TRANSPORTATION RESEARCH RECORD 862

Short-Range Transit Operations Planning and Development

TRB
TRANSPORTATION RESEARCH BOARD
NATIONAL ACADEMY OF SCIENCES

The **Transportation Research Record** series consists of collections of papers in a given subject. Most of the papers in a Transportation Research Record were originally prepared for presentation at a TRB Annual Meeting. All papers (both Annual Meeting papers and those submitted solely for publication) have been reviewed and accepted for publication by TRB's peer review process according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in these papers are those of the authors and do not necessarily reflect those of the sponsoring committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of TRB activities.

Transportation Research Records are issued irregularly; approximately 50 are released each year. Each is classified according to the modes and subject areas dealt with in the individual papers it contains. TRB publications are available on direct order from TRB, or they may 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 Academy of Sciences, 2101 Constitution Avenue, N.W., Washington, DC 20418.

TRANSPORTATION RESEARCH RECORD 862

Short-Range Transit Operations Planning and Development

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1982

Transportation Research Record 862
Price $8.60
Edited for TRB by Scott C. Herman, Naomi Kassabian, Mary McLaughlin, and Susan Singer-Bart

mode
2 public transit

subject areas
11 administration
12 planning
14 finance
54 operations and traffic control

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.

Short-range transit operations planning and development.

(Transportation research record; 862)
Papers presented at the Workshop on Short-range Transit Operations Planning and Management, held in Atlanta, Ga., Mar. 7-10, 1982, conducted by the Transportation Research Board and sponsored by the U.S. Dept. of Transportation.
1. Local transit–United States–Planning–Congresses.
I. National Research Council (U.S.). Transportation Board.
II. Workshop on Short-Range Transit Operations Planning and Management (1982: Atlanta, Ga.) IV. Series.
TE7.H5 no. 862 [HE4456] 380.5s [388.4'068] 82-18995
ISBN 0-309-03365-9 ISSN 0361-1981

Sponsorship of the Papers in This Transportation Research Record

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

Steering Committee for the Workshop on Short-Range Transit Operations Planning and Management
Jack Reilly, Capital District Authority, chairman
John Dockendorf, Ronald J. Hartman, James M. Holec, Jr., David R. Miller, John Sajovec, Gary S. Spivack, Douglas L. Wentworth, Nigel H.M. Wilson, Donald G. Yuratovac

Liaison Representatives, Brian E. McCollom and Granville E. Paules

Wm. Campbell Graeub, Transportation Research Board staff

The organizational units, officers, and members are as of December 31, 1981.

Contents

Authors of the Papers in This Record

Abrams, Edward M., Abrams-Cherwony and Associates, PSFS Building, Suite 3139, 12 South 12th Street, Philadelphia, PA 19107

Cherwony, Walter, Abrams-Cherwony and Associates, PSFS Building, Suite 3139, 12 South 12th Street, Philadelphia, PA 19107

Dockendorf, John, Mass Transit Assistance Division, Pennsylvania Department of Transportation, Transportation and Safety Building, Commonwealth and Forster Streets, Harrisburg, PA 17120

González, Sergio L., Department of Civil Engineering, University of Puerto Rico, Mayagüez, Puerto Rico 00708

Jones, Janet, Tri-Met, 4012 S.E. 17th Avenue, Portland, OR 97202

Mayworm, Patrick D., Ecosometrics, Inc., 4715 Cordell Avenue, Bethesda, MD 20814

Peskin, Robert L., Peat, Marwick, Mitchell and Co., 1900 K Street, N.W., Washington, DC 20006

Porter, Benjamin, Booz-Allen and Hamilton, Inc., 400 Market Street, Suite 900, Philadelphia, PA 19106

Sheskin, Ira M., Department of Geography, University of Miami, Coral Gables, FL 33124

Stopher, Peter R., Schimpeler-Corradino Associates, 300 Palermo Avenue, Coral Gables, FL 33134

Wilson, Nigel H.M., Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Yuratovac, Donald G., Greater Cleveland Regional Transit Authority, 615 Superior Avenue, Cleveland, OH 44113

Preface

The papers in this Record were presented at the Workshop on Short-Range Transit Operations Planning and Management held in Atlanta, Georgia, March 7-10, 1982. This workshop was conducted by the Transportation Research Board (TRB) and was sponsored by the U.S. Department of Transportation.

As a result of changing federal policy toward financing mass transit operations, short-range operations planning has become more important to transit operating and planning agencies. The papers in this Record are representative of the state of the art of short-range transit operations planning. These articles are a mix of both survey pieces and case studies. One paper presented at the workshop is not included here but has since been published by TRB in Transportation Research Record 857 (Design of Bus Transit Monitoring Program by John Attanucci, Nigel Wilson, Brian McCollum, and Imogene Burns).

These papers each reflect current concerns about operating transit services in an era of dwindling resources. For the most part they introduce methods of data collection and, more importantly, data analysis needed to support decisions on transit levels and prices. Although not all papers relate directly to the operating practices of each transit agency, much of the material in this Record can be transferred to other agencies to improve transit operating performance.

The papers are in four areas of transit operations planning: (a) service design, (b) surveillance and monitoring, (c) fare and revenue analysis, and (d) cost analysis. In a very broad sense, nearly all decisions on short-range transit operations planning involve one or more of these activities.

Methods for Service Design

NIGEL H.M. WILSON AND SERGIO L. GONZÁLEZ

This paper explores current practice in the design of bus services. It focuses on methods used to identify problems in the existing system and to design service changes. Current practice is handicapped by the lack of reliable data of a type desirable for good planning, and problem-identification activities consist mainly of flagging routes that rank low in terms of cost-effectiveness indices. As a result, only a small set of potential improvement actions are usually considered, and the (usually implicit) objectives in providing transit service are not effectively included in the process. The paper then recommends changes to the existing process that would encourage planners to look for opportunities that may exist on routes that are not flagged as substandard. Modifications are also proposed that recognize the multiple objectives that transit operators should be dealing with.

This paper is intended to provoke discussion on the effectiveness of current approaches to service design in the transit industry. Although service design has always been an important, perhaps even central, element in short-range transit planning, (SRTP) changes in the environment within which transit operates is placing increased stress on this function. During the past decade service planning took place, increasing resources were made available for transit, and the focus was on issues of service expansion. Now tighter financial constraints are forcing operators to look for ways of getting more out of existing resources, and often the question is where and how to reduce service. Approaches to service planning appropriate in the recent past may be less satisfactory in the future given this shifting emphasis.

SRTP: SCOPE AND BASIC ACTIVITIES

SRTP is the process of monitoring the operations of the transit system and planning modifications that can be implemented during the next schedule change.

An important implication of this definition is the short time frame of SRTP; in particular, some changes to the transit system are not available in SRTP. Examples of these actions include the acquisition of new vehicles, changes in the general configuration of the network (e.g., grid versus radial), changes in the fare structure, planning of major capital facilities, and the introduction of new transportation modes. Decisions related to these options are usually the domain of longer-range planning and programming that should, of course, be coordinated with SRTP.

The remaining system modifications, those feasible during SRTP, can be grouped at various levels: The system coverage level, the route structure level, the frequency level, and the control level (1). A distinction can be made at each level between actions that tend to increase cost and ridership and those that tend to decrease cost and ridership. Depending on the financial (and other) constraints that face the property, actions taken may be predominantly of one type or the other or may be a mixture, in which case the system is being fine tuned to better meet the objectives of the agency. Of course, system fine tuning may also include actions aimed at more efficient production of the same level of transit service.

At the highest level, feasible actions include implementation of a new route, extension of an existing route, replacement of a small set of routes with a new set, discontinuance of service on a route, shortening of a route, and the making of minor modifications in route alignment. Another type of action at this level is substitution of a privately provided service for the existing publicly operated fixed route. This new service might be paratransit or fixed route, the aim being to reduce cost and provide a more suitable service. Actions at this level are the most disruptive for the public and so merit the most intense scrutiny. Consequently, many of these actions are among the most time-consuming to plan and implement within the short-range planning process.

Actions at the route-structure level are the splitting of a route into two nonoverlapping segments, the splitting of a route into zones or express and local segments, the linking of two existing routes to form one new one, and the introduction of deadheading of some buses of a route. Although these actions are generally less disruptive than changes in system coverage, they do require some re-education of the public and careful planning.

At the frequency level more or less service can be provided on a given route at a specific time of day. Finally, at the control level, the following actions, usually aimed at maintaining closer adherence to the schedule, are considered: installation or removal of a control point on a route at which slack is built into the schedule, a change in the running time allowed for a route segment, and modification of the layover time (a special case of the first strategy).

Notice that this set of system changes contains general modifications that can be applied to any part of the network during any time period. Because of their generality, we refer to these changes as generic actions. An alternative is defined as the application of a generic action to a part of the transit system. For example, an alternative may consist of applying the generic action frequency change (for example, service reduced from five to four buses/h) to a specific route during the morning peak.

Based on the definitions given above, a more operational definition of SRTP can be developed. It can be viewed as the process of determining where on the existing system and during which time period generic actions should be taken to develop the most promising alternatives for implementation during the next schedule change.

The list below summarizes the generic actions in SRTP. Although the number of generic actions is small, SRTP is complex because the number of elements of the transit system to which each generic action can be applied is large, which results in an even larger set of feasible alternatives:

1. Area coverage level--new route, route extension, a small set of routes replaced by a new set, route abandonment, shortening a route, route realignment, and change of service type or operator;
2. Route structure level--route splitting, zonal service, express or local service, linking of two routes, and deadheading;
3. Frequency level--changes in route frequency; and
4. Control level--installing or removing control points, changes in layover time or positioning time; and modifying running times.

This requires that SRTP, like most complex planning problems, be structured around the following set of basic, sequential activities:

1. Problem identification,
2. Design of alternatives,
3. Analysis of each alternative, and
4. Recommendation of the most promising alternative.

Problem identification involves the gathering and review of data on individual services to determine whether or not a problem exists. The idea of problem identification clearly implies that the objectives in providing the service are not being well met and that some change in the service may be warranted. Problem identification is an ongoing process that must be supported by some type of data collection and analysis.

Once a problem has been found, one or more generic actions could be taken to alleviate it. The design of alternative actions may be quite straightforward or difficult. For example, a route that exhibits extreme crowding would obviously be considered for increased service frequency whereas a route that has very unreliable service might be a candidate for several different generic actions.

Each alternative is subject to some type of analysis to predict the impacts of adopting it. This analysis process is often largely judgmental, but it may include one or more models to predict impacts. The planner will be concerned about impacts such as

1. Changes in operating costs based on driver and vehicle requirements or
2. Changes in ridership and revenue.

More generally, also considered would be the extent to which the initial problem would be corrected and the degree to which underlying transit objectives would be furthered.

Once the impacts associated with each alternative are predicted, the most suitable alternative can be recommended based on review of the possibilities by different departments within the organization and, in many cases, with external groups as well. The extent of internal and external discussion and negotiation will, of course, depend on the generic action being considered. Typically, a lengthier process is involved in determination of the best service-reduction-type action in which the public is adversely affected than if an expansion-type action is being taken.

CURRENT PRACTICE

General statements about current practice in the transit industry are dangerous because of the diversity of methods used among operators of different size and in different parts of the country. Furthermore, in the amount of space available in this paper, it is impossible to report on all the different approaches to planning now being used for each of the basic activities identified in the preceding section. Consequently, here we will focus on the first two activities--problem identification and design of alternatives. Despite the dangers, this discussion will be couched in general terms; recognize that there are exceptions to most of the points made. The discussion of current practice that follows is based on our personal experience and information available in the transit literature (1-5).

One of the most important influences on SRTP is the type and quality of information available to the planner. Currently, the variation is great among properties in the data available in terms of the type, level of detail, frequency, and amount of data collected, and perceived quality. Recently many properties have reevaluated their data-collection programs in light of Section 15 reporting requirements of the Urban Mass Transportation Act of 1964, as amended, and the need to make tough choices on where service cuts should be made. Even so, there is still room for the improvement of most data-collection programs by more formal consideration of accuracy and sampling issues (6). Many planners, particularly those in the larger transit authorities, think that they do not obtain the type and quality of data on existing services needed to make sound planning decisions. In some cases this problem is exacerbated by tensions or poor communication within the property among those who collect the data, the planners, and the schedulers. These organizational difficulties are not as common in smaller agencies that can also often make effective use of information informally received from drivers and starters.

Typically, raw data must be summarized or processed before it is in a form useful for the planner. Here again, industry practice varies widely: Many properties rely on completely manual tabulation and file storage, and others have moved aggressively toward computerization. If data are handled manually, it can be quite difficult, particularly when urgent actions are required, as in a budget crisis, to bring disparate types of data gathered at different times to bear effectively on an analysis of a particular route or set of routes.

Whatever information is available from data-collection activities and other input, such as passenger suggestions and complaints, is used to evaluate current services and to identify problems that require attention. This type of analysis typically looks for unacceptable performance as defined in terms of measures and standards that may have been formally adopted by the property or just be based on the experience of the planner. Services may also be flagged for further study if they appear to have changed significantly over time, even if they are not substandard. This type of analysis is often hampered by lack of a composite picture of a route (a route profile) from earlier data collection and analysis cycles.

Central, then, to the problem-identification phase of current SRTP is the use of service measures and, to a lesser extent, service standards. Service measures are statistical summaries of route data such as passengers per bus hour, revenue per cost, and percentage of buses on time, whereas service standards establish a critical level for a particular service measure, such as 25 passengers/bus hour as the minimum acceptable level of productivity for a route (7).

Service measures are used by virtually all properties but, although many different measures have been proposed, rarely does a specific agency use more than three in route planning. Principal reasons for this are that planners focus on only a few problems, such as overcrowding or underuse, that can be represented by a few measures, and the quality and amount of data and limited planning resources preclude effective use of more measures. The vast majority of service measures used are ridership oriented, such as passengers per bus hour, passengers per bus mile, and peak load factor. Other service measures that are commonly used are subsidy per passenger, revenue per cost, and on-time performance. These measures accurately reflect the primary concerns of most properties with ridership and cost-effectiveness and the dominant role that schedules play in modifying the services provided.

Service standards are more often used as guidelines to indicate when a route may be in need of study than to dictate that a specific change should be made. Ridership and revenue-oriented standards are generally used to flag poorly performing routes that are well below average performance for the system as a whole. Beyond this screening role, these measures are not used effectively in planning route changes. On the other hand, measures that are directly related to possible schedule changes such as peak load factor and schedule adherence are often used directly to design changes. As expected, the actual standards used for specific measures vary greatly among properties, depending on their size and financial situation.

One of the important benefits of establishing service policies in the form of measures and standards should have been to encourage operators to think hard about the objectives behind the provision of transit service in that metropolitan area. This process was badly needed after the shift of the industry from the private to the public sector removed the profit-maximization objective. Unfortunately, it is now common in the transit industry to find goals and objectives, where they are stated at all, couched in general and vague terms. Little guidance is given from above about relative priorities and the resolution of conflicting objectives, so the planner is left very much alone in defining useful measures to identify problems and later in the evaluation of alternatives. Where standards are used, it is not clear that the levels have been chosen soundly or what the impacts of changing them would be on achieving agency objectives. Indeed, the ease with which some agencies have made significant changes in these standards suggests that the standards may be arbitrary and serve to simplify the planning problem without a sound basis.

Once a problem has been identified through the collection and analysis of data, specific alternatives must be designed to deal with it. If the problem is minor, for example, heavy loads on a route or poor schedule adherence, the solution will often be quite straightforward, such as adjusting running or recovery times and adding bus trips. These scheduling changes can generally be implemented without extensive analysis in the next driver pick provided that budgetary or vehicle fleet constraints do not prevent the obvious solution and that the required schedule changes can be accomplished in time.

For more substantial problems a planner will often be given responsibility for development of alternative actions and their analysis and evaluation. Frequently input will be obtained from drivers, supervisors, and the community, and additional data may be collected to clarify the problem and on which to base the design of solutions. In some cases design standards are used such as policy headways or route-accessibility guidelines to suggest appropriate types of changes and to disqualify other proposals. Often a planner will have a portfolio of route changes that he or she would like to make if an opportunity presents itself; some changes will originate internally and others will come from riders or the community in general. Political pressure is often a factor in the selection of routes for service reductions and most planners will think about political reactions before recommending specific change. This is a major factor in the retention of some extremely poor routes for which no useful strategies can be developed for improving performance. It is fair to say that the process of designing useful changes for routes that have been flagged as substandard in terms of cost-effectiveness is not very productive. Thus, schedule changes represent the clear majority of service changes implemented during any planning cycle.

In summarizing current short-range planning practice the following critical points should be noted:

1. The type and reliability of data available to the planner severely limit his or her ability to identify problems with the system and to design effective responses to problems that have been identified.
2. Even if reliable data were available, an important problem that needs to be addressed is how to summarize and use these data to help the planner in problem identification and design of alternatives.
3. Although service standards are effective in identifying routes that can be improved with scheduling changes, service standards that flag poor routes in terms of cost-effectiveness are not so effective in identifying routes that can be changed to better meet the multiple objectives of the agency.
4. The process of designing specific alternatives to solve problems at the route level is ad hoc and is probably not effective in identifying changes such as express, zonal, or deadheading services that might improve performance.
5. In an environment of decreasing planning resources and increasing pressure to get the most out of transit subsidies, the basic objectives in providing transit service at all must be included more directly in the design of changes in individual services.

PROPOSED SRTP PROCESS

The proposed short-range planning process is a modification of the current process that tries to address the weaknesses that were identified in the previous section. In addition, it recognizes, unlike many previous methods for system design, that SRTP occurs in a complex institution. [For a review of some of these methods, see Furth (8) and Furth and Wilson (9)]. The following characteristics of transit operating agencies are incorporated, either as constraints or guidelines, in the proposed method.

1. Multiple goals--The SRTP process must recognize that properties have multiple goals such as providing mobility for those without automobiles, reducing traffic congestion, and reducing energy consumption. These goals are associated with specific routes (that serve specific markets) and times of day (e.g., congestion occurs principally during peak periods). Thus, analysis and design must be based on data at the route and time period levels.
2. Coordination with related activities in the agency--Short-range planning is only one activity within a property and its effectiveness depends on the interfaces with other elements of the organization. For example, after approval, actions recommended by planners must be implemented by the scheduling group. Only by considering the interdependencies between SRTP and other activities can it be ensured that actions recommended by planning will be acceptable to the total organization.
3. Constraints in planning resources--Planning resources available for SRTP (principally time and personnel) are tightly constrained; thus, it is important to focus attention on services that have high potential for positive payoff. Since detailed analysis of all possible alternatives and services is impossible, a screening procedure is essential. This also implies that large-scale or radical changes that require extensive analysis and runcutting cannot usually be undertaken.
4. Changes in the agency's environment--SRTP has to be able to respond effectively to changes in the operating situation of the agency, such as sudden changes in budget.

5. Limitations of technical analysis--Since the state of the art in transit technical analysis is imperfect, quantitative methods should be used to supplement the planner's judgment and experience, not to replace it.

The presentation of the proposed SRTP process will be structured around the basic steps of SRTP that are being emphasized in this paper: problem identification and design of alternatives. In transit, the problem-identification step usually also includes a preliminary design of alternatives because a set of solutions (or generic actions) is often directly associated with each problem. For each of these steps, modifications to current practice will be proposed, both in terms of the processes followed and the required data and analytical support.

Regarding the data and analytical support for the problem-identification and design-alternative steps, note that, although the general requirement of both steps is quite similar, the levels of detail are very different. In problem identification, a large number of routes have to be screened; thus, because of the time constraints, detailed route analysis is inappropriate. For this step, simple aggregate performance measures are used. Design of alternatives, however, requires more detailed information and analyses, many times by route segment and time period. These analyses can be undertaken for the relatively small number of routes that have previously been identified as good candidates for changes.

Problem Identification

Problem identification as implemented in current practice has two basic limitations. First, the term problem is often used as a synonym for substandard performance. This narrow definition usually excludes routes that, although currently performing satisfactorily, could be improved significantly. The second limitation is that the multiple objectives of the transit agency are usually not incorporated in the problem-identification step. For example, routes that have a low revenue-to-cost ratio are usually considered substandard; however, this low ratio may be caused by a large number of elderly passengers, which is usually viewed as an asset.

Defining Problems in SRTP

For dealing with the first limitation, a better definition of problem is required. For developing this definition, recall the concept of generic action that was introduced earlier in this paper. These actions are the control variables available to the agency to modify its system in order to improve its performance. With this in mind, we define a problem route as one whose performance could be significantly improved with the application of one of the generic actions. This definition encompasses both types of problem routes of interest, those that are substandard, for example in terms of schedule adherence or productivity, and those whose efficiency in providing a given service could be improved.

We recognize that both types of problems are important for problem identification; however, because they are very different, different methods are required. Both methods, however, are based on relations between generic actions and types of problems.

The first method, referred to as the problem-centered approach, is similar to current practice. The major difference is the recognition that the generic actions that are applicable for dealing with a specific problem is a small subset of all the possible actions. This implies that, to narrow down the set of all possible changes to a small subset (which is the role of problem identification), we just need a set of performance measures that will indicate the existence of a problem in any given route.

Table 1 presents the starting point for the (traditional) problem-centered approach. In this table we present the most important performance indicators required to identify each type of problem and its possible solutions. Note, however, that in this table we are not directly incorporating the multiple objectives of transit operators. This issue and methods to deal with it are discussed later.

The second approach to problem identification is most appropriate for improving parts of the system in which heavy pressure for change does not currently exist; i.e., for routes that have no obvious problems. The key to this approach is realizing that the potential of any generic action for improving the performance of a route is dependent on the existence of a set of conditions on that route. The problem then becomes one of identifying the set of conditions that will indicate the potential for each generic action and finding measures for these conditions. Since this type of problem-identification approach is structured around the generic actions, we refer to it as generic-action centered.

Principal advantages of this generic-action-centered search are twofold. First, actions that are not usually appropriate for problem routes are included directly in the set of potential service improvements. For example, problem routes are usually characterized by low ridership and policy headways, and actions such as express or zonal routing or partial deadheading will never be of value for these routes, and hence may never be considered by the planner. Second, some routes that are not problems will be the subject of planners attention, which may result in implementation of unusual actions such as zoning or deadheading that might either free resources to tackle problem routes or improve overall service quality.

Table 2 presents the starting point for the generic-action-centered approach to problem identification. It contains each generic action and the set of conditions that indicate its potential.

Notice that, for measuring these conditions, several types of indicators may be required. For example, schedule adherence can be characterized by a numerical indicator such as percentage of trips that are late. For identifying a point on a route that has low ridership, which is required for several generic actions, a graphical load profile, similar to the one shown in Figure 1, may be most appropriate (10). For measuring the potential for a route extension, a map in which areas of new development and possible traffic generators are marked may be useful. Locations of possible bus turnaround points, which are important for route extensions and splitting, may also be plotted on a map. Verbal indicators may consist of comments from planners, supervisors, and drivers that could later be verified with data.

Of course, the choice of which measures or indicators to use, for both the problem- and generic-centered approaches, will depend on the cost, accuracy, and reliability of each type of information as well as on the data currently available to the operator. For example, suggestions or comments from drivers will cost very little, but the reliability of this method will depend on the availability of mechanisms and incentives to transmit the information accurately. Performance measures can be used only if there is an ongoing data-collection effort, which will probably provide the most reliable information, but in a more costly manner.

Multiple Objectives and Search

The other limitation of the current problem-identification process is that it does not recognize the multiple, and often conflicting, objectives of the transit agency. When dealing with multiple objectives, it is not possible to find a single measure that indicates goal attainment; usually different measures will be required for each goal. For example, a measure such as number of elderly riders could be used to evaluate the performance of a route with respect to the goal of providing service to the elderly. Another measure for this goal could be the percentage of elderly within 0.25 mile of a route. For the goal of cost efficiency of service, the traditional revenue-to-cost ratio can be used.

Since some goals are conflicting, attainment of an acceptable level in one of them sometimes results in difficulties in meeting another. For example, a route that serves many elderly will often have a low revenue-to-cost ratio because of low fares for the elderly. To deal with this problem, we propose ranking all the routes in terms of the performance measures selected for each goal. It is important to do this ranking by corridor (or area) and time period because, in this way, only similar services will be evaluated against each other and spatial and temporal equity will also be taken into account. These rankings can then be used as a screening mechanism; for example, for the goal of cost efficiency, the routes in the lowest 10 percent revenue-to-cost ratios that are not in the upper 20-30 percent in the rankings for other goals could be screened for further analysis.

To demonstrate the feasibility of this approach, a computer program to summarize the required information has been developed. The output of this program is presented in Figure 2. In this particular case, the variables used are passengers per trip for different fare categories, total passengers per trips, and revenue per trip. Of course, other measures could also have been used.

The report shows the ranking and values of the measures for each route in the system. Two summary variables that indicate the number of measures for which a route belongs in the upper and lower 15 percent are also included and routes are further cate-

Table 1. Problems and corresponding actions.

Problem	Indicator	Possible Action
Schedule adherence	Percent of trips late	Holding strategy, increase run time or layover, modify route
Unacceptable crowding	Load	Increase frequency
Poor productivity	Revenue/cost	Decrease frequency
	Load	Split route
	Passengers per vehicle hour	Short turn strategies; local, express, zonal strategies; partial deadheading
Poor vehicle use	Revenue/cost	Eliminate route segments
	Passengers per vehicle hour	Eliminate trips, extend route, modify schedule

Table 2. Generic actions and appropriate route conditions.

Generic Action	Route Condition
Holding strategy	Schedule adherence problem, long route, point on route with low through ridership
Increase running Layover time	Schedule adherence problem Low loads
Increase frequency	Unacceptable crowding, moderate rather than high ridership, even load profile
Decrease frequency	Low productivity and loads, headways below policy levels
Split route	Low productivity, uneven load profile, long route
Short turns	Tapering load profile, long route, high ridership
Express or zonal	High ridership, tapering load profile, long route, large time differentials local or express zone
Partial deadheading	Large imbalance in flows, large time differential in service and high frequencies
Eliminate route segment	Low ridership generation on segment, vehicle savings possible from elimination, higher frequency possible from elimination
Eliminate trips	Low ridership on trips, high cost savings from elimination

Figure 1. Plot of cumulative boardings and alightings for trip in a given direction.

