability of the data. These include showing purchased transportation expenses as to total amount unassigned to functions; shifting passenger security, ticketing, and fare collection to the operations function; and requiring all expenses to be allocated to modes (i.e., no unallocated "joint expenses"). These recommendations are all under further study.

Object Classes

Finally, recommendations on expense object classes (line items) would reduce the 54 most detailed categories to 25. The least detailed level now includes 21 categories; items 4, 9, 14, and 17 in the following list would be in addition to those current categories:

1. Operators' salaries and wages (currently 501.01);
2. Other salaries and wages (currently 501.02);
3. Fringe benefit costs (currently 502.15);
4. Contract maintenance costs (currently 503.05);
5. Other services (currently 503.01-503.04 and 503.06-503.99);
6. Fuel, including fuel taxes (currently 504.01 and 507.05);
7. Tires and tubes (currently 504.02);
8. Other materials and supplies (currently 504.99);
9. Propulsion power (currently 505.01);
10. Other utilities (currently 505.02);
11. Casualty costs (currently 506.01-506.10);
12. All taxes other than fuel (currently 507.01-507.04, 507.06, and 507.99);
13. Purchased transportation (currently 508.01);
14. Advertising and promotion (currently 509.08);
15. All other miscellaneous expenses (currently 509.01-509.07 and 509.99);
16. Expense transfer reclassifications (currently 510.01 and 510.02);
17. Capitalization of nonoperating costs (currently 510.03); and
18. Reconciling items (to remain the same, 511-516).

This set of proposals was extensively reviewed by transit agencies during the summer and fall of 1984. A consensus was reached on this set of function, object, and revenue categories, and general recommendations were prepared on how these categories would be represented on forms.

Nonfinancial Data Forms

Although the committee developed detailed recommendations on the nonfinancial data forms as well, many issues remained unresolved pending further analysis of specific items, such as maintenance and accident reporting. In every case, however, the same criteria were applied to screening nonfinancial data. The principal recommendations have been to reduce the number of items on the service supplied/consumed forms (406 and 407) and delete or substantially revise other items that have suffered from inconsistent reporting. Definitions of "roadcalls," for instance, are notoriously inconsistent across operators and provide misleading indications of maintenance performance. On the 406/407 forms, capacity miles and all service personnel reporting would be deleted, but other items by time period would be preserved.

Recommendations

As the APTA committee continues to develop recommendations, they will be forwarded to the UMTA committee for consideration. Depending on the outcome of those deliberations, certain recommendations will be made to UMTA, which then must decide how to respond to its advisory committee. Even though there is a long lead-time to change the reporting system substantially (including both internal UMTA review and Office of Management and Budget review), UMTA staff have already taken many actions to improve some of the early recommendations of its committee, including the following:

- Incorporating summary graphics in the FY 1982 National Report,
- Adding guidance on use of the data in FY 1982 Report,
- Revising and reducing the number of reporting forms for added clarity,
- Consolidating reference manuals into a single document,
- Conducting workshops for transit agency personnel on how to fill out the required forms, and
- Formalizing the data validation process for reporting agencies to verify any changes before publishing the annual report.

The positive response to date by UMTA staff augurs well for the future treatment of committee recommendations.

IMPLICATIONS FOR RESEARCHERS

For researchers who have relied only on the published reports, the proposed changes should improve the quality of much of the summary data that currently appear there. For those who have delved into the detailed data tapes or have gone back to the voluntary level source documents, the loss of detail will be more noticeable. The major loss in data would be in the financial items, but few systems reported consistently at the voluntary levels in the past. The primary additional item requested by researchers has been identification of fare revenue by mode. This is now included in UMTA's latest forms revision as a voluntary, optional item. Response to the voluntary item may provide clues to the feasibility of making the item mandatory in the future, but difficulties with multimodal systems, especially those with a high level of multiride pass use, must be resolved first. Researchers might want to pay attention to the practical aspects of collecting such data items because the accuracy of the data is directly tied to the ease of collection or estimation at the source.

Researchers and policy makers alike will benefit from a streamlined national reporting system. Such a system will better ensure consistent, accurate data over time for the variety of research purposes of interest. No matter what recommendations for simplifying the reporting system are eventually implemented, the burden will always be on the individual researcher to make certain that the data are not used blindly. It is the rare number that truly speaks for itself. More frequently, the number only takes on meaning when placed in an appropriate context and it will remain for the researcher to determine what that context should be.

The development of recommendations for substantially reducing the number of data items reported nationally may alarm some researchers. It is essential that the research community develop its own recommendations on Section 15 reform to provide a balance to what may have started out as an overly eager effort to cut back. Researchers must, however, be prepared to defend the necessity of data items they wish to have preserved or added. There is common ground among researchers, managers, and policy makers in wanting to improve the quality of information on U.S. transit systems. Interested researchers should make their voices heard as the various review groups develop recommendations that may affect the national transit data base for years to come.

REFERENCES

1. 49 U.S.C. 1611, Public Law 93-503 (1974), Public Law 97-424 (1983).
2. 49 C.F.R. 630, Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System, June 1977 (four volumes), and Reporting Manual and Sample Forms, July 1984.
3. 49 U.S.C. 1607a (Section 9).
4. G.J. Fielding, M.E. Brenner, and O. de la Rocha. Using Section 15 Data: Adapting and Evaluating the Magnetic Tape Version for Statistical Analysis. In Transportation Research Record 961, TRB, National Research Council, Washington, D.C., 1984, pp. 28-35.
5. G.J. Fielding, T.T. Babitsky, and M.E. Brenner. Performance Evaluation for Fixed Route Transit: the Key to Quick, Efficient and Inexpensive Analysis. Presented at 63rd Annual Meeting of the Transportation Research Board, Washington, D.C., 1984.
6. M. Vaziri and J. Deacon. Peer Comparisons in Transit Performance Evaluation. In Transportation Research Record 961, TRB, National Research Council, Washington, D.C., 1984, pp. 13-21.
7. A.G. Hobeika, C. Kanok-kantapong, and T.K. Tran. A Methodology for Comparative Transit Performance Evaluation Using UMTA Section 15 Data. In Transportation Research Record 961, TRB, National Research Council, Washington, D.C., 1984, pp. 36-43.
8. T. Patton. Transit Performance: Production Cost and Service Utilization Indicators Based on Section 15 Data. Presented at the 63rd Annual Meeting of the Transportation Research Board, Washington, D.C., 1984.
9. Reliability Assessment of Section 15 Data. UMTA, U.S. Department of Transportation, c. May 1984.

Part-Time Public Transit Operators: Experiences and Prospects

KENNETH CHOMITZ, GENEVIEVE GIULIANO, and CHARLES LAVE

ABSTRACT

Most U.S. transit agencies have begun to use part-time operators as a way to reduce operating cost. In this paper, based on five case studies, the cost savings and organizational impacts associated with this change are evaluated. Results indicate that cost savings have been small but significant in situations in which peak service expansion occurred. Where the schedule was static, contract protections for existing operators made it difficult to use part-time labor and hence savings were small or insignificant. It was found that in transit agencies with highly peaked schedules, part-time operators (PTOs) save money for two reasons: they improve schedule efficiency (the ratio of hours-paid to hours-worked) and their wages and fringe benefits are lower than those of full-time operators. In agencies with relatively flat schedules the only savings is from lower wages and fringe benefits, and it is possible that this kind of "two-tier" wage system may be bargained away over time. It was found that, for agencies with flat or static schedules, it may be more effective to concentrate on alternative strategies such as absenteeism control and extraboard staffing, which may be more beneficial and easier to implement. On the organizational side, no unusual costs associated with use of PTOs were identified. PTOs have proven to be as reliable as, or even more so than, full-time operators; they have not created unusual supervisory costs; and there have been relatively few problems between part-time and full-time operators. Instead of creating a permanent force of PTOs, as had been anticipated, most of the PTOs who were hired really wanted full-time work.

During the past decade the U.S. transit industry has made a concerted effort to contain ever-increasing operating deficits and halt the long-term decline in productivity. Faced with the alternatives of cutting service, increasing fares, or reducing service costs, transit agencies have developed and implemented a number of actions to achieve the latter. These actions have frequently focused on improving labor productivity because labor is the largest single component of transit operating costs.

The use of part-time operators (PTOs) has emerged as one of the most widely adopted, yet controversial, methods for improving productivity. Pioneered by a handful of transit agencies during the late 1970s, contract provisions allowing PTOs became nearly universal during the early 1980s. A recent American Public Transit Association (APTA) survey indicates that more than nine out of ten U.S. agencies have secured the right to employ PTOs, and the great majority currently exercise that right.

The purpose of using PTOs is to match operators to service patterns. PTOs can be assigned the short pieces of peak-period work that are extremely costly to operate with full-time operators (FTOs). Management attempts to use PTOs have met with significant opposition from transit unions, however. Unions claim that part-time labor enables management to circumvent hard-won work rules, and they fear that PTOs will replace FTOs and that this will lead to an actual loss of full-time jobs. Such opposition has resulted in strikes at several major transit agencies.

Despite the obvious importance of these issues, little is known about the actual consequences of using PTOs. The fiscal and organizational impact of the use of PTOs is examined in this paper, which is based on a national survey of the use of PTOs and in-depth case studies of five representative transit agencies. These agencies range in size from 100 to 1,100 vehicles and are located in medium to large metropolitan areas. Two of the case study agencies are suburban systems; the others operate both local and downtown-oriented commuter services. Data on scheduling, expenses, and personnel were examined, and extensive interviews with transit managers, department heads, union leaders, and operators were conducted.

WHY PART-TIME OPERATORS?

It is expensive for transit agencies to provide peak-period service because of its inherent inefficiency (1,2). The size of the transit agency (number of operators, vehicles, and garages) is determined by peak service requirements, but these resources remain underutilized during the rest of the day.

Labor is a prime example. It is inherently inefficient to assign peak service to an FTO because the operator is not needed during midday, though he is guaranteed a full day's pay. For instance, an operator responsible for two 3-hr peak assignments would also receive 2 hr of guarantee pay, along with 6 hr of pay for driving. In many cases the operator also receives an additional spread premium payment if there is a long interval between morning sign-on and evening sign-off. Consequently, there is an inefficient excess of pay-hours over driving-hours. PTOs have the potential for reducing this inefficiency because they are paid only for the hours they work; hence the cost of peak service falls. In situations in which the peak-to-base ratio is high and work rules are restrictive, PTOs can offer substantial savings (2).

Transit agencies can use the cost-saving potential of PTOs in three ways. First, PTOs can replace FTOs (perhaps through natural attrition) on existing peak-hour runs, thus reducing the operating deficit. When subsidy constraints are severe, such cost reductions reduce the need for fare increases or service cutbacks.

Second, PTOs can be used to expand peak service. Additional peak service would be prohibitively costly if FTOs were used; it is relatively less expensive to expand with PTOs. Many transit agencies consider peak service their first priority and wish to expand it whenever possible. In addition, some transit agencies have an implied mandate to provide peak service by virtue of their subsidy arrangements. In many cases, earmarked local subsidies are aimed at providing better transit services for commuters.

Third, PTOs may enable transit agencies to reduce unproductive day-base service. Many transit agencies keep excess vehicles in service at midday because there is little additional cost involved: the peak-period operator is guaranteed 8 hr pay, so he might as well be driving even if the service is not needed. Thus PTOs enable transit agencies to tailor service to actual demand.

EXTENT OF USE OF PART-TIME OPERATORS

Given the apparent promise of PTOs, it is not surprising that their use has diffused rapidly throughout the industry. In 1971 Seattle Metro became the first large district to win the right to use PTOs. By 1981 more than half of the APTA member transit agencies had obtained a PTO provision, and by 1983 the right to use part-time labor was almost universal (Table 1) (3,4). However, Seattle remains unique in the proportion of PTOs allowed (100 percent of FTOs). Excluding Seattle Metro, this proportion ranged from 1 to 40 percent in 1983, with an average of 13 percent. Thus, although almost all agencies have the right to use PTOs, the number permitted is generally quite small.

TABLE 1 Extent of Use of Part-Time Operators[a]

	1981 (%)	1983 (%)
Part-time operators allowed (as percentage of survey total)	58	92
Part-time operators allowed (but not hired)	18	13
Number of part-time operators allowed (as percentage of full-time operators)		
Average	NA	13
Range[b]	< 1 to 100	< 1 to 100
	N = 207	N = 182

[a]Compiled from APTA data (3,4).
[b]Includes Seattle Metro. Without Seattle the range is < 1 to 40 percent in 1981 and 1983.

In addition to limiting the number of PTOs, most contracts also restrict the amount and type of work they may do. To preserve the distinction between part-time and full-time operators, total work time is restricted: the limitation ranges from 15 to 40 hr, with an average of 28 hr. Where the limit is 40 hr, there are other restrictions that generally prevent the PTO from actually working 40 hr. PTOs are most commonly restricted to trippers (short pieces of work that cannot be paired together to create a full-time split run). Other assignments include charter, holiday, night, and weekend work.

To protect existing full-time jobs, many contracts (35 percent) require that PTOs be laid off first during any general cutbacks. Some contracts also require that all FTOs be rehired before any PTOs, and others require that the size of the full-time work force be guaranteed. As long as the agency is stable or expanding, these provisions cause no problem. However, if budget problems ever cause a service reduction, the agency must lay off its most productive operators, the part-timers, first.

FISCAL IMPACT OF PART-TIME OPERATORS

The use of part-time labor can reduce costs in two ways. First, substitution of part-time for full-time operators increases schedule efficiency (the ratio of pay-hours to vehicle-hours) by reducing guarantee and spread premium payments. Second, PTOs generally receive an effectively lower wage rate, and a significantly lower fringe benefit rate, than do FTOs. Against these savings must be balanced any cost-increasing bargaining concessions, such as increased wages or fringe benefits for FTOs, necessary to secure union acceptance of part-time labor provisions. These three issues are examined in turn.

Schedule Efficiency

Chomitz and Lave estimated that using PTOs could reduce operator cost by as much as 13 percent, depending on the service profile, spread limit and premium provisions, and the percentage of PTOs allowed (1). Given most "typical" work rule restrictions (spread limit of 13 hr; premium pay after 12 hr), estimated savings are 1 to 4 percent, depending on the particular peak-to-base ratio, if PTOs are limited to 10 percent of FTOs and 2 to 6 percent if PTOs are limited to 20 percent. Because operator compensation comprises about half of total costs, this translates into total cost savings of 0.5 to 3 percent.

The Chomitz and Lave estimates were based on experimental run cuts, using five actual transit schedules. The schedules were recut using the RUCUS automated run-cutting procedures, and savings estimates were based on the change in scheduled pay-hours resulting from using PTOs. The estimates assume everything else is held constant: the schedule remains unchanged; and no significant concessions, in the form of more expensive work rules, more fringe benefits, or wage increases, are given in return for the right to use PTOs.

How does actual experience compare with those experimental run cuts? It would be easy to measure the financial impact of PTOs if an agency's service schedule remained unchanged: Just calculate schedule efficiency before and after the introduction of PTOs. Unfortunately for the analyst, schedules do change--and to some extent they change as a direct consequence of the decision to employ PTOs. In many cases the motivation for adopting PTOs was a desire to expand peak service. In other cases, because the contract guarantees the jobs of existing FTOs, the agency must expand service in order to provide openings for the PTOs.

If an agency simultaneously introduces PTOs and alters the service schedule, it becomes difficult to even define, let alone measure, the savings from PTO implementation. Consider, for instance, a hypothetical transit agency that hires PTOs and changes to a more peaked service schedule at the same time. The data in Table 2 indicate that there are four possible combinations for work rules and schedules whose costs can be compared. Which comparison yields the

TABLE 2 Hypothetical Operating Cost of Schedule and Work Rule Combinations

	Service Schedule	
Work Rules	Old (flat) ($)	New (peaky) ($)
Old: pre-PTOs	10	20
New: post-PTOs	8	10

"true" value of cost savings? If the old rules and old schedule combination is compared with the new rules and new schedule combination, there is no change in costs. All the potential savings from PTOs have been spent on expanded service. To evaluate changes in labor productivity, the service schedule must be held constant (i.e., costs examined within a single column). But the left-hand column indicates a 20 percent saving, from $10 to $8, whereas the right-hand column shows a 50 percent saving, from $20 down to $10. Alternatively stated, under the old service schedule, use of PTOs could save 20 percent; under the new service schedule, if management were to give up use of PTOs, operating costs would double. (Note that columnwise comparisons require experimental run cuts, because the old/new and new/old combinations were never actually put on the street.) This is not an unreasonable example, nor is it a semantic game. The savings from part-time labor can be discussed only in the context of a given service schedule.

Case Study Results

For four of the five case study agencies, expansion of peak service was the primary motivation for using PTOs; it would have been prohibitively expensive to undertake the new service without PTOs. The fifth agency planned to expand base service. Only two of the case study agencies actually sustained the increased service. Financial problems, brought about by loss of subsidy money or by insufficient fare-box revenue, resulted in service cutbacks at the other three. The contracts at these three agencies specify that PTOs must be laid off first. Consequently, two agencies lost all their PTOs, but the third was able to keep some part-time positions through an aggressive early retirement program for FTOs.

In all cases the use of PTOs was one of several productivity improvement strategies being pursued by management during the period of this study. Other strategies included absentee reduction programs; changes in wage scales, cost-of-living adjustments, and vacation pay; reductions in extraboard staffing; work rule changes; and a host of minor policy changes. To isolate the impact of PTOs, it is necessary to take all of these other changes into account as well. Accordingly, a disaggregate approach was chosen. Cost impacts can be broken down into savings on scheduled costs and on fringe benefits. Each will be discussed in turn.

Impact on Schedule Efficiency

Schedule efficiency is expressed here as the ratio of pay-hours to platform-hours. This ratio is always greater than one because operators are paid for report and travel time. The minimum possible (e.g., if no make-up, premium, or overtime were paid) is about 1.04. There are two possible ways to estimate the impact of PTOs on schedule efficiency. One is to use actual "before" and "after" schedule data and attempt to control for service and other changes. Referring to Table 2, this is equivalent to moving diagonally from "Old/Old" to "New/New" while trying to estimate "Old/New." A second method is to use experimental run-cut data: take the new schedule, perform a run cut under the old rules, and compare the results. This gives the needed within-column comparison. However, because the new schedule would never have been adopted under the old rules, it could be argued that such a comparison may not be appropriate.

Both methods were used in this case study analysis, and the results are summarized in Table 3. Actual "before" and "after" schedule data were available from two agencies. Agency A increased service by about 40 percent, and the peak-to-base ratio increased from 2.25 to 2.65. Such increased peaking should have substantially reduced the schedule efficiency at this transit agency. However, the simultaneous adoption of PTOs, with their inherently higher productivity, overcame the negative effects of increased peaking. The overall result was a 5 percent improvement in the pay-to-platform-hour ratio, compared to the preexisting service.

Agency B reduced peak service and increased base service after hiring PTOs. The result was a decrease in the pay-to-platform-hour ratio of 2 percent. Assuming two part-time runs are equivalent to one split run and using the change in the ratio of straight to split runs resulting from the change in peak-to-base, it was estimated that about half of the pay-to-platform-hour reduction is due to PTOs.

Because of anticipated service cutbacks, two agencies had performed experimental run cuts to determine the impact of losing their PTOs. Agency C had a peak-to-base ratio of 3.5, and they were allowed 15 percent PTOs. Their run-cut simulation of the effects of losing PTOs showed a 2.9 percent decrease in schedule efficiency. This is substantially less than would have been expected for an agency with such a high peak-to-base ratio (1). The reason for the small change is that Agency C already had an exceptionally efficient schedule (1.19 pay-to-platform-hour ratio) because of pay calculation provisions that were quite favorable to management; thus PTOs could not make as much difference as might have been anticipated from their peaky schedule. Agency D had a peak-to-base ratio of 2.0, and 10 percent of its operators were PTOs. An experimental run cut was performed to see what would happen if the agency were to retain its existing service schedule while changing the labor force from 10 per-

TABLE 3 Impact of Part-Time Operators on Schedule Efficiency

Observed Data				Experimental Data			
	Service	Peak/Base	Change in Pay/ Platform-Hour (%)		Percentage Part Time	Peak/ Base	Change in Pay/ Platform-Hour (%)
Agency A	Increased	Peaky, increased	-5.0 -1.7	Agency C	15	3.5	-2.9
Agency B	Stable	Flat, decreased	-2.0 total -1.0 due to PTOs	Agency D	10	2.0	-2.5

cent PTOs to no PTOs. The results showed a 2.6 percent decrease in schedule efficiency.

Wages and Benefits of Part-Time Operators

The rationale behind use of PTOs is their ability to increase schedule efficiency. In practice, most transit agencies gain additional savings by paying PTOs lower fringe benefits and (effectively) lower wages as well. Table 4 gives the data for fringe benefits: most agencies offer no sick, holiday, or vacation pay, no health insurance, and no retirement pay to PTOs. In cases in which these benefits are provided, they are most frequently provided at reduced rates.

TABLE 4 Part-Time Operator Fringe Benefits[a]

	Same as Full Time (%)	Reduced (%)	None (%)
Sick leave	10	13	77
Holiday pay	12	17	71
Vacation	13	23	64
Health insurance	15	17	68
Retirement	21	7	72
N = 112			

[a]Compiled from APTA data *(4)*.

Additional wage cost savings are generated by the wage rate progression. In most cases PTOs have the same pay scale and progression as FTOs. Because the wage progression is based on work hours, it requires more calendar time for a PTO to reach the top of the scale. For example, the 1983 U.S. average number of work hours to reach top rate was 4,938. In full-time equivalents (e.g., 2,080 work hours per year) this is about 2.4 years, but in part-time equivalents (national average of 28 hr per week = 1,456 hr per year) this is about 3.4 years. If the tenure of PTOs is 2 or 3 years (the case study data indicate less than 2 years), most PTOs never reach the top of the pay scale.

An example from an experimental run cut can be used to show the relative contribution of each of these factors to PTO cost savings. On the basis of data from the Agency D run cuts, a 10 percent complement of PTOs reduces pay hours by 2.6 percent. The lower wage rate of PTOs contributes an additional 2 percent reduction, raising the estimated savings from 2.6 to 4.6 percent. The savings from reduced fringe benefits brings the total reduction to 5.7 percent compared to an all-full-time operator schedule. This translates to 2.5 percent of total operating cost.

Cost of Part-Time Operators

A central issue in this research is the cost (in general terms) of winning the right to use PTOs. In view of staunch union opposition, it was anticipated that management would have to give up something in return for PTOs. Only one of the case study agencies specifically identified a bargained cost: an extremely small wage rate increase. In all other cases management had identified a set of possible bargaining issues, and the cost of PTOs was an "opportunity cost": other alternative means of reducing costs were not pursued, and attention was concentrated on gaining PTOs. In at least one case detailed analyses of the relative merits of these alternatives were conducted; in other cases the choice was largely subjective. In any event, the actual outcome of using PTOs could not be accurately forecast by management because of the complexity of schedule characteristics and work rule provisions that affect PTO utilization and because of all the other changes affecting labor productivity that were implemented over the same time period.

An additional analysis of the impact of hiring PTOs on compensation rates is currently being conducted. Using data from 50 U.S. transit agencies, FTO compensation is being predicted as a function of environmental and service characteristics. By comparing predicted versus actual values, it can be determined whether agencies that obtained the right to use PTOs gave greater compensation increases. Prelimiary results show that an initial increase in fringe benefits did occur, but under later contracts benefits returned to normal levels. A similar effect on wage rates was also found, though it was not quite statistically significant.

CONSTRAINTS ON USE OF PTOs

PTO cost-savings estimates are based on the schedule not the actual operating costs of the case study agencies. To the extent that constraints on the use of PTOs come into play, these savings may not be realized. The case studies revealed that the transit agency's ability to use PTOs can be significantly constrained by a variety of contract limitations. These limitations, together with the characteristics of the service schedule, can make it impossible for an agency to use the full number of PTOs the contract allows. For example, PTOs are often restricted to runs that begin and end at a bus depot, whereas FTOs can be relieved "on the road" without taking the bus out of service. In addition, there is almost always a maximum allowable time limit for part-time runs, and there is sometimes a minimum. Pieces of work smaller than the minimum are reserved as biddable overtime for FTOs. Another common provision is that the number (or proportion) of PTOs must be the same at each division. Because the service profile usually differs among divisions, this provision limits the total proportion of operators to the number that can be used in the division with the least peaky schedule. Two case study agencies also have a provision that prohibits the splitting up of two-piece runs in order to create part-time work. In practice, this provision is unenforceable because it is almost impossible to retain the identity of specific runs over several run cuts.

Contract provisions like these tend to reduce the potential efficiency gains from use of PTOs. Transit managers who have been involved in the contract bargaining process acknowledged that the full impact of such provisions is difficult to anticipate, hence savings from using PTOs are often much lower than had been anticipated. When subsequent contracts come up for renewal, these unanticipated restrictions become focal points for bargaining.

ORGANIZATIONAL AND OPERATIONAL ISSUES

When the part-time labor issue was first raised, unions predicted a number of serious negative consequences. It was feared that PTOs would be uncommitted and unprofessional, resulting in higher accident rates, absenteeism, turnover, and passenger complaints. To a large extent these fears have been unfounded. The unions were also concerned about the impact of PTOs on the job security and overtime opportunities of incumbent FTOs. There was concern about the working conditions of the part-timers and apprehension about how well they could be integrated into the union. On these counts the record is mixed.

Impacts on Full-Time Operators

The right to use PTOs was not easily achieved by most agencies. FTOs perceived the issue as a threat to both their jobs and their working conditions. Whereas transit management has every incentive to replace FTOs with PTOs, union opposition is natural. A nearly universal compromise is to protect the jobs of incumbent FTOs. This is accomplished by requiring that no FTO be laid off until all PTOs have been dismissed or by guaranteeing a minimum number of FTOs or runs. At such agencies the only way to implement PTOs is to add new service.

A second effect on FTOs is a reduction in the types of runs available to them. Because management assigns the most costly runs (those with a high ratio of pay-hours to work-hours) to PTOs, there will be fewer runs with premium pay and overtime available to FTOs. FTOs can lose the chance to earn such pay in two ways. First, at some agencies, FTOs can volunteer to drive trippers in addition to their assigned runs. In the absence of a contractual agreement to the contrary, such trippers will generally be reassigned to PTOs. Second, FTOs can lose the chance to earn premium pay when split runs are reassigned to PTOs. (At agencies with a high peak-to-base ratio, split runs pay more than straight runs and can involve less actual driving time.) If there is a contractual minimum number of FTOs, however, the only way to reassign a split run is to add new base service, creating a new straight run for the displaced FTO.

This strategy was used at one of the agencies studied. Both peak and base service were increased; PTOs were assigned all the peak service and some of the old split runs; and the former holders of those split runs were assigned to new straight runs. Thus the number of full-time runs has remained constant, but their composition has been altered significantly: the proportion of split runs fell from 41 to 31 percent. Because the split runs had paid about 15 percent more than straight runs, this is a significant loss to the FTOs. On the other hand, the work available to FTOs is now more pleasant (e.g., a higher proportion of straight runs is available). Premium and overtime payments were originally begun as extra compensation for onerous assignments. Thus the loss of premium pay is now offset by better work.

FTOs may also end up with a less desirable selection of weekly schedules. The proportion of FTOs who can have the weekend off depends on the ratio of weekday runs to weekend runs. If PTOs, who are generally restricted to weekday peak service, supplant some FTOs, the remaining full-timers will face a lower probability of securing weekends off. On the other hand, it is hard to say how important this is to FTOs. In most instances in which management has asked to use PTOs on weekend runs, thus giving FTOS a regular weekday schedule, the unions have been adamantly opposed.

Status and Performance of Part-Time Operators

Three issues related to PTOs were explored during the case study visits: PTOs' perception of the job, relationships between part-time and full-time operators, and job performance of PTOs.

Both transit management and union members expected that those people who applied for PTO positions (e.g., college students, mothers of younger children) would be interested in permanent part-time work, but it has turned out that most PTOs are seeking full-time work. As part-time work is now scheduled it is not surprising that few PTOs are permanent. Work hours are inconvenient for those who need child care, and work schedules change too frequently for students or people working other jobs.

Both union and management officials estimate that 60 to 80 percent of PTOs would really prefer full-time work. All of the case study transit agencies have contract provisions that give preference to PTOs when full-time jobs become available. In some cases a majority of PTOs move on to full-time positions, and most full-time positions are filled this way. Thus the part-time position has become a stepping-stone to full-time employment. It should be noted that the case studies took place during a period of economic recession and high unemployment. Given that most PTOs would prefer full-time work, it remains to be seen whether part-time recruitment will become more difficult as the economy improves.

Transit managers cited several indirect benefits of the part-time to full-time progression. In effect, the part-time position becomes a longer probation period, and managers and supervisors have more opportunity to evaluate operators before they are hired in full-time positions and thus managers believe they can make better choices. In addition, an already experienced operator is hired, lessening the need for training.

A major concern in bringing PTOs into the transit agency has been whether they would be accepted by full-time operators and whether a good working relationship between part-time and full-time people could be established. In spite of the initial opposition to part-time contract provisions, no hostility appears to have carried over to PTOs themselves. Discussions with operators indicated that PTOs are not treated differently than FTOs. Some part-time people thought that the union did not seem committed to them, but no specific problems were identified.