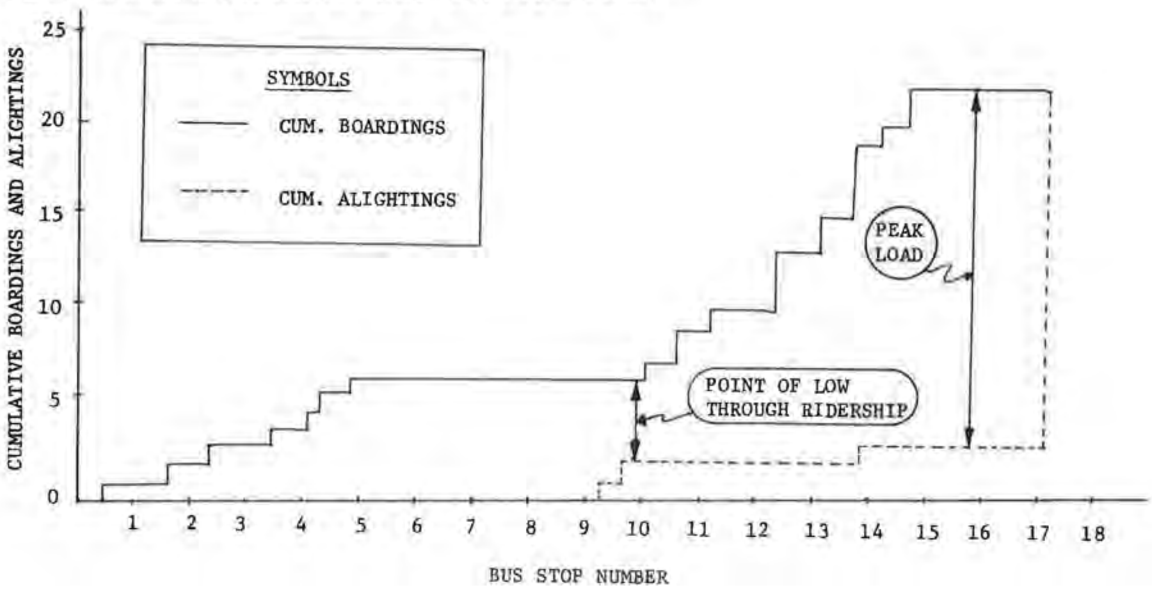


Figure 2. Multiple-objective ranking table for problem identification.

ROUTE	DAILY SCHEDULED TRIPS	THE VARIABLES BELOW ARE, FOR EACH CATEGORY, THE AVERAGE PER TRIP AND THE RANK: REVENUE AVG	REVENUE RANK	TOTAL PASSENGERS AVG	TOTAL PASSENGERS RANK	REGULAR AVG	REGULAR RANK	TRANSFER AVG	TRANSFER RANK	STUDENT AVG	STUDENT RANK	ELDERLY AVG	ELDERLY RANK	HANDICAP AVG	HANDICAP RANK	CHILDREN AVG	CHILDREN RANK	NO. OF CATEGORIES IN TOP 15%	NO. OF CATEGORIES IN BOTTOM 15%
NOTE: THE ROUTES THAT FOLLOW ARE IN THE TOP 15% FOR AT LEAST ONE CATEGORY WITHOUT BEING IN THE BOTTOM 15% OF ANY CATEGORY																			
1	104	$15.26	2	52.72	1	26.36	1	13.18	1	2.11	2	5.27	2	0.53	3	5.27	1	7	0
2	54	$15.53	1	49.54	2	26.25	2	11.89	2	0.99	6	7.43	1	0.99	1	1.98	3	6	0
3	78	$6.95	6	23.97	4	10.55	7	5.99	3	2.40	1	3.59	7	0.24	10	1.20	5	1	0
4	30	$6.84	7	19.83	7	9.72	9	1.98	8	1.98	3	4.96	3	0.59	2	0.59	8	1	0
13	26	$7.27	5	22.60	6	11.52	6	3.39	6	1.81	4	3.39	8	0.45	5	2.03	2	1	0
NOTE: THE ROUTES THAT FOLLOW ARE NEITHER IN THE TOP OR BOTTOM 15% IN ANY CATEGORY																			
5	68	$8.91	3	24.00	3	15.60	3	3.60	5	0.48	10	3.60	6	0.24	9	0.48	10	0	0
8	66	$7.75	4	23.86	5	12.17	5	4.77	4	1.19	5	4.77	4	0.48	4	0.48	11	0	0
9	24	$6.79	8	16.82	10	12.45	4	1.68	10	0.34	12	1.68	10	0.17	12	0.50	9	0	0
10	24	$5.98	9	16.90	9	9.47	10	1.69	9	0.34	11	4.23	5	0.34	7	0.85	6	0	0
12	42	$5.93	10	17.17	8	10.30	8	2.58	7	0.86	7	1.72	9	0.34	6	1.37	4	0	0
NOTE: THE ROUTES THAT FOLLOW ARE IN THE BOTTOM 15% FOR AT LEAST ONE CATEGORY WITHOUT BEING IN THE TOP 15% OF ANY CATEGORY																			
6	46	$4.76	11	12.18	11	8.28	11	1.22	12	0.85	8	1.22	11	0.24	8	0.37	13	0	1

gorized by these variables. The group on top consists of those that perform very well with respect to some measures and are at least average in the others. The last group contains routes that are the worst performers in some categories and are not good performers in any single category. These are candidates for remedial action. The analysis presented in Figure 2 can be applied to different types of data. In this particular case, monthly revenue by route was used, together with judgmental estimates of the percentages, by route, of each fare category. These data are usually available to most operators. If fare classification counts were available, these could also have been used and would have provided more accurate input to the program.

It is interesting to look at routes 4 and 13 in the table. By using traditional ranking schemes based on revenue/cost and other economic performance measures, these routes would probably have been flagged for remedial action. However, as our analysis points out, these are excellent performers with respect to other objectives. This, of course, should be considered when recommending any service changes.

Design of Alternatives

The output of the problem-identification step is a small subset of routes that have the potential for improvement by the application of one or more generic actions. The purpose of the design-of-alternatives step is to develop detailed alternative changes for these routes that can then be evaluated for possible implementation.

The analyses required for this design stage are more detailed because during it specific decisions about where and when to implement the generic actions have to be made. This more detailed analysis is possible because the number of routes now being considered is much smaller. As an example, assume that a schedule-adherence problem has been identified on a route. To develop a specific alternative solution to this problem, information on route segment level actual running times may be required. For more detailed analysis of a potential service cut, segment-level ridership data by passenger type may be needed.

Again, this information may be collected and summarized in different ways. Estimates of route segment ridership may be obtained from bus drivers. Alternatively, data collected through riding counts can be used.

To aid this design step, several computer programs to analyze and represent riding count data have been developed. One example is the graph presented in Figure 1, which shows cumulative boardings and alightings by stop. This will provide the detailed segment-level information required for design. It gives totals and percentages of route ridership by segment. Figures 3 and 4 provide time-period information about elderly and transfer passengers, respectively, that is required to assess the possible impacts of the changes on these groups.

A situation that would require a different type of process at this detailed design step is the development of alternatives to better meet the overall objectives of the agency. Unlike other situations, a set of predefined solutions (like route splitting) does not exist for this case. What is proposed is for the agency to develop in advance sets of alternatives for each of its goals. In this way, if the situation requires implementation of improvements aimed at the attainment of a specific goal, the appropriate set of alternatives can be used as a starting point. By having this preliminary design ready, the operator will be in a better position to implement these changes quickly when circumstances so dictate.

Interface of SRTP with Other Activities

SRTP is only one of the activities for which the transit agency is responsible. The other required activities of the transit operator include the following (11,12):

1. Scheduling--runcutting, driver assignment;
2. Operations--driver supervision;
3. Marketing and community relations--communication between the agency and the public; and
4. Administration.

All these different activities are interrelated with SRTP, perhaps the most important relations being among operations, scheduling, and planning. These are inextricably linked to each other in the planning and implementation of transit service. For example, it is very important that operational constraints be introduced into SRTP at an early stage to ensure that the proposed changes are acceptable in a practical sense as well as in a theoretical one. The interrelation between runcutting and cost estimation also requires close coordination between SRTP and scheduling. Only in this way can alterna-

Figure 3. Report of transfer passengers for route 2.

ALL THE STATISTICS PRESENTED BELOW ARE FOR TRANSFER PASSENGERS

									Between Day		Within Period		
TIME PERIOD	DIRECTION	DAYS SAMPLED	DAILY SCHEDULED TRIPS	EXPANDED TOTAL	PERCENT OF TOTAL	AVERAGE PER TRIP	90% CL LOWER BOUND	90% CL UPPER BOUND	VARIANCE	COEFF. OF VARIATION	VARIANCE	COEFF. OF VARIATION	90% CL ACCURACY
1	1	2	6	0.0	0.0	0.0	0.0	0.0	0.000		0.000		
1	0	2	6	208.0	68.4	34.7	29.3	40.1	5.444	0.0673	27.000	0.1499	0.1558
2	1	2	11	17.0	3.9	1.5	0.3	2.8	0.926	0.6227	2.373	0.9967	0.8068
2	0	2	11	99.0	32.8	9.0	7.8	10.2	0.067	0.0287	10.791	0.3650	0.1383
3	1	2	6	12.0	7.1	2.0	0.4	3.6	1.500	0.6124	3.238	0.8997	0.8099
3	0	2	7	45.0	27.1	6.4	4.9	8.0	0.276	0.0816	8.762	0.4605	0.2381
4	1	2	4	6.0	15.8	1.5	-1.7	4.7	4.500	1.4142	9.000	2.0000	2.1137
4	0	2	4	12.0	25.5	3.0	1.1	4.9	0.500	0.2357	6.500	0.8498	0.6322

AVERAGE DAILY PASSENGERS OF FARE CATEGORY = 399.0 OR 24.9 PERCENT OF TOTAL

PASSENGERS OF FARE CATEGORY FOR FIRST DAY = 384.7 OR 24.5 PERCENT OF TOTAL

PASSENGERS OF FARE CATEGORY FOR SECOND DAY = 417.0 OR 25.1 PERCENT OF TOTAL

Figure 4. Report of elderly passengers for route 2.

ALL THE STATISTICS PRESENTED BELOW ARE FOR ELDERLY PASSENGERS

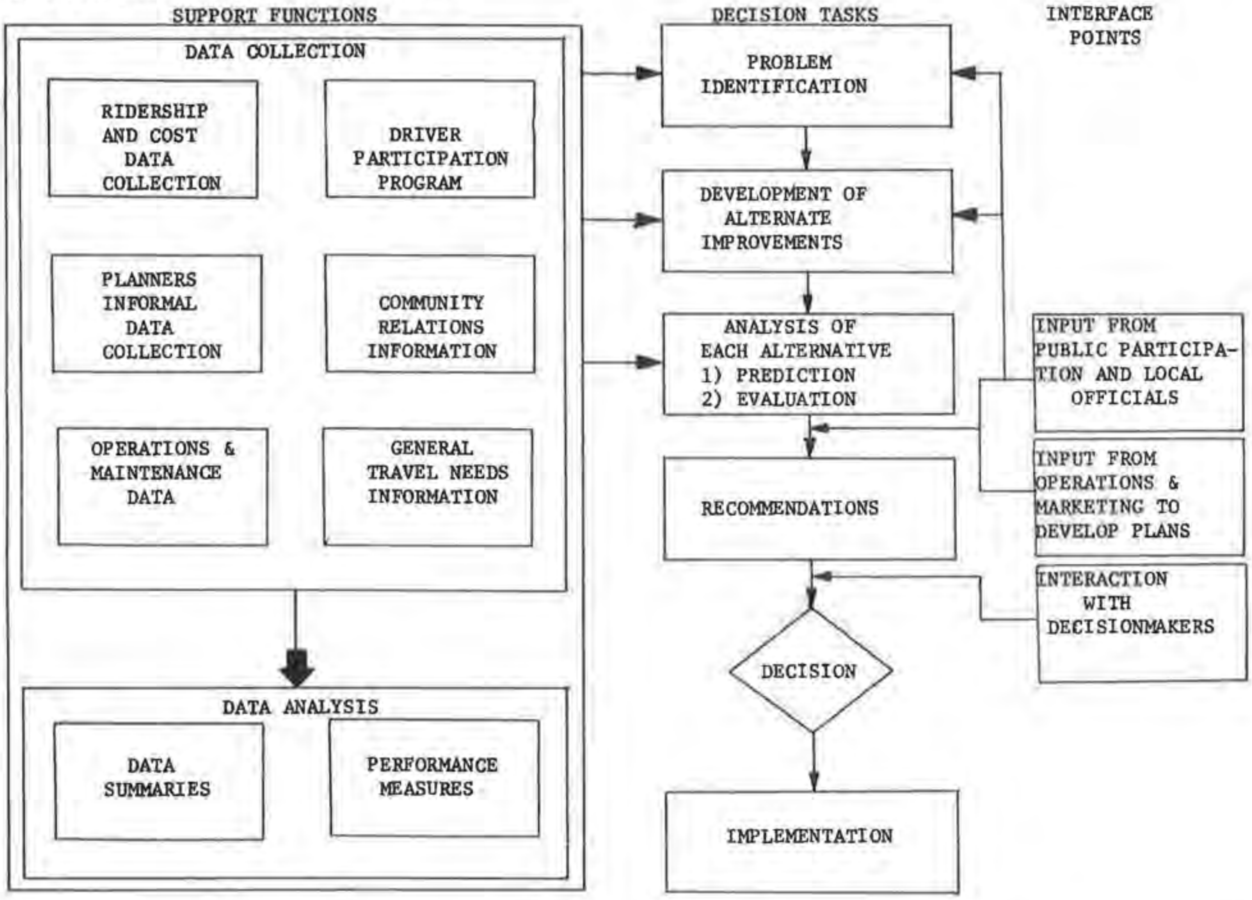
									Between Day		Within Period		
TIME PERIOD	DIRECTION	DAYS SAMPLED	DAILY SCHEDULED TRIPS	EXPANDED TOTAL	PERCENT OF TOTAL	AVERAGE PER TRIP	90% CL LOWER BOUND	90% CL UPPER BOUND	VARIANCE	COEFF. OF VARIATION	VARIANCE	COEFF. OF VARIATION	90% CL ACCURACY
1	1	2	6	14.0	9.5	2.3	1.0	3.6	0.444	0.2857	1.333	0.4949	0.5546
1	0	2	6	12.0	3.9	2.0	0.1	3.9	1.000	0.5000	3.000	0.8660	0.9705
2	1	2	11	133.0	30.9	12.:	10.1	14.1	0.159	0.0330	27.236	0.4316	0.1633
2	0	2	11	63.0	20.9	5.7	4.1	7.4	0.754	0.1516	12.800	0.6247	0.2888
3	1	2	6	18.9	11.2	3.1	1.3	5.0	0.983	0.3155	10.667	1.0392	0.5842
3	0	2	7	22.0	13.3	3.1	1.6	4.6	0.412	0.2041	7.619	0.8783	0.4790
4	1	2	4	0.0	0.0	0.0	0.0	0.0	0.000	.	0.000	.	.
4	0	2	4	0.0	0.0	0.0	0.0	0.0	0.000	.	0.000	.	.

AVERAGE DAILY PASSENGERS OF FARE CATEGORY = 262.9 OR 16.4 PERCENT OF TOTAL

PASSENGERS OF FARE CATEGORY FOR FIRST DAY = 275.5 OR 17.5 PERCENT OF TOTAL

PASSENGERS OF FARE CATEGORY FOR SECOND DAY = 250.7 OR 15.1 PERCENT OF TOTAL

Figure 5. Proposed short-range transit planning process.



tive changes be evaluated based on their true expected cost.

As has been discussed earlier, we are not in a position to prescribe the processes to be followed for incorporating these interface points into SRTP. The best approaches to doing this will depend on the capabilities of the transit organization. All we can do is indicate, for the proposed SRTP approach, the steps in which communications with other parts of the agency are important. Figure 5 (10) is a graphical representation of the proposed approach to SRTP, including the interface points with the rest of the organization.

SUMMARY

This paper has explored current practice as it applies to the design of bus services and suggested modifications that might make the planning process better suited to the needs of the transit industry in times of fiscal austerity. Perhaps the most important change suggested is to move away from an exclusive reliance on problem-centered screening of services that require study and possible change. This reliance, which is tied to the widely accepted practice of setting service standards and flagging substandard routes, may mean that the planner does not consider opportunities that may exist for improvement on acceptable routes. For example, strategies such as segmentation of service on a route into express and local portions, establishment of service zones, or having some vehicles deadhead in the lightly traveled direction to improve productivity are never likely to be feasible on problem routes, yet they may be quite useful on high-ridership corridors. By improving productivity on such routes, resources might be made available to better tackle the true problem routes. Thus, a second focus of attention to be added to the problem-centered approach would be an action-centered screening to identify opportunities for improvement on routes where no problems exist. Modifications were also proposed to recognize the multiple objectives that transit operators are striving to achieve and to deal with the problem of presentation of data in forms more directly useful in planning.

ACKNOWLEDGMENT

The research described in this paper was made possible by funding from the Urban Mass Transportation Administration through separate grants to the Massachusetts Institute of Technology and to the Greater Bridgeport Transit District. We would like to thank Brendon Hemily and Eve Wyatt for their many insightful comments on the ideas presented in this paper.

REFERENCES

1. N.H.M. Wilson. Bus Service Planning: Current Practice and New Approaches. In Compendium of Technical Papers, Institute of Transportation Engineers 51st Annual Meeting, Aug. 1981.
2. A.B. Bauer. Information Use and Decision Processes in Short Range Transit Planning. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, M.S. thesis, Sept. 1981.

3. S.L. Gonzalez. Short Range Bus Transit Planning: Demand Prediction at the Route Level. Center for Transportation Studies, Massachusetts Institute of Technology, Cambridge, Res. Rept. CTS-RAMP-80-1, 1980.
4. J.P. Attanucci, L. Jaeger, and J. Becker; Massachusetts Bay Transportation Authority. Bus Service Evaluation Procedures: A Review. Urban Mass Transportation Administration, NTIS: PB-79-296 314/8ST, March 1979.
5. R.F. Kirby and M.A. Green. Policies and Procedures for Transit Service Development. Urban Institute, Washington, DC, Working Paper 5066-7-1, 1978.
6. J. Attanucci, I. Burns, and N. Wilson; Multisystems, Inc. Bus Transit Monitoring Manual: Vol. 1, Data Collection Program Design. Urban Mass Transportation Administration, Final Rept., Aug. 1981. NTIS: PB82-122227.
7. G.J. Fielding, R.E. Glauthier, and C.A. Lave. Performance Indicators for Transit Management. Transportation, Vol. 7, Dec. 1978, pp. 365-379.
8. P.G. Furth. Designing Bus Routes in Urban Corridors. Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Ph.D. thesis, July 1981.
9. P.G. Furth and N.H.M. Wilson. Setting Frequencies on Bus Routes: Theory and Practice. TRB, Transportation Research Record 818, 1981, pp. 1-6.
10. S.L. Gonzalez and R. Richardson. Bus Performance Monitoring Study Interim Report. Greater Bridgeport Transit District, Bridgeport, CT, 1981.
11. G.M. Smerk. The Management of Public Transit. In Public Transportation: Planning, Operations, and Management (G.E. Gray and L.A. Hoel, eds.), Prentice Hall, Englewood Cliffs, NJ, 1979.
12. E.F. Klaesi. Operating a Bus System. In Canadian Transit Handbook (R.M. Soberman and H.A. Hazard, eds.), Univ. of Toronto-York Univ. Joint Program in Transportation, 1980.

Surveillance and Monitoring of a Bus System

IRA M. SHESKIN AND PETER R. STOPHER

Most transit operators occasionally conduct an on-board survey of riders. Based on experiences in Washtenaw County (Ann Arbor) Michigan, Dade County (Miami) Florida, and Honolulu, Hawaii, this paper examines three aspects of such surveys. First, a survey instrument is described that permits considerably more information to be collected than is possible from the traditional postcard type of on-board survey. Descriptions of the types of data needed to be collected on the participatory self-administered survey of riders for both systemwide surveillance and individual route monitoring are provided. In addition, it is recommended that the survey personnel record observable information (e.g., passenger volumes) on bus operations. Second, procedures are described for reducing nonresponse bias for collecting at least some information from a subgroup of riders who would otherwise be nonrespondents. Third, sampling strategies (including the necessary sample sizes) are described both for systemwide surveillance and individual route monitoring.

Most transit operators maintain the collection of a certain amount of data about the system, mainly from the perspective of the operation of the system in contrast to data about the system delivery to the actual and potential rider. Data frequently collected include revenue, load profiles (including maximum load points), vehicle hours and vehicle miles of service operated, and a variety of similar operational and financial data. All of this information is necessary to the management of a transit property and provides much needed information on the system performance; however, it is not complete because it does not measure how people use the system and, therefore, who will be affected by system changes and in what way they will be affected.

Most transit operators, from time to time, conduct some form of on-board survey of riders. This survey is typically a brief set of questions on a postcard-size form that is distributed by drivers or survey personnel who ride the buses during the survey. Postcards are usually designed either to be returned on the bus or to be mailed back later. The main drawbacks to this type of survey are the perceived restriction of limiting questions to what will fit on (usually) one side of a postcard and the usually unscientific sampling that is used. Two aspects are of concern with respect to sampling. The first is the lack of application of basic sampling procedures that would lead to near-optimal efficiency in the data collection and will usually reduce considerably the sample sizes needed to obtain data of known and calculable reliability. The second is the problem of nonresponse that occurs in a self-administered survey of this type, and that in this kind of application is generally uncontrolled and provides no information on the biases that may be caused by nonresponse (1).

In this paper, we deal with three aspects of the design of on-board surveys. First is the issue of the amount of information that can be collected. A survey instrument design is described that provides considerably more information than is possible to obtain from the simple postcard survey. Second, procedures for reducing nonresponse through instrument design are described, where these procedures also provide information about a subgroup of nonrespondents and some indicators of the potential biases that may exist. Third, sampling strategies are described for two types of situation. The first context is that of systemwide surveillance, where the desire is to obtain data about all or a large sample of routes that make up the system, but only to a sufficient degree of accuracy to describe systemwide patronage and system use and sufficient to focus attention on routes that may not be performing to the standards desired or required and may warrant further, more detailed, study. The second context is that of individual route monitoring, where the need is to obtain data of sufficient accuracy on a single route to be able to identify changes that occur in patronage patterns as a result of specific changes to the route.

For these contexts, surveillance is defined as the collection of data for systemwide profiles and information, with sufficient detail and accuracy on individual routes to be able to identify any routes that require further, more detailed study but insufficient detail to reveal specific problems or actions required at the level of individual routes. A program of surveillance would involve a periodic survey effort design to produce systemwide statistics and ridership profiles on an annual basis and would satisfy most, if not all, requirements of state and federal agencies for system performance data. Monitoring is defined as a program of repeated surveys on individual routes, designed to measure the effects over time of changes made to each such route. In this case, the emphasis is on sufficient detail in the measurement of a route to permit detection of fairly small changes in use, patronage, and performance.

DESIGN PROBLEM

The first issue in the design of any survey is to define the data needs. There are distinct differences and some similarities between the needs of surveillance and monitoring. The basic data needs for surveillance are as follows:

1. Systemwide and route-by-route passenger volume;
2. Systemwide and route-by-route passenger miles;
3. Demographic characteristics of riders;
4. Use of transfers and patterns of route use;
5. Origin and destination pattern served by the system;
6. Maximum load factors on each route and location and maximum load points, starting with an existing maximum load point assumption;
7. Service reliability and schedule adherence;
8. Run completion rate and reasons for failure to complete;
9. Proportions of different fares used; and
10. Trip purposes served.

Similarly, for a route that is to be monitored, the following data are required:

1. Passenger volume and passenger miles,
2. Demographic profile of passengers (e.g., minorities, elderly, handicapped, or carless),
3. Number or proportion of passengers who use transfers and routes (by type) transferred to or from,
4. Waiting time for transfers,
5. Passenger volume by time of day,
6. Schedule adherence and variability of running time per stop-arrival times,
7. Desired arrival times of passengers vis-a-vis bus times, and
8. Passenger attitudes to service before and after a change, including perception of travel time.

Thus, monitored routes require a more detailed level of information and, hence, a longer survey instrument.

Different levels of accuracy are required for the surveillance and monitoring data. For surveillance, the need is for a reasonably accurate picture of the functioning of the entire system, with sufficient accuracy on a route-by-route basis to determine whether special attention needs to be paid to a specific route. For monitoring, which requires a survey before and after a route is changed, the requirement is to be able to detect whether statistically significant changes have occurred between the before and after measurements. Precise sampling rates cannot be determined without estimates of the standard deviations of key variables to be measured (2). At the design stage, it is only possible to estimate order-of-magnitude differences in sampling rates and to make some assumptions about the probable sizes of standard deviations. This topic is addressed at greater length later in this paper and in a related report by the Kaiser Transit Group (3). Its significance here is mainly to note that, in addition to a more detailed survey instrument, higher sampling rates are required for monitored routes.

SURVEY METHODS

The data collection envisaged here involves two elements:

1. A participatory survey of bus passengers and
2. Recording information on bus operation and passenger volumes experienced by the participating passengers.

The participatory survey is designed as a self-administered on-board survey. For the surveillance data, a survey form that could be completed while a passenger is on the bus is desired. Therefore, on-board distribution and collection should be undertaken by using a trained survey person to distribute questionnaires, assist respondents when necessary, help ensure that completed forms are returned before passengers disembark, and also record key information (passenger volumes, time) about bus operation. Vehicle data provide a tie between riders' reports of system performance and objective measures of that performance.

For monitoring, where the data needs are not likely to be obtainable by an instrument that can be completed during a bus ride, a two-part survey instrument has been devised. This instrument consists of a short questionnaire to be completed during the bus ride that requests data about that bus ride, some key demographic characteristics, and the respondent's address. The address is requested to permit a mail follow-up scheme to be implemented for nonresponse and may be presented as a means to provide free bus passes or some other incentive to respondents, thus the request is unlikely to affect response rates negatively. The second part of the survey instrument is a long questionnaire that is designed to be completed at home and mailed back in a reply-paid envelope provided. Also, it is assumed that the action of completing the short form on the bus will help to fix that bus ride in the respondent's mind and thus make it easier to complete the longer form later.

NONRESPONSE INFORMATION

A key issue in this survey design is nonresponse. Any participatory survey is affected by nonresponse. In general, the existence of nonresponse should lead to a presumption of bias. However, in past transportation surveys, nonresponse has been assumed to be unbiased and little or no attempt was made to determine the validity of this assumption. Recent work suggests that this assumption is a dubious one and that significant and important biases do arise in any transportation survey (4,5). Unfortunately, a detailed study of nonresponse is generally infeasible within the time and cost constraints of most transportation surveys. Certainly, it seems unlikely that such a study can be included in a surveillance and monitoring activity. For a self-administered survey mechanism, the issue of nonresponse takes on additional importance and it is especially important to exercise some control over this in the monitoring study. The two-part survey design allows some limited study and control of nonresponse bias. The greatest potential for nonresponse arises with the long take-home portion of the monitoring instrument. Based on standard procedures, one might expect the take-home form to achieve no more than a 10-20 percent response rate prior to any followup process, although a good design might increase this fairly significantly. In comparison, the on-board instrument should receive a fairly high response rate because (a) it provides a diversion during the bus ride, (b) peer pressure to complete it will arise as some bus passengers decide to fill it out, and (c) there is no problem of remembering to do it, as might occur for a mail-out or take-home survey.

For this two-part survey, there are two groups of nonrespondents. One group responds neither to the short, on-board form nor to the long, take-home form. The second group responds to the short, on-board form, but not to the take-home form. (A third group responds to the take-home but not to the on-board form. Because most needed information is on the take-home form, this group need not be considered as a nonrespondent group.) The design provides no information on the first group but does provide data on the second. An analysis of the characteristics of those returning the on-board form compared with those who return the take-home form should reveal whether or not certain population groups or certain types of bus riders are under- or over-represented in the take-home responses. A bias correction can then be computed in the form of a reweighting of the data to conform with the response pattern of the on-board form. If no differences are found between the two groups, then a nonresponse bias would not be apparent. Although this procedure provides no evidence that a nonresponse bias does not exist for those who respond to neither form, it does seem to suggest that there might need to be less concern for that.

To reduce nonresponse, a three-step follow-up procedure should be included as part of the monitoring strategy. This follow-up involves a postcard reminder to respondents to the on-board form who provided addresses but who had not returned the take-home form 7-10 days after receiving the survey instrument. Seven to 10 days after that, a letter reminder with another copy of the take-home survey form would be sent. Finally, if warranted by the responses to the two previous reminders, a third postcard reminder would be sent after a further 7-10 days. This follow-up procedure should serve to reduce the magnitude of the nonresponse to the take-home portion of the survey and thereby reduce the potential biases.

As an additional mechanism to increase response, the return address can be used to generate an incentive that is mailed to those who provide their address. The most effective incentive is a pass that entitles the holder to some small number of free rides on the bus. As with any use of an incentive in a survey, care is required to select an incentive that is large enough to be effective but not large enough to bias the responses (6). An effective incentive on most transit systems would be a free pass for a week or something of similar value. This is large enough to encourage response, but is not likely to cause a significant bias in the response nor to give rise to serious attempts at forgery or the development of black markets.

Recording Information on Bus Operations and Passenger Volumes

Some information should also be collected by the survey workers. Such information includes readings of revenue and transfer meters (if these are installed on fare-collection devices), the time the bus departs from important intersections or timing points, passenger volumes, and number of riders who refuse to take the forms. Information is also required on any abnormality that occurs on the bus trip (e.g., a traffic accident or a breakdown of the surveyed bus). Other information could be asked of the respondents but is collected more accurately by a survey worker. This includes the time and location at which each passenger boards the bus. Also, not having to ask these questions on the survey form helps to shorten the form.

Thus, a log sheet has been designed to be completed by the survey worker. At locations spaced approximately one mile or more apart, survey workers record the time of day, the number of passengers on the bus at that location, and the identification number of the next form to be distributed after leaving the location. This last data item depends on the use of prenumbered forms that are handed out in strict sequence. The control and traceability offered by sequential distribution of prenumbered forms are a significant advantage that makes the small cost of this well worthwhile. Given this prenumbering, it is then possible to reconstruct approximately when and where (within one mile or so) each form was handed out. Although the recording of such information at shorter intervals would be beneficial, it will make it exceedingly difficult for survey personnel to perform all their tasks (distribution and collection of forms and pencils and the answering of questions from passengers) and fill out the log sheets at closer intervals.

From the log sheets it becomes possible to compute such information as systemwide and route-by-route passenger volumes, maximum load factors and location of maximum load points, time-of-day distribution of ridership for the system and for each route, and schedule adherence.

Sampling Design

The basic sampling unit is the bus rider. The annual ridership of the bus system is the universe being sampled for a rider survey of the type envisaged here, whether the survey is for surveillance or for monitoring. For most transit properties, this universe will be large enough that the population can be considered infinite for sampling purposes and corrections for finite populations or large sampling rates are not necessary in computing sampling error.

Given that the survey design requires distribution of a survey instrument to bus riders, a simple random sample of bus riders from the entire system on a given day would be inefficient and unduly expensive. Given the idea of an intercept survey, it makes far more sense to survey passengers on a bus trip and use at least a two-stage sampling procedure. In fact, given the mode of operation of most bus systems, the ideal procedure is the three-stage sample described in this section.

There are two primary modes of operation of a bus system and these affect the sampling method to some degree. One type of operation is based on extensive interlining of each vehicle, so that a bus will operate consecutively as two or more routes and will repeat this pattern throughout much of the day. The other type of operation uses almost no interlining, so that a specific vehicle is allocated to a route on leaving the garage and remains as that route for the entire day. The second operational mode is the one most amenable to the sampling design of an on-board survey, because a surveyor can be assigned to ride a specific vehicle and that assignment represents a sample from the route operated by that vehicle.

The initial sampling unit is the bus route. For monitoring, this is likely to be a purposeful, as opposed to random, sample of certain routes for which monitoring is needed, either to track the effects of route changes or to provide more data about a route on which changes appear to be warranted. For surveillance, this is a random sample, although the sample may, in certain cases, be designed to cover all routes.

The second stage of the sampling is to select a specific component of the scheduled operation of each route selected in the first stage. This is where the operational procedures of the system affect the form of the sampling. In any system there will usually exist a vehicle schedule that is organized differently from the published schedules and identifies what a given vehicle (and driver) will do from the time it leaves the garage until it returns at the end of its scheduled day. During this period, the vehicle may be driven by more than one driver, which is immaterial to the sampling, and may operate on more than one route, which is of primary concern to the sampling. The vehicle schedule will also include details of the deadhead runs required for the vehicle's operation and these details are needed to plan efficient allocation of survey personnel for an on-board survey.

For ease of explaining the stages in the sampling procedure, it is assumed that each vehicle in the schedule is assigned a unique identification number (as distinct from the bus number that identifies and is painted on the vehicle itself). On an interlined system, this might mean that three routes are interlined for daily operation and require seven vehicles to operate the timetabled headways of those three routes. Each of these seven vehicles is identified as a run number. In a base-vehicle system, the base vehicles on each route are assigned run numbers that provide an identical unique identification for each element of operation of the timetable. Unique identifiers are also assumed to exist for either operation for trippers, and these will usually be specific to a route in both systems, although the vehicle may interline.

For each route selected in the first-stage sample, the number of runs of that route is identified next and is kept separate between base vehicle runs and trippers. The second stage of the sampling consists of drawing a predetermined number of these base runs and trippers at random for each route in the sample. In the interlined system this selection will probably identify most or all of the individual vehicles that operate on the sampled route, and it will be necessary to build surveyor schedules that will have the surveyors transferring between buses from time to time. In the base-vehicle operation, the second-stage sampling identifies a vehicle that the survey personnel will ride all day, thus simplifying the design and administration of the survey considerably.

The third-stage sampling consists of the drawing of days for the sampled routes and runs. Ideally, this should be done on a random basis within a quota sampling scheme that produces an even distribution of the sample over the days of the week. It is desirable that, in those cases where both a tripper and a base vehicle have been sampled for a given route, a bus run and a tripper from a route are allocated the same survey day. This is particularly important when transit operations are subject to frequent minor aberrations in service delivery, such as last-minute cancellations of trippers or base vehicles and other departures from the printed schedule. Such aberrations tend to have significant impacts on the other base and tripper runs of the affected route and to distort the representativeness of the sample.

In the sampling procedure proposed here, all passengers on the sampled runs are to be included in the survey, subject to exceptions described below. Thus, the three-stage sampling technique selects the routes, the specific runs (base and tripper), and the day or days of the week for the survey and thereby identifies the sample of passengers to be surveyed. The exceptions, for a self-administered survey of this type, would generally be children under the age of 12, who would be unlikely to be able to complete the survey competently. In addition, it is usually appropriate to give the surveyor discretion about including groups of schoolchildren between the ages of 12 and 16, based on the group mood and the expectation that forms will be taken seriously and genuine attempts made to complete them.

Sample Size

Remember in determining sample size, although the measurement unit is a bus-rider trip, many fewer distinct individuals ride the bus than there are bus-rider trips. Most individuals make at least two trips per day (to and from work or shopping), for several days per week, for most of the year. Generally, individuals do not like to be subjected to repeated surveys. Therefore, the pool of riders is much smaller than the daily volume of bus-rider trips.

The determination of the required sample size is derived from the desired accuracy of the data and a knowledge of the population variance of the critical measures for the survey. This is the case for both surveillance and monitoring.

For surveillance, by using the formula for the sampling error from a simple random sample (which understates slightly the error of a multistage sample), the desired sample size is given by

$$n = (\text{standard deviation of Y})^2/(\text{required sampling error of Y})^2 \quad (1)$$

where n is the desired sample size and Y is the critical variable for the sample design.

A useful example of the application of this formula is provided by the Dade County, Florida, Metrobus system. On the 1978 Metrobus system, an average of 3.2 base vehicles and 1.2 trippers were assigned to each route. The average time of a one-way bus trip (i.e., from the origin point of a bus route to the destination point, in one primary direction) was 90 min. In 1978, an average of 211 000 passengers/day were carried on about 96 separately numbered routes. An average of 270 passenger miles were made per bus trip and the average route operated about 18 h/weekday. Most routes operated on Saturdays and about half on Sundays.

By using the figures noted above, an order-of-magnitude estimate can be made of the desired sample size, given a specified level of required accuracy. Suppose that passenger mileage is the critical variable on which the survey is designed and that the desire is to achieve a level of accuracy of ±10 percent with 95 percent confidence for each route. Suppose that, for the mean of 270 passenger miles/trip, the standard deviation is ±75 passenger miles. A 10 percent accuracy level is 27 passenger miles and requires a sampling error of

$$\text{Required sampling error} = 27/1.96 = 13.78 \quad (2)$$

By applying the formula for determining the sample size for a simple random sample, shown in Equation 1, the required sample size is about 30 vehicle trips/year per route, or about 2.5 bus runs/year per route (each bus run comprises about 12 trips) as shown in Equation 3.

$$n = (75)^2/(13.78)^2 = 29.6 \text{ vehicle trips} \quad (3)$$

For convenience, a sample of two bus runs per route per year (about 24 vehicle trips) may be selected. By using the same equation, this would provide an accuracy of ±11.1 percent with 95 percent confidence ($\pm 1.96 \times 75/\sqrt{24}$). Until data accrue to permit more accurate computation from Equation 1, this would be the recommended sample size.

Similarly, for any variable measured by the rider survey, a level of accuracy can be calculated from Equation 4.

$$\text{Percentage sampling error} = \text{percentage standard deviation of a unit}/n \quad (4)$$

For example, suppose a measure of interest has a percentage standard deviation of ±50 and the sample size is 500 passengers (the average ridership per bus run), the percentage sampling error is ±4.5 percent with 95 percent confidence--a very adequate measurement accuracy.

The key issue in monitoring is to be able to detect significant changes between the before and after periods. As a result, the relevant sample sizes (n) would be obtained from Equation 5 for the mean of a continuous variable (such as passenger miles) or from Equation 6 for a dichotomous variable (such as percentage of handicapped persons riding the route) (2).

$$n = [(s^2_1 + s^2_2)/\text{s.e.}\,(\bar{y}_1 - \bar{y}_2)]^2 \quad (5)$$

$$n = [(p_1q_1 + p_2q_2)/\text{s.e.}\,(p_1 - p_2)]^2 \quad (6)$$

where

n = required sample size,
s_1, s_2 = standard deviations of y before and after the change,
$\bar{y}_1, \bar{y}_2$ = means of y before and after the change,
p_1, p_2 = probability of finding a characteristic before or after,
$q_1 = 1 - p_1$, and
$q_2 = 1 - p_2$.

The sample size is assumed to be the same on each occasion.

Again, an example is useful to demonstrate the implications of these equations. Suppose one considers reported waiting time as a variable to be measured to within ±1 min for a change in system operation, at a 95 percent confidence level. Suppose the standard deviation is now known to be ±6 min and that this will be halved by system improvement. The number of samples (n) that will be required in each of the before and after surveys is then 173.

Given the accuracy considerations for detecting significant change and the problem of attempting to survey the same individuals more than once, the monitoring samples could be selected as follows: In a system with base vehicles and trippers, for a route that has more than three base vehicles allocated, one bus run should be selected at random for each weekday to yield a sample of five bus runs for the route. Note that, for a route with four base vehicles, one run would be surveyed twice by using this procedure. For routes that have five base vehicles or more, runs are selected randomly and repeats of a run will occur relatively infrequently. For routes with three or fewer base vehicles, three different weekdays would be selected at random and one run selected randomly for each of the selected weekdays. The weekdays should be ordered randomly and preferably should be scheduled within one calendar month for each of the before and the after surveys.

A final consideration in sample size concerns the rate of sampling bus riders. The sampling errors indicate that sampling one in two or one in three bus riders normally would be adequate. However, a number of potential disadvantages to such a procedure exists. First, it would have to be left to the on-board survey person to select passengers to receive forms. The obvious problem is that there is no way to know if the surveyor is following the rules about handing out the forms. Second, a potential public-relations problem is created if some riders receive surveys and others do not, particularly if there exists a free pass or some similar survey incentive. Third, if response rates are lower than expected, sampling errors will be higher than desired. Because the cost of printing the forms is trivial compared with the costs of the logistics of performing an on-board survey and the cost of labor, it is logical to distribute forms to all passengers to minimize the possibility of a low response rate. To avoid these problems, all riders should be given survey forms.

Thus, based on certain assumptions about standard deviations and critical variables, sample sizes for the surveillance and monitoring activities can be defined. Note that the accuracy of the surveillance activity exceeds that required by the Urban Mass Transportation Administration (UMTA) (7) but at a lower survey effort than UMTA recommends.

Given this rather small sampling requirement, selection of the time of the year for the surveys is important. Many bus systems have a specific time in the year when route changes are made (as infrequently as once a year, but more often two to four times a year) and these will dictate when the monitoring surveys take place. In general, the before survey should be done one or two months before the changes will go into effect. Depending on how many routes are to be monitored and the survey team that it is planned to maintain, these surveys may take one to three months to execute. The after survey should be done about 6 months after the changes were effected or 1 month before the next set of system changes is scheduled, whichever is sooner. Conventional wisdom maintains that 6 months are needed for a steady state to be achieved after a change is made, although intervention of other system changes may make it impossible to determine what system changes caused what passenger reactions. We have no reason to question the conventional wisdom, and the timing of the after survey is based on acceptance of this.

For surveillance, the scheduling during the year of the survey activity is an important issue because not only do seasonal changes in patronage exist but also there are probably seasonal variations in the various measures of concern. Two alternative strategies can be considered. The first involves increasing the sample size and surveying each route on several occasions throughout the year, thereby obtaining some information on seasonal variations, but at a considerably higher cost. By using available historic information, the second strategy for surveillance appears preferable, where this involves selection of two occasions during the year. In this case, again, surveying should not take place in the month in which system changes are made, because of the instability these changes may cause, not to mention the potential public-relations problems of such timing.

Once experience has accumulated on route differences, some modification may be needed for routes that have low ridership, as the sample generated may be too small to permit sufficiently accurate data to be obtained. The day of the week and the bus run should each be chosen at random for each route. A quota sample should be used for days of the week to ensure an equal representation of each weekday to provide data on the day-by-day variations in systemwide operations and to facilitate the logistics of scheduling survey workers. A simple random sample of weekdays can lead to an inordinate number of bus runs being scheduled on one weekday.

In conclusion, the careful application of sampling theory and the use of prior information on the variance of measures of importance can be used to produce small samples that meet requirements of measurement accuracy. Furthermore, the institution of a regular surveillance and monitoring program will itself reduce sample requirements as regular updates to the data base are produced and the accuracy of existing information is progressively enhanced. This argues against the common practice of occasional surveys, often years apart, that use instruments or questions that differ so markedly that updating is not possible.

SOME ILLUSTRATIVE CASES

Some extracts from three recent surveys are useful to illustrate some of the points in this paper. First, a survey of the type described here was undertaken in Washtenaw County in southern Michigan (8). This involved an on-board survey form in two parts--a card to be completed and returned on board the bus and a longer form to be taken away and mailed back. The on-board survey form was returned by 88 percent of the passengers who were handed a form, and 44 percent of these passengers (38 percent of the total number of passengers who were handed forms) returned the mail survey. Both of these response rates are higher than those that would normally be expected, even though no incentive was used. The surveys collected information that is of the order of the specifications given in this paper for the monitoring activity. From a comparison of the two forms, the extent to which the mail-back survey was biased was established by its lower response (9), and it was found that relatively little bias was present on the basis of questions asked on both forms. No comparison of the two-part survey with a mail-back only was made in the study, but both response rates compare favorably with other mail and on-board surveys.

A second example is provided from an on-board survey conducted for the Dade County Transportation Administration on the Metrobus system (10). This survey was carried out on a sample of bus routes in early fall 1980, having been postponed by four months because of the civil disturbances in Miami in May 1980. (The outbreak began the night before the on-board bus survey was to have started; no connection between these two events has been established.) Two problems of some magnitude affected this survey and reduced the performance of the instruments. First, the survey instruments were produced in both Spanish and English. Although people were offered either form, it was apparent in both the pretest and the main survey that many Spanish-speaking people opted to take an English form. It is not known to what extent this might have reduced the response rate, but the response from the Spanish-language forms was well below that from the English forms. Also, the Spanish-speaking population of Miami appears to have different community goals and loyalties and it was found to be very difficult to motivate this group to respond to the survey. Inducements to cooperate that worked on the English-speaking population were not effective on the Spanish-speaking population, judging from the difference in response rates. Second, a number of Haitians and illegal immigrants use Dade County buses. Given that many of the Haitians speak only Haitian French and that survey forms were not provided in their language, this group probably did not respond to the survey. Clearly, illegal immigrants would not complete and return surveys because they would have an expectation that this would be a means by which they could be traced.

Despite these problems, of a total of some 58 000 forms handed out in a six-week intensive survey, a little more than 13 000 (22.4 percent) of the on-board forms were returned and a little more than 9000 (16 percent) of the mail-back forms were sent back. Again, the responses from the 13 000 were used to determine the degree to which biases existed and could be corrected for in the 9000, and several bias adjustments were made (11). Although most of the adjustments were small, one group (males aged 45-54) were missing completely from the mail-back survey but not from the on-board survey; this was determined from comparing the two response sets. Also note that the on-board survey form covered both sides of a legal-size sheet of card, and that the mail-back form was a six-page questionnaire, printed on sheets midway between legal and letter size.

The third example is a recently completed on-board bus survey in Honolulu, Hawaii (12), conducted for the Oahu Metropolitan Planning Organization. In this case, the survey was only an on-board survey, of the type described in this paper for the surveillance activity. The survey form again covered both sides of a legal-size card. The bottom of the second side consisted of a reply-paid panel for return of the survey form by mail, for those unable to complete it on the bus. A total of 4928 forms were distributed and 58 percent of these (2815) were returned. No incentive was offered for completion of the survey forms, and it is estimated that a reasonable incentive, such as the bus passes used in Dade County, could have boosted the response to as much as 75-80 percent. Of the 58 percent response, 45 percent came back off the bus and 13 percent by mail. A problem that is likely to have decreased the response rate is that the Hololulu buses travel fully loaded in the peak hours. The surveyors counted loads of around 100 passengers on most peak-hour trips (on standard 49-seat buses) and one bus had around 120 passengers. Drivers assured the survey team that these were normal and expected loads. Notable in this instance is that, although the survey purpose was different from the other two (to supplement data for travel-forecasting calibration) and a system profile was not the primary motivation, this sample of 2815 returned forms provides a statistically sound description of one-third of the system's routes.

CONCLUSIONS

The issues addressed by this paper involve the design of an efficient and cost-effective measurement procedure for a bus operation that will allow system changes to be monitored and will provide annual data on systemwide operation (surveillance) for both federal and state reporting requirements as well as statistically sound data on the patronage of the system and responses of riders to system changes. Four elements of the problem are addressed: data needs, survey mechanisms, sampling procedure, and sample sizes. Monitoring and surveillance each require different measurements and sample sizes but are otherwise similar in design.

The recommended procedure is a mix of observation and participation, whereby survey personnel ride sampled buses, note various items of information on bus operation (passenger loads and times at various locations), and distribute self-administered survey forms to all passengers. For surveillance, where data needs are less extensive, the survey form can be restricted to one sheet of light card stock that can be completed by most passengers before disembarking. However, a mail-return capability is essential to this survey because otherwise the survey will be biased against those who make very short trips, those who travel on very crowded peak-hour buses, and the riders who have poor vision or other problems that make completion of the form on the bus difficult or impossible. For monitoring, a two-part survey form is recommended. The first part can be the same as the surveillance form. The second part is designed to be taken home and mailed back. To permit follow-up and tracking of nonresponse bias on the take-home form, the respondent's home address should be requested on the on-board form and a system used that allows the returned take-home and on-board forms to be matched up.

Principally, this paper has put forward a design of survey instruments that provides collection of much more data than would usually be obtained in an on-board survey, but, based on the case studies noted here, without loss of response. Second, the paper has described means of tracking part of the nonresponse and determining the extent to which the responses from a self-administered survey of this type should be weighted to correct for nonresponse bias.

The paper has provided details on the computation of sample sizes that yield significantly smaller samples that are statistically adequate. In developing this design into an annual activity for Dade County, Florida, it was estimated that a full surveillance and monitoring activity for this 99-route system with a bus fleet of nearly 700 buses (daily ridership on the order of 225 000 rides) would require 1050 surveyor-days of weekdays per year and 400 surveyor-days of weekends. By careful scheduling of the survey activities, this translates into 6-8 survey workers, working for eight months of the year in two four-month periods. The interviewers can be used in the remaining four months to assist in data processing, thereby permitting their retention as a permanent work force.

With the possibility of minor modifications to the sampling rates based on actual measured standard deviations, the sampling rates described should provide data of adequate accuracy for the required measurements. Furthermore, the basis for generating the sample here should be readily applicable to other transit properties.

ACKNOWLEDGMENT

We wish to acknowledge the assistance of Howard G. Eisenstadt in the design of these procedures and in reviewing an early draft of this paper; and the contributions of Gary S. Spivack and Richard G. Walz and other staff members from Kaiser Transit Group and the Dade County Transportation Administration in the design and testing of the procedures.

REFERENCES

1. Sozialforschung Brög. Non-response-Untersuchung zum Verkehrsverhalten von Hauschalten in Berlin (West). Unterschung in Rahmen des Verkehrsentwicklungsplan (VEP) Berlin, im Auftag des Senats von Berlin, Munich, West Germany, 1979.
2. P.R. Stopher and A.H. Meyburg. Survey Sampling and Multivariate Statistics for Social Scientists and Engineers. Lexington Books, Lexington, MA, 1979.
3. Kaiser Transit Group. Design of a Performance Measurement and Evaluation System for Both the TDP Phased Plan and the Operating System. Deliverable VIII A2, Metro Dade Office of Transportation Administration, Miami, FL, Oct. 1979.
4. M. Wermuth. Uncertainty of Household Travel Surveys in Methodological and Measurement Contexts. In New Horizons in Travel-Behavior Research (P.R. Stopher, A.H. Meyburg, and W. Brög, eds.), Lexington Books, Lexington, MA, 1981.
5. W. Brög and A.H. Meyburg. Consideration of Nonresponse Effects in Large-Scale Mobility Surveys. TRB, Transportation Research Record 807, 1981, pp. 39-46.
6. M.L. Tischer. Attitude Measures: Psychometric Modeling. In New Horizons in Travel-Behavior Research (P.R. Stopher, A.H. Meyburg, and W. Brög, eds.), Lexington Books, Lexington, MA, 1981.
7. Sampling Procedures for Obtaining Fixed Route Bus Operating Data Required Under the Section 15 Reporting System. Urban Mass Transportation Administration, UMTA Circular C2710.1, Feb. 1978.
8. Schimpeler Corradino Associates. Survey Results. Washtenaw County Transit Needs Study, Ann Arbor-Ypsilanti Urban Area Transportation Study Committee, Washtenaw County, MI, Rept. 6, Oct. 1980.
9. P.R. Stopher and I.M. Sheskin. Method for Determining and Reducing Non-Response Bias. TRB, Transportation Research Record (in preparation).
10. I.M. Sheskin, G.S. Spivack, and P.R. Stopher. The Dade County On-Board Survey. Transit Journal, Spring 1981, pp. 15-28.
11. Kaiser Transit Group. The 1980 Dade County On-Board Transit Survey. Dade County Transportation Administration, Miami, FL, Final Rept., 1981.
12. Schimpeler Corradino Associates. Data Collection. Oahu Travel Forecasting Model Update, Oahu Metropolitan Planning Organization, Honolulu, HI, Interim Rept. 2, Feb. 1982.

Bus Route-Level Demand Modeling

DONALD G. YURATOVAC

The need for improved techniques in the area of bus route-level demand modeling is discussed, and a summary of the existing state-of-the-art methods used throughout the transit industry is provided. A working example of a modeling technique for local radial bus routes in Cleveland, Ohio, which was funded by the Office of Planning Assistance of the Urban Mass Transportation Administration, is presented.

For some time now, the Urban Mass Transportation Administration (UMTA) has equipped the transit planning community with an extensive set of computer programs designed to aid in the long-range planning of multimodal urban transportation systems. The bulk of the UMTA Urban Transportation Planning System (UTPS) deals with projecting the level of patronage that will be realized from alternative system considerations. Although UTPS has been invaluable in the design and study of fixed-guideway proposals throughout the country, it has not been used to any great extent in the design or planning of short-range bus route-level improvements. Consequently, individual transit properties have been left to fend for themselves in the development of in-house techniques to predict the impact on system ridership of new and/or extended bus routes as well as changes in the level of service provided on a given route.

Recognizing the need for improved techniques for projecting bus route-level ridership, the UMTA Office of Planning Assistance has recently initiated and funded a series of four short-range ridership projection study efforts by using data from the following cities:

City	System
Albuquerque, New Mexico	SUNTRAN
Cleveland, Ohio	Greater Cleveland Regional Transit Authority
Los Angeles, California	Southern California Rapid Transit District
Portland, Oregon	Tri-County Metropolitan Transportation District of Oregon

The need for simplified and accurate techniques for estimating bus route-level patronage is reflected in the changing priorities brought on, in part, by the funding philosophy dictated by the policies of the "New Federalism". With most major highway and transit facilities in place, interest has grown in management-oriented or transportation system management improvements, which are designed to improve the efficiency and increase the productivity of existing services. Therefore, it is imperative that any future bus-route expansions to the nation's transit systems must generate a sufficient level of ridership and/or farebox recovery so as not to further deplete already dwindling transit service resources. In other words, transit systems today do not have sufficient "risk" capital to operate new or extended routes where the incremental ridership gains fall far below passenger projections and minimal service performance standards.

SURVEY FINDINGS

As part of the study effort for Los Angeles, Multisystems, Inc., researched the state of the art in route-level demand modeling. In addition to an extensive literature search on the subject, 40 transit properties in the United States and Canada were surveyed to determine what, if any, in-house techniques were used to project ridership on route-level modifications. This survey indicated that the methods currently used to project route-level ridership changes fall into one of the following four categories:

1. Professional judgment, based on the judgment of one or more of the property's operations analysts;
2. Noncommittal survey techniques, where potential riders are asked directly if they would use a proposed service;
3. Cross-sectional data techniques, which examine the relation between transit use and a range of characteristics of the service and populations to be served; and
4. Time-series data techniques, which compare changes in ridership as service changes over time.

Despite the diversity of the different ridership projection techniques that fall into these four general categories, the survey results enable one to draw the following conclusions with respect to state-of-the-art techniques used throughout the industry:

1. The accuracy of existing techniques is open to question because very few empirical tests have been performed in which estimates of ridership made before implementation of a route or route modification were compared with the actual resultant ridership.
2. The application of the various techniques is done in an informal rather than a formal manner. Consequently, it is not known whether one analyst can replicate the predictions of another, given the same data base.
3. Inasmuch as the accuracy of the various techniques is not known and the processes are not formalized or well documented, there is little opportunity for transferability of techniques from one city or geographic area to another.

Thus, the survey illustrates a need for the development of short-range ridership projection techniques that meet the following criteria: (a) demonstrated accuracy, (b) formalized application and documentation, (c) low cost of application, (d) minimal technical sophistication, and (e) transferability among urbanized areas.