When the right to use PTOs was won, unions claimed that qualified people willing to take part-time jobs would be difficult to find and that safety problems and customer complaints would consequently increase, but the performance record of PTOs has been at least as good as that of FTOs. At the case study agencies no evidence could be found to indicate that PTOs behave any differently than do FTOs on the job. As one supervisor put it, "Once they have the uniform on, there's no way to distinguish a PTO from a full-time operator."

Absenteeism

Absenteeism is one aspect of job performance that is of great concern to transit management. Table 5 gives comparative sick rates for PTOs and FTOs for the five case study agencies. The rates are based on approximately 1 year of data at each agency, and they are computed as the percentage of workdays per year when an operator calls in sick. The FTO sick

TABLE 5 Comparative Sick Rates for Part-Time and Full-Time Operators (%)

	Agency				
	A	B	C[a]	D	E
FTO sick rate[b]	3.75	3.52	2.31	4.29	3.06
PTO sick rate	1.41	1.71	1.02	1.59	1.60

[a] Agency C is an unreliably small sample.
[b] Proportion of yearly work days an operator will call in sick.

rate exceeds the PTO rate at every agency and on average is 2.3 times higher.

Why do PTOs have lower sick rates? The most obvious explanation is that PTOs do not receive sick benefits. However, at two of the agencies it is possible to compare groups of drivers with identical sick benefits. Table 6 gives a comparison at Agency E where PTOs receive no sick benefits and FTOs receive no sick benefits during their first year of employment. The rates are expressed as the percentage of workdays the operators call in sick. In small samples like this, the presence of a few random instances of major illness can substantially bias the apparent rate. Hence, Rate 1 excludes any operator who was sick more than 40 days (8 weeks), and Rate 2 sets a tougher standard by excluding any operator who was sick for more than 6 weeks. (Neither Rate 1 nor Rate 2 screening ever exclude more than 10 percent of the sample.) Because the FTOs are on probation for much of this period, their sick rates should be biased downward. Despite this bias, the data in Table 6 indicate that the PTO sick rate is lower.

TABLE 6 Comparison of Part-Time and Full-Time Operators when Neither Receive Sick Pay (percentage[a])

	Rate 1 (no sicks >40 days)	Rate 2 (no sicks >30 days)
Full time with no sick pay		
Hired in 1982, 1983 data, 18 operators	3.56	3.03
Hired in 1983, 1984 data, 18 operators	3.27	2.39
Part time with no sick pay		
Hired in 1982, 1983 data, 18 operators	1.67	1.67
Hired in 1983, 1983 data, 41 operators	1.64	1.64
Hired in 1983, 1984 data, 23 operators	1.52	1.52
Hired in 1984, 1984 data, 33 operators	1.58	1.58

[a]Proportion of yearly workdays.

Agency B has a class of PTOs who receive the same sick benefits as their FTOs. Using the Rate 2 definition, it was found that the absence rate was 3.52 percent for FTOs and 2.44 percent for PTOs. It was, therefore, concluded that PTOs have less absenteeism than do FTOs and that this effect is even true in those instances where both groups of operators receive identical sick benefits.

Accident Rates

The analysis in this section is still in progress and the results should be regarded as tentative. Table 7 gives comparative accident rates, PTO versus FTO, as a function of amount of experience for one of the case study agencies. The data show that the number of accidents per year declines with experience and that PTO accident rates are lower than those of FTOs. The table also breaks down the accidents into "chargeable" (i.e., the driver could have prevented the accident) and "nonchargeable."

TABLE 7 Comparative Accident Rates for Part-Time and Full-Time Operators[a]

	FTO	FTO	FTO	FTO	PTO	PTO
Years of experience	3.7	2.6	2.3	1.3	1.3	0.6
Accident rate						
Total	1.33	1.50	1.17	1.59	1.17	0.95
Chargeable	0.49	0.27	0.34	0.59	0.58	0.38
Nonchargeable	0.84	1.23	0.83	1.00	0.59	0.57
Sample size	9	28	18	18	23	33

[a]Accidents per year; total of all vehicle and passenger incidents. Rates are not standardized for driving exposure.

Table 8 gives comparative accident rates at a different agency, and this time the data are structured by the type of work assignment. The PTO accident rate is higher than that of FTOs who do regular runs but lower than or equal to that of regular drivers who do relief runs or extraboard work. The accidents are also broken down as preventable and nonpreventable. The PTOs are judged to have a higher proportion of preventable accidents. This might be an indication that PTOs are worse drivers or that the drivers who evaluate the accidents are biased against PTOs and thus more likely to decide that PTOs were at fault.

TABLE 8 Comparative Accident Rates, Agency D[a]

	Regular Run	Regular Relief	Extra-board	Vacation Relief	Part-Time Run
Accidents per year	0.68	2.35	2.20	1.38	1.39

[a]Potential "reporting" bias against PTOs as percentage of total accidents judged "preventable": extraboard, 45; regular drivers, 51; and PTOs, 60.

Accidents per year is not a wholly adequate statistic for judging the quality of the two driver groups because it does not take into account other factors that may affect full-time and part-time accident rates. FTOs do more driving and thus might be expected to have more accidents. On the other hand, FTOs also have more experience, and experience should lower the rate. PTOs do more driving in congested conditions in which accidents are more likely to occur. Moreover, there may be substantial differences in the drivability of the vehicles (e.g., size, age) used by the two groups. Ideally, the accident rates should be standardized for all of these different exposure factors.

In their study of accident rates at the Massachusetts Bay Transportation Authority (MBTA) in Boston, Attanucci, Wilson, and Vozzolo (5) reported that standardized PTO accident rates were clearly higher during the PTO introductory period but appear to converge with the FTO rates thereafter. Given the unusual nature of the MBTA data and the tentatively positive evidence of the data in Tables 7 and 8, evidence on accident rates appears to be quite mixed. The tentative conclusion is that FTO and PTO accident rates are roughly similar.

CONCLUSIONS

The results of this research indicate that PTOs can be used to expand peak service economically. However, few transit agencies today are in a financial position that permits service expansion. Whether part-time labor can be used to reduce the cost of a static service schedule depends a great deal on contractual restrictions. Many apparently minor restrictions can prevent full or efficient use of the nominal quota of PTOs. Above all, the efficiency gains from part-time labor depend on the existing ratio of pay-hours to platform-hours; where this ratio is high (greater than 1.15), there are significant opportunities to increase productivity. The ratio itself depends on both schedule peaking and work rules: an agency with generous work rules may have a low pay-hour-to-platform-hour ratio despite substantial peaking. Additional savings have been realized at many properties by restricting fringe benefits of PTOs. However, there is some evidence that unions are successfully pressing for increased PTO benefits: sick and vacation benefits have recently been granted to PTOs at two of the case study agencies.

Observed changes in schedule efficiency are consistent with the experimental run-cut predictions of Chomitz and Lave (1). Because the indirect effects of PTOs are negligible, experimental run cuts are an effective tool for exploring the potential cost effects of changes in labor provisions. A new generation of run-cutting software makes such experimentation feasible for transit districts with adequate computational resources.

There is no appreciable indirect effect of PTOs on absenteeism, supervision, hiring, or training costs. It may, however, become more expensive and difficult to recruit part-time workers as the burgeoning economy provides more alternatives for full-time work. There is no definitive evaluation of accident effects as yet.

Finally, there appears to be some opportunity to make better use of part-time drivers. First, the limitation of PTOs to weekday work seems to be unnecessarily restrictive. If PTOs were permitted to work weekends, FTOs would have proportionately more weekends off. In addition, PTOs would have the opportunity to work more hours, which would make the job more attractive. The weekend schedule might also be better suited to permanent part-time work. Two of the case study agencies have recently allowed a limited amount of part-time weekend work.

Second, the option to work part time on a temporary basis might be given to FTOs. Two of the case study agencies have such a provision. In one case, the distinction between full and part time was replaced with a two-class system. Class I operators can have up to 4 percent part-time positions. Class II operators are all part time and limited to 6 percent of the full-time force. Class I operators, whether full or part time, receive the same fringe benefits. The assignment of part-time work thus depends on the seniority roster. This system allows FTOs the option of choosing part-time work without loosing seniority or benefits. Many FTOs took advantage of the opportunity and chose part-time runs at the summer shake-up. At the same time, low seniority PTOs were able to work full time. Both of these alternatives provide benefits to full-time as well as part-time operators.

Use of part-time labor has not been a panacea. Although it has permitted some agencies to expand peak service or increase efficiency, it has made relatively little difference at others. Alternative strategies for increasing labor productivity may be more beneficial and easier to implement; two areas with particular promise are absenteeism control and extraboard staffing. Operator absenteeism is costly because a corps of standby operators (the extraboard) must be maintained to cover the absent operators' assigned work. Moreover, the costs of reducing absenteeism may be relatively low: improved record keeping coupled with increased supervision and counseling. The political costs of implementing an absentee control program may also be relatively small. Because a small number of operators account for a disproportionate number of absences, the majority of operators may be sympathetic to a more equitable enforcement of absence rules. Two of the case study agencies implemented absenteeism control programs with relative ease, decreasing absenteeism by 2 to 5 percent, which is more than they had saved by implementing part-time labor. The use of PTOs is thus just one of many possible strategies for increasing labor productivity.

ACKNOWLEDGMENT

This paper is based on work conducted for the Urban Mass Transportation Administration, Office of Service Methods and Demonstrations, under grant CA-06-0187, "Fiscal and Organizational Impacts of Part-Time Drivers in Public Transit." The views expressed herein are those of the authors.

REFERENCES

1. K.M. Chomitz and C.A. Lave. Part-Time Labor, Work Rules, and Transit Costs, Final Report. Report CA-11-0018-1. UMTA, U.S. Department of Transportation, 1981.
2. R.L. Oram. Peak Period Supplements: The Contemporary Economics of Urban Bus Transportation in the U.K. and the U.S.A. Progress in Planning, Vol. 12, No. 2, 1978, pp. 81-154.
3. Comparative Labor Practices: Report 7, Part-Time Operators. American Public Transit Association, Washington, D.C., July 1981.
4. Comparative Labor Practices: Report 4, Part-Time Operators. American Public Transit Association, Washington, D.C., April 1983.
5. J. Attanucci, N.H.M. Wilson, and D. Vozzolo. An Assessment of the Use of Part-Time Operators at the Massachusetts Bay Transportation Authority. In Transportation Research Record 961, TRB, National Research Council, Washington, D.C., 1984, pp. 21-28.

Review of the Use of Part-Time Transit Operators and Methods for Assigning Part-Time Work

MARTIN J. WELLS, BRIAN McCOLLOM, and THOMAS DOOLEY

ABSTRACT

The use of part-time transit operators is a subject of increasing attention as a means of controlling labor costs and improving transit productivity. Part-time operators can significantly reduce the cost of providing peak-period service because they are subject to less restrictive work rules than are their full-time counterparts: they typically receive no spread or overtime premiums, they almost always receive lower fringe benefits, and they may earn lower wages. Three of four labor contracts permit the use of part-time operators, and one of every twenty operators nationwide is a part-timer. A national perspective on the range and norms of contractual provisions affecting the use of part-time operators is offered. The methodologies used by three transit agencies to assign part-time operators on the basis of existing run cuts, in accordance with the different work rules that govern the use of part-timers at each agency, are presented. The methodologies used by two systems to incorporate part-time operators into automated run-cutting procedures are also presented.

Productivity in the transit industry has become a subject of increasing attention as capital and operating costs have risen and fare-box recovery ratios have fallen in recent years. Transportation wages and fringe benefits account for nearly half of total operating costs. Transportation salaries and wages accounted for 32 percent, and fringe benefits another 13 percent, of total 1980 transit operating expenses according to the American Public Transit Association (1). It is logical, therefore, to focus on controlling labor costs in the effort to improve transit productivity.

Operator labor costs are significantly affected by the work rule provisions that are a fundamental part of all operator-management contracts (2). These work rules were formulated in response to the peaked nature of transit demand. Approximately two-thirds of all daily transit passengers are carried during the morning and late afternoon commuter peak periods. Less than half this number of passengers is carried in the early morning, midday, and late evening periods.

The numbers of vehicles (including spares) and operators (including absence and vacation extras) are determined by peak-period passenger demand. Twice as many operators are needed in the two peak periods as in the base period. The additional operators can be provided in three ways: (a) by assigning regular operators to split runs that include both a morning and an afternoon shift and a break in between; (b) by assigning extraboard operators to short tripper assignments; or (c) by working short trippers on an overtime basis. Each approach can be costly, involving spread premiums, unproductive guarantee pay, or overtime pay.

The use of part-time operators (PTOs) can significantly reduce the cost of providing peak-period service, thereby improving labor productivity, for the following reasons:

1. PTOs are subject to less restrictive work rules than are their full-time operator (FTO) counterparts. In nearly eight of every ten transit systems, PTOs receive no guarantee pay per assignment. The median guarantee at transit systems that have one is only 2 hr per assignment, compared to a guarantee of 8 hr for FTOs.
2. PTOs typically receive no spread or overtime premiums. However, they may be subject to a maximum spread time, or effectively restricted to working only single trippers, by daily or weekly work hour limitations.
3. PTOs almost always receive lower fringe benefits than do FTOs. A transit system can save on both fixed and variable fringe benefit costs if a PTO obviates the need to hire an additional FTO.
4. PTOs earn lower wages than FTOs at two of every ten systems permitted to use PTOs.

The use of PTOs has become widespread in systems of all sizes in all regions of the nation. Three of four systems are currently permitted to use PTOs, and one of every twenty operators nationwide is a PTO.

The Urban Mass Transportation Administration (UMTA) of the U.S. Department of Transportation undertook a study in June 1984 to (a) examine the extent to which PTOs are currently used in the U.S. transit industry and (b) identify methods currently used to assign PTOs to work assignments.

The findings of that study are presented in this paper. First, a national perspective on the range and norms of contractual provisions affecting the use of PTOs is presented. Second, current methodologies used to assign PTOs work is presented. Three methods are described for assigning work on the basis of existing run cuts and two methods are presented that are used to incorporate PTOs into automated run-cutting procedures. In the third and final section the major conclusions of the study are summarized.

NATIONAL PERSPECTIVE ON THE EXTENT OF PART-TIME OPERATOR USE

An analysis of the Comparative Labor Practices Reports 3 (Number of Employees by Type) and 5 (Part-

Time Operators) compiled by the American Public Transit Association (APTA) and telephone interviews and site visits conducted as part of this study indicate that the use of PTOs is widespread, in systems of all sizes in all regions of the nation. Three of four labor contracts permit the use of PTOs, and one of every ten operators at these systems is a PTO. The typical PTO (a) is a union member, (b) is paid at or near the FTO wage scale, (c) is permitted to work a maximum of 25 to 30 hr per week, (d) receives no guarantee per assignment or premium pay, and (e) receives reduced fringe benefits compared to FTOs. PTO seniority is generally not transferable to FTO status. About half of the sample agencies require that all PTOs be laid off before any FTOs are laid off.

These findings were developed through a statistical analysis of the data assembled for 228 transit agencies in the United States. The transit systems in the sample ranged in size from agencies with fewer than 10 employees to the largest system in the country, the New York City Transit Authority, that has more than 10,000 employees. Summary numbers were computed for the major labor practices regarding PTO use. Chi-square tests were made to determine if there were differences in practices by system size or geographic region. A value of 95 percent (i.e., probability of chance = 0.05) was used as the confidence level in these tests.

System size was frequently found to be a significant variable in the chi-square tests of PTO labor practices. Labor practices in small systems (50 or fewer employees) are less restrictive regarding the use of PTOs than are those in large systems (more than 50 employees). This suggests that the labor-management climates in small and large systems may be different with respect to PTOs.

Use of PTOs

One hundred seventy-six or 77 percent of the transit agencies in the sample are permitted to use PTOs. The ability to use PTOs is statistically related to the size (measured in terms of number of FTOs employed) of the system. As the data in Tables 1 and 2 indicate, a higher percentage of systems with fewer than 50 employees than of systems with more than 50 employees is permitted to use PTOs. This may suggest that FTOs in small transit systems feel less threatened by PTOs than do their counterparts in large systems.

A statistical relationship was also found between the ability to use PTOs and geographic location. The transit systems in the APTA reports were coded by UMTA federal region and grouped into five geographic areas (see Table 2). Between 85 and 95 percent of the transit systems in the west (i.e., midwest, southwest, and west) are permitted to use PTOs compared to 60 percent of the systems in the east (i.e., northeast and southeast). This regional relationship may reflect historic differences in the relative power of management and labor in contract negotiations.

Number of PTOs

The 176 systems permitted to use PTOs employ a total of 42,173 FTOs and 4,402 PTOs. This number of PTOs is equivalent to 10.44 percent of the number of FTOs, or 9.45 percent of the total operator work force. Therefore, nearly one of every ten operators at systems permitted to use PTOs is a PTO.

The number of PTOs that can be hired is often limited in the labor contract. Almost half of the

TABLE 1 Use of Part-Time Operators by System Size

Number of Employees	Systems in Sample: Number	Systems in Sample: Percentage	Systems Permitted PTOs: Number	Systems Permitted PTOs: Percentage of System Employee Class	Systems Not Permitted PTOs: Number	Systems Not Permitted PTOs: Percentage of Employee Class
0-25	51	22.4	47	92.2	4	7.8
26-50	42	18.4	34	81.0	8	19.0
51-100	44	19.3	30	68.2	14	31.8
101-250	37	16.2	29	78.4	8	21.6
251-500	18	7.9	12	66.7	6	33.3
501-1,000	14	6.1	10	71.4	4	28.6
More than 1,000	22	9.6	14	63.6	8	36.4
Total	228	100.0	176	77.2	52	22.8

Note: Chi-square = 12.59, degrees of freedom = 6, and probability of chance = 0.05.

TABLE 2 Use of Part-Time Operators by Geographic Location

Location	UMTA Region(s)	Systems in Sample: Number	Systems in Sample: Percentage	Systems Permitted PTOs: Number	Systems Permitted PTOs: Percentage of Systems in Region	Systems not Permitted PTOs: Number	Systems not Permitted PTOs: Percentage of Systems in Region
Northeast	1, 2, 3	59	25.9	35	59.3	24	40.7
Southeast	4	32	14.0	19	59.4	13	40.6
Midwest	5, 7	62	27.2	53	85.5	9	14.5
Southwest	6	21	9.2	19	90.5	2	9.5
West	8, 9, 10	54	23.7	50	92.6	4	7.4
Total		228	100.0	176	77.2	52	22.8

Note: Chi-square = 28.27, degrees of freedom = 4, and probability of chance = 0.00.

contracts that permit the use of PTOs have an expressed provision that limits the number that can be employed. As the data given in Table 3 indicate, the most common limitation is to express the maximum number of PTOs as a percentage of the number of FTOs. Sixty-four systems have this provision with percentages ranging from 5 to 100 percent. In the remaining 19 contracts, an actual number is specified in 12 contracts and the maximum number is based on the number of scheduled runs, biddable runs, unsigned trippers, or peak-hour trippers in the contracts.

The size of the transit system was found to be statistically related to the type, if any, of limitation on the number of PTOs. In small systems (i.e., 50 employees or fewer) more than 86 percent of the contracts have no provision that limits the number of PTOs who can be hired (see Table 3). In contrast, 74 percent of the contracts in large systems (i.e., more than 50 employees) contain some type of limiting provision. This difference further supports the inference that FTOs in small systems feel less threatened by PTOs or have had less power in labor negotiations than FTOs in large systems, or both.

PTO Wages

PTOs are typically paid the same wage that is paid to FTOs. Nearly eight of every ten PTOs (79 percent) earn the same wages as FTOs. Six percent earn 50 to 74 percent of FTO wages and 14 percent receive between 75 and 99 percent of FTO wages. Only one system, the San Francisco Bay Area Rapid Transit District (BART), was found to pay a higher wage rate to PTOs. PTOs at BART receive 110 percent of the FTO wage rate. No statistical relationships were found between wage rates and system size.

PTO Fringe Benefits

Although their basic wages are generally the same, PTOs usually receive fewer fringe benefits than FTOs. At nearly nine of every ten transit agencies, PTOs are union members and are therefore represented by grievance and arbitration procedures. However, PTOs are often treated differently in terms of seniority and layoff procedures. Seniority as a PTO is transferable to full-time status at only about three of every ten systems (see Table 4). Almost one-half of the systems require that all PTOs be laid off before any FTOs are laid off. A system size relationship was again found in which a greater percentage of small systems tend to allow the transfer of PTO seniority and do not require PTOs to be laid off first.

Most transit systems do not grant sick leave, holiday, vacation, health and welfare insurance, or retirement benefits to PTOs. One-fourth of all sys-

TABLE 3 Contract Limitations on Number of Part-Time Operators

Limiting Provision	Systems Permitting PTOs: Number	Systems Permitting PTOs: Percentage	System Size: 50 Employees or Fewer: Number	50 Employees or Fewer: Percentage of Systems with 50 or Fewer Employees	More than 50 Employees: Number	More than 50 Employees: Percentage of Systems with More than 50 Employees
Percentage of FTOs	62	36.0	8	10.4	54	56.8
Other basis[a]	19	11.0	3	3.9	25	16.8
No provision	91	52.9	66	85.7	16	26.3
Total	172[b]	100.0	77	100.0	95	100.0

Note: Chi-square ≠60.27, degrees of freedom = 2, and probability of chance = 0.00.

[a] Includes the specification of an actual number and percentages based on number of scheduled runs, biddable runs, unsigned trippers, or peak-hour trippers.

[b] Four additional systems employ only PTOs and were not included in this analysis.

TABLE 4 Seniority and Layoff Procedures

	Systems Permitted PTOs: Number	Systems Permitted PTOs: Percentage	System Size: 50 Employees or Fewer: Number	50 Employees or Fewer: Percentage	More than 50 Employees: Number	More than 50 Employees: Percentage
Seniority[a]						
PTO seniority transferable to full-time status						
Yes	45	28.8	33	51.6	12	13.0
No	111	71.2	31	48.4	80	87.0
Total	156	100.0	64	100.0	92	100.0
Layoff Procedures[b]						
PTOs laid off first?						
Yes	75	48.1	18	28.1	57	62.0
No	81	51.9	46	71.9	35	38.0
Total	156	100.0	64	100.0	92	100.0

[a] Chi-square (corrected) = 25.43, degree of freedom = 1, and probability of chance = 0.00.

[b] Chi-square (corrected) = 15.97, degree of freedom = 1, and probability of chance = 0.00.

tems give PTOs full or reduced sick leave (see Table 5). About one-third of the systems give PTOs full or reduced holiday, vacation, and retirement benefits. Four of every ten PTOs receive full or reduced health and welfare insurance benefits. With the exception of retirement benefits there is a statistical relationship between system size and the granting of these benefits. Roughly one-half of the small systems grant full or reduced benefits whereas only about one-third of the large systems provide them.

PTO Work Rules

The use of PTOs can significantly reduce the cost of providing peak-period service because they are subject to less restrictive work rules than are their FTO counterparts. At nearly eight of every ten transit systems, PTOs receive no guarantee per assignment. Another 16 percent receive guarantees of 2 hr or less. No statistical pattern was found in work rule guarantees, either by system size or geographic location.

PTOs also do not receive spread premiums. Only one system, the Central Contra Costa Transit Authority in California, that pays spread premiums was identified.

Spread premiums are probably not an issue in most systems because of the maximum work hour limitations that are contained in many contracts. As the data in Table 6 indicate, more than three-fourths of the systems have work hour limitations with a median

TABLE 5 Fringe Benefits

	Systems Permitted PTOs		50 Employees or Fewer		More than 50 Employees	
	Number	Percentage	Number	Percentage	Number	Percentage
Sick Leave[a]						
No	117	74.5	44	68.8	73	78.5
Reduced	20	12.7	8	12.5	12	12.9
Full	20	12.7	12	18.8	8	8.6
Total	157	100.0	64	100.0	93	100.0
Holidays[a]						
No	104	66.2	35	54.7	69	74.2
Reduced	26	16.6	11	17.2	15	16.1
Full	27	17.2	18	28.1	9	9.7
Total	157	100.0	64	100.0	93	100.0
Vacation[b]						
No	97	61.8	31	48.4	66	71.0
Reduced	33	21.0	14	21.9	19	20.4
Full	27	17.2	19	29.7	8	8.6
Total	157	100.0	64	100.0	93	100.0
Retirement[c]						
No	105	66.9	38	59.4	67	72.0
Reduced	13	8.3	3	4.7	10	10.8
Full	39	24.8	23	35.9	16	17.2
Total	157	100.0	64	100.0	93	100.0
Health and Welfare Insurance[d]						
No	93	59.6	30	46.9	63	68.5
Reduced	33	21.2	12	18.8	21	22.8
Yes	30	19.2	22	34.4	8	8.7
Total	156	100.0	64	100.0	92	100.0

[a] Chi-square = 1.12, degrees of freedom = 2, and probability of chance = 0.57.
[b] Chi-square = 12.95, degrees of freedom = 2, and probability of chance = 0.00.
[c] Chi-square = 7.4, degrees of freedom = 2, and probability of chance = 0.02.
[d] Chi-square = 7.4, degrees of freedom = 2, and probability of chance = 0.00.

TABLE 6 Maximum Work Hour Provisions

			System Size			
	Systems Permitting PTOs		50 or Fewer Employees		More than 50 Employees	
Maximum Work Hours per Week	Number	Percentage	Number	Percentage	Number	Percentage
Less than 20	3	2.2	1	1.9	2	2.4
20-24	25	18.2	10	18.5	15	18.1
25-29	27	19.7	2	3.7	25	30.1
30-39	37	27.0	11	20.4	26	31.3
40-42	13	9.5	11	20.4	2	2.4
No maximum	32	23.4	19	35.2	13	15.7
Total	137	100.0	54	100.0	83	100.0

Note: Chi-square = 29.54, degrees of freedom = 5, and probability of chance = 0.00.

value of 25 hr. There is a statistical pattern by system size: a higher percentage of small systems than of large systems has either no or extremely large maximum work hour limits. Again this small-large system pattern suggests better acceptance of PTOs by small systems than by large systems.

METHODS OF ASSIGNING PTOs WORK

The purpose of the second part of the UMTA study was to identify methods currently used to assign PTOs work. Two types of methodologies were identified: (a) methodologies used to assign PTOs work on the basis of existing run cuts and (b) methodologies used to incorporate PTOs into automated run-cutting procedures. The five methodologies that were found in the study are discussed in the next sections.

Methods of Assigning PTOs on the Basis of Existing Run Cuts

Methodologies for assigning PTOs on the basis of existing run cuts were identified at three systems. These systems are the Washington Metropolitan Area Transit Authority (WMATA), the Southern California Rapid Transit District (SCRTD), and the Alameda-Contra Costa Transit District (AC Transit). These agencies have more part-time eligible pieces of work than they have PTOs to fill them. The question they face is: "What pieces of work should be assigned to PTOs in order to maximize cost savings?"

Work rules governing the use of PTOs influenced the development of the assignment procedure at each agency. As the data in Table 7 indicate, the definitions and values of work rules vary considerably among the three agencies. For example, there are three variations of work hour limitations of which WMATA must follow one (weekly maximum), AC Transit two (daily maximum and weekly maximum), and SCRTD all three (daily minimum, daily maximum, and weekly maximum). Because each of these rules must be considered, the PTO assignment procedures at AC Transit, SCRTD, and WMATA are different.

TABLE 7 Summary of PTO Work Rules at AC Transit, SCRTD, WMATA

Work Rule Limits	AC Transit	SCRTD	WMATA
Percentage of FTOs			
Division	15	10	None
Systemwide	10	None	10
Work hours			
Daily minimum	None	2.5	None
Daily maximum	5	5	None
Weekly maximum	25	25	30
Type of work piece permitted			
Weekday trippers	Yes	Yes	Yes
Split runs	Yes	No	Yes

Washington Metropolitan Area Transit Authority

WMATA uses a three-step approach to assigning work to PTOs. First, a.m. and p.m. trippers are rank-ordered on the basis of descending pay time. Second, the number of FTOs and PTOs working trippers off of the extraboard is determined for each division by WMATA's Schedules Section. Finally, the tripper pairs with the highest pay times are assigned to FTOs by WMATA's Operations Department; the remaining pairs are assigned to PTOs.

Full-time extraboard operators, regular FTOs, and PTOs work trippers at WMATA. The number of regular FTOs working trippers is calculated at each division as the difference between the number of a.m. and p.m. trippers at that division. For example, 29 regular FTOs worked trippers at WMATA's Four Mile Run Division that had 112 a.m. trippers and 83 p.m. trippers for the schedule effective January 24, 1983. Approximately 70 percent of the remaining trippers are assigned to PTOs; 30 percent are assigned to full-time extraboard operators (i.e., FTOs). This 70/30 split was calculated to comply with the contract provision that limits the maximum number of PTOs to 10 percent of the number of FTOs, systemwide.

FTOs and PTOs are assigned to a.m. and p.m. paired trippers (or "married" trippers) on the basis of the criterion of combined pay time. The objective of WMATA's Schedules Section is to minimize make-up time (or the difference between the 8-hr guarantee and combined pay time) paid to FTOs. Morning and p.m. trippers are each rank-ordered by descending pay time. The highest paid a.m. tripper is then "married" to the highest paid p.m. tripper, the second highest paid a.m. tripper is married to the second highest p.m. tripper, and so forth.

Figure 1 shows how the a.m. and p.m. trippers were paired at WMATA's Four Mile Run Division for the schedule effective January 24, 1983. Each a.m. and p.m. tripper is ranked by number by descending pay time. The combined pay time is shown in column 6, and difference between the 8-hr guarantee and combined pay time (i.e., make-up time) is shown in column 7.

In the case of the Four Mile Run Division, this yielded 83 married trippers effective January 24, 1983. The 20 married pairs with the greatest combined pay times (and, hence, lowest combined make-up times) were assigned to FTOs. The combined make-up time for the top 20 married trippers is then determined in order to establish a daily make-up time budget for each division. In this case, the Four Mile Run Division had a budget of 23 hr 20 min of daily make-up time effective January 24, 1983. Part-time paid hours are equal to 334 hr 16 min or 5 hr 18 min per day per PTO, on average. The overtime penalty, calculated at one-half times the number of hours worked by regular FTOs assigned single-piece trippers, also appears at the top of this sheet and is equal to 34 hr 37 min.

Southern California Rapid Transit District

SCRTD uses a two-step procedure to assign work to PTOs. First, part-time eligible pieces of work are identified at each division on the basis of the constraints of the labor agreement and other practical considerations specified by SCRTD's Transportation Department. Second, these pieces of work are rank-ordered and the highest ranked pieces are assigned to PTOs; the remaining work is assigned to FTOs. PTOs are restricted to working only single-piece, weekday tripper assignments at SCRTD.

Only certain trippers within a division are eligible to be worked by PTOs at SCRTD. They must, by contract, (a) be nonbiddable by regular FTOs, (b) have at least 2.5 hr of work, and (c) have no more than 5 hr of work. Biddable trippers are defined by the Schedules Department as short peak-period pieces of work that are worked at overtime by regular FTOs before or after their regular runs. They are generally less than 2.5 hr work time and are less costly to work at overtime than by extraboard FTOs. Unlike WMATA, SCRTD does not define the number of biddable trippers as the difference between the number of

Regular FTO Overtime Penalty

Effective: January 24, 1983
8 Hrs. Reg. MEN GUAR. 23:20
69:13 @ 1 1/2 = 34:37 O/T PENALTY
PART TIME PAY HOURS 334:16

DIVISION: Four Mile

Tripper Number	112 A.M. Pay		Tripper Number	83 P.M. Pay	Combined Pay Time	8 Hours Guaranteed
1004	3:27	Full-Time Extraboard	2043	3:51	7:18	0:42
1000	3:26		2013	3:50	7:16	0:44
1019	3:25		2004	3:49	7:14	0:46
1003	3:24		2031	3:48	7:12	0:48
1009	3:22		2003	3:45	7:07	0:53
1055	3:19		2021	3:44	7:03	0:57
1020	3:16		2002	3:40	6:56	1:04
1025	3:15		2007	3:39	6:54	1:06
1011	3:13		2007	3:36	6:49	1:11
1002	3:12		2006	3:35	6:47	1:13
1006	3:09		2008	3:35	6:44	1:16
1032	3:08		2023	3:33	6:41	1:19
1042	3:08		2011	3:32	6:40	1:20
1061	3:08		2025	3:32	6:40	1:20
1023	3:05		2042	3:32	6:37	1:23
1050	3:05		2012	3:29	6:34	1:26
1005	3:04		2018	3:29	6:33	1:27
1014	3:04		2051	3:29	6:33	1:27
1041	3:04		2017	3:28	6:32	1:28
1031	3:03		2005	3:27	6:30	1:70
1109	2:00	Part-Time Operators	2048	3:27	5:27	23:20 (Make-up Time Budget)
1103	2:00		2038	3:26	5:26	
1093	2:00		2041	3:23	5:23	
1048	2:00		2050	3:23	5:23	
1101	2:10		2054	3:23	5:33	
1090	2:11		2030	3:22	5:33	
1089	2:11		2057	3:21	5:32	
1104	2:11		2028	3:16	5:27	
1108	2:12		2014	3:15	5:27	
1094	2:14		2016	3:15	5:29	
1099	2:16		2035	3:12	5:28	
1091	2:16		2009	3:12	5:28	
1062	2:17		2015	3:11	5:28	
1037	2:17		2024	3:11	5:28	
1078	2:17		2026	3:11	5:28	
1086	2:18		2019	3:09	5:27	
1084	2:20		2000	3:08	5:28	
1092	2:21		2060	3:08	5:29	
1087	2:23		2044	3:07	5:30	
1105	2:23		2027	3:06	5:29	
1070	2:24		2047	3:06	5:30	
1033	2:24		2066	3:01	5:25	
1021	2:24		2049	3:00	5:24	
1106	2:24		2020	2:59	5:23	
1054	3:02		2037	2:57	5:59	
1053	3:01		2063	2:57	5:58	
1028	2:59		2029	2:56	5:55	
1060	2:58		2061	2:55	5:53	
1024	2:57		2039	2:53	5:50	
1012	2:56		2033	2:51	5:47	
1045	2:56		2034	2:51	5:47	
1071	2.56		2082	2:51	5:47	
1013	2:54		2043	2:49	5:43	
1018	2:54		2059	2:49	5:43	
1027	2:54		2022	2:48	5:42	

FIGURE 1 Facsimile of rank order of a.m. and p.m. trippers at WMATA's Four Mile Run Division.

a.m. and p.m. trippers. Therefore, after the biddable trippers have been assigned to regular FTOs, the numbers of a.m. and p.m. trippers may not be equal at the division. As a result, not all full-time extraboard operators are assigned an a.m./p.m. tripper pair. Some full-time extraboard operators must work an a.m. or p.m. tripper and stand extra during the remainder of the day.

SCRTD's prioritization algorithm consists of three procedural steps. First, a.m. and p.m. trippers are separated and listed by ascending sign-on and sign-off times, respectively. For example, as shown in Figure 2, Iteration 1, the a.m. tripper signing on at 4:32 a.m. is listed first and followed in order by those trippers signing on at 5:00, 5:09, 5:12, and 5:15 a.m., respectively. Thus, runs representing the most spread cost (i.e., the earliest pull-outs) are placed at the top of the list. These runs are best worked by PTOs because PTOs receive no spread premiums.

In the second step, the time savings that would result from exchanging the positions of run N with run N+1, N+2, and all subsequent runs are considered. This is done by computing the time savings of operating run N with an FTO and run N+1 with a PTO. The increase in FTO spread time is calculated at

Effective:
8 Hrs. Reg. MEN GUAR.
@ 1 1/2 = O/T PENALTY
PART TIME PAY HOURS

DIVISION: Four Mile

Tripper Number	112 A.M Pay		Tripper Number	83 P.M. Pay	Combined Pay Time	8 Hours Guaranteed
1035	2:53		2069	2:44	5:37	
1039	2:53		2064	2:44	5:37	
1040	2:53		2056	2:43	5:36	
1080	2:53		2052	2:43	5:36	
1001	2:52		2053	2:43	5:35	
1015	2:52		2058	2:42	5:34	
1079	2:52		2040	2:41	5:33	
1077	2:52		2046	2:40	5:32	
1008	2:51		2036	2:38	5:29	
1049	2:51		2055	2:36	5:27	
1017	2:50		2070	2:36	5:26	
1026	2:50		2082	2:35	5:25	
1043	2:50		2078	2:31	5:27	
1044	2:50		2065	2:24	5:14	
1068	2:50		2067	2:17	5:07	
1016	2:48		2080	2:11	4:59	
1030	2:48		2075	2:01	4:49	
1073	2:48		2081	2:00	4:48	
1067	2:47		2062	2:00	4:48	
1010	2:46		2068		4:46	
1052	2:45		2091		4:45	
1057	2:43		2072		4:43	
1007	2:42		2073		4:42	
1065	2:42		2074		4:42	
1034	2:40		2076		4:40	
1051	2:39		2077		4:39	
1095	2:38		2079		4:39	
1082	2:37		20		334:16	
1096	2:37		20			
1047	2:36		20			
1097	2:36		20			
1085	2:35		20			
1038	2:34		20			
1059	2:34		20			
1036	2:33		20			
1066	2:33		20			
1076	2:33		20			
1069	2:32		20			
1102	2:32		20			
1107	2:32		20			
1022	2:31		20			
1058	2:31		20			
1063	2:30		20			
1083	2:30		20			
1081	2:29		20			
1075	2:28		20			
1064	2:28		20			
1029	2:27		20			
1046	2:26		20			
1088	2:26		20			
1056	2:25		20			
1072	2:25		20			
1074	2:25		20			
1100	2:25		20			
1110	2:25		20			
1111	2:00					

Part-Time Operators

Regular Full-Time Operators Working Overtime

PTO Pay Hours

69:13 Full-Time

Regular FTO Work Hours

FIGURE 1 continued.

time and one-half the difference in sign-on time for a.m. trippers or sign-off time for p.m. trippers. Next, the decrease in work time paid to PTOs is calculated at straight time because PTOs are paid for only the hours they work, without overtime or spread premiums.

The increase in FTO spread is then added to the decrease in PTO work time. If the sum is negative, the decrease in PTO work time is greater than the increase in FTO spread time, and run N+1 is less costly worked by a PTO than by an FTO. In this case the positions of run N and run N+1 should be exchanged. If the sum is positive, the increase in FTO spread time more than offsets any PTO work time savings and the positions should not be exchanged.

All sums are positive relative to the 4:32 a.m.

Rank	Sign-On Time	FTO Spread Premium Difference	Sign-Off Time	Work Time	PTO Work Time Difference	Sum of Differences
Iteration 1: Examine Exchanging 4:32 a.m. Run						
1	4:32 a.m.		8:04 a.m.	3:32		
	5:00	+42	8:43	3:43	+11	+53
	5:09	+55½	8:31	3:22	-10	+45½
	5:12	+60	8:23	3:11	-21	+39
	5:15	+64½	8:36	3:21	-11	+53½
Iteration 2: Examine Exchanging 5:00 a.m. Run						
1	4:32 a.m.		8:04 a.m.	3:32		
2	5:12	+18	8:23	3:11	-32	-14
	5:09	+13½	8:31	3:22	-21	-7½
	5:00		8:43	3:43		
	5:15	+22½	8:36	3:21	-22	+½
Iteration 3: Examine Exchanging 5:09 a.m. Run						
1	4:32 a.m.		8:04 a.m.	3:32		
2	5:12		8:23	3:11		
3	5:09		8:31	3:22	+21	+7½
	5:00	-13½	8:43	3:43		
	5:15	+9	8:36	3:21	-1	+8
Iteration 4: Examine Exchanging 5:00 a.m. Run						
1	4:32 a.m.		8:04 a.m.	3:32		
2	5:12		8:23	3:11		
3	5:09		8:31	3:22		
4	5:00		8:43	3:43		
	5:15	+22½	8:36	3:21	-22	+½
Iteration 5: Final Rankings						
1	4:32 a.m.		8:04 a.m.	3:32		
2	5:12		8:23	3:11		
3	5:09		8:31	3:22		
4	5:00		8:43	3:43		
5	5:15		8:36	3:21		

FIGURE 2 Example of SCRTD's part-time operator assignment algorithm.

run in Figure 2, Iteration 1, and no changes are made. The 4:32 a.m. run is therefore ranked first among the five a.m. trippers.

In Iteration 2, the 5:00 a.m. run is compared to the 5:09, 5:12, and 5:15 a.m. runs, respectively. The 5:09 and 5:12 a.m. runs have negative sums of 7.5 and 14 min, respectively. The 5:12 a.m. run is ranked second because it has the most negative sum. That is, the 5:12 a.m. run has the highest net decrease in PTO work time compared to the 5:00 a.m. run.

These computations are repeated in Iterations 3 through 5. The 5:09 a.m. run is ranked third in Iteration 3; the 5:00 a.m. run is ranked fourth in Iteration 4. The final rankings are shown in Iteration 5. It can be observed that the same ranking would be obtained by listing the runs in ascending order of the sum of differences computed in Iteration 1.

The rank-ordered list of a.m. and p.m. trippers constitutes a priority-ordered list of PTO assignments for each of SCRTD's 13 divisions. These lists are forwarded to the Transportation Department for use on a routine daily basis in assigning work to PTOs and FTOs.

Alameda-Contra Costa Transit District

AC Transit uses a less formalized procedure for assigning operators to part-time eligible pieces of work than does SCRTD or WMATA. Under AC Transit's labor agreement, PTOs (a) are guaranteed 2 hr per day but can work no more than 5 hr per day or 25 hr per week, (b) can work only on weekdays, (c) cannot exceed 10 percent of FTOs systemwide or 15 percent at any division, and (d) must originate and terminate their assignments at a division (i.e., no on-street relief).

AC Transit's Schedules Department selects PTO assignments from among all eligible pieces of work at each division on the basis of a number of criteria derived from the labor agreement work rule provisions and other considerations. These criteria include:

1. PTOs should generally work close to the 5-hr daily limitation;
2. PTOs should generally be assigned to early pull-outs and late pull-ins to reduce spread premiums paid to full-time extraboard operators;
3. PTOs should generally work split runs instead of straight runs;
4. If PTOs work a split run, they should work the same line in the a.m. and p.m. because PTOs break in on only one line;
5. FTOs (i.e., "expensive" labor) should be assigned to contract service operated for BART and others by AC Transit;
6. FTOs must work runs that are relieved on the street, according to the labor contract; and
7. Individual PTO preferences regarding work times, work hours, and days off may also be taken into account.

The Schedule Department's suggestions are forwarded to the Operations Department, which may take these suggestions or assign PTOs to alternate runs.

Comparison of Methodologies

A comprehensive evaluation would include testing the methodologies on a common set of PTO assignment problems. Properly constructed, this testing would provide useful information on the sensitivity and accuracy of the procedures. Unfortunately, this type of testing was not included in the UMTA study. Therefore, the comparison of methodologies in the study was limited to an evaluation of the variables that are included in each procedure.

A comprehensive procedure for assigning PTOs to runs selected from existing schedules should consider PTO pay hours, FTO make-up time, FTO spread premiums and overtime, and PTO and FTO fringe benefits. As the data in Table 8 indicate, the procedures used by WMATA, SCRTD, and AC Transit consider some but not all of these variables. Each agency considers PTO pay hours; WMATA considers FTO make-up time; and both SCRTD and AC Transit consider FTO spread premiums in assigning PTOs to existing runs. None of these agencies includes FTO extraboard overtime or FTO or PTO fringe benefits in driver assignment decisions.

TABLE 8 Comparison of PTO Assignment Procedures

	Property		
	WMATA	SCRTD	AC Transit
Type of procedure			
Automated		X	
Manual	X		X
Variables considered			
PTO pay-hours	Yes	Yes	Yes
FTO make-up hours	Yes	No	No
FTO spread premiums	No	Yes	Yes
FTO overtime	No	No	No
PTO fringe benefits	No	No	No
FTO fringe benefits	No	No	No

The importance of considering full-time extraboard operator spread premiums depends on a system's spread rule provisions and service profile. Spread premiums are most onerous at systems with relatively short maximum spread times and spread penalty thresholds, relatively sharp peaks and relatively long a.m. and p.m. peak periods. Spread premiums were shown to be significant at SCRTD, but they may or may not be as important at WMATA or AC Transit. The consideration of both spread and overtime premiums is especially important at systems such as AC Transit that pay both overtime and spread penalties when applicable.

Fringe benefits are an important factor in determining which trippers to assign to regular FTOs, extraboard FTOs, and PTOs. Systems with relatively high fringe benefits may find it less costly to assign more trippers to regular FTOs on an overtime basis in order to avoid the fixed fringe benefit costs that would be incurred by hiring additional extraboard operators. A significant operator cost savings may be attributable to the lower fringe benefits received by PTOs.

Methods of Incorporating PTOs in Automated Run-Cutting Procedures

Two automated procedures were identified that consider PTOs when run cuts are made. The first procedure uses a version of RUCUS and is used by the San Francisco Municipal Railway (Muni). The second procedure is part of a computerized scheduling package called RAMCUTTER. Both Muni and AC Transit are currently experimenting with this procedure.

San Francisco Municipal Railway

Muni uses Version 5.01 of the RUCUS run-cutting package, developed by Kenneth Roberts & Associates, to schedule both PTOs and FTOs. PTOs at Muni are permitted to work "short runs" and a PTO extraboard, up to 5 hr a day, 25 hr a week. PTOs are guaranteed 3.5 hr per assignment.

The RUCUS methodology uses a seven-step procedure. In the first step, long blocks are cut into straight runs. All runs beginning before 5:50 a.m. and ending after 6:00 p.m. are made into straight runs. The minimum and maximum platform times are specified as 7 hr 14 min and 8 hr 50 min, respectively. The maximum spread time is specified as 10 hr 59 min. (The spread premium threshold is 10 hr.) These parameters are established in an iterative fashion, through repeated attempts to improve the results of previous runs by adjusting each parameter. The straight runs are then "frozen" and are not modified in subsequent steps.

In the second step, the work remaining after the straight runs are cut is divided into two nearly equal pieces with a target platform time of 4 hr. The minimum and maximum platform times are specified as 30 min and 5 hr 45 min for any piece of a two-piece run.

Next, two-piece PTO runs are cut with relatively long spread and swing times. The minimum and maximum spreads for PTO runs are specified as 10 hr 15 min and 11 hr 59 min, respectively. As a matter of policy, 11 hr 59 min is used as the maximum spread for PTO runs; 11 hr 59 min is the maximum spread for FTO runs established by Muni's labor agreement. Swing times for PTO runs are specified as 6 hr to 9 hr 30 min. The minimum and maximum platform times for PTO two-piece runs are specified as 3 hr 20 min and 4 hr 39 min, respectively, in accordance with the 3.5 hr guarantee and 5 hr work per day limitation stipulated for PTOs in Muni's labor agreement. The PTO two-piece runs are then "frozen" and are not modified in subsequent steps.

FTO two-piece runs are cut from all remaining work in the next two steps. In both steps the maximum spread time and the average platform time are specified as 11 hr 59 min and 8 hr 30 min, respectively. FTO swing time is limited to no more than 3 hr in the fourth step.

The work remaining is then cut in the fifth step into two-piece runs with a maximum swing of 4 hr. All other parameters are held constant.

An attempt is made to reduce costs in the sixth step by switching pieces between two two-piece runs output from the fourth and fifth steps. Any one-piece trippers remaining are manually worked into the cut.

AC Transit

Muni and AC Transit are currently experimenting with an automated run-cutting procedure that takes into direct account wages, fixed and variable overhead, and work rules governing the use of both PTOs and FTOs when searching for a least-cost run cut for a given service schedule. This package, called the RAMCUTTER, was developed by Research Applications for Management (RAM), Inc.

The RAMCUTTER minimizes total annualized cost incurred for schedule work time, fixed and variable overhead, and other allowances for both FTOs and PTOs, subject to a series of constraints imposed by the labor agreement and the schedule department. AC Transit uses a total of 270 input parameters that specify minimum and maximum constraints, various thresholds, penalties and bonuses, output formats, and so forth. Thirty-four of these variables affect the use of PTOs. These parameters include such "hard" constraints as maximum percentage of PTOs, hourly pay rate, and maximum pay time. The "soft" rules include penalties for runs starting before a specified time, runs ending after a specified time, pieces of work below a specified threshold size, runs with platform times less than a specified time, and so forth. The schedules department can also penalize (or reward) part-time work in general, thereby reducing or increasing the number of part-time runs cut by the RAMCUTTER. Annual fixed overhead costs for each PTO and FTO, and variable overhead expressed as a percentage of PTO and FTO pay times, are also direct inputs to the RAMCUTTER.

There are a large number of potential solutions for assigning PTOs and FTOs to a given service schedule. It is not practical to test each of these alternatives to identify the least-cost solution, even with modern electronic computers. A greater number of alternatives can be tested by the RAMCUTTER as schedulers allocate more and more computer time to the problem.

AC Transit's schedulers operate the RAMCUTTER in an iterative fashion. A few minutes of computer time are allocated to produce an initial run cut. Additional runs are then produced using the same amount of computer time by tightening or loosening certain constraints or rules in order to achieve implementable run cuts that are acceptable to the schedules department and others. When the input parameter values have been established, the schedulers can allocate a greater amount of computer time to achieve a more nearly optimal solution.

Comparison of Methodologies

The RAMCUTTER and RUCUS Version 5.01 are both automated procedures that consider both PTOs and FTOs when cutting runs for a given service schedule. Both consider differences in PTO and FTO work rules. The RAMCUTTER also considers PTO and FTO fixed and variable overhead costs, which RUCUS does not. In addition, the RAMCUTTER incorporates a greater number of constraints regarding start and end times, road reliefs, platform ties, and so forth, which help to generate acceptable, implementable run cuts. A comparison of the RAMCUTTER and RUCUS Version 5.01 by the creator of the RAMCUTTER at Tri-Met in Portland, Oregon, indicated that the procedures result in solutions of approximately equal cost.

CONCLUSIONS

The major conclusions of this study regarding the methods for determining the use of part-time operators are as follows:

The use of PTOs is widely regarded as a means of reducing the cost of providing peak-period transit service, thereby improving transit productivity, because

- PTOs are governed by less restrictive work rules than are their FTO counterparts,
- PTOs typically receive no spread or overtime premiums,
- PTOs almost always receive lower fringe benefits than do FTOs, and
- PTOs sometimes earn lower wages than do FTOs.

The use of PTOs is widespread in transit systems of all sizes in all regions of the nation. Three-fourths of all U.S. transit systems are permitted to use PTOs; one of every ten operators in these systems is a PTO.

A higher percentage of small systems (50 employees or fewer) are permitted to use PTOs than large systems (more than 50 employees). Labor contracts of small systems also tend to be less restrictive in the permitted use of PTOs.

A variety of procedures is being used to assign PTOs to pieces of work selected from existing run cuts. These procedures consider PTO pay hours, FTO make-up hours, and FTO spread premiums or FTO overtime, or both, in deciding which pieces to assign to FTOs and which to assign to PTOs. No procedure was identified that considered all of these variables or PTO or FTO fringe benefits.

RUCUS Version 5.01 and the RAMCUTTER are promising computerized procedures that incorporate PTOs directly into the run-cutting process. RUCUS is presently used at San Francisco Muni, and the RAMCUTTER is in the testing stages at both Muni and AC Transit.

REFERENCES

1. 1981 Transit Fact Book. American Public Transit Association, Washington, D.C., 1981.
2. K.M. Chomitz and C.A. Lave. Forecasting the Financial Effects of Work Rule Changes. Transportation Quarterly, July 1983.

Analysis of Bus Transit's Operating Labor Efficiency Using Section 15 Data

ATHANASSIOS K. BLADIKAS and CHARLES PAPADIMITRIOU

ABSTRACT

Operator labor costs are the biggest operating expense category for motor bus transit systems; these costs account for approximately 42 percent of total operating expenses. Thus, if operating labor becomes more cost-efficient, significant overall cost reductions are possible. Operator labor cost per platform-hour or vehicle-operating hour was used as a macro measure of cost efficiency, and this measure was built up gradually from elementary and composite factors. The effect that environmental factors have on operator costs is examined by regressing them on each of the elementary and composite cost-efficiency measures. Using the results of these analyses, transit managers will be able to diagnose and possibly remedy the causes of their labor inefficiency.

Escalating costs, declining productivity, and increasing dependence on public subsidies have been the trend in the transit industry for the past two decades (1). Furthermore, transit has been given the assignment of accomplishing an array of social objectives, ranging from energy conservation to providing mobility for the poor and the handicapped. All this has led to an increased interest in the performance evaluation of the nation's transit systems. There is no general agreement on how to define and measure the performance of a transit system because the goals to be accomplished are often vague and conflicting. However, most researchers agree that transit performance is a multidimensional concept that includes some or all of the following elements (2,3):

- Efficiency,
- Effectiveness,
- Quality of service, and
- Societal impacts.

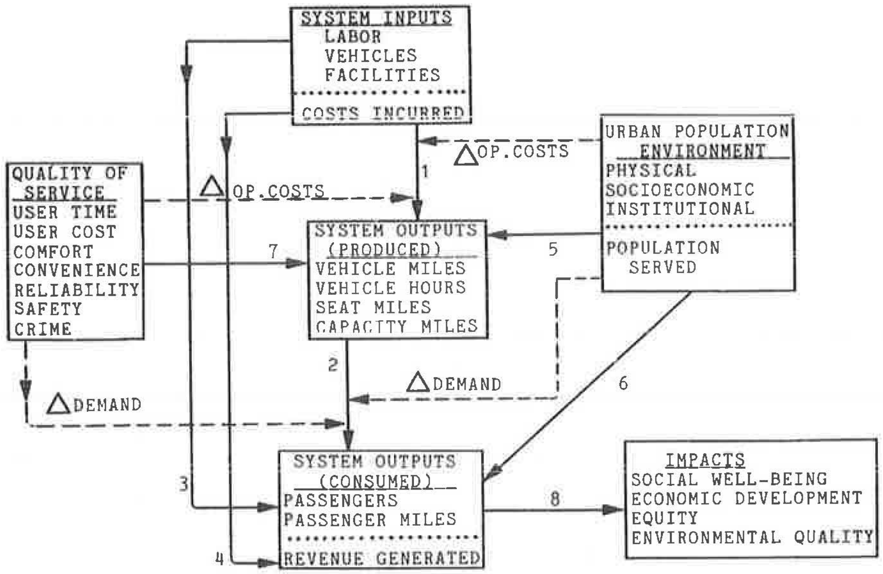
All of these elements of performance are not dealt with here; the focus here is only on the cost-efficiency concept (Link 1 in Figure 1) as it relates to the efficient use of operators in providing a vehicle-hour of service. For the purposes of this paper the term "operator" means only vehicle operators (i.e., drivers). The transit agency or firm that is responsible for the provision of service will be called system operator. Operator labor costs are the biggest system operating expense category and account for approximately 42 percent of total operating expenses (4). Thus, if operating labor becomes more cost-efficient, significant overall cost reductions are possible. Direct comparisons of systems and cross-sectional analyses are not generally useful because the major causes of operator cost variations are factors that are determined by the environment in which the system operates and are mostly outside the system operator's control. These environmental factors and their effect on operator labor costs are examined.

TRANSIT LABOR PRODUCTIVITY

Operators are paid at different hourly rates not only for the hours that they drive their buses (platform time) but for additional time as well, as the various Section 15 reporting categories given in Table 1 indicate. The definitions of all Table 1 items can be found in Volume II of the Uniform System of Accounts and Records (5). Table 2 gives abbreviated definitions of some key items as well as information on how these items are treated in labor contracts. Some of the categories in Table 1 are affected greatly by the combination of the system's service profile and its labor contract provisions (e.g., overtime and spread premiums, pull-in and pull-out times), and others are the result of contract clauses (e.g., run selection time, student training time). Because different categories are paid at different rates, the crucial question is not how many hours operators were paid in excess of their platform time but how much more they received in excess of the amount due to straight platform time service.

Before Section 15 became available, the best method for examining operating labor productivity was to look at vehicle-hours per operator. Some researchers used the ratio of revenue vehicle-hours per operator, which does not reflect only labor productivity but is also dependent on the route and network structure of the system, which is a major factor in the accumulation of deadheading time. Besides, as pointed out by Fielding et al. (6), both of these ratios have an inherent major flaw: they make transit systems that use part-time operators seem unproductive, whereas the reverse is actually true.

The concept of employee equivalents was introduced later, with one operator being equivalent to 2,000 or 2,080 hours of work, but problems persist with this measure because it does not make the distinction between hours of work and hours paid for. An operator may "work" 8 hr during a day but he may get paid for 8, 9, or 10 hr depending on when and under what circumstances those hours were clocked. Employee equivalents, therefore, tend to hide the effect of work rules that require system operators to pay premium wages for certain types or hours of work occurring outside some predetermined norms. To make matters worse, system operators (or at least those who fill out the Section 15 forms) do not appear to grasp the employee equivalent concept. The hard copy Section 15 annual reports present figures



Link interaction definitions:

1	Efficiency (cost, labor, resources)	EX: Rev.Veh.Hrs./Op.Costs
2	Service Effectiveness	EX: Passengers/Veh.Hr.
3	Cost Effectiveness	EX: Passengers/Op.Costs
4	Revenue Effectiveness	EX: Op.Costs/Fare Revenue
5	Service Coverage/Intensity	EX: Veh.Hrs/Population
6	System Effectiveness	EX: Passengers/Population
7	Service Quality	EX: Accidents/Mill. Veh. Miles
8	System Impacts	EX: Retail Sales

FIGURE 1 Transit system performance evaluation model.