CLEVELAND EXPERIENCE

Although the Cleveland study effort is still not complete, it has become apparent that no single model can be used to accurately project ridership for all of the different types of service operated in the Cleveland area. Consequently, a series of models are being developed that are individually sensitive to the unique characteristics of the different types of service under consideration--e.g., local radial, crosstown, express radial, and rapid feeder.

Rather than discuss the development and calibration of the models at great length, it may be more useful to go through a sample validation of ridership on a local radial route to demonstrate the application of the model for this type of service.

Model Application

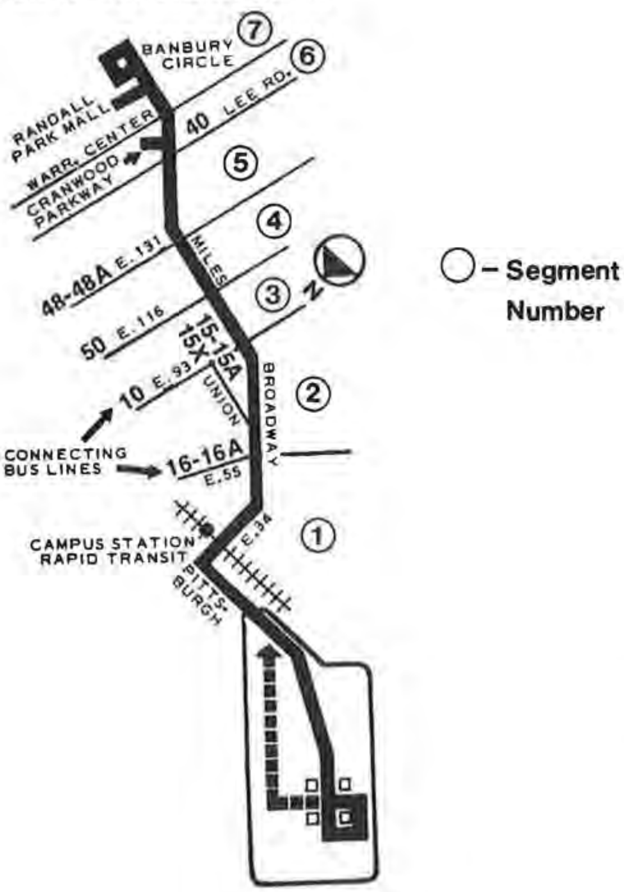
The data requirements for use of the model are as follows:

1. Map of route,
2. Socioeconomic data at traffic zone or census tract level,
3. U.S. Geological Survey (USGS) or land use maps,
4. Bus-route travel times,
5. Land use and/or employment data at traffic zone or census tract level, and
6. Schedules and ridership data for intersecting routes.

The steps in using the model are as follows:

1. Divide the route into segments,
2. Determine the market area for individual segments,
3. Determine the mean income for route-segment market areas,
4. Determine the home-based transit trip rate,
5. Calculate the home-based transit trips,
6. Calculate transfer trips, and
7. Distribute trips.

Figure 1. Route 19 bus-route segments.



The example used here is bus-route 19, Broadway-Miles, a radial route in Cleveland.

Step 1: Divide Route into Segments

In step 1, the route is divided into logical segments based on major intersections and transfer points. Segment divisions for Broadway are as follows (see Figure 1):

Segment	Boundaries
1	Public Square to East 55th Street
2	East 55th Street to East 93rd Street
3	East 93rd Street to East 116th Street
4	East 116th Street to East 131st Street
5	East 131st Street to Lee Road
6	Lee Road to Warrensville Center Road
7	Warrensville Center Road to Banbury Circle

It should be noted that intersections are assigned to the lower of the two adjacent segment numbers (i.e., passengers boarding at the Miles-East 131st Street intersection are included in segment 4).

Step 2: Determine Market Area for Individual Segments

In step 2, the market area for each route segment is determined. Northeast Ohio Areawide Coordinating Agency (NOACA) socioeconomic data or U.S. Census data at the block or tract level can be used. NOACA traffic-zone data are used in this example.

The market area for the bus route is defined as the area within 0.25 mile of the route. The number of households within that area is determined by the following procedure:

1. Determine the traffic zones that are located within 0.25 mile of each route segment.
2. If some of the zone is within 0.25 mile of the route and some of the zone is not, determine the percentage of the zone that is within the market area.
3. If a zone is partly in one route segment and partly in another, determine the percentage of the zone that is in each segment.
4. Determine the percentage of the zone within each segment by multiplying the two percentages from steps 2 and 3. In other words, if 40 percent of traffic zone i is within the route market area (within 0.25 mile) and 50 percent of that portion of the zone is within the segment A market area, then the proportion of traffic zone i households in market area of segment A = (0.5) x (0.4) = 0.2, or 20 percent. Figure 2 shows this process.

Figure 2. Traffic-zone segmentation.

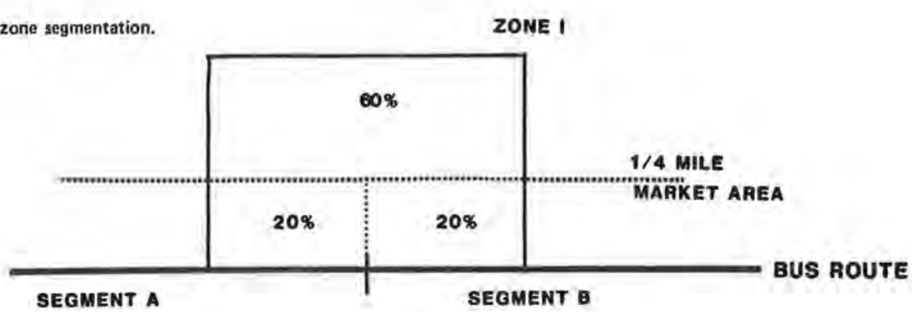


Figure 3. Partitioning of zonal land uses.

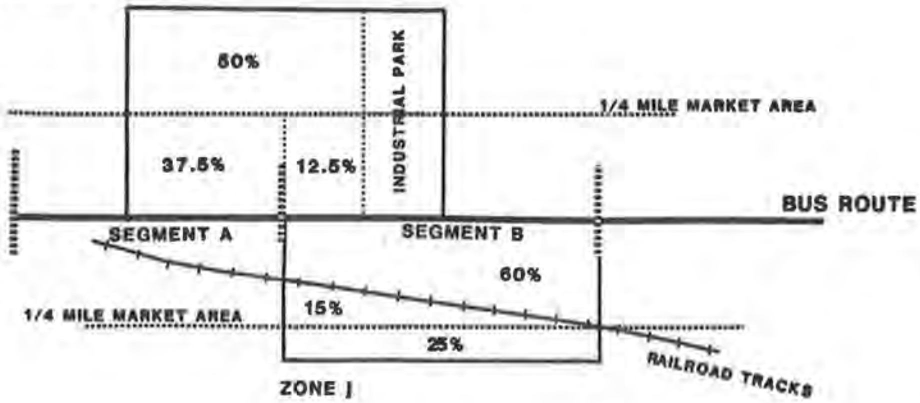


Table 1. Number of households in bus-route market areas.

Segment	Traffic Zone	No. of Households in Traffic Zone	Zone Households in Route-Segment Market Area (%)	No. of Households in Route-Segment Market Area[a]
7	655	509	100	509
6	656	511	75	383
	525	692	48	332
	661	1142	42	480
Total				1195
5	660	564	33	186
	524	106	0	0
	620	1010	40	404
	521	309	22	68
	629	661	85	562
Total				1220
4	628	0	0	0
	630	1072	100	1072
Total				1072
3	503	682	85	580
	502	1287	83	1068
Total				1648
2	501	425	85	361
	507	1656	50	828
	500	105	100	105
	499	877	75	658
	491	1483	40	593
	492	1138	50	569
	493	719	90	647
	508	960	15	144
	494	797	30	239
	490	74	100	+74
Total				4218

[a]Determined by multiplying columns 3 and 4.

5. Use USGS maps to make sure that the percentage of residences in the route-segment market area is accurate. Empty land, industrial land, or major barriers such as highways or railroad tracks might require a revision of the percentage derived in step 4. This type of problem is illustrated in Figure 3. In zone i, 25 percent of the land area consists of industrial land use. Therefore, the division of residential land between the two segments must be modified. Fifty percent of the residential area is within the market area but, due to the location of the industrial park, 75 percent of that 50 percent is in segment A and 25 percent is in segment B. The calculation is therefore as follows: Segment A = (0.75) (0.50) = 0.375, or 37.5 percent of households in zone i; segment B = (0.25) (0.50) = 0.125, or 12.5 percent of households in zone i. In the case of zone j, 75 percent of the residential land in the zone is within the market area for segment B. However, 20 percent of that 75 percent is located beyond a railroad embankment and is inaccessible to the route. The percentage of zone j actually accessible to the route is (0.80) (0.75) = 0.60, or 60 percent of zone j is in the bus-route market area.

6. Residential market areas are calculated for all route segments outside the central business district (CBD). The CBD is always designated as route segment 1, and residential market areas are not calculated for the CBD area. This is because the CBD has little residential land use and also has its own circulation system, the loop bus. Thus, for route 19, the number of households in the bus-route market area for bus-route segments 2 through 7 is as given in Table 1.

Step 3: Determine Average Income for Route Segments

The income of residents in the bus-route market area will affect the rate of transit tripmaking. By using NOACA data on average income at the traffic-zone level for 1980, the average income for each route segment can be determined.

1. Find average income from the NOACA tables for each traffic zone. For bus-route segment 6, the figures would be as follows:

Traffic Zone	Mean Income ($)
656	14 010
525	14 314
661	14 607

2. Calculate a weighted average for the segment based on the number of households in each traffic zone, as given below for segment 6:

Traffic Zone	No. of Households in Bus-Route Market Area
656	383
525	332
661	480

The formula used is [(income x number of households for zone 1) + (income x number of households for zone 2) + ... + (income x number of households for zone N)] ÷ total number of households (zones 1-N); or, for route 19, segment 6,

[(383 x 14 010) + (332 x 14 314) + (480 x 14 607)]/ (383 + 332 + 480) = 14 334

Step 4: Determine Home-Based Transit Trip Rate

The number of home-based transit trips for each segment of the route is based on the average income of the segment and the frequency of service provided. The home-based transit trip rate is determined as follows:

1. Determine the income category of each segment based on the following breakdown:

Level	Amount ($)
Low	<10 000
Middle	10 000 to 14 000
High	>14 000

For route 19, average income and income level by segment are as follows:

Segment	Avg Income ($)	Level
7	11 414	Middle
6	14 334	High
5	10 945	Middle
4	10 164	Middle
3	10 126	Middle
2	9 085	Low

2. Determine the combined peak and off-peak service frequency for each route segment by using the following formula: combined frequency - (0.67 x peak frequency) + (0.33 x off-peak frequency). For segments 2-4, the average peak headway is 13 min and the average off-peak headway is 14 min. Segments 5-7 have less frequent service in the peak period (22 min) but run at 14-min headways in the off-peak. This unusual service pattern on segments 5-7 exists because the route serves a major suburban mall. For segments 1-4, the combined frequency is 13.3 = (0.67 x 13) + (0.33 x 14); for segments 5-7, the combined frequency is 19.4 = (0.67 x 22) + (0.33 x 14).

3. Equations are used to generate a home-based trip rate for each route segment. Based on the income category of the segment, the following equations are used for radial routes (in the final model, graphs will be available to determine the home-based trip rate):

Income Level	Equation
Low	0.78 - (0.221 x natural log combined frequency)
Middle	0.65 - (0.0232 x combined frequency)
High	0.015 - (0.0013 x combined frequency)

The rates for each segment are calculated in Table 2.

Table 2. Home-based trip rate by route segment.

Segment	Income Level	Combined Service Frequency	Calculation
2	Low	13.3	0.78 - (0.221 ln 13.3) = 0.78 - (0.221 x 2.588) = 0.238
3	Middle	13.3	0.65 - (0.0232 x 13.3) = 0.341
4	Middle	13.3	0.65 - (0.0232 x 13.3) = 0.341
5	Middle	19.4	0.65 - (0.0232 x 19.4) = 0.20
6	High	19.4	0.105 - (0.0013 x 19.4) = 0.08
7	Middle	19.4	0.65 - (0.0232 x 19.4) = 0.20

Step 5: Determine Number of Home-Based Trips

For each segment, the trip rate is multiplied by the number of households in the market to obtain the number of home-based trips in each segment:

Segment	Trip Rate	No. of Households	No. of Home-Based Trips
2	0.238	4218	1004
3	0.341	1648	562
4	0.341	1072	366
5	0.200	1220	244
6	0.080	1195	96
7	0.200	509	102

Step 6: Determine Number of Transfer Trips

In step 6, the number of passengers transferring onto the route in each segment is determined. It is assumed that passengers will transfer from crosstown routes to radial routes only, not from a radial to a radial. The number of passengers transferring is a function of (a) the total number of passengers on the crosstown bus at the transfer point and (b) the combined frequencies of the two routes. When a transfer point is located at the intersection of two route segments, the transferring passengers are loaded onto the segment closest to the CBD.

To demonstrate this technique, the transfers from routes 16-16A to route 19 (boarding in segment 1) are calculated:

1. Calculate the number of passengers on the crosstown route at the transfer point:

Route	No. of Passengers
16 Northbound	370
16-16A Southbound	680
16A Northbound	192
Total	1242

2. The combined service frequency for route 19, segment 1, was previously determined to be 13.3. The combined frequency for routes 16-16A can be determined from Greater Cleveland Regional Transit Authority (GCRTA) schedules. At Broadway the peak headway is 12 min and the off-peak headway is 15 min. The combined frequency is therefore (0.67 x 12) + (0.33 x 15) = 13.

The sum of the combined frequencies is 26.3, and is used in the following equation to determine the transfer rate.

Table 3. Transfer rate by route transfer points.

Segment	Crossing Route	Crossing-Route Frequency	Combined Frequency	No. of Passengers on Bus	Transfer Rate	No. of Transferring Passengers
2	10	18	31.3	635	0.070	45
3	50	18	31.3	335	0.070	24
4	48-48A	13.67	27.0	1548	0.089	137
5	40	16.3	32.6	267	0.065	17
6	41	25.0	41.3	466	0.036	17

Transfer rate = 0.498 - (0.1242 x natural log combined frequencies) = 0.498 - (0.1242 x ln 26.3)

or 0.092 = 0.498 - (0.1242 x 3.27). By taking the transfer rate (0.092) and multiplying by the number of passengers on the bus (1242), the number of transferring passengers is obtained: 0.092 x 1242 = 114. Thus, 114 passengers will transfer from route 16-16A and board the Broadway bus. Table 3 gives the number of transfers for each of the transfer points along route 19.

3. Transferring passengers are added to home-based boardings for each segment to obtain total one-way boardings:

Segment	Home-Based Boardings	Transfers	Total One-Way Boardings
1	0	114	114
2	1004	45	1049
3	562	24	586
4	366	137	503
5	244	17	261
6	96	17	113
7	102	0	102
Total			2728

Step 7: Distribute Trips to Other Segments

The next step is to determine where the boarding passengers are going. This will eventually make it possible to obtain total two-way boardings by route segment. The distribution is a function of the distance between segments and the level of employment in each route-segment market area.

1. For radial routes, the number of trips bound for the CBD is a function of travel time from the CBD. For each segment, the following equation is used:

Percentage of trips to CBD = 72.7 - (0.718 x travel time to CBD)

For segment 2, the travel time from the center of the segment to downtown is estimated at 16 min. The calculation therefore is 72.7 - (0.718 x 16), or 61.2 percent = 72.7 - (0.718 x 16). Thus, 61.2 percent of the passengers boarding in zone 2 will alight in zone 1 (the CBD). The other 38.8 percent will alight in one of the other zones. The CBD-bound trips for all zones are given in Table 4.

Table 4. CBD-bound trips.

Segment	Total One-Way Boardings	Travel Time to CBD (min)	Passengers to CBD Percent	Passengers to CBD No.	No. of Passengers to Other Segments
1	144	14	90	62.6	54
2	1049	16	642	61.2	407
3	586	22	333	56.9	253
4	503	26	272	54.0	231
5	261	30	134	51.2	127
6	113	35	54	47.6	59
7	102	42	43	42.5	59
Total	2758		1568	56.8	1190

2. Each alighting will become a boarding later in the day. In other words, it is assumed that a person who travels to the CBD from segment 2 will make a return trip later in the day, boarding in the CBD and alighting in segment 2.

3. To determine the distribution of the non-CBD trips, the level of employment in each route-segment market area must be determined. The procedure used is similar to that used in step 2 to estimate the route-segment residential market area. The NOACA data base included a breakdown of commercial, industrial, and retail land use acreage for each traffic zone. NOACA also has data on employment densities for each zone and for different land uses.

Traffic-zone and USGS maps were used to determine the acreage within the 0.25-mile bus-route market area. With the data available, it was not possible to separate retail and commercial land uses. Therefore, two categories, industrial and retail-commercial, were used. USGS maps were used to locate industrial areas. Employment densities for the zone were then multiplied by land use acreage to determine the number of employees in the segment market area. An example for segment 3 of route 19 is given in Table 5.

In some cases, more specific information can be obtained. If a traffic zone includes only a major shopping mall, for example, specific employment data may be available from the mall. Where specific land uses can be identified, NOACA data on land-use-specific employment densities can be used. Strip commercial development, for example, is estimated to have 25-30 employees/acre, a measure that can be used instead of zone-specific densities. For route 19, estimates of employment for each route segment are given in Table 6.

4. The travel time between each segment on the route can easily be determined from the GCRTA schedules. For each segment, determine the travel time to all other segments except the CBD (segment 1). For segment 2, these times to other segments are as follows:

Segment	Time (min)
3	6
4	10
5	14
6	19
7	26

Travel within a segment is generally not estimated unless the segment is unusually large (>10-min travel time from end to end) because walking is competitive with transit for short trips.

5. For each segment, divide employment in all

Table 5. Land use acreage and employment density for segment 3.

Traffic Zone	No. of Acres: Industrial	No. of Acres: Retail-Commercial	Acres in Market Area: Industrial No.	Acres in Market Area: Industrial Percent	Acres in Market Area: Retail-Commercial No.	Acres in Market Area: Retail-Commercial Percent	Total Acres	Traffic-Zone Employees per Acre	Employees in Segment Market Area
503	9.5	9.9	9.5	100	9.9	100	19.4	23.6	458
502	43.2	13.8	21.6	50	6.9	50	28.5	26.1	744
Total									1202

other segments by travel time to those segments. For segment 2, the results are as follows:

Segment	No. of Employees	Travel Time (min)	Employment/ Travel Time
3	1202	6	200.3
4	1358	10	135.8
5	1531	14	109.4
6	2011	19	105.8
7	5482	26	210.9
Total			762.2

6. The non-CBD trips are then distributed by dividing the employment travel time for each segment by the sum of employment/travel time for all segments. This fraction is multiplied by the number of non-CBD trips for each zone. From segment 2, there are 407 non-CBD trips, which are divided as in Table 7. As the table indicates, 107 passengers boarding in segment 2 will go to zone 3, 72 will go to zone 4, and so on. These will become reverse trips later in the day. The 107 passengers going from segment 2 to segment 3 will reverse the trip, boarding in segment 3 and alighting in segment 2 later in the day.

7. The trip distribution table for non-CBD trips along route 19 can be structured as in Table 8.

8. The trip distribution table is completed by adding CBD trips and reverse trips. Reverse trips are calculated by adding paired cells. For example, segment 3 to segment 2 has 135 trips and segment 2 to segment 3 has 107. Reversing the trips would result in addition of these two numbers so that both cells have 242 (see Table 9).

Table 6. Employment by route segment.

Segment	No. of Acres: Industrial	Retail-Commercial	Total	Employees per Acre	No. of Employees
7	23	403	426	12.9	5284
6	183	91	274	7.3	2011
5	55	45	100	15.3	1531
4	7	59	66	20.6	1358
3	31	17	48	25.0	1202
2	117	154	271	21.7	5880

Results

As Table 9 indicates, the model predicted total daily passenger boardings of 5500. This is compared in the table below with an actual 1980 boarding count of 5777 for a margin of error of -5 percent:

Segment	Boardings Estimated by Model	Ridership Count	Error (%)
1 (CBD)	1712	2084	-18
2	1346	1124	+19
3	749	649	+15
4	649	838	-23
5	400	457	-12
6	274	156	+76
7	370	469	-21
Total	5500	5777	-5

Similar validation efforts for other routes in the Cleveland area have produced projections in the range of +2 to +8 percent of actual observed ridership.

At the moment, reliable models have been developed for local radial and crosstown service and are in the process of being validated through the use of historical ridership data. Work is still progressing on models for express radial service and rapid feeder service. The calibration of those two models should be completed shortly so that their validation can begin and a final report on the entire range of models developed as a result of this study effort can be published in the very near future.

FUTURE DIRECTIONS

Admittedly, the four ridership projection model study efforts mentioned earlier were designed to suit local conditions and needs and to be compatible with local data bases. This is not to say that the models in their present form cannot be used in areas other than that for which they were initially designed. This is to say that the transferability of any or all of the models is not known and can only be ascertained through further study.

To that end, it would appear reasonable that some effort over and above normal technical report dis-

Table 7. Non-CBD trip distribution for segment 2.

Segment	Employment/ Travel Time		Sum of Employment/ Travel Time				Non-CBD Trips		Trip Distribution
3	200.3	÷	762.2	=	0.263	x	407	=	107
4	135.8	÷	762.2	=	0.178	x	407	=	72
5	109.4	÷	762.2	=	0.144	x	407	=	58
6	105.8	÷	762.2	=	0.139	x	407	=	56
7	210.9	÷	762.2	=	0.277	x	407	=	114
Total									407

Table 8. Non-CBD trip distribution for all segments.

Origin Segment	No. of Trips to Segment: 2	3	4	5	6	7	Variance
1	30	6	4	3	3	8	-54
2	-	107	72	58	56	114	-407
3	135	-	38	24	20	36	-253
4	82	33	-	43	28	45	-231
5	32	10	21	-	30	34	-127
6	11	3	5	11	-	31	-59
7	17	4	6	10	22	-	-59
Total	307	163	146	149	159	268	

Table 9. Distribution of reverse trips for all segments.

Origin Segment	No. of Trips to Segment: 1	2	3	4	5	6	7	Total
1[a]	180	672	339	276	137	57	51	1712
2	672	-	242	154	80	67	131	1346
3	339	242	-	71	34	23	40	749
4	276	154	71	-	64	33	51	649
5	137	80	34	64	-	41	44	400
6	57	67	23	33	41	-	53	274
7	51	131	40	51	44	53	-	370
Total								5500

[a]CBD.

tribution is warranted to disseminate the findings of these study efforts to the transit community. Just as UMTA has offered training sessions for existing and potential users of the UTPS software packages, a similar effort could be undertaken by either the Transportation Research Board, UMTA, or the American Public Transit Association for the short-range demand estimation planners.

Transit Fare Development Procedures and Policies

EDWARD M. ABRAMS

Transit properties throughout the country will be facing difficult policy decisions in the next several years, principally due to the phasing out of federal Section 5 funds (Urban Mass Transportation Act of 1964, as amended). Furthermore, local communities have shown a strong resistance for further increases in their tax assessments. Although in the past transit was heavily subsidized by public funds, these shifts are causing more of the financial burden of transit to be placed on user charges. Fares, however, are a sensitive and visible element of transit services. The transit rider is constantly reminded of the cost of the journey each time he or she boards a bus. These riders will be hard pressed to accept the reasons for the shift in the financial burden for transit. Therefore, a fare-pricing policy must be cognizant of this attitude and attempt to mitigate it through innovative approaches and marketing programs. In the past, transit properties generally increased their fares by a uniform increase throughout the fare structure. Little attention was paid to each element of the structure in terms of effect on ridership and revenue. However, recent research has found that fare-change impacts are not uniform throughout the various submarkets within a transit property. Therefore, transit properties are faced with initiating a high level of fare increase more frequently and without a complete understanding of how the various submarkets will be affected. This paper presents an approach for addressing this problem by delineating a comprehensive development process for making transit fare changes. The process has at its foundation the development of policy guidelines to identify what is expected from the fare change. Two major analysis procedures are defined. First, a technique is presented whereby individual submarket elements of the fare structure can be changed and their ridership and revenue impact readily determined. Second, the development process relies on a building-block approach, whereby changes to each unit of the structure are tested with respect to ridership and revenue impacts and then combined into overall fare-structure alternatives. The procedures contained in this paper were developed as part of a study sponsored by the Pennsylvania Department of Transportation to offer guidance to small and medium-sized transit properties in making service and fare-structure changes.

This paper contains a suggested approach for making fare-structure changes in order to deal with the different ways each submarket of a system is affected by the change. It relies on the process developed as part of the Transit System Performance Evaluation and Service Change Manual for the Pennsylvania Department of Transportation (PennDOT) (1). The process of fare-change development suggested here contains seven major steps, as listed below:

1. Define evaluation procedures,
2. Develop analytical tools,
3. Describe fare actions,
4. Determine ridership and revenue impacts,
5. Develop alternative fare structures,
6. Evaluate alternative structures, and
7. Select and implement preferred alternative.

The procedures suggested for each, as well as how each works together to form a total fare-structure development program, are described in detail below.

DEFINE EVALUATION PROCEDURES

In planning for fare-structure changes, the first and perhaps most important step is to determine what are the objectives to be accomplished by the changes. Is it to increase revenue by a certain percentage? Is it to simplify the structure? These and other questions must be answered to guide the transit fare-development process.

The process for answering these questions could involve the delineation of criteria that have significance levels assigned to each. Such criteria may involve local priorities with respect to six factors associated with transit fares, such as the following:

1. Fiscal integrity: With cutbacks in federal operating support and with the general concern to minimize local financial support, fiscal integrity is probably the most important priority. It describes the financial objectives to be achieved by the fare changes that may be measured by the amount or percentage of a revenue increase, or the percentage of expenses recovered by the farebox.
2. Fare-structure simplification: A major advantage of transit is its low cost. In order to market this advantage, the fare structure should not be overly confusing to hinder its use.
3. Fare promotion programs: Other things being equal, the fare structure that can attract the most transit trips is preferred. Fare promotion programs, at relatively little revenue loss, can be designed to draw attention to new or improved services and also try to establish a transit riding habit.
4. Passenger equity: Although transit fare equity is hard to define, it generally can be considered in three categories: riding distance, quality of service, and patron's ability to pay. A zone structure is usually established to equalize the patron fare based on distance traveled. Premium services, which offer the patron a higher standard of dependability, speed, comfort, or convenience, may command a higher price. Finally, the relative importance of the fare to different user groups should be considered before setting the same price for everyone.
5. Ease of administration: A fare structure must lend itself to easy (low administrative cost) collection of, and accounting for, route revenues. Security of revenues is also a consideration.
6. Effect on energy and the environment: Transit can play a major role in energy conservation and in improving the environment. Therefore, the fare structure may be a key element to influence a modal shift from private automobile to transit.

What is probably most obvious from these six criteria is that they do not work together; that is, a change to enhance satisfaction of one criterion may affect another criterion detrimentally. Thus, there are trade-offs to be made, so that a balance may be struck between the different criteria. These trade-offs can be made by assigning significance levels to each by ranking each as to its relative importance in making fare changes.