TABLE 1 Operators' Wages Subsidiary Schedule Data Elements

Form No. 321

TIME CLASSIFICATION	DOLLARS	HOURS
1. OPERATING TIME		
1.01 Report time (Pull out)		
1.02 Turn-in time (Pull in)		
1.03 Travel time		
1.04 Platform time-line service		
1.05 Platform time-charter & special service		
1.06 Intervening time		
1.07 Paid breaks & meal allowance		
1.08 Min. guarantee for call out		
1.09 Minimum guarante-daily		
1.10 Minimum guarantee-weekly		
1.11 Overtime premium-scheduled		
1.12 Overtime premium-unscheduled		
1.13 Spread time premium		
1.14 Shift premium		
1.15 Other operating premium		
1.00 TOTAL OPERATING TIME		
2. NONOPERATING PAID WORK TIME		
2.01 Instructor premium for operator training		
2.02 Student training time		
2.03 Accident reporting time		
2.04 Witness time		
2.05 Stand-by time		
2.06 Time spent on uniform functions		
2.07 Run selection time		
2.08 Other time spent in transportation administration		
2.09 Time spent in revenue vehicle movement control		
2.10 Time spent in ticketing and fare collection		
2.11 Time spent in customer service		
2.12 Time spent in other nonoperating functions		
2.00 TOTAL NONOPERATING PAID WORK TIME		
3.00 TOTAL OPERATING AND NONOPERATING TIME		

TABLE 2 Definitions of Selected Terms

Term	Item No.	Definition
Report time Turn-in time	1.01 1.02	Covers payments for the time allowed an operator to report to the dispatcher and receive instructions at the beginning and end of a piece of work.
Travel time	1.03	Covers payment for the time allowed an operator to travel between the operation station and the point where he relieves or is relieved by another operator.
Platform time-line plus charter and special service	1.04 1.05	Covers payments for time during which an operator operates the revenue vehicle in line or charter and special service, respectively. Deadheading and layover time is included.
Intervening time	1.06	Covers payment for the time between any two pieces of a run that is made up of more than two pieces.
Paid breaks and meal allowances	1.07	Covers payment for break time, other than layover time and intervening time, and allowances for company-paid meals.
Minimum guarantee for call-out	1.08	Covers payment for the time beyond that associated with the performance of a work piece, in order to bring the total amount paid up to the guaranteed minimum for the call-out.
Spread time premium	1.13	Is the bonus above straight-time pay for hours worked after a specified number of hours from the start of the operator's day. An operator who works two shifts, say from 7 to 11 a.m. and from 3 to 7 p.m., actually works 8 hr but "spreads" them over a 12-hr period in two "split" shifts, each one consisting of 4 hr of continuous working time. The contract may provide that operators are entitled to spread premiums after the tenth hour of work in a spread. Thus, this operator will get paid the spread premium for the last 2 hr of work. In addition, work assignments may be prohibited beyond a specified maximum spread time (e.g., 12 or 13 hr). Some contracts may also provide for an upper limit on the runs that could be spread (e.g., it may be stipulated that straight runs should be at least 60 percent of the total runs). Furthermore, a minimum work piece in a split shift may also be stipulated. This usually requires that each piece of a split shift be at least 2 or 2.5 hr long. Thus, an operator may not work just 1 hr in the morning and 7 hr in the afternoon and evening.
Shift premiums	1.14	Cover bonuses for working during times of the day that are subject to special time differentials (e.g., night or weekend service).

on vehicle-hours per operator and revenue vehicle-hours per operator. It might be thought impossible that these ratios exceed 2,000; however, the Year 4 annual report has 97 cases in which vehicle-hours per operator exceed 2,000 and some go higher than 5,000.

The results of recent research efforts in the area of transit labor productivity illustrate the problems encountered when trying to explain the labor efficiency of transit systems. Most of the studies were done using a relatively small and regionally segregated number of systems (7), and the data were not generated by a uniform reporting system. This is the reason for the conflicting results that were produced on some occasions. For example, in Giuliano's work (8) the positive sign of the variable average wage rate indicates that higher wages induce a more efficient use of labor, but Barnum's equation (9) shows the opposite because the variable enters with a positive sign also, but it explains what is effectively the inverse of Giuliano's variable. Some attempted to explain their efficiency measures by including a number of subsidy variables (9,10) that failed to increase significantly the explanatory power of the regressions. Using mostly contract provision variables in his equation, Wilson was somewhat successful in predicting pay-hours per bus-hour (11).

The operator efficiency analysis presented in this paper differs from all previous studies in most, if not all, of the following respects:

- Data are used for the first time that were produced from uniform, consistent, and precise reporting procedures and classifications;
- The transit systems in the sample represent all geographic regions of the country;
- The analysis is carried to the most detailed level possible, thus pinpointing the causes of inefficiency; and
- Greater emphasis is placed on the service characteristics of the systems.

OPERATOR LABOR COSTS

The indicator that best describes the total costs associated with operators is operator labor cost per platform-hour or vehicle-operating hour. However, this ratio is a macro measure, and it is possible to dissect it and build it up from elementary or composite factors as follows:

- Element A: Unproductivity factor ($ paid for total salaries/$ paid at the base rate for platform service),
- Element B: Average base wage rate ($ paid at the base rate for platform service/platform-hours),
- Element C: Total salaries per platform-hour (product of A • B),
- Element D: Fringe benefits per platform-hour, and
- Element E: Operator labor cost per platform-hour [C + D or A•(B + D)].

If transit system managers want to evaluate their performance, analyses have to be performed not only on the macro measure but on all of its component elements as well. A transit system, for example, may appear to be doing well in terms of the macro ratio, but a closer examination may reveal that its unproductivity ratio is well above average and that it is paying rather low wages. Obviously, the corrective action suggested in this case would be entirely different from that for a situation in which the reverse is true (i.e., low unproductivity ratio and higher wages). The approach taken in this paper is to build up the macro measure from its component parts analyzing each one of them along the way.

DATA SOURCES

Data on operator labor costs, characteristics of service supplied and consumed, and generated revenues were obtained from the fourth year (FY ending

June 82) of Section 15 data (4). Data relating to socioeconomic and physical variables were obtained from the 1983 City and County Data Book (12), and additional information on wages for city employees was extracted from Bureau of the Census statistics (13).

Section 15 Data Elements

The Operators Wage Subsidiary Schedule (Form 321 and machine readable file OWSS, required for systems operating more than 25 revenue vehicles and effectively duplicated in Table 1) was used to obtain values for the first two elementary factors that can be used for the calculation of the macro measure. If the average base wage rate is defined as

Dollars for platform (line + charter) service/Hours of platform (line + charter) service,

the Form 321 items needed to compute it are

Dollars (1.04 + 1.05)/Hours (1.04 + 1.05).

In computing the unproductivity factor, the question may legitimately be raised whether Items 2.08 through 2.11 (time spent on a temporary basis on nonoperating functions) should be counted as productive or unproductive time. Operators getting paid for these items may be unproductive because they are not performing their major function (i.e., driving a bus). On the other hand, it may be noted that they perform at least some service to the system's public and therefore they are productive. To resolve this problem, three major unproductivity ratios were computed in terms of the following Table 1 items:

Y1 = Item 3.00/Items (1.04 + 1.05),
Y2 = Item 3.00/Items (1.04 + 1.05 + 2.08 + 2.09 + 2.10 + 2.11), and
Y3 = Items [3.00 - (2.08 + 2.09 + 2.10 + 2.11)] ÷ Items (1.04 + 1.05).

The first of these unproductivity factors considers Items 2.08 through 2.11 as unproductive time, the second considers them as productive, and finally the third simply ignores them completely by not including them in the computation. The unproductivity factor is a dimensionless ratio and either hours or dollars can be used for its derivation. Dollar amounts were chosen as more representative because the objective is not to reduce operator hours per se but to reduce the cost associated with those hours. A transit system, for example, may pay 1.1 times the base wage rate for night shift service and twice the base rate for unscheduled overtime. Therefore, paying an operator for an hour of unscheduled overtime results in overpayments that are equal to 10 hr of night shift premiums.

A close examination of the items in Table 1 reveals that some of them are greatly dependent on a system's service characteristics (e.g., Items 1.06, 1.13, 1.14) and that others simply reflect system policy (e.g., Items 1.01, 1.02, 2.01). To perform a more detailed analysis of a system's unproductivity, dollar amounts from file OWSS were used to produce the following six partial ratios:

Y4 = Items (1.06 through 1.15 + 2.05)/Items (1.04 + 1.05);
Y5 = Items (1.03 + 1.06 through 1.15 + 2.05)/Items (1.04 + 1.05);
Y6 = Items (1.01 + 1.02 + 1.03 + 1.06 through 1.15 + 2.05)/Items (1.04 + 1.05); and
Y7 = } Y8 = } Y9 = } Numerators as in Y4, Y5, and Y6, respectively, and denominator = Items (1.04 + 1.05 + 2.08 + 2.09 + 2.10 + 2.11).

The major unproductivity ratios (Y1 to Y3) are always greater than one, but the six partial ratios (Y4 to Y9) have values that are less than one and cannot be used in their presented raw form as a building block for the computation of total operator labor cost per platform-hour. Variables Y4 to Y9 are in effect ratios of unproductive to productive time, whereas Y1, Y2, and Y3 are ratios of total to productive time. Payments for the unquestionably productive times (platform time-line plus charter and special service) are the denominator of Y4, Y5, and Y6, and the questionable Items 2.08 to 2.11 are added to the denominators of Y7, Y8, and Y9. The numerators of Y4 and Y7 contain only items that are influenced mainly by a system's service characteristics. Travel time is added to the numerators of Y5 and Y8, and, finally, report and turn-in times are also included in the numerators of Y6 and Y9. The three major and six partial unproductivity ratios constitute a total set of nine unproductivity measures that were examined individually in order to determine whether they are affected differently by the various environmental and service characteristics factors.

Fringe benefit data were obtained from the Transit System Employee Count Schedule (Form 404 and file EMPSCH) and Expenses Classified by Function (Form 301 and file XTFO). File XTFO contains aggregate data by function (Vehicle Operations, Vehicle Maintenance, Nonvehicle Maintenance, and General Administration), and it is not possible to isolate operators' benefits. File EMPSCH contains employee equivalents for the following classes in the vehicle operations function:

11. Transportation executive, professional, and supervisory personnel;
12. Transportation support personnel; and
13. Revenue-vehicle operators.

In terms of these classes, the fraction of vehicle operating personnel that actually operates vehicles is 13/(11 + 12 + 13). Therefore, by taking the product of this fraction times total fringe benefits for the vehicle operations function, an approximate figure for operators' benefits can be obtained.

The Transit System Service Supplied, Service Consumed and Service Personnel Schedule (Form 406 and file NRSTDY) provides data by time period for vehicles in operation, vehicle-hours, vehicle-miles, and full-time plus part-time operators. Service period durations were obtained from the Transit System Service Period Schedule (Form 401 and file WDSPSC), and system revenue information was extracted from the Revenue Summary Schedule (Form 201 and file REVSCH). Systems were considered to be privately owned if they reported in their balance sheets (Form 101 and file CAPSCH) capital for private corporation or noncorporate ownership. Finally, the population of the urbanized area in which the system operates was obtained from file UAREA.

Only single-mode motor bus transit systems were analyzed in order to avoid problems with joint expenses. Because the use of file OWSS was necessary, only systems with more than 25 vehicles were included. There are 108 such systems, but 20 of them had to be eliminated because they either did not file Form 321 or had zero entries in file WDSPSC. Two additional systems were excluded due to missing data on other variables. Thus 86 systems with valid data were available for the analyses. These systems represent 29 states and range in size from 26 to 2,960 revenue vehicles.

Census and Other Data Sources

The 1983 edition of the County and City Data Book (12) was used to extract information on

- Percentage of persons using public transportation for the work trip for both the county and city area,
- Percentage of the civilian labor force unemployed in the county and city,
- Percentage of area (county) that is urbanized,
- Mean temperature in January and July, and
- Heating and cooling degree-days in a year.

Data on average monthly earnings of city employees were derived from Government Statistics Reports on City Employment (13). These reports provide data for the month of October of each year. Reports for 1980, 1981, and 1982 were used to extrapolate data and make them coincidental with the sixth month of each system's fiscal year. The Directory of Regularly Scheduled, Fixed Route, Local Public Transportation Service (14) was used to identify the systems that are managed by private contract management firms.

OPERATOR LABOR COST ELEMENT ANALYSIS

Unproductivity Factor (Element A)

Factors Hypothesized to Influence this Variable

The service profile of a transit system influences greatly the payment amounts for some of the categories given in Table 1. The variables derived from the service profile and expected to be proportional with the unproductivity factor are as follows:

- Vehicles high peak to vehicles midday,
- Shoulder-to-shoulder time (start of a.m. peak to end of p.m. peak),
- Midday duration (end of a.m. peak to start of p.m. peak), and
- Vehicles high peak to vehicles low peak.

In addition, the unproductivity factor should be influenced by the following:

1. Size of the transit system. Union strength and bargaining power should be greater in larger systems, but, on the other hand, larger systems may have better scheduling techniques that may result in the reduction of the unproductivity factor. The variables used to represent system size were
   - Number of employees,
   - Number of revenue vehicles (total and during each service period),
   - Weekday hours of operation,
   - Annual hours of operation,
   - Annual vehicle-miles, and
   - Annual vehicle-hours.
2. The relative wealth of the system's area of operation represented by
   - County income per capita and
   - Average monthly earnings of city employees.
3. The system's ability to generate revenue represented by the variable passenger revenues per platform-hour.
4. The system's organizational or management structure or whether it is
   - Public or private,
   - Managed under contract by a private firm, or
   - A transit authority or transit district.
5. The fraction of operators that work full time. The use of part-time operators enables a system to alleviate some of the problems arising from its peaking characteristics, and it would be expected that unproductivity would increase as this fraction approaches 1.0.

Results Obtained

For these as well as all other variables, the Statistical Package for the Social Sciences (SPSS) was used to test regression equations of various linear and nonlinear functional forms. Variables were checked for multicollinearity problems and they were included in the equations only if they entered at a 0.05 level of significance or better. The number f cases (N) is 86 for all regressions, and the standardized regression coefficient along with the F-value of each independent variable are presented in brackets and parentheses, respectively. The equations that predict best the nine unproductivity indices are the following:

$$\begin{gathered} Y1 = 0.342 + \underset{\substack{\{0.53\}\\(33.7)}}{0.705\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.21\}\\(5.4)}}{0.119\cdot 10^{-4}\cdot X3} + \underset{\substack{\{0.30\}\\(16.2)}}{0.609\cdot X4} \\ R^2 = 0.561\ (\text{adjusted} = 0.545) \end{gathered} \tag{1}$$

$$\begin{gathered} Y2 = 0.386 + \underset{\substack{\{0.53\}\\(36.8)}}{0.653\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.24\}\\(7.5)}}{0.125\cdot 10^{-4}\cdot X3} + \underset{\substack{\{0.30\}\\(17.1)}}{0.556\cdot X4} \\ R^2 = 0.593\ (\text{adjusted} = 0.578) \end{gathered} \tag{2}$$

$$\begin{gathered} Y3 = 0.369 + \underset{\substack{\{0.54\}\\(37.6)}}{0.669\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.24\}\\(7.4)}}{0.124\cdot 10^{-4}\cdot X3} + \underset{\substack{\{0.30\}\\(17.4)}}{0.574\cdot X4} \\ R^2 = 0.596\ (\text{adjusted} = 0.581) \end{gathered} \tag{3}$$

$$\begin{gathered} Y4 = -0.423 + \underset{\substack{\{0.39\}\\(23.2)}}{0.367\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.36\}\\(20.7)}}{0.694\cdot 10^{-2}\cdot X2} \\ + \underset{\substack{\{0.25\}\\(10.3)}}{0.353\cdot X4} + \underset{\substack{\{0.16\}\\(4.1)}}{0.394\cdot 10^{-4}\cdot X5} \\ R^2 = 0.587\ (\text{adjusted} = 0.567) \end{gathered} \tag{4}$$

$$\begin{gathered} Y5 = -0.394 + \underset{\substack{\{0.55\}\\(42.9)}}{0.641\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.18\}\\(5.6)}}{0.417\cdot 10^{-2}\cdot X2} \\ + \underset{\substack{\{0.20\}\\(5.8)}}{0.982\cdot 10^{-5}\cdot X3} + \underset{\substack{\{0.17\}\\(5.1)}}{0.292\cdot X4} \\ R^2 = 0.640\ (\text{adjusted} = 0.622) \end{gathered} \tag{5}$$

$$\begin{gathered} Y6 = -0.374 + \underset{\substack{\{0.55\}\\(40.7)}}{0.645\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.20\}\\(6.5)}}{0.463\cdot 10^{-2}\cdot X2} \\ + \underset{\substack{\{0.17\}\\(4.1)}}{0.854\cdot 10^{-5}\cdot X3} + \underset{\substack{\{0.17\}\\(5.4)}}{0.310\cdot X4} \\ R^2 = 0.625\ (\text{adjusted} = 0.607) \end{gathered} \tag{6}$$

$$\begin{gathered} Y7 = -0.410 + \underset{\substack{\{0.40\}\\(23.4)}}{0.360\cdot 10^{-2}\cdot X1} + \underset{\substack{\{0.35\}\\(19.2)}}{0.654\cdot 10^{-2}\cdot X2} \\ + \underset{\substack{\{0.25\}\\(10.4)}}{0.346\cdot X4} + \underset{\substack{\{0.16\}\\(4.0)}}{0.378\cdot 10^{-4}\cdot X5} \\ R^2 = 0.581\ (\text{adjusted} = 0.567) \end{gathered} \tag{7}$$

$$
\begin{aligned}
Y8 = -0.387 &+ \underset{\substack{\{0.55\}\\(42.3)}}{0.629 \cdot 10^{-2} \cdot X1} + \underset{\substack{\{0.16\}\\(4.6)}}{0.376 \cdot 10^{-2} \cdot X2} \\
&+ \underset{\substack{\{0.21\}\\(6.1)}}{1.00 \cdot 10^{-4} \cdot X3} + \underset{\substack{\{0.17\}\\(5.1)}}{0.288 \cdot X4} \\
&R^2 = 0.612 \text{ (adjusted = 0.598)}
\end{aligned} \qquad (8)
$$

$$
\begin{aligned}
Y9 = -0.367 &+ \underset{\substack{\{0.55\}\\(39.8)}}{0.632 \cdot 10^{-2} \cdot X1} + \underset{\substack{\{0.16\}\\(5.4)}}{0.419 \cdot 10^{-2} \cdot X2} \\
&+ \underset{\substack{\{0.21\}\\(4.4)}}{0.877 \cdot 10^{-5} \cdot X3} + \underset{\substack{\{0.17\}\\(5.3)}}{0.305 \cdot X4} \\
&R^2 = 0.618 \text{ (adjusted = 0.599)}
\end{aligned} \qquad (9)
$$

where

X1 = passenger fares/platform-hour,
X2 = duration midday x fraction of full-time operators x high peak-to-base ratio,
X3 = county income per capita (1981),
X4 = vehicles high peak/vehicles low peak, and
X5 = average monthly earnings for city employees (adjusted for system FYs).

The signs of the various regression coefficients indicate that the factors hypothesized to influence the unproductivity factor really do so, although not all of them managed to be included in the equations. The relatively low coefficients of determination (R^2) are somewhat disappointing. However, this is because the explanatory variables used here did not try to forecast costs using actual contract provisions but rather to investigate the influences that the system's operating environment has on the various measures of unproductivity. Entries for Table 1 Items 2.08 to 2.11 are supplied by few systems and in extremely small amounts. This is the reason for the practically identical fits and coefficients of the equations that predict the three major unproductivity indices, although Equation 3 has a slight edge over the previous two. Considering that any work done on Items 2.08 to 2.11 is by definition only on a temporary basis and only a minute fraction of the total labor hours, it appears that Y3 should be the most appropriate unproductivity factor.

The revenue-generating ability variable, passenger fares per platform-hour, which also reflects service use intensity, proved to be the variable with the highest explanatory power for all nine indices. The service characteristics factors, duration midday, fraction of operators working full time, and high peak-to-base ratio, were represented in a single variable (X2), which had a better explanatory power than the sum of its three individual components. This probably reflects the fact that service characteristics factors have a dynamic and multiplicative influence on each other. It is interesting to note that variable X2 explains the six partial unproductivity ratios (Y4 to Y9) that focus on payment categories that are greatly influenced by the system's service profile, but it does not enter the equations explaining the three major unproductivity factors. The only service profile variable that entered into all equations was the vehicles high peak-to-vehicles low peak ratio. The hypothesis that operating labor will obtain more generous contract provisions in wealthier areas is supported by the finding that county income per capita (X3) is an explanatory variable for all three major and most of the minor unproductivity factors. In the only two cases (Y4 and Y7), where X3 does not enter the equations, another measure of area wealth, the average monthly earnings of city employees (X5), takes its place. System size variables have some individual correlation with the unproductivity factors (r = 0.25 to 0.40) but they did not enter the equations at the required level of significance. Shoulder-to-shoulder time proved to be insignificant also, possibly because of definitional ambiguities and erroneous Section 15 reporting by system operators. Private ownership or contract management dummy variables were not well correlated with the unproductivity factors.

Average Base Wage Rate (Element B)

The average base wage rate represents hourly payments for regular, straight operating time and its exact derivation was presented previously.

Factors Hypothesized to Influence this Variable

A variety of system operating characteristics, policy, and environmental factors was considered to influence the wage rate as follows:

1. City employee wages in the system's area of operation.
2. Income per capita in the county of operation.
3. Transit system size.
4. Public transportation's predominance in the area as measured by the percentage of work trips made by public transportation in the system's city and county of operation.
5. The fraction of operators that work full time. This factor influences the wage rate because of the following:
   - Part-timers may be getting paid at a lower rate;
   - The mere allowance of part-time operator use might indicate a diminished union strength, which in turn implies that full-time operators may be forced to accept lower wages; and
   - The union may allow part-timer use as a trade-off for higher full-time operator wages.
6. The ratio of line service hours to total (line plus charter and special) service hours. Driving in a regular line service environment requires more effort than charter and special services do. Operators may, therefore, ask for higher wages as this ratio increases.
7. Average vehicle capacity. The larger the vehicle an operator drives, the more likely it is that he would want to get paid more for his services.
8. Intensity of system use (or utilization) factors such as
   - Passenger-miles/vehicle-mile,
   - Passengers/vehicle-mile,
   - Passenger-miles/vehicle-hour, and
   - Passengers/vehicle-hour.
9. Regional characteristics. The Section 15 variable UMTA Population was used to indicate the size of the urban area, and, to avoid the use of dummy variables, the following were used to describe regional differences:
   - Mean July temperature (degrees Fahrenheit),
   - Mean January temperature (degrees Fahrenheit),
   - Heating degree-days in a year, and
   - Cooling degree-days in a year.
10. The system's organizational and management structure.

Results Obtained

The regression equation that best predicted the average wage rate was

$$Y10 = -11.24 + \underset{\substack{\{0.47\}\\(40.1)}}{0.242 \cdot 10^{-2} \cdot X5} + \underset{\substack{\{0.30\}\\(17.0)}}{0.515 \cdot X6} + \underset{\substack{\{0.26\}\\(14.4)}}{0.212 \cdot 10^{-3} \cdot X7} + \underset{\substack{\{0.22\}\\(10.3)}}{3.22 \cdot X8} + \underset{\substack{\{0.15\}\\(4.4)}}{9.64 \cdot X9}$$

$$R^2 = 0.648 \text{ (adjusted = 0.628)} \qquad (10)$$

where

- Y10 = average base wage rate,
- X5 = average monthly earnings for city employees (adjusted for system FYs),
- X6 = Ln (vehicles operating in the p.m. peak),
- X7 = heating degree-days per year,
- X8 = fraction of operators working full time, and
- X9 = line service hours/total service hours.

The coefficient signs of the independent variables are in agreement with the hypotheses. The natural log of vehicles operated during the p.m. peak (X6) is the only variable associated with system size that entered the equation. The positive sign of X8 indicates that wages are higher when full-time operators predominate in a system. This implies that, in Item 5 in the preceding list, the first and second hypotheses, but not the third, are correct. The variable heating degree-days per year (X7) acts as a proxy for all regional characteristics descriptors. The positive sign of its coefficient indicates that, other things being equal, systems in the north are paying higher wages. Historical reasons may be the cause of this because transit was developed first in the old, northern, industrial cities. Vehicle capacity, the intensity of use factors, and city population are all reasonably well correlated with the base wage rate ($r = 0.30$ to 0.40), but they become insignificant when entering the equation along with the other variables. The percentage of work trips made by public transportation in the city was well correlated ($r = 0.50$) with the wage rate. However, it was also correlated with other independent variables such as system size and city employee wages, and this is the reason for its exclusion from the equation. Variables describing the system's organizational and management structure were not well correlated with the wage rate.

Total Salaries per Platform Hour (Element C)

The total amount paid for salaries per platform hour is a composite variable obtained by the product of the wage rate times the unproductivity factor. This composite variable should be influenced by the same factors that influence its two component parts.

The regression equation that predicted total salaries per platform hour best was

$$Y11 = -45.68 + \underset{\substack{\{0.44\}\\(30.0)}}{0.32 \cdot 10^{-2} \cdot X5} + \underset{\substack{\{0.17\}\\(7.2)}}{3.46 \cdot X8} + \underset{\substack{\{0.17\}\\(5.7)}}{15.38 \cdot X9} + \underset{\substack{\{0.18\}\\(7.8)}}{0.740 \cdot X10} + \underset{\substack{\{0.19\}\\(7.2)}}{0.138 \cdot X11} + \underset{\substack{\{0.23\}\\(8.4)}}{3.03 \cdot X12}$$

$$R^2 = 0.720 \text{ (adjusted = 0.70)} \qquad (11)$$

where

- Y11 = total salaries per platform-hour;
- X5 = average monthly earnings for city employees (adjusted for system FYs);
- X8 = fraction of operators working full time;
- X9 = line service hours/total service hours;
- X10 = high peak-to-base ratio;
- X11 = weekday hours of operation; and
- X12 = Ln (county income per capita, 1981).

The results are as expected. Most of the variables that explain this composite ratio appeared also as explanatory variables of either or both of its two component elements (X8 and X9 for the wage rate, X5 for both, and X12 for the unproductivity factors but in a different functional form). The high peak-to-base ratio (X10) replaced X4 (and partly X2) as a service profile characteristic variable, and the weekday hours of operation (X11) entered as a substitute for X6. In addition, X11 serves as a service characteristic descriptor because it leads to higher unproductivity through the payment of night shift premiums.

Fringe Benefits per Platform-Hour (Element D)

This variable is greatly affected by the total salaries per platform-hour because most fringes (FICA, pensions, and so forth) are by legal or contractual provisions in direct proportion to paid salaries. Even fringe categories such as vacation and holiday pay will be directly proportional to wages (the partial correlation between wages per platform-hour and fringes per platform-hour is 0.8). Therefore, fringe benefits per platform-hour would be affected by the same factors that influence the unproductivity factor and the average base wage rate.

The regression equation that best predicts fringe benefits per platform-hour is

$$Y12 = -16.32 + \underset{\substack{\{0.48\}\\(48.0)}}{0.208 \cdot 10^{-2} \cdot X5} + \underset{\substack{\{0.16\}\\(5.9)}}{1.90 \cdot X8} + \underset{\substack{\{0.23\}\\(11.7)}}{12.23 \cdot X9} + \underset{\substack{\{0.32\}\\(21.2)}}{0.759 \cdot X10} + \underset{\substack{\{0.20\}\\(7.4)}}{0.088 \cdot X11}$$

$$R^2 = 0.682 \text{ (adjusted = 0.662)} \qquad (12)$$

where

- Y12 = fringe benefits per platform-hour,
- X5 = average monthly earnings for city employees (adjusted for system FYs),
- X8 = fraction of operators working full time,
- X9 = line service hours/total service hours,
- X10 = high peak-to-base ratio, and
- X11 = weekday hours of operation.

The strong relationship between total salaries and fringe benefits is confirmed by the fact that all the independent variables of Equation 12 were also used in Equation 11.

Total Operator Cost per Platform-Hour (Element E)

This is the overall composite measure of the operators' cost efficiency and it should be influenced by the combined effects of the variables appearing in all previous equations.

The regression equation that best predicts total operator cost per platform hour is

$$Y13 = -34.34 + \underset{\substack{\{0.55\}\\(75.0)}}{0.616 \cdot 10^{-2} \cdot X5} + \underset{\substack{\{0.24\}\\(12.2)}}{6.88 \cdot X8} + \underset{\substack{\{0.22\}\\(13.7)}}{26.56 \cdot X9} + \underset{\substack{\{0.19\}\\(7.9)}}{1.21 \cdot X10} + \underset{\substack{\{0.19\}\\(9.7)}}{0.273 \cdot X11} - \underset{\substack{\{0.14\}\\(4.8)}}{0.421 \cdot X13}$$

$$R^2 = 0.746 \text{ (adjusted = 0.726)} \qquad (13)$$

where

Y13 = total operator cost per platform-hour,
X5 = average monthly earnings for city employees (adjusted for system FYs)
X8 = fraction of operators working full time,
X9 = line service hours/total service hours,
X10 = high peak-to-base ratio,
X11 = weekday hours of operation,
X13 = mean January temperature (degrees Fahrenheit).

With the exception of X13 all other variables have already been used to explain some of the component parts of this final, composite, efficiency indicator. The mean January temperature entered the equation replacing the number of heating degree-days per year (X7) that was used previously as the geographic region descriptor. Although both X13 and X7 are temperature-related variables, they are negatively correlated and this is the reason why X13 enters Equation 13 with a negative sign, whereas X7 had a positive sign in Equation 10.

SUMMARY OF RESULTS

The results indicate that a major portion of the variation in the operator unproductivity factors, the base wage rate, and the consequent composite operators' efficiency ratios, can be explained by the socioeconomic, regional, revenue, and service characteristics variables that constitute the environment in which a transit system operates. These findings make it possible to make useful and meaningful comparisons among transit systems by accounting for the exogenous factors that affect their efficiency, and thus moving a step further along the difficult, crucial process, which is of interest to every transit manager, of exploring and determining the sources of unit cost variations among transit systems.

Explaining the total cost variations is, without a doubt, a much easier process than explaining unit cost variations, as the following model demonstrates:

$$Y14 = 0.757 + \underset{\substack{(0.98)\\(1,828)}}{1.148 \cdot X14} \qquad R^2 = 0.956 \text{ (adjusted = 0.956)} \tag{14}$$

$$Y14 = -0.0247 + \underset{\substack{(0.86)\\(19,524)}}{1.013 \cdot X14} + \underset{\substack{(0.20)\\(1,017)}}{1.085 \cdot X15} + \underset{\substack{(0.08)\\(181)}}{1.072 \cdot X16} \qquad R^2 = 0.999 \text{ (adjusted = 0.998)} \tag{15}$$

where

Y14 = Ln (total operators' cost),
X14 = Ln (platform-hours),
X15 = Ln (average base wage rate), and
X16 = Ln (unproductivity factor Y1).

Platform-hours explain almost all the variation in total operators' cost (Equation 14), and the addition of two more variables produces a perfect fit. However, this is a rather trivial exercise because Equations 14 and 15 can only be used for forecasting purposes and are completely useless as diagnostic tools. On the other hand, Equations 1 through 13 can be used for diagnostic purposes and can pinpoint the sources of operators' inefficiencies.

OBSERVATIONS AND CONCLUSIONS

Although the diagnostic equations presented here have reasonably good fits, their applicability, validity, and accuracy are a function of the following considerations that should be kept in mind before system comparisons are undertaken:

1. Most variables were derived from Section 15 data, the uniformity of which provides unique research opportunities. However, Section 15 data are far from perfect. Detailed examinations of each transit system had to be performed to ensure data validity (see paper by Bladikas and Papadimitriou in this Record). Missing data are a minor problem compared with possible definitional ambiguities in the data definitional elements that cause erroneous entries that are harder, and occasionally impossible, to identify.
2. Layover time--the time spent at the end of a route before the commencement of another run--is included in the platform-hours of service, and, therefore, it is counted as productive time even though vehicle and driver are idle and are not serving passengers. Layover time provides leeway for variations in the running time and is used to maintain schedule adherence. There is, therefore, a trade-off between efficiency and quality of service because long layover times practically guarantee strict schedule adherence, whereas short layover times imply a high risk of scheduling abnormalities.
3. Deadheading hours--traveling to and from the first (last) passenger stop from (to) the bus yard--are included in the platform-hours. Although deadheading hours should be minimized, the problem here is not one of driver productivity but one of route structure and garage location.
4. The wage rates calculated and used here include only 1 year of observations, although most labor contracts are in effect for more than a year. It may, therefore, make a difference if the wage rates used represent wages that a contract stipulates for its first or last year. Pooled wages from 3 consecutive years of data could be used in further research.
5. The wage rate also affects the quality of hired and retained personnel. It is safe to assume that low wages will not attract good drivers and will also induce high turnover rates, thus creating unproductive times during the training of new drivers. Employee dissatisfaction and absenteeism could also be the product of low wage rates.
6. The service characteristics variables used represented supplied and not demanded service. For example, the actual high peak-to-base passenger demand may be three, but the high peak-to-base ratio calculated from service supplied data is two. This is because transit managers find it more cost effective to run vehicles during the midday period and incur the running costs than to pay operators to remain idle.

In spite of the limitations that are inherent in these models, the results indicate that the addition of a limited number of environmental variables to the Section 15 data is sufficient to analyze the factors that influence variations in operators' costs. The variables identified and used in the equations are to a large extent and for all practical purposes outside the system operator's control. This provides the opportunity to make valid comparisons of transit systems because inefficiencies are diagnosed in terms of variables that cannot be affected by the system operators' managerial skills. However, a system operator is not completely help-less, even if the operator's efficiency is a function of mostly exogenous variables.

Using the models presented here, transit managers can determine their efficiency relative to other systems and thus know if they are above or below industry norms. Although no indexing measure has been developed to determine the exact position of a transit system among the rest, it is sufficient to know at least if a system is over- or underperforming. With the assistance of the models that diagnose each individual cost component, a system operator may take corrective action to improve efficiency in any of the cost components. Future labor contracts could be less generous with clauses that affect efficiency, and wage increases could be tied to productivity improvements. Steps can also be taken to reduce the detrimental effects of peaking characteristics by using more part-time labor, using operators for other functions during the midday, and purchasing transportation during peak periods.

It is difficult to solve or even discuss the extremely delicate problems of labor efficiency. The issue is not only highly political, it also deals with human resources that cannot be manipulated or treated like inanimate objects. However, in view of the financial difficulties of the transit industry, it behooves both labor and management to improve the system's efficiency. The proper diagnosis and understanding of the problem is in the best interest of all parties concerned with the viability and survivability of public transportation.

ACKNOWLEDGMENT

The research the results of which are presented in this paper was supported partly by a grant from the Eno Foundation.

REFERENCES

1. D.H. Pickrell. Sources of Rising Operating Deficits in Urban Bus Transit. In Transportation Research Record 915, TRB, National Research Council, Washington, D.C., 1983, pp. 18-24.
2. S.C. Anderson and G.J. Fielding. Comparative Analysis of Transit Performance. UMTA-CA-11-0020-82-1. Institute of Transportation Studies and School of Social Sciences, University of California, Irvine, Jan. 1982.
3. J.S. Dajani and G. Gilbert. Measuring the Performance of Transit Systems. Transportation Planning and Technology, Vol. 4, No. 2, Jan. 1978, pp. 97-103.
4. National Urban Mass Transportation Statistics: 1982 Section 15 Annual Report. UMTA-MA-06-0107-84-1. UMTA, U.S. Department of Transportation, Nov. 1982.
5. Urban Mass Transportation Industry Uniform System of Accounts and Records and Reporting System: Volume II, Uniform System of Accounts and Records. UMTA, U.S. Department of Transportation, 1977.
6. G.J. Fielding, R.E. Glauthier, and C.A. Lave. Development of Performance Indicators for Transit. UMTA-CA-11-0014-78-1. University of California, Irvine, Dec. 1977.
7. K.C. Sinha and B.D. Dorby, Jr. A Comprehensive Analysis of Urban Bus Transit Efficiency and Productivity: Part II, Labor Aspects of Urban Transit Productivity. UMTA-IN-11-0003-79-3. School of Civil Engineering, Purdue University, West Lafayette, Ind., Dec. 1978.
8. G.M. Giuliano. Transit Performance: The Effect of Environmental Factors. Ph.D. dissertation. University of California, Irvine, 1980.
9. D.T. Barnum and J.M. Gleason. Measuring the Influence of Subsidies on Transit Efficiency and Effectiveness. UMTA-NE-11-0002-80-4. Indiana University, Gary; Nebraska University, Omaha, June 1979.
10. J. Pucher. Redesigning Federal Transit Subsidies to Control Costs and to Increase the Effectiveness of the Transit Program. UMTA-NJ-11-0011-82-2. Department of Urban Planning, Rutgers University, New Brunswick, N.J., Dec. 1982.
11. H.G. Wilson. Forecasting Bus Transit Operating Costs. UMTA-PA-11-0010. Pennsylvania Transportation Institute, Pennsylvania State University, University Park, Nov. 1975.
12. County and City Data Book: A Statistical Abstract Supplement. Bureau of the Census, U.S. Department of Commerce, 1983.
13. Government Employment (Series GE): No. 2, City Employment in 1980, (and 1981, 1982). Bureau of the Census, U.S. Department of Commerce, 1980-1982.
14. A Directory of Regularly Scheduled, Fixed Route, Local Public Transportation Service in Urbanized Areas Over 50,000 Population. UMTA, U.S. Department of Transportation, Aug. 1981.
15. N.H. Nie et al. Statistical Package for the Social Sciences, 2nd ed. McGraw-Hill Book Co., New York, 1975.

Getting Control Over Operating Budgets: Methodology for Evaluating Productivity Alternatives

JEFFREY E. PURDY, DOUGLAS W. CARTER, and WILLIAM R. STEINMETZ

ABSTRACT

Productivity improvements have become a key element in the struggle to preserve the integrity of transit service in an environment of diminishing public funding. The need for increasing operating efficiency through productivity improvements is not unique to transit operators in the United States. Transit systems in the United Kingdom are also being pressed to realize operating savings by reductions in funding from the central and regional governments. One agency, London Transport-Rail, has incorporated productivity improvement programs as a fundamental element of daily operations. In doing so, they have encountered a number of problems related to productivity project development, analysis, and control. These problems provided the impetus for conducting a study that developed a standardized process and analytic technique for preparing a productivity program. The methodology, detailed in this paper, allows for consistent analysis and evaluation of a wide variety of productivity tasks that require varying degrees of capital or operating expenditures, or both, and savings over incongruent time periods. The analytic technique is simple to apply and uses net present value analysis to compare diverse projects. The standardized productivity development process allows anyone in the organization to present a "good idea" for improving efficiency and provides a consistent evaluation of these ideas to facilitate sound decision making by top management.

Operating in an era of diminished public funding, transit agencies are pressed to find ways to reduce or constrain expenditure of limited operating dollars. Many systems facing severe financial constraints have already initiated substantial service reductions and fare increases to balance costs with available revenues. Although this response may be appropriate for averting an immediate crisis, it often serves to undermine transit's goal of meeting local transportation needs and contributes to the downward spiral of service utilization. Transit managers cannot rely on service and fare changes as the sole means of budget control and are now seeking improvements in productivity as a means to balance budgets while preserving service integrity.

Productivity options take a number of different forms, each with unique requirements in terms of cost and time. Transit systems may use a capital investment to reduce operating costs. Procurement of computers and automation of labor-intensive activities have proven successful in many agencies. Other systems have engineered improvements of equipment and capital to reduce operating and maintenance requirements. Transit systems may also invest operating dollars to gain productivity improvements. Such investments may focus on staff training or purchasing improved materials, which are then leveraged to achieve even greater savings. Other productivity programs may focus on tactical changes such as disbanding an internal function and contracting those services out. There are many options available to transit agencies for improving resource utilization ranging from the sublime to the ridiculous. The difference between one extreme and the other often translates to what program will bring the greatest net return (i.e., reduction in overall expenditures) if implemented. This relationship is not always apparent because productivity options require varying degrees of investment of funds and energy and provide different returns over incongruous time periods.

Transit systems need a structured approach for identifying and evaluating productivity options to meet successfully the current financial challenge. The results of a study that developed a standardized process for identifying, reviewing, and evaluating productivity options for London Transport-Rail (LT-Rail) are described. Although the methodology was developed to address the specific needs of LT-Rail, many parallels exist with agencies in the United States, and the evaluation framework has broad applications elsewhere.

CASE STUDY: LONDON TRANSPORT-RAIL

LT-Rail has fully recovered operating expenditures from passenger fare revenues (i.e., achieved a 1.00 fare-box recovery ratio) during the last decade. However, during the past 2 years, achieving a favorable fare-box recovery ratio (i.e., the percentage of operating expenses met with passenger revenues) has been increasingly difficult. Factors that have been contributing to the problem include

- Declining central London population and employment,
- Increasing staff levels due in part to a shorter work week,
- Operating costs that are increasing at a pace that exceeds the Retail Price Index, and
- A tendency to increase fares to compensate for increased costs and declining passenger levels.

At the same time, the LT-Rail infrastructure was increasingly in need of major capital investment. Stations, trackform, workshops, and the fare collection system were in need of modernization and attendant infusion of capital investment funds. However, because total funding for transport has been limited, a new approach for revitalizing the LT-Rail system was needed.

In May 1983 the LT-Rail directors approved their first comprehensive strategic plan. The plan committed LT-Rail to an ambitious capital investment program that would modernize and revitalize the huge underground rail system. The key to the program's success was the more efficient use of capital funds and the reduction of operating expenditures to be achieved through new procedures and reduced staffing levels wherever possible. One area of obvious opportunity for reduced operating cost was the maintenance budget, which currently stands at £100 million ($130 million) per year in expenditures for maintenance activities. The plan required a reduction in this operating expenditure of 2 percent per year, or roughly £2 million ($2.6 million) per year, for 2 years. The challenge was to achieve this operating cost reduction while improving the condition of the infrastructure through judicious capital investment and maintaining the quality standards for which LT-Rail is famous worldwide.

Two types of initiatives are now being undertaken by LT-Rail to implement their approved strategic plan. The first is capital projects that are developed by a formal process conducted by dedicated staff, including financial evaluation and top management review and direction resulting in successful project implementation. The second set of initiatives is productivity tasks for maintenance and engineering functions. The objective of these tasks is to reduce the operating expenditure for rail maintenance through a wide variety of procedural, organizational, and operational changes. Departments are given targets in the strategic plan for reducing operating budgets and are directed to develop programs to reach those targets. Productivity options are developed by any number of staff members in a given department. Some options are a response to a visible problem (e.g., an equipment reliability problem), although most are the result of a "good idea" generated by someone who has identified a better (i.e., more efficient) way to perform existing functions. There are now more than 300 productivity tasks that were developed to reduce maintenance expenditure. The implementation of these tasks is under way and has contributed more than £2 million per year in cost reduction to date. The challenge is to keep these tasks progressing and contributing to operating cost reduction.

A system entitled value analysis was developed to assist the manager of each productivity task in the development and implementation of his productivity tasks. Many of these tasks require significant capital investment to achieve the operating expenditure reduction desired. Moreover, productivity tasks generally have alternative approaches that must be evaluated. All productivity tasks require review by several layers of the organization, including the line manager responsible for day-to-day operation, financial staff, and top management. LT-Rail's experience in developing capital projects needed significant improvement to meet the ambitious objectives of the new strategic plan. The objective of reducing maintenance expenses required a new process that was efficient and easy to use for the managers involved and the reviewers. The resulting value analysis system, in both its manual and microcomputer versions, meets these objectives.

INITIAL PROBLEM

Although the response to the need for productivity improvements has been tremendous, several problems have occurred in comparing potential tasks and developing implementation priorities. The problems fall into three areas: the process, the analysis, and the documentation of productivity tasks. Each of these is discussed hereafter.

The approach to productivity task development is decentralized--virtually any staff member can suggest a potential improvement. The review and authorization of productivity tasks, however, are centralized and require that top management review each task. At the onset, there was no accepted standard for developing tasks or soliciting input into task development by affected parties. Poor communications in the initial proposal development process led to incomplete identification of alternative ways to implement a project and to interpersonal conflicts when affected parties were not consulted during proposal development. Further, the staff members responsible for developing projects did not fully comprehend finance department and top management expectations in terms of alternatives development, analysis, and documentation. Several engineers found the productivity proposal process to be frustrating and could not determine the criteria that the finance department used to evaluate projects. The staff members responsible for proposal development were becoming increasingly frustrated by the seemingly ad hoc productivity development and review process.

In addition to inadequate communications, the analysis of alternatives varied considerably among departments and individuals. Capital and operating costs (savings) were estimated at various levels of aggregation and accuracy. The time period required to implement programs, and realize the savings, was not typically considered in the analysis. Also, no indication of the relative accuracy or range of potential savings was shown in proposals. The result was that managers responsible for project authorization and cost reduction were comparing incomparable numbers to make key productivity decisions.

Another problem occurred in documenting projects. There was no accepted format for documentation and even the best projects often lacked sufficient documentation for authorization. Two common problems were that alternative implementation plans that were dropped during the analytical phase and relevant assumptions or constraints were omitted from the proposals. Frequently, reviewers would request that additional analysis be done on those alternatives already dismissed by the proposal author; this resulted in substantial delays.

The deficiencies inherent in the initial productivity development process, shown in Figure 1, led to substantial problems at all levels of the productivity program. The results included extensive delays in authorizing projects and failure to realize the maximum savings anticipated from many projects. Recognizing the shortfall, LT-Rail and their consultant conducted a study to revise and standardize the productivity evaluation program.

PROJECT OBJECTIVE

The objective of the project was to develop a productivity process and analytic methodology that

- Assesses capital, operating, and maintenance costs of productivity projects;
- Applies to capital investment, tactical (e.g., contracting out), and productivity proposals;
- Is consistent with and complementary to existing finance department principles and practices;
- Produces reliable and comparable results through consistent analysis of all projects;

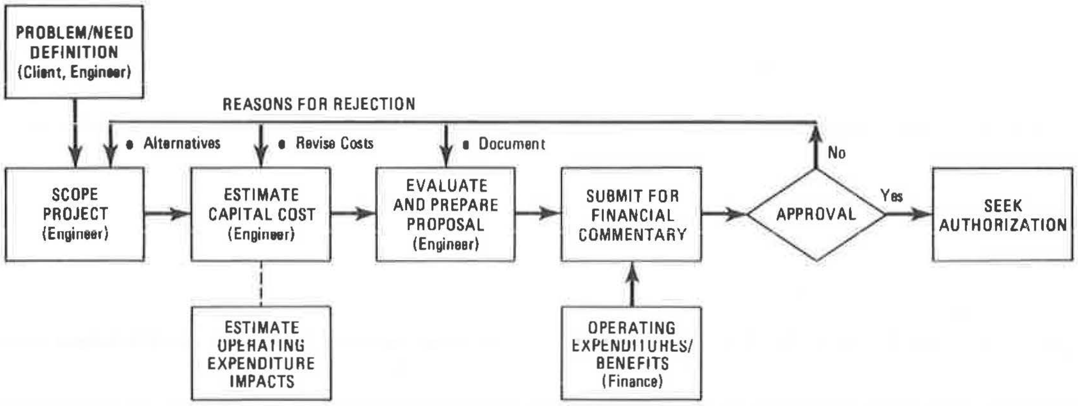


FIGURE 1 Overview of initial task/project development process.

- Promotes effective alternative project identification and analysis; and
- Is easy to use and understand.

To satisfy these diverse criteria and address the problems identified in the review of existing procedures, a two-level approach to value analysis was formulated.

VALUE ANALYSIS FRAMEWORK

The analytical process and procedure developed to address LT-Rail's need was entitled "value analysis." The process consists of two levels, a preliminary and a final value analysis, as shown in Figure 2.

The value analysis methodology has several key features:

- Capital and operating cost and savings implications of engineering productivity projects are expressed in whole-life costs, or net present value;
- Generation of alternative implementation strategies is encouraged early in the process through use of a quick and simple level of analysis;
- Input from the finance department and other affected parties is solicited early in the process; and
- The value analysis technique serves as a standard to guide all productivity projects, thus ensuring consistent and comparable program analysis.

The preliminary analysis step is intended to develop many alternatives in brief detail without developing rigorous specifications for the task's operating characteristics. Aggregate cost estimates are developed for capital requirements and the impact on operating expenses. A simplified net present value analysis, which assumes an even annual cash flow, is performed so that alternatives can be compared and evaluated. The preliminary proposal is submitted for initial review by all affected parties and identifies the most promising alternatives for inclusion in the final analysis. Reviewers may pro-

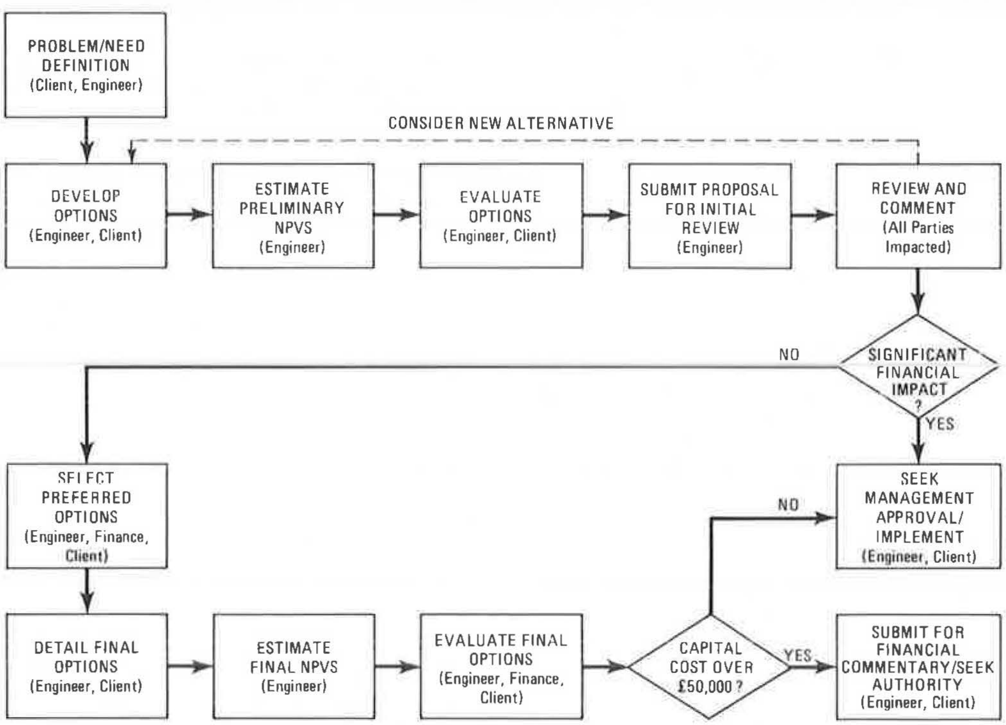


FIGURE 2 Value analysis framework.

vide new alternatives or suggest a revision to existing options before the detailed analysis begins.

The final value analysis structure is similar to that of the preliminary value analysis except that it requires a more rigorous analysis of fewer alternatives. The procedure commences with the selection of preferred options that are arrived at by project review with the affected parties.

The preferred option specifications are detailed and comprehensive capital and operating cost savings are estimated. A net present value analysis is performed for each preferred option and incorporates implementation schedules for realization of capital and operating costs and ensuing cost reductions. The results are compared, a recommendation is developed, and a final proposal is written and submitted for financial commentary before authority is sought.

The financial commentary is currently a step in the existing procedure and acts as a check to ensure that the estimates are reasonable and the analytic techniques are acceptable. In this activity, the finance department serves as an "outside auditor" or watchdog over financial estimates. Following financial commentary, the proposal is submitted to top management for the authorization decision.

APPLICATION OF THE PROCEDURE

The value analysis procedure is applicable to virtually all types of capital investment, tactical, and productivity projects. By using net present value analysis, capital investment, tactical, and productivity alternatives can be equitably compared with each other to optimize the return of investment to the transit system. The procedure should be applied to all projects that realize both capital cost and operating expense or savings. It is particularly important to use the procedure with projects that have phased capital costs or uneven annual streams of operating expense or savings, or both.

However, the procedure does not necessarily apply to all projects. The procedure need not apply to projects that realize constant annual operating cost reductions without incurring capital expenses. An additional case in which the procedure may not be appropriate is when the cost savings of a particular project are relatively small (i.e., the level of effort required might exceed the benefit received).

Flexibility can be built into the procedure. The preliminary value analysis should be applied to all projects regardless of the magnitude of savings. However, a cut-off point can be established to avoid excessive work on minor productivity issues. The cut-off point could be a net present value or capital cost threshold that, if exceeded, would require application of the final value analysis step. If the threshold is not exceeded, recommendations would be developed from the results of the preliminary value analysis.

In the study for LT-Rail, a threshold of £25,000 for operating cost (savings) or a capital cost (savings) of £50,000, or both, was determined as having a significant impact on the organization. Below these thresholds only the preliminary value analysis is required. The full two-level value analysis is to be applied to all projects that exceed these cost or savings figures.

PRELIMINARY VALUE ANALYSIS PROCEDURES

The preliminary value analysis focuses on developing and screening productivity task options. The preliminary analysis commences with definition of a problem or need, or simply a good idea for doing things more efficiently. On the basis of the need defined, the task manager must develop options for meeting the need. The task manager can rely on his own ingenuity, assistance from affected parties, and knowledge of similar situations to develop an initial set of options.

Options are alternative approaches to solving productivity problems and reducing operating expenditures. They may involve a capital expenditure to achieve a greater operating expense reduction and involve an immediate or phased-in implementation schedule. Options can include technology investments, organizational change, incentive schemes, new production techniques, or procedural revisions. Any process can be considered an option at this stage of the analysis regardless of reasonableness or feasibility (qualitative measures of cost-effectiveness). The most obvious and common solution may not be the most cost-effective.

The process of option generation relies on the different engineering disciplines and experience, which often bring varying solutions and contributions. The process of developing options encourages the generation of the "bright idea" and unconventional approaches to problem resolution.

When options have been identified, the task manager proceeds to conduct the preliminary financial analysis, which is relatively uncomplicated and requires a nominal effort to apply properly. The operating savings and capital cost estimates are provisional in nature and are intended for option comparison and screening purposes. The preliminary financial analysis is guided by a series of standard procedures. The result of the preliminary analysis is an estimate of the net financial impact, and value range, for each option.

Options are then compared and evaluated by the task manager and the client. Evaluation should focus on both the financial and the nonfinancial implications of each option. The annual operating cost reductions, capital cost, and net present value provide a sound basis for financial evaluation. Anticipated impacts on organizational effectiveness must be considered and documented as well. The initial proposal and analysis documentation are completed and submitted for review by other interested parties (e.g., department manager and finance director).

The preliminary value analysis is intended to be an uncomplicated, but sound, financial evaluation of task options. The analysis is guided by four standard procedures:

- Capital cost estimation,
- Incremental operating cost/savings,
- Net present value calculation, and
- Initial proposal documentation.

Each procedure develops a specific element of the preliminary value analysis that, when completed, provides adequate documentation of the analytic result.

Capital Cost Estimation

The procedure for preparing the preliminary capital cost estimate assumes that all capital costs will occur in the first year of project implementation. High and low capital costs are estimated to define a value range. The specific requirements for estimating cost are as follows:

- On the basis of the options identified, major and ancillary equipment items are to be defined in terms of quantity and unit cost, the product of which is total cost.

- Installation and site costs may be estimated using a single value amount (i.e., lump sum).
- Training and manuals may be needed for some task options. These labor and materials costs can be estimated in one lump sum.
- The preliminary capital cost estimate should include a single amount for the initial supply of maintenance materials and supplies (e.g., spare parts).
- Each capital cost area is summed for best and worst cases.
- The useful life of the equipment is required for net present value analysis.
- All sources should be listed and may include telephone inquiries to sales outlets, manufacturers' price lists, examination of similar projects, and previous experience.

The results of the capital cost estimate for each option are documented for inclusion in the net present value analysis and in the initial proposal.

Annual Incremental Operating Cost

The procedure for preparing the preliminary incremental operating cost/saving estimate is conceptually simple. The operating cost estimate represents the expected change in current annual operating costs if the proposed option were implemented. High and low estimates of the expected effect on revenue expenditures are to be prepared and recorded for each option. The specific requirements for estimating the change in revenue expenditures are

- Direct labor costs are estimated on the basis of the change in the number of employees by position and average annual salary for the respective position.
- An estimate should be made for those elements in overhead expenses that vary with the number of staff employed or person-hours worked, and the variable overhead factor should be determined.
- Total labor expense/savings is calculated by summing direct labor costs for all positions and multiplying the sum by the variable overhead factor.
- The change in the cost of materials and services can be expressed as a lump sum based on past experience, or built up based on major elements (e.g., spare parts, utilities, paper and ribbons, bearings).
- The expected change in total expenses is found by summing total labor and materials costs.

The present value of the expected change in annual revenue expenditures is also calculated. Both best and worst case revenue estimates are analyzed as follows:

- The previously determined expected annual operating savings/expense is inserted.
- The analysis period is identified consistent with the longest useful equipment life if a capital expenditure is involved in any task option. If no capital expense is required for any option, the analysis period can be relatively short (e.g., 5 years). The same analysis period must be used for all options in a single task.
- The simple present value factor for an annuity is determined from standard annuity tables and must correspond with the analysis period.
- The present value of operating expense/savings is the product of annual operating expense/savings and the present value factor.

The results of the present value analysis and preceding steps are documented. The sources for each operating cost estimate are recorded. This will assist in answering reviewer comments and in refining the preliminary estimates in the final value analysis.

The discount (i.e., interest) rate has been established as a matter of policy at LT-Rail. The figure used by London Transport is 7 percent and represents the actual cost of money to LT, given its sources of funds and avenues for borrowing and investing public monies. U.S. agencies faced with determining a discount rate might consider using current interest rates for U.S. Treasury Bills, which include some allowance for inflation, or some similar rate to determine the cost of funds. It is important to note that the LT-Rail value analysis procedure is applied by individual engineers, and as such does not include an algorithm for incorporating inflation into productivity estimates. The finance department, however, does use a complex computer model to incorporate the anticipated implications of inflation in the actual budgeting process.

Net Present Value

The net present value for a given option is defined as the sum of capital costs and the present value of the change in annual operating expenses over the effective life of the task. The procedure for determining net present value is simple and relies on estimates developed in the preceding two steps. The steps in the analysis are as follows:

- The mean of the capital cost estimate is found by summing the best and worst case capital cost and dividing by two.
- The mean of the present value of operating cost/savings is found by summing the best and worst case operating expenses and dividing by two.
- The net present value for best, worst, and mean cases is found by summing capital cost and the present value of operating expenses.
- The value range from the mean is found by dividing either best or worst case net present value by the mean and expressing the result as a percentage difference from the mean.
- The average annual operating expense/savings is determined by summing best and worst cases and dividing by two.