DEVELOP ANALYTICAL TOOLS

Use of mass transit service is a function of several factors, including demand characteristics (such as population density and socioeconomic characteristics of the area residents) and supply characteristics (such as route spacing, headways, area coverage, type of service, price of service, etc.). Demand for transit service, however, is closely related to two important supply characteristics--price and level of service. Variation in price, level of service, or both generates change in the use of transit.

An increase in price causes a decrease in ridership. The relation between price and transit use is known as the demand function and is measured by a price-elasticity formula. The literature on transit price elasticity generally refers to two elasticity measures--arc elasticity and shrinkage ratio--although others are mentioned. Historically, the shrinkage ratio was the most commonly used price-elasticity tool (2). However, as a result of extensive research in this field, the arc-elasticity method appears to be more accurate. It is calculated as a ratio of the percentage change in transit use divided by the percentage change in price when the base of the price percentage change is the average of the before and after values. For small changes in ridership, the arc-elasticity formula can be expressed as follows:

$$\text{Arc elasticity} = [(R_2 - R_1)/R_1] \div \{(F_2 - F_1)/[(F_1 + F_2)/2]\} \quad (1)$$

where

R_1 = average daily ridership before fare change,
R_2 = average daily ridership after fare change,
F_1 = average fare before fare change, and
F_2 = average fare after fare change.

The wide range of elasticities reported in the literature points out that riders in different cities and among various transit submarkets within the same city react differently to pricing (3). This evidence leads to the conclusion that ridership and revenue impacts should be accomplished on a disaggregate level. During the preparation of the fare-change manual for PennDOT, data were collected on elasticity values for different fare categories. The values from this research of medium-sized properties located in cities in Pennsylvania indicated that variations in elasticity values can be significant, as seen below:

Item	Elasticity Values
Local routes	-0.20 to -0.30
Express routes	-0.30 to -0.40
Elderly and handicapped riders	-0.30 to -0.35
Student riders	-0.30 to -0.40
Peak ridership	-0.10 to -0.20
Off-peak ridership	-0.30 to -0.40
Transfer riders	-0.30 to -0.40

For these elasticity factors, ridership and revenue estimates were developed in graphic form and are shown in Figures 1 and 2, respectively. These charts are applicable to fare changes that use the arc-elasticity method when they are keyed to the local elasticity values.

DESCRIBE FARE ACTIONS

A transit fare structure is a composite of a number of elements, including various levels of fare with respect to types of service, various methods of payment of fare, and various procedures used to identify eligibility of a passenger to use a given type of fare.

A transit fare structure generally consists of a variety of programs that include fares for the elderly during off-peak periods, reduced fares for handicapped persons, weekly passes, a transfer fee, etc. In theory, revisions to the fare structure could involve almost limitless permutations. At the same time, consideration must be given to the effect of revisions on individual user groups, not only the bottom line, since various combinations of changes may yield similar ridership and revenue total results.

To analyze potential revisions to a fare structure, a building-block approach is suggested. In this way, alternative revisions are evaluated, one by one, as discrete actions; these revisions are changes that might be made to individual elements of the fare structure. The emphasis should be placed on formulating numerous individual or discrete actions for altering the fare, each of which can be analyzed independently. Examples of discrete actions would include items such as peak pricing, off-peak pricing, zone fares, transfer fees, express surcharges, multiride tickets, changing pass discounts or multiples, and promotional activities.

The final step would be to construct an array of potential changes by each fare category. For example, a hypothetical array of fare-change actions is given in Table 1. For each type of fare category, the existing fare structure is described along with several alternative ways the category could be changed.

DETERMINE RIDERSHIP AND REVENUE IMPACTS

Fare revisions must be evaluated to determine their impacts on different rider groups and on systemwide revenue. This step, therefore, involves using the elasticity formulas to determine the ridership and revenue impacts of each discrete action.

Of course, the completion of this process requires additional data to perform the analysis. These would primarily include existing ridership levels by each discrete-action category. For example, if a fare change to increase the transfer fee is contemplated, it is necessary to know the number of transfer passengers before the fare change can be properly evaluated. Results from this step will be an array of changes in ridership and revenues due to each discrete action.

DEVELOP ALTERNATE FARE STRUCTURES

Once discrete actions for fare revisions are identified, fare-structure alternatives can be formulated to fulfill wide-ranging objectives. This can be accomplished by combining the most attractive building blocks into the overall fare programs.

Alternative fare structures might embrace concepts such as fiscal integrity, simplicity, or equity. Each fare program may also contain elements that are not exclusive to a particular structure. For example, Table 2 sets forth three hypothetical alternative structures, which in several instances have the same discrete fare change for more than one alternative.

Figure 1. Percentage of fare increase versus ridership loss.

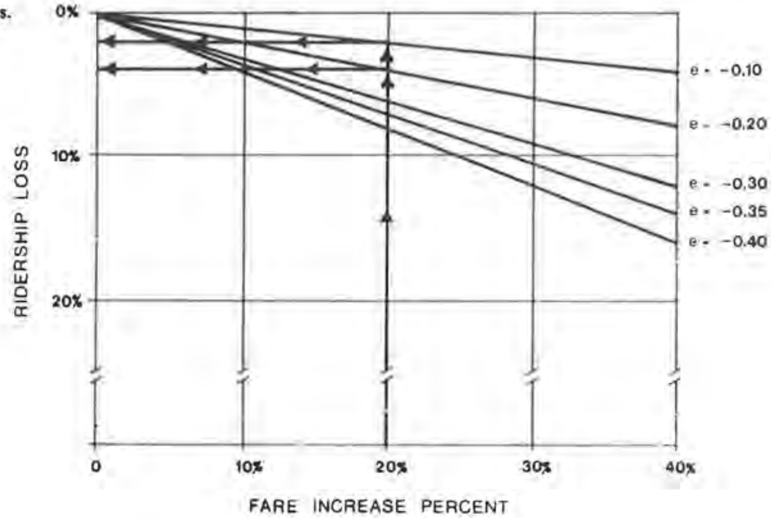


Figure 2. Percentage of fare increase versus revenue gain.

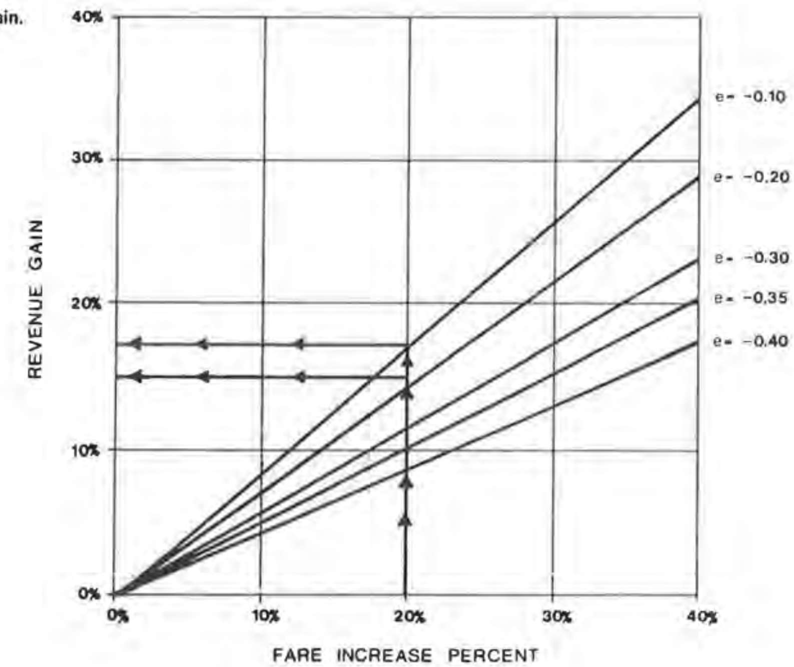


Table 1. Hypothetical array of discrete fare-change actions.

Item	Current Hypothetical Fare Structure	Revenue Impact of Discrete Fare-Change Actions		
		Minimum	Intermediate	Maximum
Peak adult fare ($)	0.45	+0.05	+0.10	+0.15
Off-peak adult fare ($)	0.35	Existing	+0.05	+0.10
Zone fare (per zone) ($)	0.30	Existing	+0.10	+0.20
Transfers ($)	Free	Free	+0.05	+0.10
Express surcharge ($)	0	0	+0.05	+0.10
Student fares ($)	0.35	Existing	+0.05	+0.10
Monthly pass discount (%)	10	Existing	5	0

EVALUATE ALTERNATIVE STRUCTURES

Evaluation criteria were established earlier to set the policy guidelines for making a fare-structure change. In this step, those criteria would be used to assess each fare-structure alternative to determine which alternative should be selected.

The process of evaluating the alternative structures is fairly straightforward. It involves multiplying the relative importance score for each of the six criteria (i.e., fare structure simplification, fiscal integrity, fare promotion programs, passenger equity, ease of administration, and effect on energy and the environment) times the degree to which the alternative satisfies those specific criteria to equal a weighted score. (Note that the score for relative importance and for degree of satisfaction is between 0 and 5, where 0 represents the least satisfaction and 5 the greatest.) The weighted score for each criterion would be summed to arrive at the composite score for the particular alternative being evaluated.

SELECT AND IMPLEMENT PREFERRED ALTERNATIVE

The final step in the development process of fare changes is the ranking of alternatives by their respective scores obtained in the prior step. At this

Table 2. Example of fare program alternatives: fiscal integrity focus.

Item	Alternative A: Minimal			Alternative B: Selective			Alternative C: Major		
	Fare	Ridership Change	Revenue Change	Fare	Ridership Change	Revenue Change	Fare	Ridership Change	Revenue Change
Peak period adult ($)	+0.05			+0.15			+0.15		
Off-peak adult ($)	+0.05			+0.05			+0.10		
Zone fare ($)	Existing			Existing			+0.20		
Transfers ($)	Existing			+0.05			+0.10		
Express surcharge	+0.05			Existing			+0.10		
Student fares ($)	+0.05			Existing			+0.10		
Monthly pass discount (%)	Existing			5			0		

Note: This table is set up as a guide for transit property administrators. The ridership and revenue columns are left blank to show the items to be completed by administrators, so as to see how each fare alternative would alter both ridership and revenue.

point, the fare alternative that has the highest score should be reviewed to ensure that it satisfies the overall objectives. If it does not, then it may be changed by adding or deleting certain elements. If it does, it should be moved forward to implementation.

REFERENCES

1. Simpson and Curtin; University of Pennsylvania. Transit System Performance Evaluation and Service Change Manual. Pennsylvania Department of Transportation, Harrisburg, Feb. 1981.
2. J.F. Curtin. Effect of Fares on Transit Riding. HRB, Highway Research Record 213, 1968, pp. 8-20.
3. M.G. Ferreri. Transit Price Elasticities Revisited. ITE Transit Journal, Summer 1979.

Pennsylvania's Urban Operating-Assistance Grant Formula Methodology

JOHN DOCKENDORF

The Pennsylvania Department of Transportation (PennDOT) has provided state transit operating assistance since 1968. For many years, the Bureau of Mass Transit Systems, PennDOT, managed this activity as a discretionary grant program. However, due to the growth of the program and the increasing complexity of these grants, the Bureau determined that a better grant methodology was needed to ensure that state operating-assistance grants were adequate, equitable, and predictable. In FY 1976-1977, the Bureau began experimenting with a formula grant methodology that was based on financial need and system performance. After two years of experimentation and refinement, the Bureau developed a formula grant methodology that it believed achieved the objectives of predictability, equity, and adequacy. This concept was accepted by the state's transit industry as a reasonable and fair method to determine state operating-assistance grants, and efforts began in FY 1979-1980 to achieve passage of state transit legislation based on this grant methodology. This effort was successful and culminated in the passage of the Pennsylvania Urban Mass Transportation Law (Act 101) on July 10, 1980. The key elements of Act 101 and how they are applied to state operating-assistance grants are described. PennDOT believes this methodology has application for other transit funding programs and hopes other agencies can use the concepts in administering their transit programs.

Ever since the passage of Act 8 (Pennsylvania Urban Mass Transportation Assistance Law of 1967), Pennsylvania has participated in providing transit operating assistance to urban and nonurban transit systems. Originally, this program was administered by the Department of Community Affairs. However, in 1970 this function was transferred to the newly created Bureau of Mass Transit Systems when the Pennsylvania Department of Transportation (PennDOT) was established.

Act 8 authorized the state to fund up to two-thirds of operating losses, and localities were responsible for providing the remaining one-third match. With the introduction of the federal Section 5 operating-assistance program in 1974-1975 (Urban Mass Transportation Act of 1964, as amended), this policy was modified to authorize the state to fund up to two-thirds of the nonfederal share of operating deficits. State funds were allowed as matching funds for Section 5 grants.

This policy remained in effect until FY 1980-1981, when Act 101 (Pennsylvania Urban Mass Transportation Law) was passed. This legislation authorized PennDOT to provide a state subsidy of at least two-thirds, but not more than three-quarters, of its constrained deficit. The constrained deficit was defined as an amount equal to constrained operating cost reduced by assumed revenues and federal operating subsidies. These concepts are defined and discussed later in the paper.

Until the passage of Act 101, the Bureau determined operating-assistance grants in a discretionary manner by relying on rules of thumb, past state funding experience, and anticipated budgets. In the early years of the program, there were not a large number of program applicants, so it was possible to review projects on a line-item, as well as an overall, basis. The Bureau developed an extensive transit data-reporting system, which included an annual questionnaire, a detailed project application, and quarterly progress reports. Therefore, the staff had reasonably good knowledge of the participating transit systems, and the discretionary grant program worked fairly well for many years.

However, in the mid-1970s, the state operating-assistance program grew significantly, and it became increasingly more difficult to administer the program on a discretionary basis. Systems frequently did not understand how their grants were determined, oftentimes they did not permit the staff to take into account all relevant information in assessing financial need, and there was a growing perception that state transit grants were inequitable. Small systems believed the large systems received a disproportionate amount of state aid, while both of the state's two largest systems--Southeastern Pennsylvania Transit Authority (SEPTA) and Port Authority of Allegheny County (PAT)--thought the other was receiving a disproportionate amount of state aid.

Compounding this problem was the ever-widening gap between financial need and the overall state appropriation for transit operating assistance. Nationally, transit expenses increased an average of 11 percent annually from 1967 to 1975, while transit revenues only increased around 3 percent annually during this same period. This resulted in an average annual increase in transit deficits of more than 15 percent. Pennsylvania's transit industry experienced similar financial trends. At the same time, the statewide transit operating-assistance appropriation only increased at an average annual rate of 3 percent. From FY 1974-1975 through FY 1976-1977, the state transit appropriation was a constant $74.2 million. In the short run, this funding shortfall drained working capital accounts and required large increases in local subsidies over the levels required to match state grants. Gradually, this situation led to fare increases and service decreases. In the long run it was obvious that, unless this trend could be stopped, there would be a repeat of the cycle of higher fares, lower ridership, and reduced service that led to the collapse of most large transit systems in the 1960s and early 1970s, both statewide and nationally.

DEVELOPMENT OF FORMULA GRANT METHODOLOGY FOR DETERMINING STATE OPERATING ASSISTANCE

To overcome the problems of (a) inadequate overall transit funding, (b) uncertainty as to expected levels of individual system state transit funding, and (c) achievement of more equity in state transit funding, the Bureau began developing a formula grant methodology to distribute state transit operating assistance in FY 1976-1977. A comprehensive review was made of state and federal transit financial and operating trends as well as a study of probable future financial need and likely available state financial resources to meet transit needs. The outcome of this six-month in-house study was the development of a state transit formula methodology that was first used on an experimental basis in FY 1977-1978 for both the determination of individual state transit grants and for determining the required level of overall state transit assistance in FY 1978-1979. (The experimental methodology for FY 1977-1978 grants began after the statewide transit appropriation of $79 million was already established, so the procedures were limited to allocating the $79 million to individual participating transit authorities. A different methodology was tested in FY 1978-1979, which addressed not only the distribution of state funds but also the magnitude of the statewide transit appropriation.) This original formula methodology was refined and improved on for two years and ultimately became part of the state law with the passage of Act 101 in FY 1980-1981.

The calculation of state transit grants under Act 101 is divided into two parts: (a) the determination of financial need and (b) the assessment of transit system performance. The first part reflects a recognition that the primary justification for annual increases in state transit operating assistance is to cope with inflation. Therefore, 90 percent of the state's annual transit appropriation is earmarked to meet basic financial need. This part of the grant is known as the basic grant. [The 90 percent ratio is derived from the portion of the state grant dedicated to fund financial need (66.67 percent of the constrained nonfederal deficit) relative to the maximum possible state grant equal to 75 percent of the constrained nonfederal deficit. The 10 percent portion of the grant for financial need is based on the remaining 8.50 percent funding of the nonfederal constrained deficit relative to the maximum 75 percent state funding level.]

The remaining 10 percent of the state annual appropriation is earmarked to reward transit systems for improved performance. This element of the grant-determination methodology is needed to provide a financial incentive for transit systems to improve overall efficiency, effectiveness, and use that, in the long run, will reduce financial need.

Both concepts were designed to overcome what were regarded as major shortcomings in the discretionary state operating-assistance grant program. The lack of parameters in defining financial need gave the perception that there were no limits on state transit aid. This blank-check perception was a serious deterrent to obtaining legislative approval for increases in the annual state operating-assistance appropriation.

In addition, the lack of any financial incentives to reward improved transit performance made the entire state grant a function of financial need. This had the effect of serving as a disincentive for improved transit productivity, as relatively inefficient transit systems generated higher operating assistance when compared with more efficient transit systems. Although the Act 101 transit aid methodology does not eliminate this outcome, it does at least moderate this counterproductive trend.

The determination of financial need is based on three important concepts. The first is the application of a maximum expense factor, which represents the percentage ceiling on the increase in transit operating expenses. The second concept is the use of a ratio of minimum revenue to expense, which represents the lowest allowable percentage of operating expenses that revenues are expected to cover for maximum state funding. This ratio served as a floor for operating revenue levels in a given year. The final concept is the constrained nonfederal operating deficit, which is the difference between the maximum allowable transit operating expenses based on the maximum expense factor and the minimum required transit operating revenue based on the ratio of minimum revenue to expense and after estimated federal funding has been deducted. The constrained nonfederal operating deficit is the bottom-line variable in determining state transit operating assistance.

Use of the constrained nonfederal operating deficit is only applied in cases where the projected deficit submitted by a transit authority exceeds the value of the constrained nonfederal operating deficit, as Act 101 stipulates that state reimbursement shall not exceed the difference between actual operating costs less actual revenues and federal subsidies for any fiscal year. For convenience, it will be assumed that the constrained nonfederal operating deficit is less than the projected nonfederal deficit in the remainder of this paper.

The maximum expense factor was devised to help ensure that financial need projected by transit systems is not excessive relative to inflation. It is derived by calculating the percentage increase in aggregate transit operating expenses for the previous fiscal year versus aggregate transit operating expenses for the prior fiscal year. This percentage change is then multiplied by 1.15 to reflect anticipated inflation. For example, the percentage increase in aggregate transit operating expenses in FY 1980-1981 versus FY 1979-1980 was 10.94 percent. This 10.94 percent increase was multiplied by 1.15, and the result of 12.58 percent represented the maximum expense factor for FY 1981-1982.

After the maximum level of projected operating expenses is determined for each transit system, this value is multiplied by the ratio of minimum revenue to expense to determine the minimum required level of transit operating revenue.

The ratio of minimum revenue to expense was developed to help ensure that increases in transit revenue kept pace with increases in transit expenses over time. An underlying premise is that users should bear a reasonable share of financing constrained financial need through periodic fare increases that are commensurate with inflation.

Intially, it was found that the statewide average transit revenue/expense ratio in Pennsylvania was approximately 50 percent for urbanized transit systems in FY 1979-1980. A review of national and statewide financial trends indicated that this ratio had declined approximately 5 percent per year during the early and mid-1970s. This was due to annual increases in transit revenue of only 3 percent per year, while increases in transit expenses averaged 11 percent per year. It was clear that a continuation of this trend would result in transit financial need increasing beyond reasonable and realistic levels of available government transit assistance.

A compromise revenue/expense ratio policy was adopted that permitted a downward sliding scale of ratios of minimum revenue to expense over time. In effect, transit revenue is expected to increase at approximately one-half of the annual rate of the maximum expense factor. This compromise was made to avoid the likely outcome of annual fare increases. Instead, a policy of encouraging periodic (but not annual) fare increases was established. The long-run issue of how long the downward sliding scale of revenue/expense ratios could be permitted was deferred as there was an expectation that gasoline shortages and higher gasoline prices would ultimately result in an upsurge in both transit ridership and revenue that would ultimately level off, if not reverse, the past historical trend of the revenue/expense ratio.

Unlike the maximum expense factor, where the lower of constrained expenses or projected expenses is used in calculating the constrained operating deficit, Act 101 only requires that transit revenue equal to the minimum level derived by applying the ratio of minimum revenue to expense to constrained expenses be used in calculating the constrained operating deficit. This concept is known as assumed revenue. Projected transit revenue in excess of the minimum level of required revenue may be used to partly offset projected transit expenses in excess of the maximum allowable expenses. This compromise was necessary to achieve the passage of Act 101. Table 1 summarizes the maximum expense factors and the ratio of minimum revenue to expense used to date.

After the maximum level of transit expenses and minimum level of transit revenue have been determined, the constrained operating deficit is determined by calculating the difference between these respective values. The projected level of federal Section 5 funding is then deducted to derive the appropriate constrained nonfederal deficit. Transit systems are ensured that they will receive 66.67 percent of the constrained nonfederal deficit based on financial need. This represents approximately 90 percent of the maximum authorized state funding level of 75 percent of the constrained nonfederal deficit.

The remaining 10 percent of potential state transit aid is based on improved transit performance. Again, the concept of bottom-line indicators is used. Four bottom-line ratios are used as determinants of improved system performance. They are as follows:

1. Improved ridership (revenue passengers) per vehicle hour,
2. Improved operating revenue per vehicle hour,
3. Reasonable operating expense per vehicle hour, and
4. Reasonable operating revenue/expense ratio.

These ratios were adopted as proxies of transit system performance, since the overall goal of most transit programs and policies is to maximize ridership and revenue per unit of service and to minimize operating expenses per unit of service. Virtually all improvements in efficiency, effectiveness, and/or use should positively affect one or more of these selected ratios. Vehicle hours were included in the first three ratios to reflect the amount of transit service provided. Thus, a demand and supply relation was established in the first two ratios and an efficiency relation developed in the third ratio. Vehicle hours, rather than vehicle miles, were used as a proxy of transit supply due to the relatively greater cost sensitivity of the former variable.

The actual determination of which of these four possible bonuses each urban transit system qualifies for (if any) is made by comparing these ratios between the previous and prior years. Lagging is necessary because budgeting rules require that transit appropriation projections be made approximately one year in advance. In addition, the Bureau wants to make sure that each bonus awarded is based on actual, rather than projected, data.

For the bonuses for improved ridership and improved operating revenue per vehicle hour, transit

Table 1. Summary of maximum expense factors and ratio of minimum revenue to expense requirements actually used for determining financial need.

	Percentage of Revenue Requirement by Fiscal Year							
Item	1977-1978	1978-1979	1979-1980	1980-1981[a]	1981-1982[a]	1982-1983	1983-1984	1984-1985
Maximum expense factor	-	8.13	9.80	12.58	12.58	9.78	NA	NA
Ratio of minimum revenue to expense requirement	-	45	46	48	48	46	44	42

[a]Percentages reflected the assumed revenue requirements as stipulated in Act 101. The duplicate values in FY 1980-1981 and FY 1981-1982 reflect the fact that sections of Act 101 became effective in the middle of FY 1980-1981 (January 1, 1981); therefore, these values were used for both fiscal years. Values prior to FY 1980-1981 were derived through pre-Act 101 methodology.

systems qualify if this ratio in the previous year equals or exceeds this ratio in the prior year. This amounts to an all-or-nothing approach, as the degree to which this ratio improved is not taken into account in determining the magnitude of this financial bonus.

The bonuses for reasonable operating expense per vehicle hour and reasonable operating revenue/expense ratio are handled by linking them to the financial need determinants. Systems qualify for the former award if their operating expense per vehicle hour in the previous year versus the prior year increased by less than the maximum expense factor. It is accepted that this ratio is likely to increase over time due to inflation. However, systems are expected to control this increase to levels below the actual rate of inflation if they expect to receive this financial award.

Similarly, it is expected that the operating revenue/expense ratio will decline over time. Therefore, to qualify for the bonus for the reasonable operating revenue/expense ratio, this ratio must decline in the previous year versus the prior year by less than the 2 percent annual decrease allowed for in Act 101.

Systems qualify for an amount equal to 2 percent of the constrained nonfederal operating deficit for each of the ridership, revenue, and reasonable-expense bonuses. The bonus for ridership per vehicle hour is weighted slightly higher at 2.33 percent of the constrained nonfederal operating deficit. This higher weight is needed to cover the residual 0.33 percent between the basic grant level of 66.67 percent of the constrained nonfederal deficit and the maximum state authorized share of 75 percent of the constrained nonfederal deficit. (This residual assumes that the potential 6 percent of the constrained nonfederal deficit for the three other bonuses has been added to the 66.67 percent basic grant level, which results in funding of up to 72.67 percent of the constrained nonfederal deficit.)

After the financial need and bonus awards have been determined for each transit system, the individual grant amounts are totaled and an appropriate transit budget request is made based on this aggregate figure. Due to the need to do this one year in advance of the project year for which these funds will be used, it is essential that the budget information on projected expenses and revenue be highly accurate. Also, this approach requires lagging of bonus awards, as these bonuses need to be known at this time so they can be included in the budget.

The Pennsylvania General Assembly is then called on to approve the mass transit budget appropriation request based on these guidelines. Assuming both houses of the Assembly concur (House and Senate), the bill is sent to the Governor for signature and the budget becomes effective for the next fiscal year. In cases where the actual approved transit appropriation is less than the amount based on the Act 101 formula, the law requires that the state prorate each grant downward based on the constrained nonfederal deficit derived through the methodology outlined in this paper.

OVERALL ASSESSMENT OF ACT 101 FORMULA GRANT METHODOLOGY

PennDOT believes that the formula grant methodology has achieved most of the intended objectives. For example, there is evidence that the maximum expense factor and minimum assumed revenue requirement are constraining financial need. Also there is evidence that the performance bonuses are influencing revenue generation.

In addition, state transit operating assistance is more adequate than before. The average annual increase in transit operating-assistance appropriations has increased dramatically since the passage of Act 101. State transit aid is also more predictable. Most systems are able to estimate within 90-95 percent accuracy the magnitude of their final state grant 6-12 months in advance of their award. Finally, the methodology is more equitable. The Bureau maintains work sheets that document how each individual grant was calculated, so that systems may review them in order to verify the accuracy of the calculations and to confirm that their grant was determined in accordance with the law.

Basing state transit aid on the constrained nonfederal deficit is another advantage of the Act 101 formula grant methodology. It is a more direct and relevant measure of transit financial need than indirect measures such as population, population density, and variables such as ridership, vehicle miles, and peak vehicles. Also, the use of financial bonuses to reward systems for improved performance is helpful in providing an incentive to improve efficiency, effectiveness, and productivity. Use of this tool at least moderates the usual practice of basing government grants entirely on financial need. In addition, it offers the long-term potential of reducing the level of financial need as increases in ridership and revenue and improved cost control are achieved.