When the net present value of each option has been calculated, an evaluation of each alternative is performed and the initial proposal summary is prepared. The proposal summary indicates which options are recommended for final value analysis.

Initial Proposal Summary

The initial proposal summary briefly presents an overview of the task development and analysis results, as shown in Figure 3. The key issues reported are

- A brief summary of the problem, need, or objective that the task options try to address.
- The option or options recommended by the task manager for final value analysis or management approval should be noted. Reviewer recommendations are required as well.
- A summary of each option is to be recorded, including option title, mean capital cost, mean annual operating expense change, and mean net present value. The task manager is to fill in the comments

INITIAL PROPOSAL FOR

____

(PROJECT TITLE)

TO DEPT/REVIEWER: ______________ TASK NO: ______
DEPT/TASK MGR: ______________ PROJECT NO: ______
DEPT/CLIENT: ______________ DATE: ______

PROBLEM STATEMENT:

OPTIONS RECOMMENDED FOR FINAL VALUE ANALYSIS/MANAGEMENT APPROVAL
STAFF: 1 2 3 4 5
REVIEWER: 1 2 3 4 5

OPTION SUMMARY

OPTION 1:	Capital Investment £______
Annual Rev. Exp. Change £______	Net Present Value £______
Comments:	

OPTION 5:	Capital Investment £______
Annual Rev. Exp. Change £______	Net Present Value £______
Comments:	

FIGURE 3 Initial proposal summary.

section with a brief description of the option and any pertinent nonfinancial issues related to the option.

The initial proposal summary, with the documentation developed at each step of the preliminary value analysis procedure, is to be circulated to appropriate reviewers (e.g., department managers, client, finance director) to solicit their comments and suggestions.

FINAL VALUE ANALYSIS PROCEDURES

The final value analysis is intended to be a detailed evaluation of a limited number of potentially beneficial options. It builds on, and refines, the preliminary value analysis. The documentation from the preliminary analysis will serve as the starting point for further option evaluation. The final analysis procedure is guided by four standard procedures:

- Estimate final capital cost,
- Estimate final incremental operating expense,
- Calculate final net present value, and
- Develop and submit final proposal summary.

Each procedure provides detailed documentation of the analysis and results. The information on cost/savings, tasks, and implementation schedules is to be of a quality appropriate for inclusion in operating budgets, formal cost estimates, and project implementation.

The approach requires a disaggregate cost/savings buildup with documentation on sources of estimates. The disaggregate procedure incorporates phased capital expenditures and lags in realization of operating expense reductions. The net present value analysis combines the lifetime incremental operating and capital cost/savings in a single amount.

Final Capital Cost Estimate

The final capital cost estimate is to be built up from individual line item categories (e.g., major equipment items, site costs). It is intended to produce a cost figure suitable for budgeting purposes, inclusion in a formal cost estimate, and revision of capital and operating estimates for productivity within the monitoring system. The analysis procedure begins with a review of the preliminary capital cost estimate and culminates with completion of the final capital cost estimate.

A separate capital cost estimate should be prepared for best and worst cases when any significant degree of uncertainty exists in either quantity or cost elements of the estimate. In cases in which the option is well defined and the degree of uncertainty is minimal, a most likely estimate alone will suffice. The specific procedures for final capital cost estimation follow:

- The major equipment items listed in the preliminary capital cost estimate should be examined and revised, as appropriate. The project manager should examine and refine the level of detail for line items (e.g., a finer breakdown of equipment items may be appropriate), the quantities of equipment by type, and unit costs, as needed.
- The site and installation costs should be reviewed and disaggregated, as appropriate. It is often easier to develop a more accurate cost estimate by examining individual cost elements instead of a single lump sum.
- The training and manuals cost, previously addressed as a lump sum, should be itemized to allow a better cost estimate. The number of person-days required in training and required training materials (e.g., manuals) may be an appropriate level of detail.
- The initial supply of materials should be refined, as needed. It may be appropriate to note major maintenance materials and supplies by category.
- The total capital cost is found by summing cost estimates for each element.
- The useful equipment life should be reviewed and revised, as appropriate. This figure is to represent the number of years the equipment is expected to be used and maintained in a cost-efficient manner. This may well be shorter than the total life expectancy of the hardware if at some point it becomes too expensive to maintain or too costly to operate.

The level of detail contained in the final capital cost estimate is the most significant difference from the preliminary capital cost estimate. In addition, capital cost phasing, or the capital cost realization schedule, is also valued in the final estimate. This is particularly important for projects with large capital costs that require new facilities and, therefore, take time for construction and for the procurement of equipment.

Final Incremental Operating Expense Estimate

The final incremental operating expense/savings estimate is to be built up from individual line item categories and should be suitable for decision-making and budgeting purposes. The analysis procedure begins with a review of the preliminary estimate. Like the capital cost estimate, the operating cost estimate must be time relevant; that is, if it requires more than 1 year to realize the maximum revenue savings potential (e.g., labor reductions through attrition or if annual operating cost reductions are expected to decrease over time), this must be reflected in the cost estimate. If any significant degree of uncertainty exists, best and worst cases should be estimated. The specific procedures for the final incremental operating expense estimation are as follows:

- Direct labor should be estimated by position and by the period when the savings/expense occurs. If a labor savings or expense occurs partway through the year, this should be indicated. The analysis should use 13 equivalent (i.e., 4-week) periods per annum as the basis for allocation of labor savings (i.e., an employee expected to leave the rail business after six periods would be indicated by multiplying one employee by 6/13 by the annual salary for that position to determine the annual cost reduction).
- Allowance should be made for those elements in overhead expenses that vary with the number of staff employed or person-hours worked. This should be done with the department's costing office.
- Total labor expense/savings is calculated by summing direct labor costs for all positions and multiplying the sum by the variable overhead factor.
- The change in annual materials and services costs should be itemized to facilitate an accurate cost buildup. Cost categories might include maintenance materials, office supplies, utilities, and fuel.
- Other costs should be estimated by item as well. An example of a cost in this category might be the cost of contracting out maintenance services or the use of a typing service.
- Total annual expense/savings is the sum of the previously mentioned operating expenditures.
- All quantity and unit cost elements should be accompanied by a source reference to ensure accountability.

Following these steps and recording the analyses, assumptions, and sources of data will result in documentation that supports project recommendations in the final proposal summary.

Calculation of Final Net Present Value

The net present value for a given option is defined as the sum of capital costs and present value of the change in annual operating expenses over the effective life of the task option. The procedure for determining the net present value is relatively straightforward, though somewhat repetitive. Information needed for the calculation is obtained from the previous final estimates of capital costs and incremental operating expenses. The procedure for calculating final net present value, shown in Figure 4, is as follows:

- From the final capital cost estimate, extract the capital cost by year. The first year is expressed in current dollars and all the remaining years are in present value terms as calculated in the estimate.
- From the final incremental operating cost estimates, enter the cost estimates by year up to the effective life of the project. Each amount is expressed in current dollars calculated in the estimate.
- Determine the simple present value factors from appropriate financial tables made available for the analysis, for years 1 to 20. To simplify the analysis, incremental costs/savings occurring in the years beyond the 20th year are treated as uniform annual amounts or annuities. Select the annuity present value factor to express the value of the annuities in the 20th year.
- Calculate the present value of annual incremental operating costs/savings by multiplying the simple present value factor. For years 21 and beyond, multiply the costs/savings by the annuity present value factor and then by the simple present value factor for a lump sum payment in the 20th year.
- Calculate the net present value by summing all capital costs and operating costs and then subtracting total operating costs from the total capital costs.

The results of the net present value should be documented, showing all steps taken, analysis period, and interest (i.e., discount) rate used. If best and worst cases are developed, net present value should be calculated for both along with the

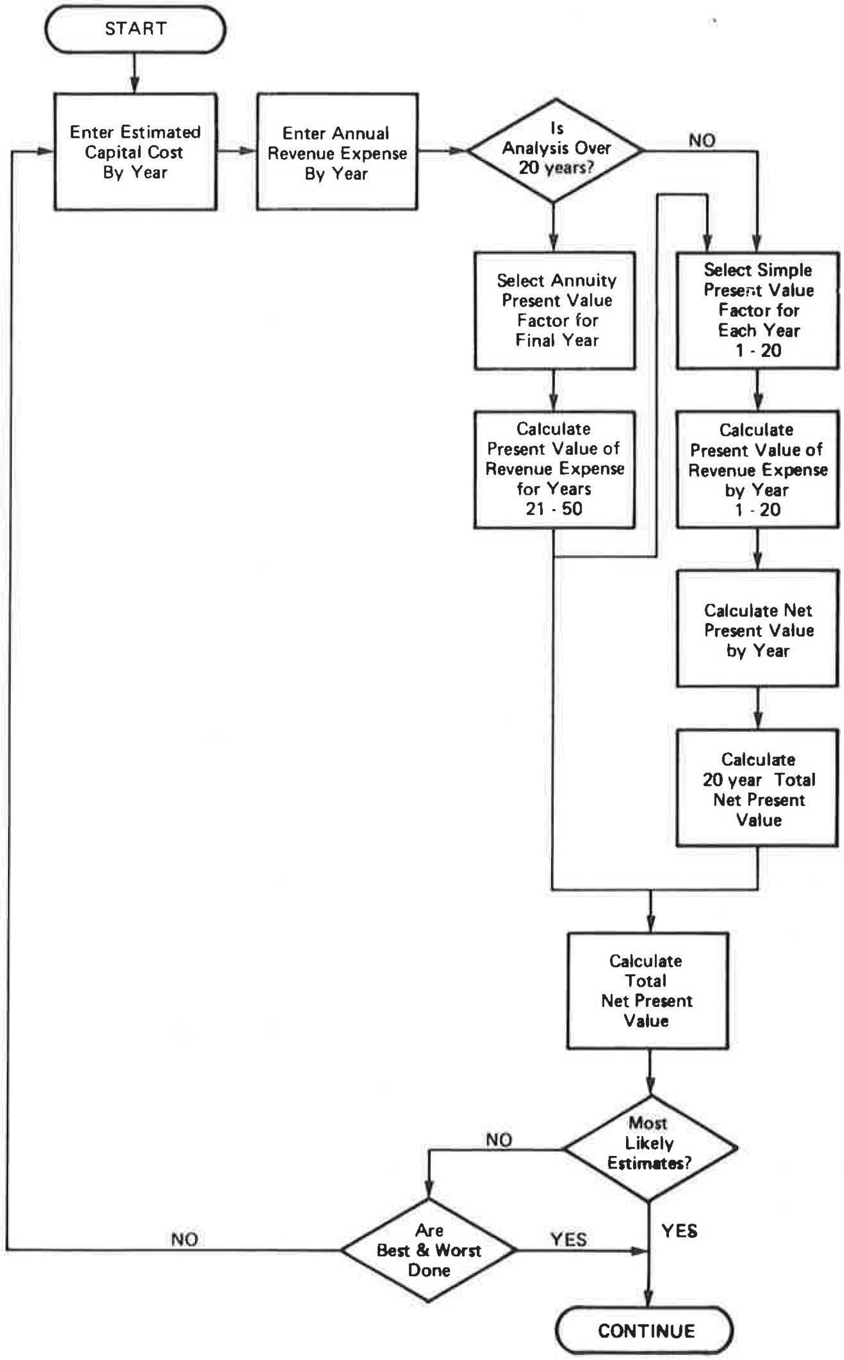


FIGURE 4 Procedure for calculating final net present value.

mean net present value. A value range, which represents the percentage difference from the mean, should then be calculated and recorded.

Final Proposal Summary

The final proposal summary briefly presents an overview of the complete value analysis, from the preliminary value analysis through the final value analysis. The key issues reported are

- A brief summary of the problem, need, or objective that the task options are intended to address.
- The action recommended by the project manager for management approval, for authorization, or for both.
- The schedule recommended for implementation that should correspond to the schedule developed for the capital investment and realization of revenue savings in the analysis.
- A brief discussion of the reasons for recommending one option and rejecting other options, or rejecting all options in some cases.
- A summary of each option evaluated in the final analysis is to be recorded, including option title, capital investments, maximum operating expense change, net present value, and the value range (best and worst from the net present value).

The final proposal summary, shown in Figure 5,

TO DEPT/REVIEWER: TASK NO:
DEPT/TASK MGR: PROJECT NO:
DEPT/CLIENT: DATE:

PROBLEM STATEMENT:

ACTION RECOMMENDED:

SCHEDULE RECOMMENDED: Start Date ____________ Completion Date ____________

REASONS AND COMMENTS:

FINAL OPTIONS SUMMARY

OPTION 1: Capital Investment £ ____________

Max. Rev. Exp. Change £ ________ Net Present Value £ ________ Value Range Best £ ________ Worst £ ________

Comments:

OPTION 2: Capital Investment £ ____________

Max. Rev. Exp. Change £ ________ Net Present Value £ ________ Value Range Best £ ________ Worst £ ________

Comments:

OPTION 3: Capital Investment £ ____________

Max. Rev. Exp. Change £ ________ Net Present Value £ ________ Value Range Best £ ________ Worst £ ________

Comments:

OTHER OPTIONS REJECTED IN PRELIMINARY ANALYSIS

OPTION: REASON:
OPTION: REASON:
OPTION: REASON:
OPTION: REASON:

FIGURE 5 Final proposal summary.

should be accompanied by all supporting documentation developed during the final value analysis procedure.

BENEFITS OF THE TWO-STEP PROCESS

The value analysis of options presents unique benefits to LT-Rail, Engineering. Before this technique was developed, the engineers responsible for developing proposals and performing cost estimates were unaware of the analytic techniques used by the finance department, such as net present value analysis. The procedure removes the mystery of financial analysis and demonstrates the simplicity of the net present value analysis technique. The finance department's work load is shifted to the engineers who are more knowledgeable about the impacts that new projects have on operating costs in their respective engineering disciplines. Furthermore, by having the net present value analysis performed by the engineers, the trade-off between capital costs and operating expense becomes more apparent to the personnel directly responsible for project implementation decisions.

Another benefit of the process is that more alternatives can be considered in developing solutions to productivity problems. The value analysis procedure screens alternatives and focuses the expenditure of resources on the most promising options--those with optimal net present value and highest return to the system.

The process is developed for line managers. The line managers are responsible for operating budgets and have been directed to achieve productivity improvements. The value analysis process generates implementation schedules during the final analysis that can be used to

- Monitor project implementation and the achievement of scheduled efficiencies and
- Develop better cost and schedule estimations for future project proposals.

The procedure improves the likelihood of authorization for proposals that change standard operating procedures, staffing levels, and job responsibilities. The procedure was conceived with direct input from the line managers, finance, and top management and recognizes that they are the users and beneficiaries of the procedure and its results. A user's manual and microcomputer program were developed for applying value analysis as a result of the study.

The value analysis procedure has been adopted by the LT-Rail engineering departments and is currently in use. The procedure has been successfully implemented, and productivity tasks evaluated using the technique are now in the implementation phase. The standardized value analysis procedure is contributing to more effective communications between staff engineers, finance, and top management. It is still too early to compare actual-to-anticipated savings estimated using the technique, but this should be known in the coming year as projects are implemented and monitored at LT.

CONCLUSIONS

Although the value analysis methodology for developing, analyzing, and reviewing productivity alternatives was devised to meet the specific needs of London Transport-Rail, it offers broad applications to transit operators in the United States. The impetus for developing a program of continued productivity improvements is apparent in both nations. Use of a standard productivity development, analysis, and review procedure offers transit agencies several key advantages:

- Ideas for improving productivity can be generated by line staff, clerical personnel, supervisors, or managers in any discipline and still be compared and analyzed on a consistent basis.
- Interaction and early communications with others encourage wider development of productivity alternatives and identification of different implementation strategies. Further, soliciting involvement of other interested parties may broaden support for productivity projects.
- Use of an uncomplicated analysis technique, like the one presented in this paper, takes the mystery out of financial analysis and the time value of money.
- Net present value can be used to compare productivity options that require varying degrees of investment of funds and energy and provide different returns over incongruous time periods on an equitable basis.
- The results of the value analysis provide a sound basis for decision making and define a plan for monitoring implementation and achievement of cost savings.

The value analysis technique, as presented herein, would require some minor revisions for application at other agencies. The primary areas of modification are the reviewing groups, the minimum threshold for a detailed value analysis, and decision-making authority. The sequential development and review process and the two-level net present value technique remain valid for practical application anywhere. Development of a productivity program, such as LT's value analysis, is a major step toward making efficiency improvements a keystone in the transit operating environment.

Improving Section 15 Passenger Data Collection Techniques

ROBERT L. SMITH, Jr.

ABSTRACT

The Urban Mass Transportation Administration requires all transit systems that receive federal funds to collect basic data on transit ridership. The transit systems are required annually to furnish estimates of systemwide passengers, passenger-miles, and, until recently, passenger-minutes under the Urban Mass Transportation Act Section 15 reporting requirements. Many transit operators have complained that the collection of Section 15 passenger data is an unwarranted burden. Modern statistical sampling techniques, however, provide the opportunity for somewhat reducing the effort required by Section 15. The specific objectives of this study were (a) to identify the range of techniques used by large transit properties to collect Section 15 passenger data and (b) to identify and evaluate improved techniques for collecting Section 15 data. A review of the literature showed little application of statistical sampling techniques to bus passenger data collection until 1977. Two recent statistical approaches that are relevant to Section 15 data collection are reviewed. An inventory of 58 transit systems in the United States with 100 or more peak-hour buses showed four main Section 15 data collection techniques: (a) standard Wells random sample, (b) sample from extensive ride checks, (c) two-stage estimation, and (d) extensive ride checks. One or more case study transit systems were selected for each technique. Statistical sampling theory was applied to develop alternative sampling plans and estimation procedures. The analysis showed substantial potential for improvements in accuracy or reductions in data collection costs, or both.

The Urban Mass Transportation Administration (UMTA) requires all transit systems that receive federal funds to collect basic data on transit ridership. The transit systems are required annually to furnish estimates of systemwide passengers, passenger-miles, and, until recently, passenger-minutes under the Section 15 reporting requirements. The procedure recommended by UMTA for collecting the ridership data is to conduct ride checks (on/off counts) on three randomly selected one-way trips every other day. The recommended procedure is designed to achieve a precision of 10 percent for a 95 percent confidence interval. Other statistical sampling plans are acceptable as long as the required level of accuracy is achieved.

Since the introduction of the Section 15 reporting requirements in 1978, many transit operators have complained that the collection of the Section 15 passenger data is an unwarranted burden. Some operators assert that ride-check data are not at all useful to them, and others argue that the random sample obtained under Section 15 is too small and scattered to be useful for operational planning. The latter group may conduct extensive ride checks but not use a random sampling procedure so that following the UMTA sampling procedure requires an additional commitment of checker time.

The concerns of transit operators about UMTA's Section 15 "passenger service consumed" reporting requirements have been addressed as part of a broader study of transit passenger data collection techniques (1). In terms of Section 15, the specific objectives of the study were (a) to identify the range of techniques used by large transit properties to collect Section 15 passenger data and (b) to identify and evaluate improved techniques for collecting Section 15 data.

RESEARCH APPROACH

To provide results that transit managers can easily relate to their own needs, a case study approach was selected for the research. The first step was to document the data collection procedures that are currently used by transit properties in the United States. To keep the data collection effort manageable, only systems with 100 or more peak-hour buses were included. Both published reports and direct telephone contacts were used to document the full range of bus passenger data collection procedures used by each system. Only the procedures for collecting Section 15 passenger data are documented in this paper.

The second step was to develop classification schemes for the Section 15 data collection procedures and to classify the systems accordingly. The classification scheme was then used as the basis for selecting at least one case study system for each classification. Statistical sampling and analytical techniques were applied to identify the potential for improvements and the limitations of each class of procedures.

LITERATURE REVIEW

Overview

Until recently transit managers have had little technical documentation of appropriate passenger data collection procedures available to them. The recently reprinted 1947 American Transit Association Bus Scheduling Manual: Traffic Checking and Schedule Preparation (2) did provide guidelines for the frequency of maximum load point checks, but no statistical analysis was presented to justify the guidelines nor was any indication given of the accuracy of the resulting point-check data. The more recent (1976) fourth edition of the Manual of Traffic Engineering Studies (3) also gives no real guidance on how frequently ride checks should be made. Reference is made to an appendix on statistical procedures for determining the accuracy of a sample, but only the case of a simple random sample is considered.

Modern statistical sampling theory was first applied to the collection of bus passenger data by the Wells Research Company in 1977 (4). John Wells developed alternative sampling procedures for collecting Section 15 data. One problem faced by Wells was the lack of historical and even current data on the day-to-day and within-day variations in bus ridership. Two subsequent studies provided information on the data collection procedures used in the transit industry, but no indication was given that modern statistical sampling methods were being applied, nor were any accuracy measures presented (5,6). In 1977 Stone and others at the University of Utah applied statistical quality control techniques to monitor route-level performance on the Denver and Salt Lake City transit systems (7). Stone also measured the accuracy of the monthly passenger counts taken by Salt Lake City bus drivers.

A comprehensive, statistically based methodology for bus passenger data collection is presented in the recently published two-volume Bus Transit Monitoring Manual (8). Although the focus of the manual is on point- and ride-check data collection, the sampling procedure could be applied to driver-based data collection programs as well. Case study applications of the manual are in progress.

Contacting transit properties for this research revealed that several systems have contracted with consultants to produce statistically based data collection plans. The reports documenting the data collection plans are either "in-house" reports or reports that have received only limited distribution. To date none of this work has been reported in the standard technical literature. Similarly, statistically based work by local staff has generally gone unreported. One exception is a paper by Susan Phifer (9). Two key technical reports that are relevant for

Section 15 data collection procedures are reviewed next.

Wells Memorandum

In the Wells memorandum (4), John Wells of the Wells Research Company documents the sampling procedures recommended for Section 15 passenger data collection in UMTA Circular 2710.1. The basic design criterion was that annual total passenger-miles are to be estimated at a 95 percent confidence level with a precision of at least 10 percent. To minimize the potential impact of periodic variations by day of week and seasonal variations, a frequent and systematic sample of days throughout the year should be selected.

Wells developed alternative sampling plans based on the formula for the relative variance of a two-stage cluster sample and cited Hansen, Hurwitz, and Madow (10) as the reference. The formula is based on taking a random sample of days in the year in the first stage of the sample and then a random sample of one-way trips to be ride checked in each of the sample days. Wells assumes that a systematic sample, such as every other day or every third day, is equivalent to a simple random sample. The basic formula is

$$CV^2_{pm} = [(M - m)/M](CV^2_b/m) + [(N - n)/N](CV^2_w/mn) \quad (1)$$

where CV^2_{pm} is the relative variance of annual total passenger-miles; CV^2_b and CV^2_w are the between and within-day relative variances of total passenger-miles, respectively; M and N are the population sizes for number of days in the sample period and number of one-way trips per day, respectively; and m and n are the sample sizes for days and one-way trips, respectively.

In computing the maximum relative variance allowed by the 95 percent confidence level and a precision of 10 percent, Wells rounded the t value of 1.96 to 2.00, thus providing a slightly more conservative estimate of the accuracy of the sampling plans. The resulting coefficient of variation is $CV_{pm} = 0.1/2.0 = 0.05$ giving $CV^2_{pm} = 0.0025$. To evaluate alternative combinations of numbers of days and trips sampled, assumptions about the between- and within-day relative variances were required. On the basis of limited data, Wells chose conservative values of $CV_b = 0.1$ and $CV_w = 1.0$. Analysis of passenger-mile data for this research indicates that Wells' assumptions are nearly always quite conservative.

The limiting case for number of days sampled per year is 365. Using Equation 1, only the within-day relative variance term remains. Thus, with m = 365 and sampling one trip each day (n = 1) so that (N - n)/N is approximately one, $CV^2_{pm} = CV^2_w/365 = (1.0)^2/365 = 0.00274$, which is equivalent to a precision of 10.5 percent for t = 2.0 or 10.3 percent if t = 1.96. Wells rejects this sampling plan because the specified precision of 10 percent is exceeded. A sampling plan with three trips every other day (n = 3 and m = 183) gives $CV^2 = 0.002093$, which gives a precision of 9.1 percent. The sampling plan requires 3 x 183 or 549 sample trips, which is fewer than the other sample plans that have a greater interval between sample days.

On the basis of an analysis of Section 15 data for Albany, New York; Madison, Wisconsin; and Omaha, Nebraska, it appears that for most transit properties the within-day coefficient of variation (CV_w) and hence CV^2_w are likely to be somewhat less than 1.0. Consequently most properties can justify a sample of one trip every day, which is a saving of about one-third over the minimum recommended sampling plan.

Bus Transit Monitoring Manual

The Bus Transit Monitoring Manual (BTMM) (8) provides a methodology for developing a statistically based route-level transit monitoring program. The focus of the sampling plan is on point and ride checks conducted by traffic checkers. The same two-stage cluster sampling plan used by Wells for Section 15 passenger data collection is applied in the BTMM methodology (see Equation 1 in the previous section). The time period over which the sample is to be taken will vary depending on whether monthly, quarterly, or annual performance estimates are wanted.

The development of a sampling plan involves specification of the desired confidence level and precision and then selection of the number of sample days and the number of sample trips per day that meet the accuracy specifications. The maximum number of trips sampled per day is constrained by the number of checkers available. The trade-off between number of days and trips sampled per day is a function of the between-day and within-day coefficients of variation for the data item of interest, usually total passengers, passenger-miles, or running time.

To develop a sampling plan, estimates of the between-day and within-day coefficients of variation (CV_b and CV_w) are required. An intensive data collection effort is proposed in the BTMM so that route-level CV_b and CV_w can be computed. Other data sources such as counts by drivers can also be used.

The primary limitation of the BTMM is that it does not address a full range of sampling options. In particular, potential applications of stratified and systematic sampling are not considered. The manual, however, does provide a starting point for transit systems that do not have any expertise in developing statistically based sampling procedures.

STATISTICAL SAMPLING PROCEDURES

Reasons for Sampling

Sampling involves selecting a number of observations or elements of a population and using the characteristics of the sample to make inferences about the characteristics of the population. For example, if an estimate of the number of passengers using a particular bus route is wanted, the passengers on a random sample of one-way bus trips during a given day could be counted. Multiplication of the total passengers counted in the sample by the inverse of the sampling rate (expansion factor) will give an estimate of the total passengers that rode on all the one-way trips during that day. Statistical sampling theory provides measures of the accuracy of the sample estimate.

Sampling provides many potential advantages: (a) reduced cost, (b) greater accuracy, (c) more information, and (d) speed in processing. If the costs of collecting data on the entire population are high, then collecting data on only a fraction of the population can reduce costs substantially. The additional costs of preparing, monitoring, and analyzing the sample results must also be included.

Basic Sampling Techniques

The two basic methods for collecting the data on passenger-miles required by Section 15 are ride checks and automatic passenger counters. Because both methods are relatively costly, sampling techniques are essential to provide the required data. The four basic sampling techniques are (a) simple random sample, (b) stratified random sample, (c) cluster sample, and (d) systematic sample. Only a brief overview of each is presented here. More detailed treatments of survey sampling theory and practice are found in standard texts. Cochran (11) provides lucid explanations of the theory as well as of the advantages and disadvantages of the various techniques.

Simple Random Sampling

A simple random sample is a method of selecting n units out of a population of N units such that every unit in the population has an equal chance of being selected. If the sample is truly random, then the sample estimate of some characteristic of the population will be unbiased; that is, the mean value of the estimate taken over all possible random samples of size n is equal to the population value. A simple random sample also has the desirable property that the sample variance is an unbiased estimate of the population variance.

The key formula for a simple random sample is the variance of the sample mean $(\bar{y})$ given by

$$v(\bar{y}) = (s^2/n)(1 - n/N) \quad (2)$$

where s^2 is the sample variance, n is the sample size, and N is the population size. The term in parenthesis (1 - n/N) is the finite population correction factor (fpc) for sampling without replacement in a finite population of size N. Clearly, as the sample size approaches the population size, the fpc approaches zero and the variance of the sample mean also approaches zero. If n = N, there is no longer any sampling error. The population mean and variance are known.

Equation 2 shows that the sampling error in estimating the population value Y or $\bar{Y}$ can be made as small as desired simply by increasing the sample size. Confidence limits for the estimate of $\bar{Y}$ are given by $\bar{y} \pm ts_{\bar{y}}$ where $s_{\bar{y}} = v(\bar{y})^{1/2}$ is the standard error of the sample mean. For a 95 percent confidence level the t-value for large n is 1.96 or about 2.0. The meaning of the confidence interval is that in repeated sampling 95 times out of 100 the population mean will fall within the confidence interval.

The precision (r) of the sample for a given confidence level is given by the error in $\bar{y}$ divided by the sample mean, so $r = ts_{\bar{y}}/\bar{y}$. Thus, using Equation 2, the precision is

$$r = [t(s/\bar{y})/n^{1/2}](1 - n/N)^{1/2} \quad (3)$$

Because $s/\bar{y}$ is the coefficient of variation (CV) Equation 3 provides a convenient means of estimating the sample size given a desired precision (level of accuracy) and an estimate of CV. An initial estimate of CV may be available from a previous sample or from data for a similar transit property.

Stratified Random Sampling

Stratification is a method of dividing a population of N units into L distinct subunits or strata, N_1, $N_2, \ldots, N_L$. The sum of the units in the L strata must equal N. For stratified random sampling a simple random sample is drawn in each stratum with sizes $n_1, n_2, \ldots, n_L$. Stratification may be used for administrative convenience or to increase the precision of the population estimate by dividing a heterogeneous population into homogeneous subgroups.

For stratified random sampling, an unbiased estimate of the variance of the sample mean $(\bar{y})$ is given by

$$v(\bar{y}) = \sum_{h=1}^{L} W_h^2(1 - n_h/N)(s_n^2/n_h) \quad (4)$$

where W_h equals N_h/N. Thus, $v(\bar{y})$ is simply the sum of variances for a simple random sample in each stratum weighted by the square of the relative size of the stratum.

Cluster Sampling

Surveys in which the sampling unit is a group or "cluster" of smaller units are called cluster surveys. The primary reason for sampling clusters is to reduce the cost of sampling. For a given size sample a smaller sampling unit usually gives a more precise estimate than a larger sampling unit. If the costs of collecting data for the large sampling unit are much less, then the sample size can be increased enough to offset the reduction in precision from using the cluster.

For a simple random sample of n clusters, each containing M elements (subunits), drawn from N clusters in the population, the variance of the sample mean per element is

$$V(\bar{y}) \approx [(1 - f)/nM]S^2[1 + (M - 1)\rho] \quad (5)$$

where S^2 is the variance among the elements, f is the sampling fraction (n/N), and ρ is the intracluster correlation coefficient. The intracluster correlation coefficient is a measure of the homogeneity of the clusters. If within each cluster the values for y are similar, then $\rho \approx 1$ and the variance of the mean is obtained by dividing the population variance by n. When the values for y within a cluster are as diverse as the entire population, then $\rho \approx 0$ and the variance of the mean is essentially the same as if a random sample of n x M elements had been selected from the population.

Systematic Sampling

Selection of a systematic sample of size n from a population with N units ordered from 1 to N involves selection of every kth unit with a random start from among the first k units. This is called an "every kth systematic sample" with nk = N. The primary advantage of systematic sampling is that the sample is usually easy to draw and can be done accurately. Also, because the systematic sample is spread evenly over the population, systematic sampling may be more precise than simple random sampling.

One problem with systematic sampling is that estimation of the sample variance requires knowledge of the population variance. With unknown populations systematic sampling should be used with caution. If there is high correlation among the units within a sample, then the sample estimate may be an extremely poor estimate of the population mean. This is true if unsuspected periodicity is present in the population. In contrast, if the population is essentially in random order, then the systematic sample will contain the same information as a simple random sample of size n and have the same variance.

Extension of Basic Sampling Techniques

The most common extension of the basic sampling techniques is the use of multiple levels or stages of sampling. The basic sampling technique at one stage may be different from the sampling technique at the next stage. The two-stage cluster sample used by Wells and the Bus Transit Monitoring Manual is the most useful extension for Section 15 purposes. Another useful extension of the basic sampling techniques is the use of ratio or regression estimation. Ratio and regression estimation can improve the precision of an estimate when there is a high correlation between the variable of interest and a second variable and there is independent information available for estimating the second variable.

With ratio estimates an auxiliary variable (x_i), which is correlated with y_i, is obtained for each unit in the sample. The population total (X) of the x_i is known so that the population total (Y) can be estimated as

$$Y_R = (\bar{y}/\bar{x})X = RX \quad (6)$$

The increase in precision obtained from the ratio estimate depends on the relative variance (or coefficient of variation squared) of the ratio (R) given by

$$CV_R^2 = [(1 - f)/n](CV_y^2 + CV_x^2 - 2\rho CV_y CV_x) \quad (7)$$

which is equivalent to

$$CV_R^2 = CV_{\bar{y}}^2 + CV_{\bar{x}}^2 - 2\rho CV_{\bar{y}} CV_{\bar{x}} \quad (8)$$

If $CV_{\bar{y}}$ and $CV_{\bar{x}}$ are approximately equal, then it is easy to show that the precision of R results in an increase in the precision of Y_R when $\rho \geq 0.5$.

The ratio estimate is useful when the x_i are much cheaper to sample than the y_i so that X can be determined directly or when X is available from some other source. For Section 15 the ratio estimate of average passenger trip length can be used to estimate total passenger-miles on the basis of total passengers at the system level.

CURRENT DATA COLLECTION PROCEDURES

The concerns that transit managers have raised about the need for Section 15 passenger data reporting could possibly indicate a lack of interest in ride-check data collection and even a more general low level of interest in all checker-based data collection. The inventory of the passenger data collection techniques used by 58 transit systems with 100 or more peak-hour buses, however, revealed that there is a substantial commitment to ride checks. As the data given in the following table, which gives the distribution of transit systems by checker effort devoted to ride checks, indicate, all but 10 percent of the transit systems devote at least some of their checking staff time to ride checks.

Percentage of Checker Effort	Percentage Distribution
0	10
1 to 33	33
34 to 66	28
67 to 99	24
100	5

More than one-quarter of the systems devote at least two-thirds of their checking staff time to ride checks.

The large transit systems also devote substantial resources to passenger data collection in the form of a regular checking staff. As the data in the following table, which gives distribution of transit systems by size of transit system, indicate, only 17 percent of the systems had less than one checker per 100 peak-hour buses.

Size of Checking Staff (staff per 100 peak buses)	Percentage Distribution
Fewer than 1.0	17
1 to 1.9	42
2 to 2.9	22
3.0 or more	19

Clearly, most transit systems have the checking staff required to conduct the ride checks under the Section 15 data collection plan recommended by UMTA.

The inventory of the 58 large transit systems revealed a surprising diversity in the procedures used to obtain Section 15 passenger data. As the data in Table 1 indicate, only about 60 percent of the systems use the standard random sample ride-check procedures developed by Wells and recommended by UMTA in Circular 2710.1. Most of these systems also have an extensive ride-check program that may be partly or wholly integrated with the random sampling procedure.

TABLE 1 Distribution of Section 15 Passenger Data Collection Procedures

Procedure	Distribution of Properties [number (percentage)]
Standard Wells (random sample)	35 (60.3)
Minimum level only	12 (20.7)
Extensive ride-check program	23 (39.6)
Sample from extensive ride checks	11 (19.0)
Year-long program	8 (13.8)
Short intensive program	3 (5.2)
Two-stage program	9 (15.5)
Extensive program	3 (5.2)
Total	58 (100)

Eleven properties meet the Section 15 requirement by selecting a random sample of ride checks from their regular, extensive ride-check program. In general, these properties ride check all of the daily one-way trips in the system at least once a year. Thus, the sample will be unbiased in terms of coverage, but it may be biased as the result of seasonal and secular trends.

The two-stage procedure involves multiplying estimates of passenger-miles per passenger by total passengers to obtain the required estimate of passenger-miles required by Section 15. The estimate of passenger-miles per passenger (average trip length) may be obtained from ride checks or from passenger surveys. With proper statistical sampling the procedure can be even more accurate than either of the first two procedures. In practice, however, the estimates of average trip length are based on whatever data are available.

A few properties such as Metro Area Transit in Omaha, Nebraska, have extensive ride-check programs that are based on a large random sample. The result is an accurate estimate of passengers and passenger-miles at the route level. Omaha uses traffic checkers. Automatic passenger counters (APCs) are also being used for extensive ride-check programs. A number of APC-based counting programs are currently being developed and more can be expected in the future as more experience is gained with the technology (12).

The four basic categories of Section 15 data collection procedures given in Table 1 provided the framework within which case study transit systems were selected to illustrate the potential for improvements and possible limitations of a particular procedure. Three of the six transit systems selected for the case studies were selected because the local transit staff had either developed an improved Section 15 data collection procedure or had analyzed Section 15 data in order to identify possible improvements. The other three case studies were selected as the result of data availability or the need to cover each of the four categories. Three case study transit systems were selected in the "standard Wells" category and one system was selected from each of the other categories given in Table 1.

CASE STUDY DATA COLLECTION PROCEDURES

Standard Wells: Madison, Wisconsin, Case Study

To develop the basic data needed to analyze alternative sampling strategies for Madison, the tabulation of the Section 15 ride-check data was computerized. The resulting data for 183 ride checks conducted during the first half of 1982 are summarized in Table 2. The computer file also included data on route, day of week, and time of day.

TABLE 2 Section 15 Sample Data for Madison Metro

Variable	Mean[a]	Standard Deviation	Coefficient of Variation	Precision[b]
Boardings	43.8	26.9	0.61	0.089
Maximum load	25.4	16.4	0.65	
Passenger-miles	157.0	110.0	0.70	0.102
Passenger-minutes	740.8	560.0	0.76	0.110
Average passenger-miles	3.47	1.25	0.36	
Average passenger-minutes	15.7	4.57	0.29	
Average speed	3.1	2.81	0.21	

[a]Based on 183 observations.
[b]For a 95 percent confidence level.

At the system level the primary interest is in reducing the sampling rate within the UMTA-imposed constraint of estimating total passenger-miles within 10 percent at the 95 percent confidence level. If the data in Table 2 are assumed to be based on a simple random sample of size 183, then the precision of the estimate of passenger-miles is almost equal to the required 10 percent level, and the estimate of average boardings per trip is even more precise (precision of 0.089). If the coefficient of variation does not increase significantly for the sample covering the entire year, then a sample of about 200 one-way trips instead of 546 trips would be adequate. Although not recommended, a minimal sample of 182 trips could be obtained by taking a simple random sample of one trip from every two-day period during the year.

With Madison Metro's current sample of one trip per day the annual sample of 365 trips will result in a precision for average passenger-miles per trip of 7.2 percent assuming a coefficient of variation of 0.70. The coefficient of variation could increase to nearly 1.0 and still meet UMTA's accuracy requirement. Thus, it is clear that Madison Metro's current sampling rate is more than adequate to cover substantially greater variations in average passenger-miles than currently occur.

Some minor improvement in the accuracy of the system-level estimate of total passenger-miles is made possible by stratifying the sample by day of the week. The average passenger-miles per trip for weekdays is 171 compared with 130 for Saturday and Sunday. This difference in stratum means appears to be large enough to make the additional effort of stratifying the sample worthwhile.

Other possible stratifications include time of day and route. One-way analysis of variance of average boardings stratified by time of day and considering only the five main routes on weekdays showed a highly significant difference between the evening period from 6 p.m. to 12 midnight and the three other periods (a.m. peak, midday, and p.m. peak). In stratifying by route there is a significant difference in passenger-miles per trip between the short university routes and the main-line routes. Although some improvement in the accuracy of the estimates of passenger-miles and total boardings could be achieved through stratification by time of day and route, the additional complexity of the resulting sampling plan would no doubt outweigh the benefits of the accuracy improvements.

Because Madison Metro now is obtaining 100 percent counts of passengers on a daily basis, the accuracy of the annual estimate of total passengers is only limited by the accuracy of the drivers' counts and clerical errors in recording the data. If the total passenger counts are assumed to be highly accurate, then substantial improvement in the accuracy of the annual estimate of passenger-miles can be obtained by using a ratio estimate of passenger-miles per passenger. Using Equation 7, the relative variance of the ratio estimate is given by

$$CV^2_{pm/p} = [(1 - f)/n](CV^2_{pm} + CV^2_p - 2\rho_{pm,p} \cdot CV_{pm}CV_b) \quad (9)$$

where CV_{pm} and CV_b are the coefficients of variation of passenger-miles and total passengers, respectively, and $\rho_{pm,p}$ is the correlation between passenger-miles and total passengers. For the Madison data $\rho_{pm,p}$ is quite high (0.916). Thus, using Equation 9, $CV_{pm/p}$ is found to be 0.021 on the basis of the available sample size of 183. The resulting precision of the ratio estimate of passenger-miles is 4.1 percent. If a precision of 10 percent is all that is required and $CV_{pm/p}$ is assumed to remain the same except for the sample size, then a sample of only 32 ride checks would be needed. Thus, use of the ratio estimate results in a reduction of the sample size required to give a precision of 10 percent from 183 to 32, or more than an 80 percent reduction.

The ratio estimate of passenger-miles per passenger can only be used to improve the precision of the estimate of total annual passenger-miles if the precision of an independent estimate of total annual passengers is about as good as or better than the ratio estimate. If an independent estimate of total annual passengers is not available, then the ratio estimate is likely to be less useful as a means of improving the precision of the passenger-miles estimate. At this point the dramatic increase in the precision of the ratio estimate compared with the direct estimate of passenger-miles should not be used to reduce the sampling rate because then the accuracy of the Section 15 estimate of total passengers would be reduced as well. The Section 15 estimate is needed to provide a check on the accuracy of the driver counts.

Standard Wells: St. Louis Case Study

As is the case with many systems, St. Louis collects Section 15 ride-check data independently of its regular ride-checking program. Because the Section 15 data have not been integrated into the regular data collection effort, there is an incentive to reduce the effort devoted to Section 15. St. Louis used Section 15 data on passenger-miles for each one-way trip stratified by time period and day of the week to develop a new sampling plan.

The first step in developing the new sampling plan was to assume a simple random sample over the entire year. On the basis of the observed coefficient of variation of passenger-miles for the 1981-1982 fiscal year of 0.971, a sample size of 362 is required to meet UMTA's accuracy requirements. This is the same basic calculation that was made using the Madison Metro data. There is clearly much greater variation in the St. Louis data, which probably reflects the much greater emphasis on express service and lower usage during midday in St. Louis. Nevertheless, St. Louis could have adopted Madison's one trip per day sample plan. Instead, St. Louis used an optimal allocation of trips to the four time periods of the day followed by a proportional allocation of trips to each day of the week. The objective of using a more complicated sampling plan apparently was not to reduce the sample size further but to provide a safety factor in case the new sampling plan is not as accurate as predicted by the available data.

The main question in evaluating St. Louis' two-stage sampling plan is whether the additional effort of stratifying the sample by time of day and day of the week is justified by the expected increase in accuracy. In analyzing the raw data obtained from St. Louis, a slightly lower coefficient of variation of 0.948 was obtained, which gives a lower sample size of 346. The required sample size is highly sensitive to the coefficient of variation.

In evaluating the accuracy of St. Louis' two-stage sampling plan, the order of the stratification was reversed and optimal allocation was used in both stages. A detailed calculation of the precision of the new sampling plan (new calculations based on two optimal allocations) resulted in a precision of 0.094. The gain in precision over a simple random sample is only 0.006, which is clearly not worth the additional effort required for stratification. The gain in precision can be translated into a reduction in sample size of about 30 trips.

Standard Wells: Denver Case Study

The Denver Section 15 ride-check procedure is of interest because application of the standard Wells sampling plan will not meet the desired precision of 10 percent for passenger-miles. As the data in Table 3 indicate, the precision of the passenger-miles estimate is only 11.6 percent. The lack of accuracy in the passenger-miles estimate is explained in part by the approximations used in computing the distance between stops. Because an up-to-date stops file with complete distances between stops was not available, the Denver Regional Transit District (RTD) allocated the known distance between time points equally to all stop-level segments. To the extent that passenger loads are concentrated on route segments that have closer spacing of stops, the assumption of equal distances between stops will tend to overestimate passenger-miles.

Because Denver does have an independent estimate of total passengers, the ratio estimate of passenger-miles per passenger can be used to increase the precision of the passenger-miles estimate. Although the correlation between passenger-miles and total passengers is relatively low ($\rho = 0.587$), the coefficient of variation of the ratio estimate from Equation 9 is low enough ($CV_{pm/p} = 0.0478$) to give a precision for the ratio estimate (9.4 percent) that is within the UMTA guidelines. Thus, if the independent estimate of passengers is accurate, the product of the ratio of passenger-miles per passenger and total passengers will give an estimate of passenger-miles that is within the UMTA guidelines.

Substantial improvements in the accuracy of the estimate of passenger-miles should be possible through stratification by route type. As the data in Table 3 indicate, there are large differences among the average passenger-miles per trip by route type. For example, Denver local service generates only 156 passenger-miles per trip whereas express and intercity routes generate more than three times as many passenger-miles per trip. Stratification by route type will eliminate that part of the total variance that is the result of the difference between the route-type mean and the overall mean. As the data in Table 4 indicate, the precision of both the passenger-miles and the passenger-minutes estimates is improved substantially by a stratified sample. In contrast, the precision of the passengers estimate is essentially the same, which is explained by the relatively small differences among the means of the route types.

TABLE 3 Denver Section 15 Data Stratified by Route Type[a]

Route Classification	Sample Size	Passengers		Passenger-Miles		Passenger-Minutes	
		Mean	CV	Mean	CV	Mean	CV
Weekday Only							
Denver local	260	40.9	0.631	156	0.872	692	0.843
Boulder local	15	21.1	0.829	63.8	1.28	262	1.046
Longmont local	12	4.2	1.07	8.3	1.23	22.0	0.658
Denver circulator	34	11.4	1.24	25.2	1.30	125	1.465
Intercity	16	28.3	0.88	552	1.13	1,118	1.097
Express	37	45.6	0.40	490	0.467	1,462	0.491
All routes	381	35.8	0.723	188	1.26	695	0.997
All Days							
All routes	538	31.6	0.796	157	1.37	582	1.09
Precision[b] (percentage)			6.3		11.6		9.2

[a]Derived from data in Beuthel *(13)*.
[b]Based on a simple random sample.

TABLE 4 Precision of Simple Versus Stratified Random Sample for Denver, Weekday Only[a]

Sample Type	Passengers	Passenger-Miles	Passenger-Minutes
Simple random sample	0.073	0.127	0.100
Stratified random sample (proportional allocation by route classification)	0.076	0.104	0.083

[a]Derived from data in Beuthel (13).

The ratio estimate could also be applied to the coefficients of variation obtained from the stratified sample. The resulting precision of the ratio estimate of passenger-miles per passenger should be substantially less than 10 percent and thus well within UMTA's guidelines.

The potential for stratification by a number of variables can be determined from Table 5. The between-class variation is eliminated by stratification. Thus, the greatest reduction in variance is achieved by selecting the classification with the highest between-day coefficient of variation. Stratification by route type is best for passenger-miles, whereas stratification by day is best for passengers and passenger-minutes. Based on the weekday data, stratification by day in Denver with a random sample of one trip each day would be adequate for both passenger and passenger-minutes estimates. A ratio estimate of passenger-miles should come close to giving a precision of 10 percent. A daily sample of one trip would represent a reduction of about one-third from the current sample rate of two trips every other day.

Sample from Extensive Ride Checks Case Study

Overview

A number of transit properties have extensive ride-checking programs in which every daily one-way trip is ride checked at least once during the year. At the end of the year ride-check data for at least one day are available for all routes so that comparisons of route performance can be made. One problem with spreading the ride checks over the entire year is that route-level comparisons are biased by seasonal variations and ridership trends. To avoid this problem, some properties concentrate their ride checks in a short period of a few weeks. In some cases both spring and fall checks are made.

A number of properties have used their 100 percent ride checks as the data base for satisfying Section 15 reporting requirements. Because all of these properties have considerably more daily one-way trips than the 546 trips required by the standard Wells minimum random sample, a random sample of 546 trips typically is drawn from the 100 percent ride checks. This two-stage sampling procedure is valid statistically if the first stage (the 100 percent ride-check data) gives an unbiased estimate of passenger-miles, passenger-minutes, and total passengers. The ride checks that are spread over the entire year should provide unbiased estimates as long as the checks for each route type such as local and express routes are distributed reasonably uniformly over the year. The ride checks that are concentrated in a short period, however, are likely to produce biased estimates because seasonal variations and secular trends are omitted. A partial solution in this case is to use a ratio estimate of passenger-miles per passenger as the basis for estimating annual passenger-miles. Passenger-miles per passenger measures average passenger trip length, which should be more stable over time than either passenger-miles or total passengers.

Milwaukee Case Study

A test of the validity of using 100 percent systemwide ride checks collected over a short period as the basis for Section 15 reporting requires data on average passenger trip length over time, which in turn requires an extensive ride-check data base. Milwaukee County Transit in Milwaukee, Wisconsin, has the required data base in machine readable format. Time series ride-check data for one crosstown feeder line, Route 55, are given in Table 6.

TABLE 6 Variations in Average Passenger-Miles per Passenger on Milwaukee Route 55

Month	Sample Size	Passenger-Miles per Passenger	Coefficient of Variation	Precision (r)
January	64	6.07	0.382	0.096
October	62	6.38	0.392	0.100
November	63	6.49	0.357	0.082

The passenger-miles per trip estimates given in Table 8 are not true ratio estimates but the average of the ratio of passenger-miles per passenger over the 100 percent ride check for that time period. The precision of the estimate, thus, is a measure of variation in the average of the ratios and is only an indirect measure of the variation in average passenger-miles per passenger over the entire day. Clearly, the average of the average trip lengths is

TABLE 5 Within- and Between-Strata Coefficients of Variation for Denver Weekday Section 15 Data[a]

			Time of Day			
Variable[bc]	Day	Day of Week	All	Denver Local	Peak vs Off-Peak	Route Type
Passengers	0.582[c]	0.714	0.681	0.476	0.696	0.647
(CV = 0.724)[d]	0.430[e]	0.113	0.242	0.209	0.198	0.323
Passenger-miles	1.01	1.24	1.17	0.659	1.18	0.972
(CV = 1.27)[d]	0.758	0.265	0.480	0.279	0.454	0.809
Passenger-minutes	0.821	0.982	0.905	0.627	0.917	0.849
(CV = 0.997)[d]	0.564	0.163	0.415	0.299	0.389	0.522

[a]Derived from data in Beuthel (13).
[b]Mean values: Passengers = 35.8, passenger-miles = 188, and passenger-minutes = 695.
[c]Within-day coefficient of variation, $S_w/\bar{X}$.
[d]Coefficient of variation.
[e]Between-day coefficient of variation, $S_b/\bar{X}$.

reasonably stable over time at least for Route 55. Pairwise tests for equality of the means between months showed that the null hypothesis of equality could not be rejected at the 5 percent level.

Although a sample of three 100 percent ride checks on one route over an 11-month period does not provide a definitive test of the hypothesis that average trip lengths are stable over time, the results for Milwaukee are encouraging. A complete evaluation of the hypothesis would require time series data on systemwide average trip lengths as well. The trip lengths for individual routes may not be constant over time but the systemwide average could still be relatively stable.

Two-Stage Estimation Case Study

Overview

As outlined in the Madison, Wisconsin, case study, a ratio estimate of average trip length (passenger-miles per passenger) can improve the precision of the estimate of passenger-miles if the correlation between passenger-miles and passengers is high and if an accurate, independent estimate of total passengers is available. Total passenger-miles are computed as the product of average trip length and total passengers. If a random sample of ride checks is used to compute the ratio estimate, then the ratio and the resulting estimate of total passenger-miles will be unbiased.

In some cases ratio estimates of average trip length have been based on intensive ride checks conducted over a short period of a few weeks. The assumption is made that the ratio estimates are stable over time. The limited time series data on average trip length for one route in Milwaukee that were presented in the previous section support the time stability assumption, but systemwide data were not available. Although the assumption of time stability may be a good assumption, there is always the possibility for change. Thus, a random or otherwise unbiased sample over the time period of interest is the only method for assuring that an accurate estimate of total passenger-miles is obtained.

Albany Case Study

As part of its transit route performance monitoring study, the Capital District Transit Authority (CDTA) implemented a new ride-check program using the sampling techniques presented in the Bus Transit Monitoring Manual (8,14). The primary purpose of the ride-check program was to estimate route-level average trip length and revenue per passenger within 15 and 10 percent, respectively. Because accurate estimates of route-level total boardings are available from 100 percent counts by drivers, it was estimated that system-level estimates of average trip length and revenue per passenger would be within 5 and 2 percent, respectively.

The number of trips per route that is to be sampled for the ride-check program varies from a low of 2 to a high of 16 for weekdays. For weekends, the range is from 4 to 13 trips. The results of ride checks for three routes are given in Table 7. The variation in average trip length (passenger-miles per passenger) as measured by the coefficient of variation is less than that for passengers and for passenger-miles. The precision of the estimates of average trip length is computed in two ways: (a) direct computation based on the coefficient of variation of the trip lengths for each observation and (b) ratio estimate based on the coefficients of variation of passenger-miles and passengers and the correlation between passenger-miles and passengers. The direct estimates of precision are within the specified 15 percent level, and two of the ratio estimates are slightly above the 15 percent level.

The CDTA's total ride-check program requires 503 weekday and 233 weekend ride checks. The total program of 736 ride checks is considerably larger than the minimum Wells' sample of 546. The advantage of CDTA's approach is that precise estimates of trip lengths and average fare are obtained at the route level. If only systemwide estimates are of interest, then the sample size can be reduced substantially particularly when the ratio estimate for average fare is used as was shown for the Madison, Wisconsin, case study.

Extensive Ride-Check Case Study

Overview

An extensive ride-check program in which every daily one-way trip is checked at least once a year can be conducted either manually or with automatic passenger counters (APCs). Properties with extensive manual count programs are generally interested in total passengers, load profiles, and running times. Passenger-miles are usually not of direct interest so they are usually computed only for a sample of trips for Section 15 reporting purposes. In contrast, with APCs passenger-miles can easily be computed for all sample trips with essentially no extra effort.

TABLE 7 Estimation of Average Trip Lengths from Ride Checks—Albany[a]

Route	No. of Observations	Passengers: Mean (CV)[b]	Passenger-Miles: Mean (CV)	Passenger-Miles per Passenger: Mean[c] (CV)	Precision[d]: Direct	Precision[d]: Ratio
9	16	24.4 (0.450)	44.8 (0.531)	1.86 (0.331)	0.145	0.161
12	15	16.1 (0.442)	42.4 (0.710)	2.47 (0.299)	0.136	0.151
8	9	19.3 (0.304)	40.0 (0.355)	2.05 (0.180)	0.112	0.126

[a]Derived from data in Transit Route Performance Monitoring Study *(14)*.
[b]Coefficient of variation.
[c]Average of the ratio of passenger-miles to passengers. The unbiased estimates of passenger-miles per passenger for the three routes are 1.84, 2.63, and 2.07, respectively.
[d]Direct estimate of precision is based on the formula for a simple random sample. The ratio estimate is based on a ratio estimate using passenger-miles and passengers; that is, $CV^2_{pm/p} = CV^2_{pm} + CV^2_p - 2\rho CV_{pm} CV_p$.

Because of the expense of the APC counting units, only a small fraction of the fleet can be equipped. Various sampling strategies can be developed depending on the data needs of the property and the logistics of assigning buses.

Columbus, Ohio, Case Study

The Central Ohio Transit Authority (COTA) contracted with a consultant to undertake a comprehensive ride-check program using APCs. The APCs provide route-level data on passengers, passenger-miles, passenger-minutes, vehicle-hours, vehicle-miles, and running times so that both route- and system-level performance measures can be computed. A weekday productivity analysis report is produced every 4 months.

Because passenger-miles are obtained routinely from the APCs, the Section 15 reporting requirements are easily met. The six APC-equipped buses could cover the entire fleet of 234 peak-hour buses in about 8 weeks. During the course of a year each vehicle block could be sampled five or six times. Thus, the primary concern is not the sample size but the need for an unbiased sample. Although the precise details of the sampling procedure used by COTA were not available, the sampling procedure appears to involve a trade-off between coverage and route- or corridor-level problem solving. All vehicle blocks have been covered at least once. Multiple checks have been made for a few select routes.

For system-level estimates a random sample of vehicle blocks each day will give the best results. In contrast, for evaluating route operational characteristics, 100 percent checks done route by route are likely to be more useful. The result is a point estimate of demand at the route level. The system-level estimates should still be reasonably accurate as long as the routes are selected in random order.

SUMMARY AND CONCLUSIONS

The three cases of alternatives to the standard Wells random sample approach demonstrate that substantial reductions in sample size are possible within the UMTA-imposed constraint of a precision of 10 percent. In Madison, Wisconsin, a reduction of the sample rate to one trip per day still provided a more than adequate level of accuracy. Further improvements in accuracy were made possible by using a ratio estimate of passenger-miles per passenger. Stratified sampling with stratification by day of the week and route type appeared to have some potential.

St. Louis chose a fairly complicated two-stage stratification by time period and day of the week to reduce the required sample size. Analysis indicated that a simple random sample or a random sample stratified by day would be nearly as accurate. In contrast, the standard Wells sample for Denver did not quite meet the UMTA accuracy requirements. By using the ratio estimate, however, the 10 percent precision level could just be met. Stratification by route type resulted in a further increase in precision.

The second major group of Section 15 data collection procedures, the sample from extensive ride checks, requires the assumption of time stability of average trip length when the extensive ride checks are concentrated in a short period. Ride-check data available from Milwaukee on one route indicated that average trip lengths were reasonably stable over a period of 11 months.

The two-stage estimation approach also uses average trip length to improve accuracy or reduce sample size. In Albany, New York, application of the approach at the route level required only modest sample sizes to achieve a precision of 15 percent for a 90 percent confidence level. The total sample size of more than 700 trips could be reduced substantially if only system-level estimates were of interest.