Naturally, the use of this formula grant methodology has some disadvantages. The Bureau of Public Transit Systems has less flexibility in modifying grants to reflect changing conditions, emergencies, and other unanticipated events. Also, the combination of requiring extremely accurate financial projections and considerable lagging in determining bonus awards makes the methodology less flexible than a discretionary grant program. Finally, the political compromises needed to achieve passage of Act 101, such as the application of assumed revenue in lieu of projected revenue and the requirement for across-the-board prorating when transit appropriation levels are less than the amounts needed to fully fund Act 101, make it less predictable and less constraining than was intended when the formula grant methodology was developed. However, the Bureau does not believe these disadvantages come close to outweighing the advantages of this grant methodology.

The major unknown relates to the impact of the elimination of federal Section 5 operating assistance on Act 101. At no time during the development of Act 101 was it anticipated that federal Section 5 funds would be terminated. Therefore, the cornerstone of financial need was the calculation of the nonfederal deficit. Without federal aid, both state and local shares of the constrained nonfederal deficit will increase dramatically. It remains to be seen if Pennsylvania and the local governments that sponsor the urbanized systems have sufficient financial resources to make up this lost federal transit aid. If not, amendments to Act 101 will probably be needed in order to cope with this potential shortfall in overall state and local transit operating assistance due to the elimination of federal transit aid.

Transit Fare Elasticity: Role in Fare Policy and Planning

PATRICK D. MAYWORM

With the planned phaseout of federal operating assistance over the next five years, transit managers across the country are looking to the farebox to raise the necessary revenues to maintain current levels of service. Since raising fares is a necessary but politically sensitive means of boosting operating revenues, much more attention is being placed on developing accurate ridership and revenue models and on identifying fare policies that will increase revenues with minimal effect on ridership. Fare elasticity of demand and its role in ridership and revenue planning and in developing fare policy are discussed. The fare elasticity is a useful concept because it provides information on how riders respond to fare changes. Since fare elasticities usually vary significantly by trip distance, time of day, and quality of service, transit managers can take advantage of these differences by differentially pricing their services in order to increase revenues and ridership. Differential pricing, however, does have its political and monetary cost. Whether the revenue-generation advantages as indicated by the variation in fare elasticities are worth the increased administrative costs will be answered by each system in the next few years as the financial pressures increase and when the revenue-generation potential of differentially priced transit service becomes of paramount importance.

The reasons behind the current financial problem in transit are very clear. Between 1975 and 1980, total transit operating expenses increased at an average annual rate of 11.8 percent while farebox revenues increased at only half that rate, or 5.8 percent annually (1, pp. 46-47). This in itself has not been a problem, since operating subsidies grew at a much faster annual rate of almost 21 percent over this same five-year period; the largest growth in operating assistance came from the federal government. Now, however, the federal government is planning to eliminate all operating assistance in the next five years.

Fortunately, the federal government's share of total operating revenue is only 17.3 percent, or slightly more than the level contributed by the states. If we assume that state and local contributions to transit service operations will continue to grow but at slightly slower rates, the loss in federal operating dollars will have to be met by higher farebox revenues and/or reductions in operating costs in order to maintain the balance between revenue and cost. Typically in the wake of subsidy shortfalls, fares are arbitrarily increased and service levels are reduced so that available revenues cover the operating expenses after all remaining operating subsidies are committed. Although such decisions can provide temporary solutions to financially troubled transit companies, more rational short-range policies that fit into a long-term approach to transit financial planning should be adopted soon if transit companies are going to remain solvent in the near future and maintain the political support they require.

Aside from determining how high the fares should be with respect to the subsidy level (i.e., what percentage of total operating revenues should come from the farebox), the principal problem with transit planning today is that fare and service-level decisions are hardly ever jointly planned and considered despite the fact that fares and service levels are intrinsically related. If a greater proportion of operating revenues is going to have to come out of the farebox with minimum losses of ridership, these fare levels will have to be determined in conjunction with the quantity, quality, and cost of the service provided. In addition, less-traditional fare and service concepts should be given serious consideration when major policy changes are under review. The financial crisis and high inflation we all face today should not divert our attention from the need to choose wisely and from a wide range of alternatives. In fact, the current financial situation should highlight the importance of making the right decision as a result of a careful analysis of choices in relation to specific operating objectives.

This paper focuses on one particular element or factor that enters into the equation when making the range of important trade-offs suggested above--the fare elasticity of demand. It is an important concept because it describes how individuals or groups of individuals react to fare changes. In a more generalized sense, the fare elasticity of demand also tells us something about how important the fare level is with respect to the total cost of travel (i.e., including wait, walk, and in-vehicle time costs). This paper presents new information on the fare elasticity of demand and suggests how the elasticity can be used in ridership and revenue analysis and for developing fare policy.

FARE ELASTICITY OF DEMAND: A DEFINITION

The demand for public transit is influenced by many factors, including the level of transit fares, the quality and quantity of service provided, and other factors outside the control of the transit company. The elasticity of demand is a convenient measure of the relative responsiveness of transit ridership to changes in these individual factors. As a quantitative measure of relative change, the elasticity of demand is defined as the ratio of the proportional change in transit demand to the proportional change in the factor being observed. Thus, the transit fare elasticity will indicate the percentage change in transit ridership resulting from a 1 percent change in fares. Since the percentage change in ridership, fares, and services is independent of the units in which each is measured, the ratio of percentage changes--the demand elasticity--is also dimensionless. Therefore, one may compare, for example, the fare elasticities observed in England with those observed in the United States.

Transportation analysts have used several methods for computing the elasticity of demand; each results in slightly different numerical values. It is obviously beyond the scope of this paper to provide a detailed review of the four principal mathematical relationships used to compute a fare elasticity, and the reader is referred to Grey (2) or to Mayworm, Lago, and McEnroe (3) for a more comprehensive discussion. Nevertheless, it is relevant to at least identify the four measures that appear most often in the literature:

Point elasticity:

$$\eta_{pt} = (\partial Q/\partial F) \cdot (F/Q)$$

Shrinkage ratio:

$$\eta_{sr} = [(Q_2 - Q_1)/Q_1] \div [(F_2 - F_1)/F_1] = (\Delta Q/Q_1)/(\Delta F/F_1)$$

Midpoint elasticity:

$$\eta_{mid} = \{(Q_2 - Q_1)/[(Q_2 + Q_1)/2]\} \div \{(F_2 - F_1)/[(F_2 + F_1)/2]\} = [\Delta Q/(Q_1 + Q_2)]/[\Delta F/(F_1 + F_2)]$$

Arc elasticity:

$$\eta_{arc} = (\log Q_2 - \log Q_1)/(\log F_2 - \log F_1)$$

where

Q_1 = initial level of ridership,
Q_2 = new level of ridership,
F_1 = initial fare level, and
F_2 = new fare level.

The point elasticity is derived from the actual transit ridership demand curve and can be evaluated at any point along the curve. Although it is perhaps the most useful measure for ridership planning, since it is derived from the demand model, many transit analysts do not have enough information from which to develop such functions for groups of riders let alone for the system as a whole.

The three remaining measures, therefore, are used most often to estimate elasticities from ridership and fare-level data corresponding to periods before and after a fare change. Of these, the shrinkage ratio or loss ratio is perhaps the most common measure. Although there are numerous advantages and disadvantages to using all three elasticity measures, the midpoint and arc elasticity definitions will yield more consistent results both for a transit company and across sites, especially for large fare changes such as those occurring today.

NEW INFORMATION ON FARE ELASTICITIES

During the year after Ecosometrics published its compilation of demand elasticities (3), several studies were released that added to the body of literature on this subject. Perhaps the most comprehensive and professional study on the demand for public transportation was recently published by the Transport and Road Research Laboratory (TRRL) in 1980 (4). This international collaborative study reviews all the factors affecting public transit demand, including fares. Based on a wide variety of studies from all over the world, this study reconfirms the fact that fare elasticities are low; they range from -0.10 to -0.60. Thus, increases in fare levels will almost always lead to increases in revenues. Off-peak travel is about twice as elastic as peak travel and those with an automobile available are more sensitive to fare changes than those captive to the public transit system for most of their travel needs. In addition, persons traveling short distances are more responsive to fare changes than those traveling long distances. However, there is some evidence to suggest that the fare elasticity will once again rise with very long trips as passengers have the opportunity to find alternative destinations.

In addition to this excellent document, several British analysts have recently published results of studies on passenger demand. Oldfield and Tyler (5) and Stark (6) looked at passenger response to changes in suburban rail fares by using time-series data from the 1970s. Stark analyzed passenger response in the suburbs of Glasgow, Scotland, during this period and estimated a fare elasticity of -0.45. Oldfield and Tyler looked at British Rail fares on services connecting London to its suburbs and found that commuters were less responsive than reduced-fare patrons to fare changes and that there was no systematic variation in the fare elasticity for commuters according to distance. The fare elasticity for reduced-fare riders, however, grew as the trip distance increased. The general results of Oldfield and Tyler's study of the elasticity of medium-distance rail travel are presented below (5):

Fare Category	Fare Elasticity ± Standard Error
Full fare	-0.20 ± 0.04
Pass	-0.50 ± 0.07
Reduced fare	-0.65 ± 0.06

Another fare study from England was performed by Urquhart and Buchanan and published by TRRL in 1981 (7). This study focused on the effects of fare and service changes on shopping and nonshopping travel in Telford, England. Although there are questions surrounding the methodology used to measure passenger response to the fare change while extensive management and service changes were taking place, the analysts were able to conclude that the fare elasticity for shopping trips was between -0.58 and -0.80; the mean for three model formulations was -0.65. Nonshopping travel was found to be less elastic; the fare elasticity varied from -0.32 to -0.46. The mean fare elasticity for nonshopping trips computed from six model formulations was -0.40. This study also found that many shopping trips taken by bus were redistributed among various shopping centers in the Telford area in response to the fare changes.

Very few new studies have been published in the United States presenting new evidence on transit fare elasticities of demand. Although based on fare changes that occurred in 1976, two studies by Knudson and Kemp provide reliable evidence of how transit riders respond to fare changes (8,9). In September 1976, the Erie Metropolitan Transit Authority (EMTA) raised the cash fare and adult and student tokens. The study results indicate that adult token riders are less elastic than cash users and that students are very sensitive to fare increases, as shown below. The systemwide point elasticity was calculated to be -0.33 (8):

Fare Category	Fare Change ($)	Point Elasticity
Adult token	0.28-0.31	-0.25
Cash	0.30-0.35	-0.38
Student token	0.20-0.25	-0.72
Systemwide	0.29-0.34	-0.33

In a second study performed by Knudson and Kemp (9), fare elasticities were estimated following a November 1976 systemwide fare change in the Kentucky suburban counties of Cincinnati. The base fare on the Transit Authority of Northern Kentucky (TANK) system increased 60 percent from $0.25 to $0.40, which resulted in a fare elasticity of only -0.15. A summary of the fare elasticities by fare category is presented below (9):

Fare Category	Fare Change ($)	Point Elasticity
Base	0.25-0.40	-0.15
Elderly and handicapped off-peak	0.10-0.20	-0.26
Student cash	0.20-0.25	-0.63
Systemwide	0.23-0.37	-0.12

Although the results are in general agreement with other studies, the very low aggregate fare elasticity indicates that riders were very insensitive to the large fare increase. Knudson and Kemp attribute this low response to the captive nature of the riders (i.e., generally low-income commuters, females, and the elderly) and the timing of the fare change (i.e., just before Thanksgiving and the Christmas season). These two studies as well as other studies performed by Kemp (10) and by Goodman, Green, and Beesley (11) should be read by all transit-pricing analysts because of the careful methodology used in attempting to isolate the effects of fares on transit ridership levels.

Other studies have also been published that have computed fare elasticities; they are not discussed here because they add little to what is already known. Unfortunately, there has yet to be published any information on how riders are reacting to the massive and frequent fare changes that have occurred during the past two years. There is very little empirical evidence to suggest that these fare increases would result in significantly different elasticities than those already reported. There are, however, theoretical arguments purporting that indeed ridership becomes more sensitive to fare changes as the fare increases. This is discussed in more detail later in this paper.

FARE ELASTICITY IN RIDERSHIP AND REVENUE PLANNING

In the past few years there has been a resurgent interest in the concept and use of the elasticity of demand--specifically, the fare elasticity of demand. Although the values obtained from quantitative analyses or borrowed from the analyses performed on other systems are an important factor in ridership and revenue planning, the amount of information that they can provide is limited. As briefly mentioned earlier, the demand for transit is affected by a large number of factors, of which fares is only one. Service quality, wait and walk time, and reliability are a few additional factors that are known to influence ridership levels. Moreover, there are factors beyond the control of a transit manager that will influence the number of people using transit and their frequency of use. In a recent paper, for example, Ulberg (12) found that the supply of gasoline and the level of employment in the Seattle metropolitan area were very important factors influencing ridership on the Seattle Metro transit system. Fare levels, therefore, can only give us part of the information we need for accurate ridership and revenue analysis.

It is important to note this limitation of the fare elasticity and of the overall influence of fares on ridership. During the two-year period 1977-1979, fares were being increased in transit systems across the country while ridership was also increasing. Many people, in fact, felt the fare increases had little or no impact on ridership. In the last two years, however, ridership has fallen dramatically as fares continued to increase. Obviously other factors, such as the level of employment and gasoline prices and supply, played important roles in influencing ridership. Thus, there is no reason to suspect that the ridership response due to the fare increase alone differed during these two periods. For an excellent summary and guide in understanding and interpreting fare elasticity information, the reader is referred to a recent working paper by Kemp (13).

It is also important to recall that given the same conditions, ridership response will be different in different cities, for different transit services and levels of service, for different periods of the day, for different trip lengths, and perhaps for different fare levels. All of this suggests that a single fare elasticity value is of little use to most transit companies if accurate ridership and revenue forecasts are required. The fare elasticity will not only differ by user group (e.g., commuter), but it may also change within the same ridership group as other factors change; that is, the fare elasticity is most likely not constant as is often alleged in much of the modeling work done.

Finally, as pointed out by Kemp (13), fare elasticities alone should not be used for forecasting ridership and revenues. Instead, revenue or ridership models should be developed that incorporate the fare variable and its elasticity as well as the other variables that influence ridership. The elasticities used in these models should in fact be derived from these models if sufficient data are available. If not, elasticities, especially elasticities for specific user groups or time periods, can be borrowed from other systems for this purpose if attention is paid to the values selected (14). Although usually enough data are available at most transit companies to develop simple models and compute elasticities based on aggregate ridership data, such information is of little use in serious ridership and revenue forecasting since most transit companies have multiple fare categories. Thus, aggregate elasticities computed from an analysis of fare changes occurring within one's own transit system can be modified for initial modeling purposes by using the relative values of fare elasticities observed in other locations. Moreover, with the advent of the microcomputer, there is little reason why most transit companies cannot begin to develop data-base monitoring systems and ridership models that are dynamic, that is, models wherein the elasticities and other parameters can be recalibrated as conditions change or better information is provided.

FARE ELASTICITY IN DEVELOPING FARE POLICY

Although the fare elasticity plays a limited but nevertheless important role in ridership and revenue planning, it does provide us with very useful information on how responsive riders are to fare changes. Since many ridership groups respond differently faced with the same fare change, the fare elasticity becomes a useful measure for comparison purposes. For example, the fare elasticities computed in one transit system can be compared with those observed in other systems for the purpose of gaining credibility of one's modeling results. Similarly, fare elasticities can be used to compare recent fare adjustments with fare changes that occurred in the past. Most importantly, however, the fare elasticity can be the measure used to compare different user groups or markets within the same transit system. This last role is perhaps the most important role the fare elasticity can play since it can and should influence fare policy.

Fare policy in the broadest sense refers to the level of fare charged and how that fare varies by distance, time of day, user group, or other classification. What fare structure to adopt and how high the fares should be are difficult decisions that must be made in consideration of the specific objectives of the transit company, the levels of service provided and their costs, the political and subsidy constraints, and the characteristics of the transit system and its users. The fare elasticity enters into the decisionmaking process since it tells us something about the characteristics of the users. It is in this way that the fare elasticity can influence or guide fare policy.

Setting Fare Level

The reactions of many citizens to the large and frequent fare increases that have taken place in many cities across the country have caused many managers and administrators to question whether the actions being taken are the best, given the objective of raising needed revenues. If we had raised the fare higher, what would have been the ridership and revenue impacts? Have we reached the point where future fare increases will not lead to increases in revenues? Should we be implementing small, frequent fare increases or should we hit the public with one large fare increase every two years? These are important and legitimate questions of policy for which there are economically and politically correct answers. Unfortunately, there is no research completed or currently being performed that directly addresses these issues. There are, however, some guidelines that can be presented based on theory and a few empirical findings related to the fare elasticity of demand.

How High Can We Raise Fares?

Fare elasticities estimated in the past have usually been based on relatively small fare changes, both in percentage and absolute terms. For example, the highest fare level for an urban bus system evaluated by Mayworm, Lago, and McEnroe (3) was \$0.65 for New York City. Today, this is a typical fare level and many cities charge \$1.00 or more for a one-way transit trip. Consequently, the very low fare elasticities computed in the past may not be appropriate for the fare levels being implemented today. What is being suggested is that the fare elasticity may increase with the level of the fare. In fact, many analysts have argued that the higher the fare level, the greater the ridership response to subsequent fare changes.

To date, however, the empirical work on fare elasticities has not found evidence to support this view. Dygert, Holec, and Hill (15) reviewed the data reported by the American Public Transit Association and concluded that the magnitude of the average fare before the fare increase did not affect the size of the fare elasticity. Bly (16) also could find no significant relationship between fare level and size of the fare elasticity.

There is, nevertheless, theoretical support for this hypothesis since the elasticities derived from many models are themselves functions of the fare level and other variables in the model. The findings of several demand analysts of the London Transport (17) suggest that the demand for transit is nonlinear with respect to fares. In all these cases, the fare elasticity rises with fare level. What, then, is the revenue-maximizing fare level? Where is the point at which further increases in fares will lose so many riders that the net result will not be an increase in revenues? This point is reached when the fare elasticity reaches -1.0.

Following the formulation of their model of transit demand based on the TANK base-fare increase from \$0.25 to \$0.40, Knudson and Kemp (9) computed the level of the fare that would have maximized gross revenues in 1976, holding all the remaining variables constant. The model predicted a fare on the order of \$1.30 to \$1.35 compared with the \$0.40 fare that riders were paying.

Finally, the concept of generalized cost suggests that the fare elasticity will increase as the fare proportion of the total travel cost increases (assuming the generalized cost elasticity remains constant).

Mathematically, the fare elasticity is related as follows:

$$\eta_f = (F/GC)\,\eta_{GC} \qquad (1)$$

where

$$GC = F + \sum_i v_i t_i$$

and

η_f = fare elasticity of demand,
F = fare level,
GC = generalized cost,
η_{GC} = generalized cost elasticity,
v_i = value of time associated with time component i, and
t_i = time spent during trip in time component i.

Thus, if our current fare is \$0.40, which represents 30 percent of total generalized cost [see the report by Oldfield (18)], and the initial fare elasticity of demand is -0.40, then the fare would have to rise to \$2.82 in order to reach the point where further fare increases would result in revenue losses (i.e., where $\eta_f = -1.0$). The relationship between fare level and fare elasticity for this example is presented in Figure 1.

The actual value of the revenue-maximizing fare will change significantly as the assumptions change. For example, if the value of time in transit is greater than originally assumed, then the generalized cost will be larger, all other factors remaining constant. Thus, by using the same assumptions identified above but with a larger value of time component and a generalized cost elasticity of -2.0, the revenue-maximizing fare would be \$1.60, a much smaller figure. However, even at this fare level, off-peak and short-distance travel would disappear. If higher fares are going to be charged, then distance-based or time-of-day fare structures should be adopted.

Although there is no evidence to suggest that any transit company in this country has reached the point where the fare elasticity is equal to unity nor is it possible to tell where that point may lie, it is clear that the economic limits to raising fares are beyond the political limits, except perhaps for very long-distance trips where the fare elasticities and fares are already very high. In Chicago, for example, I am told that we are already beginning to see many commuters switching from commuter rail lines to less expensive private paratransit and subscription-bus operations.

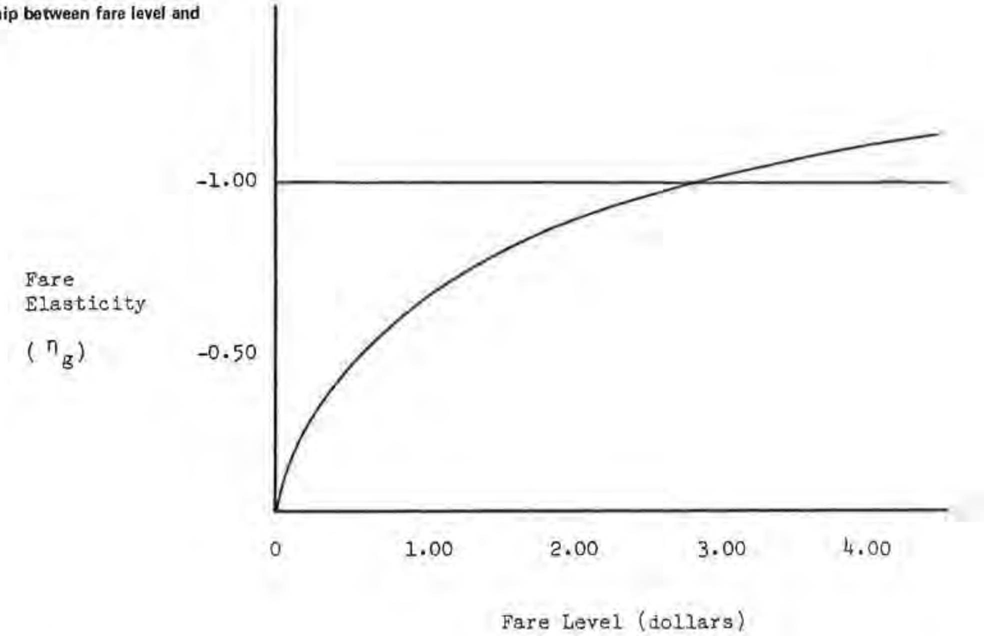
How Frequently Should We Raise Fares?

Another issue facing many transit managers concerns the size and frequency of fare changes. During the 1970s, many transit companies did not raise fares for five or six years; most, in fact, reduced their fares. Today, however, inflation is taking a heavier toll on costs, and the growth in deficits is not being offset by the subsidies provided. Since the fare level is being reviewed more frequently to fill this gap, is there any evidence that would favor semiannual, annual, or biannual fare reviews? Do riders respond differently to frequent small fare increases than to infrequent large fare increases?

Many analysts again argue that the greater the fare increase, the greater the rate of decline in transit riding. As in the case described above, the fare elasticity should increase as the fare increases to become a large portion of the total generalized cost. Since the value of time should increase over time with inflation, small fare increases that keep up with the value of time components should result in no significant change in the fare elasticity over time. Similarly, Bly (16) contends that even the large but infrequent fare changes should not affect the elasticity over time since most large fare increases are only imposed when the initial fare has become a relatively small fraction of the generalized cost. Thus, the fare is perhaps the same proportion of total generalized cost when viewed over entire periods of constant fares.

Very frequent fare changes, however, should be

Figure 1. Relationship between fare level and fare elasticity.



avoided, since most of the evidence on fare elasticities suggests that it takes six to nine months to feel the full effect of a fare change. In general, transit riders will not change their travel habits in the very short term. Thus, fare elasticities can be expected to increase slightly over this six- to nine-month period. Hensher and Paterson showed in a 1972 study (19) that the elasticity for work trips by rail in Sydney, Australia, jumped from -0.30 three months after the fare change to -0.60 six months after the fare change.

In the analyses performed by Knudson and Kemp (8,9) and by Kemp (10), the fare elasticities differed only slightly over time; most of the final ridership response occurred within six to nine months of the fare change. In the case of the 1976 fare increase in Erie, Pennsylvania, however, the analysts found that the fare elasticity was smaller in the long term. They account for this fact by suggesting that some of the ridership lost as a result of the initial impact of the fare increase may have been recaptured. Results of all three studies are presented below (8-10):

Months After Fare Change	Atlanta, GA (1972)	Northern Kentucky (1976)	Erie, PA (1976)
3	-0.16	-0.08	-0.37
6	-0.17	-0.11	-0.32
9	-	-0.12	-0.32
12	-0.18	-0.12	-0.32

Although many factors will affect the decision on when to implement a fare change, this short analysis suggests that time is not a major factor. Fares should not be changed more frequently than every six months, since it takes at least that long for the impact of the previous fare change to be realized. Transit managers should also rule out biannual or very long periods between changes since inflation is so high today. Moreover, most transit riders understand that fares have to keep up with inflation and would accept an annual fare review and adjustment period.

Designing Fare Structure

Perhaps the most important role the fare elasticity can play is with regard to developing and designing the fare structure of a transit system. It is important and useful because it indicates the way and the degree to which we should differentially price our services. Although we should be setting different prices for different services based on the costs of providing the service, by comparing fare elasticities of demand, we can determine whether we are taking full advantage of what individuals are willing to pay.

Designing fare structures is essentially a task of determining the degree to which we should differentially price. For economic efficiency arguments, this essentially involves equalizing the fare elasticities of demand for the specific markets in question. Since differential pricing results in higher revenues with no net loss in ridership, a transit manager should weigh these higher revenues against the costs incurred in creating a zone system, peak-period surcharge, or whatever differential pricing scheme is under consideration. Thus, fare elasticities are very useful because they tell us how much we can expect to gain.

There are many different types of fare structures that can be designed and each transit system will have its own combination. However, the three most common forms of differential pricing are distance-based fares, time-of-day fares, and quality-based fares. These methods are the most common forms of differential pricing for three reasons. First, the cost of providing transit service differs significantly for short and long trips, by time of day, and by service quality. Second, these forms of pricing are relatively easy to administer. Finally, the fare elasticities differ significantly within each group so that we can charge riders higher fares for long-distance, peak-period, and express service without affecting overall ridership levels.

Revenue generation is the main purpose for differentially pricing transit service, and revenues will only be generated if there are significant differences in the fare elasticities. Express-bus riders are known to place a higher priority on travel time, safety, and comfort than on the fares paid. Thus, the fare elasticities calculated for express-bus and local service users should be significantly different and different fares can be charged.

Table 1. Elements in fare structure trade-off analysis.

Fare Structure	Revenue Efficiency	Equity	Passenger Convenience	Effect on Costs: Collection Costs and Complexity	Effect on Costs: Effect on Operating Costs and Boarding Times	Fraud Avoidance
Flat	Poor	Poor	Excellent	Excellent	Excellent	Excellent
Time of day	Excellent	Good	Good	Good	Good	Good
Distance based	Excellent	Excellent	Poor	Poor	Poor	Poor
Quality based	Excellent	Fair	Good	Good	Excellent	Excellent

Similarly, peak-period riders are known to be less responsive to fare changes than off-peak riders. Although there is little evidence to suggest that peak-period demand will shift to the less expensive off-peak period by differentially pricing peak and off-peak service, such a pricing scheme can lead to higher revenues with no net loss in ridership.

Perhaps the most important differential pricing option open to a transit company concerns how the service is priced by trip distance. When British Rail fares were examined during the 1970s, Oldfield and Tyler (5) found that the fare elasticities for full-fare and pass riders did not vary by trip length. Since fares are graduated on the suburban London service, the fact that there was no systematic variation in the fare elasticity may suggest that the fare portion of the users' total generalized cost increases proportionally with distance. They did find, however, that the fare elasticities for reduced-fare riders increased with distance, which suggests that the fares are perhaps increasing too rapidly with distance for this group. Since the rail lines included in the study extend as far as 72 miles from London, many reduced-fare riders can obviously find alternative destinations for their trips.

In a recent study on intercity bus demand, Burkhardt and Riese (20) found that the fare elasticity decreased with distance as shown below:

Route Length (miles)	Fare Elasticity
20-60	-0.645
20-120	-0.352
120+	-0.268

Since the average fare paid per mile did not differ significantly for each group, the longer-distance routes perhaps provided service for individuals for whom alternative modes of travel are not available.

For any type of transit system, the choice and appropriateness of some form of distance-based fares will depend on the distribution of trip lengths and the variations of fare elasticities by trip length. Small transit systems without major differences in fare elasticities by trip distance should opt for flat fares. Under these conditions, the gains in revenue as well as equity [see report by Cervero and others (21)] are simply not worth the extra cost and inconvenience of distance-based fares. Experiments conducted on London's suburban routes in Harrow and Havering as reported by Richardson and Fairhurst (22) and by Fairhurst (23) concerning conversion back to flat fares resulted in both greater revenues and passenger miles of travel. These experiments showed that complex fare systems create opportunities for fraud and that if the differential between fare elasticities for short and long trips is not large, then flat-fare systems are relatively efficient. Based partly on these experiments, London Transport is converting to flat fares in some of its suburban bus systems. However, this scenario is by no means representative of all American settings where significant differences in fare elasticities exist.

The role of the fare elasticity in fare structure design is to provide information on the revenue-generation potential of alternative schemes so that important trade-offs can be made between revenue generation and equity on one side and convenience and cost on the other (Table 1). In large systems where peak-period travel is important and fare elasticities vary by trip length, distance-based and time-of-day fare structures are probably superior to flat fares in terms of revenue generation. Whether this advantage in revenue generation as well as equity is worth the increased administrative cost will be answered in the immediate future as the financial pressure on transit increases and when the revenue-generation potential of alternative fare structures becomes of paramount importance.

CONCLUSIONS

With subsidy shortfalls predicted for the next few years, transit managers across the country are looking to the farebox to raise the required revenues to keep their services operating. The political sensitivity of fares as a means of boosting operating revenues has caused many managers to require their staffs to provide more accurate ridership and revenue projections and to present a wider range of choices on different fare structures that are designed to increase revenues. This paper has presented a discussion of the fare elasticity of demand in terms of its role in ridership and revenue planning and in fare structure design.

Although the fare elasticity of demand is a useful concept because it provides information on how riders respond to fare changes, ridership and revenue planning must acknowledge the myriad factors that affect patronage, of which fare is only one. The fare elasticity, however, is an extremely useful measure that can provide information and guidance in developing fare policy so that we can begin to capture the farebox revenues we need with minimal effect on ridership levels. As noted in this paper and in other studies (3,4), fare elasticities often vary significantly by trip distance, time of day, and quality of service. If we are going to take advantage of the increased revenue and ridership opportunities afforded by the differences in fare elasticities across transit markets, the reliance on flat fares will have to be abandoned and more attention will have to be placed on how we price transit services.

REFERENCES

1. Transit Fact Book. American Public Transit Association, Washington, DC, 1981.
2. A. Grey. Urban Fares Policy. D.C. Heath, Lexington, MA, 1975.

3. P.D. Mayworm, A.M. Lago, and J. M. McEnroe. Patronage Impacts of Changes in Transit Fares and Services. Ecosometrics, Inc., Bethesda, MD, 1980.
4. The Demand for Public Transport: Report of the International Collaborative Study of the Factors Affecting Public Transport Patronage. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1980.
5. R.H. Oldfield and E. Tyler. The Elasticity of Medium-Distance Rail Travel. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Rept. LR 993, 1981.
6. D.C. Stark. Time Series Analysis of Glasgow Suburban Rail Patronage. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Rept. SR 649, 1981.
7. G.B. Urquhart and C.M. Buchanan. The Elasticity of Passenger Demand for Bus Services: A Case Study Telford. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Rept. SR 641, 1981.
8. B. Knudson and M.A. Kemp. The Effects of a 1976 Bus Fare Increase in Erie, Pennsylvania. Urban Institute, Washington, DC, Working Paper 1428-01, April 1980.
9. B. Knudson and M.A. Kemp. The Effects of a 1976 Bus Fare Increase in the Kentucky Suburbs of Cincinnati. Urban Institute, Washington, DC, Working Paper 1428-02, May 1980.
10. M.A. Kemp. Transit Improvements in Atlanta--The Effects of Fare and Service Changes. Urban Institute, Washington, DC, UI Rept. 1212-2-1, June 1974.
11. K.M. Goodman, M.A. Green, and M.E. Beesley. The San Diego Transit Corporation: The Impacts of Fare and Service Changes in Ridership and Deficits, 1972-1975. Urban Institute, Washington, DC, UI Rept. 5066-5-1, May 1977.
12. C. Ulberg. Short-Term Ridership-Projection Model. TRB, Transportation Research Record 854, 1982, pp. 12-16.
13. M.A. Kemp. Planning for Fare Changes: A Guide to Interpreting and Using Fare Elasticity Information for Transit Planners. Urban Institute, Washington, DC, Working Paper 1428-05, Dec. 1980.
14. A.M. Lago and P.D. Mayworm. Transit Fare Elasticities by Fare Structure Elements and Ridership Submarkets. Transit Journal, Vol. 7, No. 2, Spring 1981.
15. P. Dygert, J. Holec, and D. Hill. Public Transportation Fare Policy. Peat, Marwick, Mitchell and Co., Washington, DC, 1977.
16. P.H. Bly. The Effect of Fares on Bus Patronage. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Rept. LR 733, 1976.
17. M.H. Fairhurst and P.J. Morris. Variations in the Demand for Bus and Rail Travel up to 1974. London Transport Executive, London, United Kingdom, Economic Res. Rept. R210, April 1974.
18. R.H. Oldfield. Elasticities of Demand For Travel. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, TRRL Rept. SR 116 UC, 1974.
19. D.A. Hensher and J. Paterson. Price Elasticity of Demand for Urban Public Transport. Paterson Urban Systems, Melbourne, Australia, 1972.
20. J.E. Burkhardt and J.I. Riese. Estimating Travel Demands for Intercity Bus Routes. Presented at the 61st Annual Meeting, TRB, 1982.
21. R.B. Cervero, M. Wachs, R. Berlin, and R.J. Gephart. Efficiency and Equity Implications of Alternative Transit Fare Policies. UMTA, 1980.
22. A.P.S. Richardson and M.H. Fairhurst. The Harrow and Havering Flat Fare Schemes: Passenger Demand Assessment. London Transport Executive, London, United Kingdom, Economic Research Memorandum M 386, Aug. 1980.
23. M.H. Fairhurst. Why Simplify? A Case for Simplified Fares. London Transport Executive, London, United Kingdom, Economic Research Rept. R244, Jan. 1981.

Scheduling-Based Marginal Cost-Estimating Procedure

WALTER CHERWONY AND BENJAMIN PORTER

With changing policies regarding transit funding at all levels of government, transit planners will be required to more carefully monitor existing bus systems as well as intensively examine the net cost or savings of proposed service changes. In the past, research has focused on only one side of the equation—demand, hence revenue estimation. In the near future, more effort will be directed to operating-cost estimation and the underlying relationships that impact expenditures. Although a variety of cost-estimation techniques have been developed, there is little agreement as to which one best estimates cost. The purpose of this project was to develop a technique that is complex enough to capture the salient cost characteristics of a change in transit service. The cost model presented here is sensitive only to those line items that typically vary in response to changes in the scale or characteristics of fixed-route service. These are termed variable costs. A major variable cost component is driver cost, which is treated by the model in some detail. Driver cost is assumed to be a function of the number of drivers required to operate scheduled service, along with exceptions that normally occur in daily operations. These perturbations are captured through simulations of scheduling and dispatching processes. These are described as a set of calibrated ratios and percentages that assume no dramatic departure from the norm. Other variable costs (e.g., fuel and insurance) are estimated through a typical cost-allocation approach. This model is currently being tested along with several other prominent costing approaches. A variety of small service changes are being used as the basis for comparison. No results on the models' comparative performances are available at this time.

The current decade will represent a period of dramatic change for most transit agencies as they respond to an era of limited resources. Many systems, facing severe financial constraints, have already made substantial service changes to balance transit costs with available funds. This new direction in the transit industry will place greater demands on transit planners to forecast, with reasonable accuracy, the financial implications of service changes. Unfortunately, no single technique or procedure has been established that transit planners can readily use. Recognizing this deficiency, the Urban Mass Transportation Administration (UMTA) has commissioned a research effort to develop a bus-route costing procedure.

This study has consisted of several interrelated steps. The initial step was to review techniques now used in the industry as well as procedures identified in the technical literature. Following an assessment of these procedures and the requirements of transit planners, a proposed method was designed. The current work element is the testing of the proposed procedure and other prominent models by using actual and hypothetical service changes in the Minneapolis-St. Paul (Twin Cities) urban area. Based on the test results, the proposed method will be modified as appropriate. The concluding step will be to document the proposed costing technique and prepare training materials to encourage its use throughout the transit industry.

MODEL OVERVIEW

The primary objective of the bus-costing procedure is to design a marginal cost model capable of estimating the operating-cost impacts of service changes. A review of available procedures indicates a wide range of techniques and capabilities (1). Simple procedures (e.g., average cost per mile or hour), although easily applied, are usually too coarse and insensitive to produce accurate results. More elaborate and sophisticated techniques that specifically address driver work assignments and benefit costs are normally too time-consuming for most planning applications. For this reason, the proposed model must strike a careful balance between simplicity and sensitivity to factors that influence bus operating costs.

One common feature of many cost analyses is the use of a cost-allocation model approach. Typically, the cost of providing transit services is related to different transit resource levels (e.g., vehicle hours, miles, and peak vehicles). With this approach, each expense item is assigned or allocated to a particular resource. This allocation process is normally logic-driven, although some statistical analysis has been performed to demonstrate the cost-resource relationships. For example, expenditures for tires, tubes, and fuel are logically a function of vehicle miles. An illustration of the cost-allocation approach is shown below based on the development and calculation of this procedure for the New York City Transit Authority surface operations (2):

Allocation Basis	Allocated Costs ($)	Percentage of Total Costs	Operating Statistic
Vehicle hours (H)	147 196	67.1	7 786 600
Vehicle miles (M)	54 006	24.6	66 108 000
Peak vehicles (V)	18 047	8.2	2 084
Total	219 249		

Cost = 18.90 * H + 0.82 * M + 8659.62 * V.

This three-variable model, although a commonly employed technique, overstates the cost of service changes when it includes fixed, or overhead, costs. An enhancement of the cost-allocation model is to segregate line-item costs into those affected by varying scales of service change. Most of the cost analysis performed in Great Britain, for instance, distinguishes between fixed and variable expenditures (3). Recent research efforts have been oriented to developing cost models that focus intensively on driver wages and benefits--the largest single cost of providing bus service. Although approaches differ, the intent is to incorporate procedures or variables that accurately reflect driver assignments and their associated costs. In view of the complex work rules and arrangements governing the computation of driver wages and benefits, this is no simple task.

Two interesting and contrasting approaches have been developed to gauge driver costs of service changes. The first technique, a peak-base model allocation, was developed as part of the I-35W Urban Corridor Demonstration Project to test the impact of freeway-ramp metering and expanded express-bus service (4). Within the overall framework of a cost-allocation model, the researchers quantified separate vehicle-hour unit-cost factors for peak and off-peak service. With this approach, two indices are computed, which are then used to adjust the standard cost model for both service periods. The first index (labor productivity) measures the relative ratio of pay hours and vehicle hours for peak- and off-peak service. The second index (service) measures the extent of peaking by comparing vehicle hours for both time periods. Both indices are developed by a calibration process in which recent payroll and operating data are used. The result of this process is the computation of two coefficients that adjust a systemwide cost per vehicle hour and that reflect peak and base differentials. The attractive feature of this approach is its simplicity and ease of use. This is accomplished at the expense of sensitivity to many diverse factors such as driver types, work assignments, and specific labor provisions. While the calibration measures the overall impact of these factors, it does not deal individually with each.

An alternative approach developed in England (5) and subsequently applied in Australia (6) differs markedly from the above model. It is somewhat more complex in that driver costs are estimated in a bottom-up approach rather than through cost allocation. One attractive feature of this work is the formulation of a driver-scheduling model that transcribes buses in service by time period into driver work assignments. In turn, various assignments are costed with respect to driver wages. The research also provides considerable insight as to how service is scheduled and the resulting cost implications by time of day and day of the week.

Although the two approaches briefly described above are only a sample drawn from the literature, they are instructive in that they contrast two different solutions to deal with the same problem. For this reason, the suggested approach attempts to incorporate attractive features of each without introducing undue complexity.

MODEL FRAMEWORK

The review of commonly employed costing methods yielded three guiding principles for the proposed model's development. First, only variable costs should be included. For many service changes, there is relatively little overhead, or fixed, cost incurred that would not have ordinarily been realized. Examples of fixed costs include administrative salaries, building maintenance, and some operations salaries such as that for the transportation manager. Second, driver cost must be computed with respect to the temporal service distribution. That is, it should be scheduling-based to the extent possible with a nonautomated approach. Third, nondriver costs should be estimated via cost allocation. Generally, these costs are affected only by service scale (e.g., net vehicle miles) and are not so sensitive to temporal characteristics as is driver cost.

The resulting model adheres to these principles. All cost categories included and excluded by the model were defined on the basis of the Section 15 accounting structure. Each function-code/object-code combination was used to identify all expense accounts. Consideration of the factors influencing a particular expense account plus the magnitude of expenditures was used in developing the suggested approach. The overall approach was to develop a

Table 1. Operating cost by major category.

Item	Amount ($)	Total Operating Costs (%)	Variable Operating Costs (%)
Variable costs			
Estimated by special analysis	32 946 460	51	60
Estimated by cost allocation			
Miles	19 179 677	29	35
Hours	2 419 229	4	5
Total	21 598 906	33	40
Total variable costs	54 545 366	84	
Fixed costs	10 636 972	16	-
Total operating costs	65 182 338		-

conventional cost-allocation model with only certain expenses subjected to special analysis. For this reason, only driver wages and fringe benefits are subject to special analysis. All other variable expenses relied on a traditional cost-allocation model. In the suggested procedure, only two variables are used--vehicle hours and miles. Peak vehicles was deleted as a variable since expenses allocated to this resource level are often fixed costs.

Initially, other expense items were considered for special analysis, particularly where variables other than miles or hours would affect variation in cost. For example, fuel economy and costs are a function of operating speed. From a research perspective, it would appear beneficial to develop a costing technique that could distinguish between local, central-business-district shuttle, and express-route service changes. However, it was felt that the additional effort to develop and apply a model with this capability would be burdensome. Further, the potential cost sensitivity would be relatively limited in comparison with other cost items (e.g., driver wages) and total system costs.

The classification of transit-system expenses, as explained above, demonstrates the logic of the proposed model. By using the Metropolitan Transit Commission (Twin Cities) as an example, as shown in Table 1, it can be seen that one-half of total operating expenses and 60 percent of variable costs are attributed to driver wages and fringe benefits. This alone demonstrates the need for detailed scrutiny. The remaining variable costs can be allocated to either hours or miles; the resulting cost model is

$$C = 1.13 * H + 0.63 * M + \text{(special analysis)}.$$

DRIVER COST ESTIMATION

Driver cost is composed of two major components--wages and benefits. Both these components encompass a number of discrete cost categories, many of which are influenced by dissimilar causal (i.e., independent) variables. There are two problems with producing a reasonably accurate driver cost estimate. The first is in aggregating these categories into groups the cost of which is directly tied to a single causal variable. A second problem is in calibrating a value for that causal variable. For many cost categories, the causal variable is a product of the run-cutting process. Scheduled wages, for instance, are largely determined by the number of runs in a schedule. Because of the many decisions, both objective and subjective, encountered in the scheduling process, it is difficult to predict the results.

Most cost models developed to date avoid these problems by simplifying their approach. The most common solution is to use an average cost per platform hour to estimate driver wages and benefits. This is not without statistical justification because of the strong linear relationship existing between driver cost and platform hours. Some models go a step further by adjusting the cost per hour for differences in efficiency between peak and base periods. This approach is somewhat simpler than apportioning wage and benefit cost but is less complete.

Because there are so many variables affecting driver cost, there is some degree of noise in any estimation technique. Due to uncertainties in the creation and dispatching of driver assignments, it is impossible to estimate cost exactly. However, when one can identify the components of wages and benefits and then with reasonable accuracy and ease estimate the coefficients for the variables explaining their cost, it follows that the accuracy and resiliency of the resulting cost estimate are improved.

The technique described below is an attempt to incorporate the effects of scheduling and dispatching practices on the cost of a change in service. Its development proceeded from first identifying the variables that "drive" wage and benefit cost and then defining a process to estimate these independent variables' values. Cost estimation is then a simple process of applying these variables in a given formula.

Components of Driver Cost

Driver wages and benefits are merely the total of a number of identifiable cost categories. The key to accurately estimating their cost while retaining some degree of simplicity is to determine which categories can be aggregated without sacrificing the quality of the estimate.

Driver Wages

Wages represent about 70 percent of total driver cost. There are many categories of wages at any given transit agency, which relate to specific labor agreement clauses. However, because nearly all full-time drivers receive an eight-hour daily guarantee, wages fall neatly into two classifications--those paid as part of the guarantee and those paid exclusive of the guarantee. Any distinction between pay categories that contributes to the eight-hour guarantee can be ignored. It can then be assumed that each full-time driver who is working on a given day will receive at least eight hours of pay.

Wages paid exclusive of or in addition to the guarantee are generally composed of overtime and spread premiums. The latter is a premium paid when elapsed time at work exceeds some limit (e.g., 10.5 h). Either category can be contained in the schedule or can result from the way work is dispatched. For instance, an extraboard operator working short a.m. and p.m. assignments may receive a spread premium that was not contained in the schedule.

Driver Benefits

Benefits represent about 30 percent of driver cost. Like wages, benefits are composed of many categories, but these can be aggregated into three classifications. First, paid leave accounts for a substantial portion of benefit cost. This includes vacations, holidays, and sick leave and is a subset of total absences. Paid leave is a function of the number of drivers. Second, variable benefits is that portion of benefit cost that is relative to wages earned. These include Social Security and employer pension contributions. Variable benefits are usually paid as a percent of total wages and most

Figure 1. Use of cost model components.

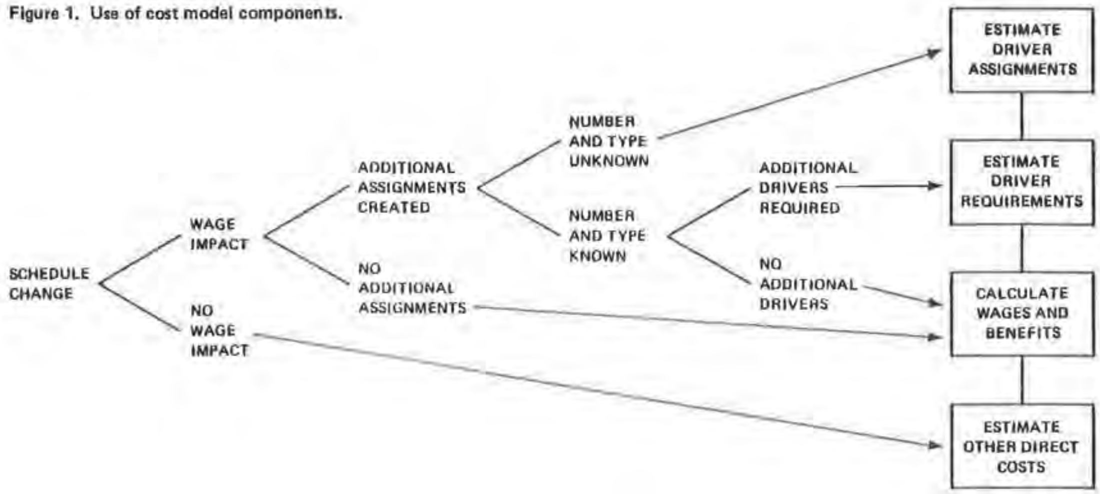
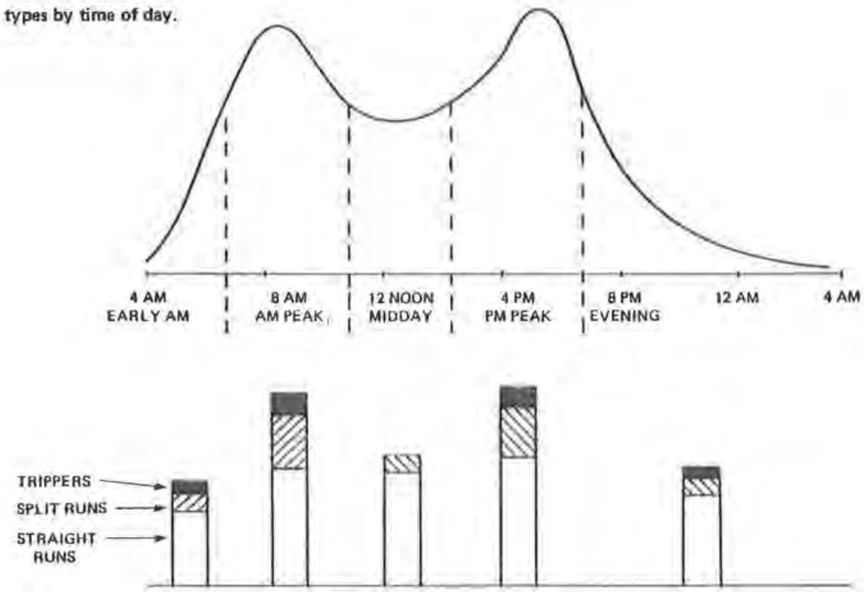


Figure 2. Assignment types by time of day.



paid absences. Third, fixed benefits include health plans, insurance plans, and uniforms. These are typically paid on a per-employee premium basis.

Model Description

Every service-planning decision has one of three driver cost impacts. First, the service change may fit within existing driver assignments so that no additional pay is required. An example of this is reducing layover time. Second, the service change may necessitate more or fewer driver assignments but may not require additional drivers. For instance, supplemental peak service could be added that is absorbed at overtime or with existing drivers. Third, the service change may require more or fewer drivers.

The structure of this cost model contains four components that relate to these potential outcomes (see Figure 1). The model is designed to conform to the level of information known about a proposed service change. If a rather large service change were proposed, for instance, the user may want to employ the full model. On the other hand, if a small change were being made and the run-cut impact known (e.g., deleting a tripper), the user may need only the wage and benefit components. Each of the components is explained more fully below.

Driver Assignments

This component uses existing run-cut information to project the number and type of driver assignments and scheduled premium hours existing after a service change. Four types of assignments (i.e., runs) are produced--straight runs, split runs, a.m. trippers, and p.m. trippers. In addition, the number of scheduled overtime and spread premium hours is forecast.

Runs are estimated by calculating the total platform hours allocated to a run type (e.g., straight run) for the day and dividing by that type's average platform time. The total platform hours allocated to one type of run are calibrated by examining the existing driver assignment data from the applicable operating base (i.e., depot, garage). The calibration calculates the proportion of hours allocated to one type of run for each of five periods in the day (see Figure 2).

Trippers are further defined based on the way in which they are usually assigned. The model is calibrated for the proportion of a.m. and p.m. trippers normally allocated to part-time drivers (if applicable), assigned at overtime, or paired and assigned to the extraboard. Pay hours for both scheduled and unscheduled overtime and spread premium are calculated based on the average hours per type of assignment.

Driver Requirements

This component estimates the number of drivers working on a given day and the total number of drivers required for a week's schedule. A by-product of this process is the average number of absences.

A driver use ratio is the basis for determining driver requirements. It is computed to establish a relationship between the number of drivers required on a given day and the number of full-time assignments to be filled. In this case, full-time driver assignments include straight and split runs as well as tripper combinations assigned to the extraboard. Daily working drivers are calculated as the difference between total driver requirements and average absences. Weekly driver requirements are determined by summing the daily driver requirements for a week's time and then dividing by 5.

Driver Wages

Once daily premium pay hours and daily working drivers have been defined, driver wages are easily calculated with the following formula:

$$\sum_{i=1}^{3} [(D_i * 8) + P_i] * S_i * W \qquad (1)$$

where

- i = type of schedule (weekday, Saturday, Sunday),
- D = working drivers,
- P = premium hours,
- S = days of operation, and
- W = weighted average wage.

Driver Benefits

The costs of three benefit categories are calculated as shown below:

$$\text{Paid leave} = \left\{\left[\sum_{i=1}^{3} (A_i * PA)\right] + (TD * H)\right\} * 8 * W \qquad (2)$$

$$\text{Variable benefits} = (GW + PL) * VR \qquad (3)$$

$$\text{Fixed benefits} = TD * FR \qquad (4)$$

where

- A = absences,
- PA = proportion of absences paid,
- TD = weekly drivers required,
- H = scheduled and personal holidays,
- GW = gross wages,
- PL = paid leave,
- VR = variable benefit rate, and
- FR = fixed benefit rate.

MODEL APPLICATION

The formulas, calibration measures, and parameters defined above can be applied to service-change data to yield corresponding cost estimates. There are two basic steps--nondriver cost estimation and driver cost estimation.

Nondriver costs are estimated by applying the cost-allocation model defined earlier. This simply involves calculating the net vehicle hours and vehicle miles resulting from a proposed service change.

Driver cost estimation is somewhat different in that net costs are estimated by comparison with a baseline. This baseline is established by applying the model to existing data. For transit systems with multiple operating bases, it is advisable that this procedure be repeated for each base. Post-service-change costs are then estimated based on the entirety of service hours existing after the service change occurs. This is because driver assignments are usually created for the whole of an operating base. Therefore, examining the net change alone, particularly on a route basis, may not yield appropriate results.

CONCLUSIONS

The previous discussion has provided an overview of the need for and previous research in cost-estimating procedures. A brief description of a proposed approach has been presented. Some key conclusions from this analysis are as follows:

1. The suggested method must balance ease of use with requirements for accuracy and sensitivity.
2. A traditional cost-allocation model appears well suited to estimate all variable operating costs, excluding driver wages and benefits.
3. Driver costs associated with service changes require a two-step process--calibration and application.
4. A calibration approach can measure various indices and statistics that influence wages and benefits.
5. The model will not accurately respond to conditions that could produce significantly different run-cut results.
6. The various calibration measures are organized in such a manner and related to causal factors to permit a modular approach in which part or all of the model can be used.
7. The proposed approach is sufficiently flexible to be applied to any unique site-specific situation.
8. Because a calibration approach is used and applied to future service changes, the model does not optimize but rather reflects continuation of previous practices.