The last approach is the extensive ride-check program that is of interest for Section 15 purposes when passenger-miles are computed for all ride checks. Full computation of passenger-miles is more likely for APC-based programs than for manual ride-check programs. As shown by the Columbus, Ohio, APC program, a large sample is obtained. A primary concern of the sampling plan for Section 15 purposes should be to give an unbiased estimate of annual passengers, passenger-miles, and other data of interest.

The case studies show that there are many opportunities for improving the accuracy of the standard Section 15 data collection programs. If the only concern is with meeting the minimum accuracy requirements, then improved sampling plans can be used to reduce the sample size. The potential for simple stratification or ratio estimates can be evaluated using the existing Section 15 data base. Use of ratio estimates is possible if an accurate, independent estimate of total passengers is available.

Additional research is needed to determine how Section 15 data collection requirements can best be met with the alternatives to the standard Wells sampling procedure. Identification of the variability in average trip length is a key problem that requires an extensive ride-check data base. Such data bases are currently available in only a few systems.

From a broader perspective the UMTA requirement for estimating annual passenger-miles with a precision of 10 percent needs to be reviewed. If the focus of Section 15 were on developing accurate estimates of total annual passengers, the data could possibly be made more relevant to transit operators. Additional research is needed to explore alternative means of integrating passenger counts obtained from ride checks with other techniques for estimating total passengers.

REFERENCES

1. R.L. Smith, Jr. Innovation and Incremental Improvement in Bus Passenger Data Collection. University of Wisconsin-Madison, Dec. 1983.
2. W.S. Rainville, Jr. Bus Scheduling Manual: Traffic Checking and Schedule Preparation. American Transit Association, Aug. 1947. (Reprinted by the American Public Transit Association and U.S. Department of Transportation, July 1982.)
3. P.C. Box and J.C. Oppenlander. Manual of Traffic Engineering Studies, 4th ed. Institute of Transportation Engineers, Arlington, Va., 1976.
4. J.D. Wells. Technical Basis for Sampling Procedures Used in the Bus Transit and Demand Response System Manuals. Memorandum to A.B. Hallman, UMTA, U.S. Department of Transportation. Wells Research Company, Silver Spring, Md., Nov. 25, 1977.
5. M. Flusberg, J. Kruger, and J. Curry. Bus Transit Monitoring Study: Interim Report 1, Data Requirements and Collection Techniques. Multisystems, Inc., Cambridge, Mass., and ATE Management and Service Co., Cincinnati, Ohio; UMTA, U.S. Department of Transportation, April 1979.
6. J.P. Attanucci, L. Jaeger, and J. Becker. Bus Service Evaluation Procedures: A Review. Massachusetts Bay Transportation Authority and Tide-

water Transportation District Commission; UMTA, U.S. Department of Transportation, March 1979.
7. T.J. Stone, J.A. Austin, R.L. Siegel, and A. Taylor-Harris. Application of Transit Performance Indicators. University of Utah; UMTA, U.S. Department of Transportation, Sept. 1979.
8. J. Attanucci, I. Burns, and N. Wilson. Bus Transit Monitoring Manual: Vol. 1, Data Collection Design. Multisystems Inc., Cambridge, Mass., and ATE Management and Service Co., Cincinnati, Ohio; UMTA, U.S. Department of Transportation, Aug. 1981.
9. S.P. Phifer. Use of Sampling in Bus-Line Data Collection. In Transportation Research Record 877, TRB, National Research Council, Washington, D.C., 1982, pp. 41-44.
10. M.H. Hansen, W.N. Hurwitz, and W.G. Madow. Sample Survey Methods and Theory: Vol. I, Methods and Applications. John Wiley and Sons, Inc., New York, 1956.
11. W.G. Cochran. Sampling Techniques, 3rd ed. John Wiley and Sons, Inc., New York, 1977.
12. D. Vozzolo and J. Attanucci. An Assessment of Automatic Passenger Counters. Multisystems Inc., Cambridge, Mass.; UMTA, U.S. Department of Transportation, Sept. 1982.
13. B. Beuthel. Section 15 Ridership Documentation 1981. Regional Transportation District, Office of Management and Budget, Denver, Colo., Aug. 1982.
14. Transit Route Performance Monitoring Study--Capital District Transportation Authority. Draft of March 1983. Comsis Corporation, Wheaton, Md., 1983.

Transit Operator Performance Evaluation: Study Group Review at Muni

LARRY S. ENGLISHER, MARTIN J. MORGENBESSER, and JOHN P. ATTANUCCI

ABSTRACT

The results of a study group review of employee performance evaluation at the San Francisco Municipal Railway are outlined. The review was undertaken as one step of a demonstration funded by the Urban Mass Transportation Administration, which is aimed at improving the reliability of transit service. Muni is currently implementing study group recommendations as part of the ongoing demonstration. The study group process was used in a previous study of safety issues at Muni and proved quite helpful. By bringing together representatives of other transit properties that have been addressing similar problems, the study group was able to "brainstorm" and exchange ideas. Both Muni staff and the representatives from other properties left the 4-day session with new insights and ideas. The study group addressed several components of a performance standards and motivation program, including measurement of performance, setting targets, establishing incentive and awards programs, ongoing procedures for appraisal and communication, and outlining and operating under a system of discipline. Among the aspects of performance discussed were attendance and punctuality; adherence to schedule; safety; courtesy and appearance; stress and substance abuse; and general conformance to rules, procedures, and directives.

The San Francisco Municipal Railway (Muni) has undertaken a transit service reliability demonstration under a Service and Methods Demonstration grant from the Urban Mass Transportation Administration. The objective of the demonstration is to improve the reliability of service delivered to transit passengers by applying a variety of management and operational strategies. Among the primary strategies are an operator performance evaluation and motivation program, an attendance management program, and on-street supervision and control strategies.

Multisystems initiated the project in December 1983, preparing summary papers to generate discussion on possible demonstration strategies. The papers reviewed Muni's current performance evaluation procedures and the approaches of several other transit authorities, including Metro-Dade, Houston Metro, Seattle Metro, MTC (Twin Cities), Flint MTA, Chittenden County (Vermont), and San Diego Transit. The papers also included summaries of the following approaches to improving productivity and motivation outlined in research performed by the Urban Institute (1): monetary incentives, performance appraisal, performance targeting, job enrichment, and employee assistance programs. Research on the causes of absenteeism among transit employees was also reviewed, including studies of the impacts of extraboard scheduling and workers' compensation by Peat, Marwick, Mitchell and Company (2), MacDorman and MacDorman Associates (3,4), and the U.S. Department of Transportation (5), and on the role of stress by the University of California at Irvine (6,7). Recent developments in employee assistance programs (8), participatory management (9,10) and attendance policies (11) were also reviewed and presented to Muni.

On the basis of recent experience with an accident and safety program, Harold Geissenheimer, Muni's General Manager, suggested that a meeting with managers from other transit properties would help Muni to formulate an effective program design. Therefore, on April 11-13, 1984, representatives of Muni labor and management and the management of six other transit properties met to discuss approaches to operator performance standards and motivation that could be applied to Muni. The six properties were selected to represent a variety of approaches and on the basis of recent innovations they had undertaken. (The study group participants are listed in the acknowledgments at the end of the paper.) The recommendations of the study group are summarized in this paper.

STUDY GROUP RECOMMENDATIONS

The study group discussions focused on three major components of a performance standards and motivation program, as outlined in research by the Urban Institute:

- Measurement and targeting;
- Incentives, awards, and discipline; and
- Appraisal and communication.

Within these three categories, six aspects of performance were considered:

1. Attendance and punctuality;
2. Schedule adherence;
3. Safety;
4. Appearance and courtesy;
5. Substance abuse and dealing with stress; and
6. Observance of rules, operating procedures, and directives.

These six aspects were selected by Muni and Multisystems. They were addressed both as a group and individually in the context of a performance standards and motivation program. The study group's recommendations are generally applicable to all aspects, although individual references are made in special cases at the end of the paper.

In the course of the study group's discussions, Muni representatives reported on Muni's activities to date in this area. As a result, the study group's recommendations built on those elements of a performance standards and motivation program already in place. The recommendations of the study group are discussed in the following sections.

Performance Measurement and Targeting

Measurement

The measurement of performance provides the foundation on which a performance standards and motivation program can be built. Although the extent to which performance can be quantified depends on the aspect of performance being measured, the study group made a number of general recommendations.

Ideally, performance data should be accurate and timely; an automated management information system would facilitate this. As a first step, Muni should establish a limited number of simple performance measures that can be managed manually and are clearly understandable to supervisors, operators, and all management levels. The use of a personal computer was suggested by some members of the study group; a pilot program using a personal computer could be part of a longer term effort to develop a performance monitoring system.

Measures should ideally be expressed in terms of performance per operator as opposed to systemwide statistics (i.e., absences per operator per year instead of the currently used percentage of operators not present on an average day). Expressing the measures in this way makes them more readily understandable and usable to supervisors and operators who are trying to achieve established goals.

Measurements should be done on a rolling basis; "per year" should mean the immediately preceding 12 months not the calendar year. This approach removes artificial time constraints and allows for a more meaningful time period (i.e., 1 year instead of the current 3 months). It also addresses an issue raised by Muni regarding the use of an individual's long-term record in the application of discipline. The use of a rolling year also departs from Muni's current practice of allowing a certain number of miss-outs per quarter.

On the basis of these general recommendations and the specific discussions at the meeting, a sample set of performance measures has been prepared:

1. Attendance
   - Unscheduled absences per operator,
   - Miss-outs per operator,
   - Sick and occupational illness absences per operator, and
   - Late reports per operator;
2. Safety
   - Chargeable accidents per operator;
   - Chargeable accidents per 100,000 miles by mode, division, and type of accident;
   - Safety commendations per operator; and
   - Safety violations per operator;
3. Schedule adherence
   - Early departures (1 min or more) as percentage of departures checked for system,
   - Percentage on-time arrivals (to be defined), and
   - Ahead-of-schedule violations per operator;
4. Appearance and courtesy
   - Passenger service reports per operator,
   - Commendations per operator, and
   - Uniform violations per operator; and
5. General adherence to rules
   - Rule infraction write-ups per operator and
   - Commendations for general adherence to rules.

Note that because performance characteristics may vary by mode, performance measurement should be performed separately for each mode, where applicable.

Performance measurement is not in and of itself a motivational factor. However, making average "per operator" statistics available, and graphically presenting trends in such statistics (for the system and by division), can provide a clear picture to operators and their supervisors of movement toward achieving goals.

Similarly, the manner in which performance measurements are implemented and articulated can significantly affect morale. A negative emphasis and complicated, imposing forms could alienate and demotivate the majority of operators who want to perform well and do little to improve the performance of others.

A positive approach to performance measurement should be presented using clear, agreed-upon definitions of the measures to be used. Furthermore, at the start of implementation, it should be understood that, as for anything new, there will be a shakedown period during which the measurement system can be adapted to the particular situation before being institutionalized. All parties involved in the initial design of the measures should have an opportunity to provide input to any revision process. Such an approach addresses the natural concerns about new procedures and provides the initial flexibility for developing a "tailor-made" system. (After the shakedown period, there should be periodic reevaluation, but changes should be infrequent.)

Targeting

Targets can be set for each of the established measures. Instead of transfering standards from other transit properties to Muni, it was recommended that Muni assess current performance and establish targets for percentage increases or decreases in selected performance statistics. The targets should be achievable and arrived at in a participatory manner; getting supervisors and operators involved in the process of establishing targets fosters commitment to achieving them. The targets should be reevaluated and revised periodically.

Multisystems is currently assisting Muni in implementing a set of performance measures and a microcomputer-based monitoring system, and in establishing targets based on improvements over current performance.

Incentives, Awards, and Discipline

Incentives and Awards

When a measurement system has been established, programs can be designed to reward employees for their superior performance. Incentives, awards, and, in a larger sense, recognition are all positive motivators. They can be divided into three distinct categories:

1. Pay incentives, which are a significant percentage of the total paycheck (i.e., 5 to 15 percent), are monetary payments (i.e., bonuses, incentive-based pay scales) directly tied to performance. Such pay incentives may be offered to individual employees or to groups (divisions) that meet the criteria.
2. Nonpay awards and recognition may be a small monetary award (less than $200), a nonmonetary gift (trip, dinner, trophy), preferential parking, dedicated bus with driver's name on it, or social activities. Publicizing the award (ceremonies, articles in newspaper) is also a form of recognition that enhances the impact of the award itself.
3. Time off is a unique type of nonmonetary reward that can be used to reward superior attendance and punctuality. Depending on its application, it can discourage the abuse of sick leave and transform a large number of expensive unscheduled absences into a smaller number of less expensive scheduled absences.

The second category, "nonpay awards and recognition" was cited by the study group as offering the greatest return in terms of being both relatively inexpensive and a strong performance motivator. Furthermore, it appears that these are implementable by Muni now, in contrast to pay incentives. Although pay incentives were cited as effective (in Flint and Houston), the increased data processing requirements and complication of pay structures might make them difficult for Muni to implement in the immediate future. More important, San Francisco City Charter laws prohibiting "give-away of city funds" would constitute a greater barrier to pay incentives than to nonpay incentives. Thus, pay incentives should be considered only a longer term possibility. Although the law may also apply to monetary nonpay awards (i.e., cash prizes), the smaller magnitude and special nature of these awards might make them easier to structure to come within the law. Furthermore, because the cost is relatively low, it may be practical to establish special funds, perhaps raised by nontransit activities (advertising, business contributions), that would not be subject to the restriction. In the long run it may be advisable to seek legal advice on how to remove the city restrictions.

In addition to these constraints, there was clear sentiment expressed at the meeting by both management and labor that nonpay awards, in particular publicized recognition, would be stronger motivators than pay incentives. This led the study group to recommend that Muni give a higher priority to nonpay incentives than to pay incentives for the immediate future.

Two philosophies were identified in designing an incentive and award system: awards can be large and go to a few operators or awards can be smaller and go to many operators. Rewarding a few results in focused recognition; positive "examples" are created. Spreading the rewards, however, spreads recognition and makes rewards more attainable; this increased attainability is consistent with the philosophy of setting achievable (not necessarily easy) goals--success is a strong motivator.

The Metropolitan Transit Commission (MTC) of Minneapolis-St. Paul struck a compromise between these two philosophies in designing its incentive system. Using a two-tiered approach, several levels of achievement with increasing rewards were established. As many as 40 to 60 percent of operators qualified for the first level; a much smaller percentage reached the highest level. This appears to be a good compromise and is recommended to Muni. It will enable Muni to recognize and give a "pat on the back" to the large group of operators doing a good job while further rewarding those doing a superlative job. (Another aspect of the MTC system recommended to Muni is that several aspects of performance are considered so that each level of achievement represents a complete performance profile.)

Muni should also widen award distribution by giving them out separately for each division or mode (i.e., instead of one driver of the month there would be several from various divisions or modes.) Alternatively, divisional awards can be given to create positive competition among divisions and esprit de corps within. This approach has been applied successfully in Houston.

There was considerable discussion at the study group meeting about which incentives are the strongest motivators; the recommendation was that this question be addressed to the employees themselves. Accordingly, a survey of operators is being undertaken. The study group stressed the importance of follow-up and action based on the survey. Otherwise, expectations might be raised but not realized, which could adversely affect morale.

Another method for selecting the strongest motivators is, where practical, to build some flexibility in the award programs. For example, an operator might be offered a choice between a cash award and some time off.

Finally, in selecting operators for awards and recognition, selection criteria should be made clear to ensure credibility. These criteria should be tied in to the performance measurement system. Additional credibility can be achieved by involving operators in the selection process (an approach used at MTC).

Discipline

Discipline works in concert with incentives and awards to encourage good performance. Although it can be considered a "negative incentive," if properly implemented, it is a positive force.

Two distinct philosophies of discipline were discussed at the meeting. One, the traditional approach, views discipline as punishment for violation. The other approach, which has been successfully applied at Seattle Metro, is dubbed "positive discipline" and views discipline as a means of clearly articulating what is expected of the operator and obtaining the operator's commitment to meeting the expectations. When these accepted expectations are not met, discipline is applied progressively (i.e., in accordance with severity and past record) along with counseling to help the operator make a commitment to and meet the expectations in the future. Should the point of termination be approached, the operator is given a decision-making leave--time to decide whether to make a commitment to meeting performance expectations or to accept termination.

The study group recommended the positive discipline approach. The success of such a system requires trust between management and labor; regular conferences with division supervisors; and the existence of an established, working grievance procedure. Muni generally possesses these prerequisites, although a greater emphasis on conferences was recommended.

The key aspects of the recommended discipline program include the following:

- Clear, understandable, agreed-upon rules. This may involve management-labor discussion and rule book revision. The current rule book, in place for some time, was deemed by both labor and management to be confusing and out-of-date. When rules are clear, understood, and agreed-on, operator commitment can be obtained.
- Fair, consistent application of specific discipline balanced by appropriate management discretion. For discipline to be respected and upheld, fairness, consistency, and specificity are essential. However, in order to (a) allow managers to manage, (b) accommodate special circumstances, and (c) consider an individual's total record, a certain amount of managerial judgment must exist. Such leeway can be built into the system and fairness and consistency can be preserved, provided there is extensive training and review of supervisors charged with administering discipline. (At MTC, a range of discipline choices is built into the system and managers are trained to make appropriate choices.)
- Discipline should be progressive. Discipline for minor infractions should be keyed to the severity of the problem and the individual's past record (over a rolling year period).
- An individual's overall performance should be considered. This will prevent demotivation of good operators who may have erred but whose past record indicates acceptable or superior performance.
- Counseling at each stage of the disciplinary process. The operator should be counseled about what performance is required and how the requirement can be met. This necessitates a management structure that allows managers administering discipline the time, information, training, and incentive to constructively counsel operators. The study group recommends that the division supervisor conference program now in effect at Muni devote more time to conferences and improve the quality of the conferences.
- Employee assistance program referrals. If, in the disciplinary process, it becomes apparent that there may be personal, stress, or substance abuse problems, the individual should be referred to the employee assistance program for appropriate counseling and rehabilitation. In the case of substance abuse or other safety threatening conditions, it is also necessary to take the operator out of service (substance abuse is discussed further later in this paper).

To successfully implement the program described, it is crucial that all managers involved in the disciplinary process receive proper training and that the program be reviewed.

Appraisal and Communication

Performance appraisal and communication were identified as crucial elements of a performance standards and motivation program. Using the information provided by performance measurement, the goals that have been targeted, and the motivational tools of incentive, award, and discipline, the manager should communicate to the operator what is expected and whether those expectations are being met and should let the operator know how to progress toward them when they are not met. The manager must also recognize and reward superior performance. The operator, in turn, needs to communicate an understanding of expectations and a commitment to them. There must be a channel open for the operator to communicate questions and problems.

To facilitate this two-way communication, all operators should have conferences with their supervisors on a regular basis. Muni already has a program of conferences by division supervisors and assistant division managers; the following paragraphs outline guidelines for conducting and reviewing conferences that may lead to improvements.

The supervisor must have the time, information, training, and incentive to conduct a constructive conference. Division managers should periodically review the conferences. It should be made clear to operators, in advance, that these conferences are separate and distinct from the disciplinary process, although disciplinary matters may be discussed. This a priori understanding will allow better communication.

The time interval between such conferences should be short enough for the conference to be remembered and thereby have an effect on performance and morale, but it should be long enough so that supervisors and operators are not overwhelmed by constant conferences. A longer interval will allow supervisors to expand the time spent with the operator and do the preparation needed to have a meaningful conference. Six months was suggested by the study group as a good interval. This represents an increase from the 3-month interval previously in place at Muni. More time for preparation should facilitate longer, higher quality conferences. As of this writing, Muni has already implemented the 6-month interval.

In addition to these formal conferences, other channels of communication can be opened or strengthened:

- Company newsletter. The study group recommends using a newsletter to acquaint operators and supervisors with each other, publicize performance trends, recognize exceptional operators, and encourage "letters to the editor" to raise questions, identify problems, and suggest solutions. Both operators and supervisors should be involved in the newsletter. [The Washington, D.C., Metropolitan Area Transit Authority (WMATA) and Seattle Metro publish newsletters and claim they are highly beneficial.] As of this writing, Muni has implemented a newsletter.

• Informal "rap sessions." Such sessions involving operators and supervisors should be used to discuss achievements, problems, and solutions.

• Joint labor-management board (JLMB). The JLMB should be continued and its role strengthened as a channel of communication and an avenue for operator participation and job enrichment.

• Intermanager communication. Improved communication within management, especially including first-line supervisors, will benefit supervisor-operator communication as well.

Aspects of Performance

The aspects of performance identified at the beginning of this paper have been referred to implicitly (and occasionally explicitly) in the preceding discussion of recommendations for an operator performance evaluation program at Muni. This section contains comments of specific pertinence to the individual aspects of performance.

Attendance and Punctuality

There are really two management goals with regard to attendance. The first is to reduce overall absenteeism. The second goal is, for a given level of absenteeism, to create incentives that will encourage the appropriate use of scheduled time off (vacation, individual personal days) in lieu of unscheduled absence (sick days) and lateness. This facilitates planning, which helps maintain schedule reliability and results in less expense than do unscheduled absences.

To achieve these goals, data collection and analysis should be undertaken so that understanding of the causes of absenteeism can be improved. Research at other properties and absenteeism records will be helpful in this task. Management should understand the different needs for different types of time off and should examine how its policies affect attendance and punctuality. Such policies include

• Size of extraboard and resulting extent of overtime use,

• Limitations on vacation or personal days (scheduled time off) that encourage sick leave abuse (unscheduled time off),

• Operator selection criteria (Are operators selected who can handle stress of the job?),

• Operator training and communication,

• Reduction of unnecessary stress, and

• Equipment availability to match operator commitment.

It is also necessary to educate operators and supervisors about the relationship of attendance and punctuality to service reliability. This will help develop a commitment to good attendance and punctuality.

Schedule Adherence

It is recognized that running ahead of schedule is more within the control of the individual operator than is running late. It should therefore be the primary focus in measuring individual operator performance. Running late, however, should not be ignored. The limited extent of driver control over lateness should be explored, along with the full range of causes, including equipment availability, scheduling, traffic, weather and road conditions, and load factors.

In measuring schedule adherence, it is important that those performing the measurement be objective. If, for example, street supervisors are charged with gathering data on schedule adherence and those same street supervisors are evaluated on the basis of schedule adherence, there may be a problem with objectivity.

Safety

Safety training and awareness should be emphasized. In conjunction with efforts to set standards and motivate safety, it is important to clearly determine what is and what is not an operator preventable accident and to investigate each accident promptly and thoroughly to establish cause.

The study group was made aware of a separate "Accident Peer Group Review" that addressed safety at Muni. Its recommendations provide additional input to the development of measures, targets, incentives, discipline, and so forth.

Appearance and Courtesy

Instilling pride and a sense of commitment in general should produce specific gains in this area. Awards and recognition are particularly appropriate to reward courteous service.

Involving operators in the design of uniforms was recommended to encourage their subsequent use. (This approach has been used in Seattle.) A process for union input to the design of uniforms is already in place at Muni.

Substance Abuse and Dealing with Stress

Although substance abuse is difficult to quantify, it is an extremely important area because it directly affects safety. A strong, accessible, and trusted employee assistance program (EAP) is a major preventive asset. One approach to encouraging preventive use of the EAP before an accident occurs is to offer the following (a version of which is in place at WMATA):

• If an operator has a drug or alcohol problem that affects the ability to drive safely, and so notifies the supervisor, the operator will be put on a noncritical job (e.g., cleaning) and sent to the EAP, drug rehabilitation program, or other appropriate source of help. After successful completion of the prescribed program, the driver will be returned to service with the understanding that recurrence of the problem will necessitate dismissal. (Follow-up counseling will be available, however, to help the operator avoid recurrence.)

• If operators with drug or alcohol problems fail to take advantage of this opportunity to seek help, however, they are subject to normal disciplinary procedure (i.e., termination of employment). (At WMATA, every employee involved in an accident is required to submit to a blood test to determine the presence of alcohol or drugs.)

In all cases the emphasis should be on getting the unsafe driver off the road before an accident occurs. Correct the condition, if possible, but first get the operator off the road.

The EAP can also provide assistance for other stress-related problems (e.g., divorce, depression) that may impair job performance; all Muni employees should be encouraged to make use of this resource before a stressful condition results in serious deterioration of performance.

Observance of Rules, Operating Procedures, and Directives

This aspect of performance includes miscellaneous yet important items, such as running with proper signs, calling out stops, stopping at designated locations, responding to passenger inquiries, following supervisor directives, and awareness of notices, all of which affect the quality of service and are performance indicators in themselves. Improved communication and appropriate use of disciplinary procedures should produce gains in this area. Many of the aforementioned communications improvements, such as newsletters and more conferences, can be used in efforts to increase observance of rules. Specific incentives and awards may be created to reward attention to these aspects of quality of service.

CONCLUSIONS

The study group review process enabled Muni to benefit from the experiences of other transit properties in tackling the issues related to evaluating employee performance and motivating employees to improve service. The study group initiated a mutually beneficial exchange of ideas among representatives of several properties, brought new ideas to Muni staff at several levels, and fostered a dialogue between union and management that will be beneficial to the special demonstration project and to Muni employees. As a result, Muni was able to reevaluate a number of its programs and to plan enhancements for both the short and the long term. The developments at Muni in implementing these enhancements, as well as the guidelines for employee performance evaluation and motivation programs, should prove useful to other transit properties.

EPILOGUE

As of this writing, Muni has implemented a number of the recommendations of the study group. Current activities include installation of an operator performance monitoring system that has been developed for use on a microcomputer using commercially available data base management software, setting of performance targets, review of attendence policies, expanded recognition programs, revisions to the operator rule book, and analysis of street supervision activities to design an experiment for implementation in the coming year.

ACKNOWLEDGMENTS

This research was sponsored by the Urban Mass Transportation Administration Service and Methods Demonstration Program; the UMTA technical monitor is Joseph Goodman. Multisystems is assisting under a contract with the Transportation Systems Center (TSC); the project manager at TSC is Robert Casey. The views expressed in the paper are not necessarily those of the sponsoring agencies.

The contribution of the following key study group participants along with that of other active observers is gratefully acknowledged. Transit industry representatives: Shirley Delibero, Director of Bus Services, Washington Metropolitan Area Transit Authority; Robert Foy, Acting General Manager, Mass Transportation Authority, Flint, Michigan; Phil Garey, Director of Transportation, Regional Transit Authority, New Orleans, Louisiana; Dan Graczyk, Manager of Base Operations, Seattle Metro; Martin Reiner, Executive Assistant to the General Manager, Metropolitan Transit Authority of Harris County, Houston, Texas; and Tom Vida, Assistant Director of Transportation, Metropolitan Transit Commission, Minneapolis-St. Paul, Minnesota. San Francisco Municipal Railway participants: Ray Antonio, Executive Vice President, Transport Workers Union, Local 250A; Bobbie Brown, Potrero Division Chairman, Transport Workers Union, Local 250A; John Hopper, Business Representative, Transport Workers Union, Local 200; George Newkirk, Manager, Field Support; Dianne Nygaard, (former) Manager, Training; Dick Sinigiani, (former) Manager, Field Operations; Paul Tolliver, Deputy General Manager of Operations; and Alvin White, Vice President, Transport Workers Union, Local 200.

REFERENCES

1. J.M. Greiner et. al. Productivity and Motivation: A Review of State and Local Government Initiatives. The Urban Institute Press, Washington, D.C., 1981.
2. H.S. Baker and O. Schueftan. Study of Operator Absenteeism and Workers' Compensation Trends in the Urban Mass Transportation Industry. UMTA-PA-0050-80. Peat, Marwick, Mitchell and Co.; U.S. Department of Transportation, 1980.
3. L.C. MacDorman and J.C. MacDorman. The Transit Extraboard: Some Opportunities for Cost Saving. Presented at Annual Meeting of the American Public Transit Association, Boston, Mass., Oct. 1982.
4. L.C. MacDorman. Labor Management Performance and Its Effect on Labor Productivity. Presented at New York Statewide Transit Meeting, June 1983.
5. C. Perin. Vehicle Operator Absenteeism and Transit Productivity. Staff study. Transportation Systems Center, U.S. Department of Transportation, Cambridge, Mass., Feb. 1982.
6. J.L. Perry. A Study of Extraboard Scheduling, Workers' Compensation and Operator Stress in California Public Transit. Institute of Transportation Studies and Graduate School of Management, University of California, Irvine, March 1983.
7. L. Long and J.L. Perry. Economic and Occupational Causes of Transit Operator Absenteeism: A Review of Research. Institute of Transportation Studies and Graduate School of Management, University of California, Irvine, March 1984.
8. J.L. Williams. Employee Assistance: A Low-Cost Alternative for Increasing Work Force Productivity. Presented at the Annual Meeting of the American Public Transit Association, Oct. 1982.
9. R. Snoble. Participatory Management: Creating an Environment that Allows Employees to Help Attain Excellence. Presented at the Western Meeting of the American Public Transit Association, April 1983.
10. J.M. English. Improved Labor Productivity Through a Comprehensive Management Approach. Presented at the Annual Meeting of the American Public Transit Association, Boston, Mass., Oct. 1982.
11. D.J. Fitzgerald. Absenteeism Controls at the Capital District Transportation Authority. Presented at the Annual Meeting of the American Public Transit Association, Boston, Mass., Oct. 1982.