REFERENCES

1. W. Cherwony, B.D. Porter, S.R. Mundle, and G.R. Gleichman. Review of Bus Costing Procedures. TRB, Transportation Research Record 854, 1982, pp. 54-60.
2. W. Cherwony and S.R. Mundle. Consequences of Express Bus Service on Transit System's Operating and Financial Performance--New York Case Study. In Compendium of Technical Papers, Institute of Transportation Engineers, Arlington, VA, 1980.
3. Symposium on the Costing of Bus Operations. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, Supplementary Rept. 180UC, 1975.
4. W. Cherwony and S.R. Mundle. Peak-Base Cost Allocation Models. TRB, Transportation Research Record 663, 1978, pp. 52-56.
5. R. Travers Morgan and Partners. Bradford Bus Study: Final Report. West Yorkshire Metropolitan County Council and West Yorkshire Passenger Transport Executive, England, 1976.
6. R. Travers Morgan and Partners. Adelaide Bus Costing Study. Director General of Transport, Adelaide, South Australia, Final Rept., 1978.

Development and Testing of a Cost-Allocation-Based Cost-Estimating Method

ROBERT L. PESKIN

The status of the development of a transit operating cost model for the Financial Analysis Portion of the Corridor Refinement Study currently being undertaken for the Metropolitan Transit Authority of Harris County, the public transit operator in the Houston, Texas, region, is described. The approach used to develop the cost model and a recent application in financial analysis in Houston are detailed.

The development of a transit operating cost model is part of a continuing transit planning effort following the original Transitway Alternatives Analysis Study begun in 1979 for the Metropolitan Transit Authority (MTA) of Harris County, the public transit operator in the Houston, Texas, region. To date, the study has identified a priority corridor and has begun preliminary engineering and financing for rail rapid transit and busways in that corridor. In addition, a regionwide program of bus service expansion, feeder bus routes, and development of other busways with the Texas State Department of Highways and Public Transportation is being considered.

The financial analysis will integrate all estimates--operating and capital expenses as well as projected operating revenues and local sales tax, state, and federal funding. The operating cost model will project the costs for all expense portions of the operating budget--wages, salaries, fringe benefits, parts, diesel fuel, electricity, claims, insurance, and taxes.

The cost model is structured around a carefully defined set of assumptions regarding the transit technologies used and institutional and administrative considerations. The fundamental consideration was that both bus and rail operations and maintenance activities would be handled by MTA.

The cost model is structured in such a way that, once annual bus and rail operating statistics have been determined, the annual costs can be quickly computed. The primary input to the bus and rail models, traditionally developed in the urban transportation planning process, include peak vehicles, annual vehicle hours, and annual vehicle miles. In addition, the rail model requires descriptors of the physical characteristics of the system including stations, route miles, and yards.

The cost models are intended to be used in evaluating alternative regional bus and rail transit systems. The computerized version projects costs in both base year (1982) and inflated dollars.

APPROACH

This operating cost model was built on the original cost model developed by Peat Marwick for the Houston Transitway Alternatives Analysis (1) and later expanded and refined for the following studies: Washington Metropolitan Area Transit Authority (WMATA) FY80 Net Income Analysis (2), New Orleans Cable-Suspended Transit System Study (3), St. Louis Light Rail Study (4), and Detroit Woodward Corridor Light Rail Transit Study (5).

The cost model, built on this experience and described in this paper, adds several improvements that make the model suitable for both short-range budget analysis and long-range system planning.

1. Costs modeled by responsibility center. To structure the model with a minimum of preliminary data preparation, the fundamental administrative units modeled were the responsibility centers defined by the MTA Office of Budget Systems. Budget information was the major source for data; this proved a convenient way to arrange costs. For each responsibility center, the budget identified the costs for union wages, nonunion salaries, and major categories of direct costs. A determination was made for each responsibility center--whether the costs were fixed, related indirectly to some measure of system size, related indirectly to overall growth in service, or required modeling of specific labor categories.

2. Costs modeled for each union position. Recognizing that union labor costs account for approximately one-half of total transit costs, the cost model was structured so that each union position was modeled. The number of employees in each position was made a function of a specific measure of system size or quantity of service provided. Average annual wages were then applied.

3. Costs modeled for each front-line supervisory position. The salaries for each front-line supervisor (foreman, street supervisor, etc.) for the union positions were also separately modeled. Generally, the costs were computed for staffing levels (number of union employees for each front-line supervisor) or shift coverage.

4. Electricity costs modeled according to utility rate structure. The costs for traction, yard, station, and chiller plant electricity were modeled on the Houston Lighting and Power Company rate structure for Large General Service (LGS). The rate structure was applied assuming all connected loads were through the traction power substations.

5. Rail operations costs based on system operating plan. The designers of the rail system, Houston Transit Consultants, provided a staffing plan and organizational structure for developing cost equations for the rail operations.

The cost model is composed of a series of equations that project costs as a function of the quantity of service provided. Costs were computed for each responsibility center, union position, front-line supervisory position, and major type of direct cost. Specific costs were identified as either fixed or variable, driven by specific descriptors of service or physical characteristics of the system. Figure 1 presents the system characteristics used to drive the costs.

The cost equations were organized generally around the current MTA management structure assuming that, in the case of the rail alternatives, a separate Rail Operations Division, parallel with the Bus Operations Division, would be created. Figures 2, 3, and 4 present a list of the equations used in the cost model. Figure 5 presents an example of the worksheets used to manually compute the costs. It presents the various cost factors, their values, and data sources. Worksheets such as these were used to formulate the model, create computer code, and check the computer-derived cost projections. A complete description of the equations, including values and sources of information for each cost factor, is documented in various study working papers (6).

INFLATION CONSIDERATIONS

All of the unit cost values in the operating cost models are in 1982 dollars. They are based on either current (1982) experience or are historical unit costs inflated to 1982 dollars. These unit costs will be used in the financial analysis to project future costs. Operating costs will be presented in both future-year dollars and in 1982 dollars.

Data Resources, Inc. (DRI), prepared the following four series of inflation projections for MTA:

1. Houston consumer price index (CPI)--the "baseline" rate of inflation, against which all other inflation projections were compared. It was assumed that wages and salaries inflate at this rate (thus, there was no "real" rate of inflation for wages and salaries).
2. U.S. average CPI--the assumed rate of inflation for other direct costs and insurance. Insurance costs could inflate at a substantially higher rate as the insurance market becomes less "soft" as a result of lower interest rates.
3. Houston diesel fuel--the rate for MTA bus diesel fuel was assumed to be equal to the rate for the Houston region. The projection considered price, demand, and availability.
4. Houston electricity--the rate for HL&P industrial customers.

Figure 1. Driving variables.

BUS SYSTEM CHARACTERISTICS

- Regular Buses: standard 40-foot coaches required for peak-period schedule service. Includes current GMC "New Look," RTS-II, RTS-04, Grumman 870, and Eagle coaches.
- Articulated Buses: Sixty-foot M.A.N. -AM General articulated coaches or equivalents, required for peak period scheduled service.
- Regular Bus Surface Street Miles: annual miles travelled by regular buses on surface streets (as coded in the transit network).
- Articulated Bus Surface Street Miles: annual miles travelled by articulated buses on surface streets (as coded in the transit network.)
- Articulated Bus Guideway Miles: annual miles travelled by articulated buses on guideways (as coded in the transit network).
- Platform Hours: annual scheduled hours of service (including revenue, layover, and deadheading hours, factored on a system-wide basis).
- Operating Garages: operating bases from which scheduled buses are dispatched and where light routine maintenance and cleaning are performed.
- Sq. Feet Maintenance Facilities: floor area of offices, shops, and storerooms of operating garages and central shops.
- Park and Ride Lots: parking lots for both express buses and rail transit stations.
- Service Areas: districts defined for use in assigning street supervisors.
- Contra Flow Busways: the I-45 (North) Contra flow lane
- Gulf Freeway-Type Busways: one-way, reversible busways.
- Other Busways: priority-corridor, two-way busways. Busways from CBD to West Belt and from CBD to the North Corridor are considered separate busways.
- Total Revenue: fare box revenue for both bus and rail systems.

RAIL SYSTEM CHARACTERISTICS

- Peak Vehicles: vehicles required during peak period service.
- Base Trains: trains operated midday.
- Early/Late Trains: trains operating before the AM peak and after the PM peak.
- Total Vehicle Miles: annual vehicle-miles travelled in revenue service including deadheading.
- Route-Miles: double-track miles between station center lines on portions of the rail system in revenue service.
- Surface Stations: stations located at-grade or on elevated structures.
- Subway Stations: stations located below grade.
- Mezzanines: station entrances with a station agent and automated fare collection equipment.
- Service and Inspection (S&I) Yards: yards with storage capacity and service and inspection maintenance capability.
- Traction Substations: HL&P connection points where high voltage AC is converted to low voltage DC at MTA owned and operated facilities for use in powering trains.
- Chiller Plants: air conditioning units used to cool CBD stations and tunnels.
- Rail Passengers: annual rail passenger boardings.

Figure 2. Cost equations: bus operations.

DIRECTOR AND STAFF

- Assistant Executive Director and Staff
- Bus Operations Director and Staff
- Safety
  - Fixed
  - Quality Assurance Inspectors
  - Fluid Testing
- Labor Relations

VEHICLE MAINTENANCE

- Manager and Staff
- Garage Superintendents and Staff
  - Operating Buses
  - Central Shops
- Repairmen
  - Regular Buses
  - Articulated Buses
- Mechanical Foremen
- Service Attendants
- Custodians
- Cleaner Foremen
- Parts, Supplies, and Services
  - Regular Buses
  - Articulated Buses
- Diesel Fuel
  - Regular Buses
    - Surface Street Miles
    - Guideway Miles
  - Articulated Buses
  - Surface Street Miles
  - Guideway Miles
- Communications
  - Administration
  - Repairmen
  - Parts and Supplies
- Other Direct Costs

FACILITY MAINTENANCE

- Manager and Staff
- Building Maintenance
  - Superintendent and Staff
  - Repairmen
  - Cleaners
  - Direct Expenses
- Public Facilities
  - Administration
  - Sign and Shelter Repairmen
  - Park and Ride Lot Cleaners
  - Direct Expenses

TRANSPORTATION

- Manager and Staff
- Garage Superintendents and Staff
- Starters
- Scheduled Operators
- Extra Board Operators
- Operator Trainee Wages
- Road Operations Superintendents and Staff
- Street Supervisors
  - Base Service Areas
  - Expanded Service Areas
  - Busways

PLANNING AND SCHEDULING

- Manager and Staff
- Service Planning
- Telephone Information
- Scheduling
  - Supervisor and Staff
  - Schedule Makers
  - Traffic Checkers

EMPLOYEE DEVELOPMENT

- Manager and Staff
- Instructors

TRANSPORTATION PROGRAMS

- Manager and Staff
- Contra Flow
- Busways
  - Supervisors
  - Operations Deployment Personnel
  - Maintenance
  - Controllers
- Charter
- Car Share
- Metro Lift
- Para-Transit
  - Supervisors
  - Operators
  - Direct Expenses

GENERAL OPERATING COSTS

- Payroll Taxes - FICA
- Union Pension
- Life Insurance Plans
- Payroll Taxes - State Employment
- Workers Compensation Insurance
- Work Injury Payments
- Sick Leave
- Uniform and Tool Allowance
  - Maintenance
  - Transportation
- Longevity Award
- Benefit Trust Contribution
- Utilities
- Physical Damage Premium
- Casualty Claims Payments
- Workers Compensation Claims Payments
  - Operators
  - Maintenance
  - Administrative Employees
- Other Insurance Premiums
- Diesel Fuel Tax
- Gasoline Tax
- Rent
- Bodily Injury and Property Damage Insurance Premiums

Figure 3. Cost equations: rail operations.

DIRECTOR AND STAFF

- Director and Staff

TRANSPORATION

- Superintendent and Staff
- Central Control:
  - Assistant Superintendent and Staff
  - Supervisors
  - Controllers
    - Peak
    - Off-Peak
- Train Operations:
  - Assistant Superintendent and Staff
  - Supervisors
  - Train Operators
  - Yard Controllers
- Station Operations:
  - Assistant Superintendent and Staff
  - Station Agents
  - Supervisors

MAINTENANCE

- Superintendent and Staff
- Vehicle Maintenance:
  - Assistant Superintendent and Staff
  - Supervisors - Service and Inspection
    - S&I Repairmen
    - Car Cleaners
    - S&I Foremen
  - Supervisors - Component and Heavy Repair
    - Heavy Vehicle Repairmen
    - Heavy Repair Foremen
    - Component Repairmen
      - Vehicle
      - Wayside

MAINTENANCE (Con't)

      - Component Repair Foremen
    - Parts, Supplies, and Service
- Maintenance Control
  - Supervisor and Staff
  - Vehicle Repair Stores Clerk
  - Maintenance of Way Stores Clerk
  - Stores Foremen
  - Schedulers
- Maintenance of Way
  - Assistant Superintendent and Staff
  - Supervisor and Staff - Track and Structures
    - Building Maintainers
    - Station Janitors
    - Building and Structure Foremen
    - Track Maintainers
    - Track Foremen
    - Maintenance of Way Shop Repairmen
    - Maintenance of Way Shop Foremen
- Supervisor and Staff - Wayside Equipment
  - Train Control System Maintainers
  - Train Power System Maintainers
  - Communication Systems Maintainers
  - Fare Collection Equipment Maintainers
  - Wayside Equipment Foremen
- Parts, Supplies, and Services
  - Stations
  - Track and Structures
  - Train Control
  - Communications
  - Power
  - Fare Collection
- Electricity
  - Traction
  - Yards
  - At-Grade/Elevated Stations
  - Subway Stations
  - Chiller Plants

SAFETY AND ASSURANCE

- Superintendent and Staff
- Vehicle Inspectors
- Maintenance of Way Inspectors

GENERAL OPERATING EXPENSES

- Payroll Taxes - FICA
- Union Pension
- Life Insurance Plans
- Payroll Taxes - State Employment
- Workers Compensation Insurance
- Work Injury Payments
- Sick Leave
- Uniform and Tool Allowance
  - Maintenance
  - Transportation
- Longevity Award
- Benefit Trust Contribution
- Physical Damage Premium
- Casualty Claims Payments
- Workers Compensation Claims Payments
  - Station Agents
  - Maintenance
  - Other Rail Employees
- Other Insurance Premiums
- Railroad Insurance Premium
- Bodily Injury and Property Damage Premium

These projections are shown in Table 1 as the annual percentage rate of increase for each year from 1982 through 2000.

Inflation rates for bus and rail maintenance parts, supplies, and services are based on near-term WMATA budget assumptions (5 percent above the CPI rate of inflation). After five years, the incremental rate is assumed to drop to 3 percent above the CPI rate.

APPLICATION

The cost model described above has been applied to four regional transit alternatives in the course of

Figure 4. Cost equations: administration.

EXECUTIVE OFFICE
- Director and Staff
- Legal
- Internal Audit

FINANCE
- Director and Staff
- Treasury Services
  - Manager and Staff
  - Ticket and Pass
  - Cashier Operations
  - Encoding Machine Operators
  - Revenue Attendants
  - Farecards
- Accounting
- Management Information Systems
- Risk Management
  - Risk Management Staff
  - Claims Adjusters
  - Claims Chief and Staff
  - Direct Expenses
- Budget and Systems

TRANSIT SYSTEMS DEVELOPMENT
- Director and Staff
- Program Control
- Engineering and Construction Manager and Staff
- Engineering
- Construction
- Project Development
- Capital Programming
- System Planning
- Right-of-way

ADMINISTRATIVE SERVICES
- Director and Staff
- Purchasing and Stores
  - Manager and Staff
  - Warehouse and Storeroom Clerks
- Office Services
- Contracts Administration
- Grants Administration
- Human Resources
- Security
  - Officers
    - Patrol
      - Buses
      - Busways
      - Train/Station
    - Revenue Protection
  - Field Supervisors
  - Investigators
    - Bus
    - Rail
  - Administration Staff - Transit Police
  - Security Guards
    - Garages
    - Rail Yards
  - Zone Monitors
    - Park and Ride Lots
    - Stations
  - Security Supervisors
  - Administrative Staff Security
  - Direct Expenses
- General Overhead
  - Rent for Administration Building
  - Payroll Taxes
  - Non-Union Pension
  - Hospital, Surgical, Medical Plans
  - Travel

PUBLIC SERVICES
- Government and Public Affairs
- Affirmative Action
- Community Relations
- Marketing

GENERAL OPERATING COSTS
- Union Pension
- Life Insurance Plans
- Work Injury Payments
- Sick Leave
- Longevity Award
- Benefit Trust Contribution
- Workers Compensation Claims Payments
  - Maintenance
  - Administration

Figure 5. Worksheet for sample cost equation.

ALTERNATIVE BASE LINE RAIL FY2000	RAIL OPERATIONS / MAINTENANCE OF WAY (Continued)	FY2000 COSTS IN 1982 DOLLARS	Peat, Marwick, Mitchell & Co. PAGE 14 OF 27
Track Maintainers	Route-Miles x Maintainers/Route-Mile x Avg Maintainers Wage/Main-Year		
$526,575	17.5 1.5 $20,060		
Track Foreman	(Track Maintainers) ÷ Maintainers/Foreman x Avg Foreman Salary/Man-Year		
$135,476	23.6 6 $30,907		
Maintenance of Way Shop Repairmen	Route-Miles x Repairman/Route-Mile x Avg Repairman Wage/Man-Year		
$422,503	17.5 1.0 $24,143		
Maintenance of Way Shop Foreman	Foreman/Shift x Shifts/Days x Days/Week ÷ Shifts Worked/Week x Avg Foreman Salary/Man-Year		
$34,805	1 1 5 4.44 $30,907		
Supervisor + Staff-Wayside Equipment	Salaries (Fixed)		
$43,500	$43,500		

the financial analysis. These alternatives were defined during the travel-demand analysis and include high-capacity facilities in travel corridors with the highest peak-period volumes. In addition, the alternatives include many non-MTA facilities to be built by the Texas State Department of Highways and Public Transportation. These alternatives are

Table 1. Inflation projection: annual percentage change.

Year	Houston SMSA CPI	U.S. CPI	Houston Diesel Fuel	Houston Electricity
1983	5.8	6.9	3.4	14.7
1984	6.2	6.8	9.6	14.7
1985	6.5	7.0	13.5	14.7
1986	6.8	7.5	15.5	14.7
1987	6.6	7.2	13.7	14.7
1988	6.5	7.0	14.5	14.7
1989	6.5	6.8	14.5	14.7
1990	6.5	6.9	13.6	14.7
1991	6.3	6.7	11.7	8.4
1992	6.1	6.5	10.8	8.4
1993	6.0	6.3	10.6	8.4
1994	5.9	6.2	10.1	8.4
1995	5.9	6.1	9.6	8.4
1996	6.0	6.2	8.8	8.4
1997	5.9	6.3	8.6	8.4
1998	5.9	6.3	8.2	8.4
1999	5.9	6.3	8.3	8.4
2000	5.8	6.2	8.1	8.4

1. Base bus contains the FY2000 all-bus system with one-way busways on the North, Gulf, and Katy Freeways. Bus routes feeding these high-speed facilities are included in the regional bus system.
2. Busway contains the FY2000 all-bus system with bidirectional busways in the Priority Corridor and one-way busways on the North, Gulf, and Katy Freeways. Bus routes feeding these high-speed facilities are included in the regional bus system.
3. Baseline rail contains a rapid rail line from the West Loop, through the central business district, and out to Crosstimbers. One-way busways are on the North, Gulf, and Katy Freeways. Bus feeder routes serving the rail stations are included in the regional bus system estimates.
4. Rail-to-North Belt contains the rapid rail line from the Base Rail alternative extended past Crosstimbers to the North Belt. The addition of feeder bus service to the new stations and the consequent reduction of line-haul bus routes are reflected in the regional bus system.

Figure 6 identifies the alignment of the various transitways studied.

The travel-demand analysis determined, for each alternative, the quantity of service to be provided and the resulting patronage in each analysis year from FY1982 (the base year) through FY2000 (the design year). Table 2 presents the FY2000 system characteristics. These include both service statis-

Figure 6. Transit way corridors.

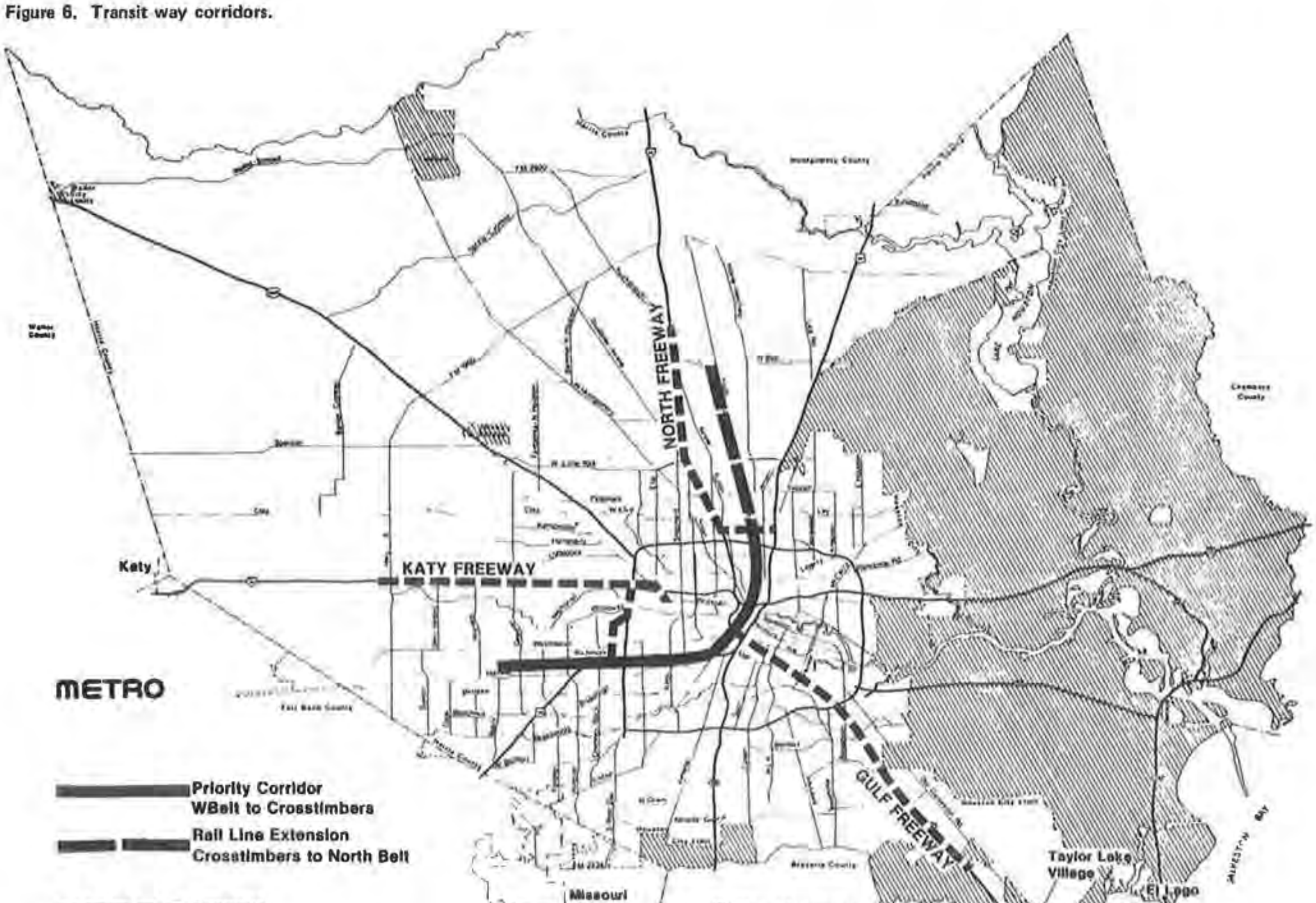


Table 2. FY2000 system characteristics.

System Characteristic	Base Bus	Busway	Base Line Rail	Rail to North Belt
Bus				
Regular buses	1885.0	1872.0	1507.0	1440.0
Articulated buses				
Total buses (peak period)	1885.0	1872.0	1507.0	1440.0
Regular bus surface vehicle miles	71.574	58.673	52.874	52.670
Regular bus guideway vehicle miles	29.609	63.253	31.761	24.670
Total regular bus vehicle miles	101.183	121.926	84.635	77.340
Articulated bus surface vehicle miles	0.0	0.0	0.0	0.0
Articulated bus guideway vehicle miles	0.0	0.0	0.0	0.0
Total articulated bus vehicle miles	0.0	0.0	0.0	0.0
Total vehicle miles (millions)	101.183	121.926	84.635	77.340
Platform hours (millions)	6.286	6.395	4.923	4.729
Operating garages	9.0	9.0	8.0	7.0
Maintenance facilities (ft^2 000 000s)	1.345	1.345	0.943	0.977
Park-and-ride lots	36.0	36.0	37.0	38.0
Service areas	13.0	13.0	13.0	13.0
Contraflow busways	0.0	0.0	0.0	0.0
Gulf freeway-type busways	3.0	3.0	4.0	3.0
Other busways	0.0	2.0	0.0	0.0
Total revenue ($000 000s)	98.406	118.018	130.191	131.157
Contract bus hours (millions)	0.0	0.0	0.0	0.0
Rail				
Peak vehicles			108.0	144.0
Peak trains			18.0	24.0
Base trains			8.0	11.0
Early/late trains			6.0	7.0
Total vehicle miles (millions)			8.728	12.518
Route miles			17.5	25.1
At-grade/elevated stations			14.0	18.0
Subway stations			3.0	3.0
Mezzanines			17.0	21.0
Service and inspection yards			1.0	1.0
Traction substations			17.0	21.0
Chiller plants			1.0	1.0
Rail passengers (millions)			85.619	86.510

Table 3. FY2000 costs by division (millions of dollars).

	Base Bus			Busway			Base Line Rail			Rail to North Belt		
Costs by Division	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation
Director and staff	2.4	7.0	2.4	2.4	7.1	2.4	2.2	6.5	2.2	2.2	6.4	2.2
Vehicle maintenance	166.1	488.0	100.6	185.5	545.0	112.4	135.3	397.6	82.0	126.1	370.6	76.5
Facility maintenance	6.6	19.3	6.3	6.6	19.3	6.3	4.9	14.3	4.7	5.0	14.6	4.8
Transportation	94.4	277.5	94.4	96.1	282.4	96.1	74.9	220.1	74.9	71.8	211.0	71.8
Planning and scheduling	2.5	7.3	2.5	2.5	7.4	2.5	2.1	6.2	2.1	2.1	6.1	2.1
Employee development	1.0	2.9	1.0	1.0	2.9	1.0	1.0	2.9	1.0	1.0	2.9	1.0
Transportation programs	5.9	17.4	5.6	6.2	18.3	5.9	6.1	18.0	5.8	5.9	17.4	5.6
General operating costs	43.6	128.0	40.7	45.8	134.5	42.8	34.7	101.8	32.4	33.1	97.2	30.9
Bus operations total	322.4	947.5	253.5	346.0	1016.9	269.4	261.2	767.6	205.1	247.2	726.2	194.8
Director and staff							0.2	0.5	0.2	0.2	0.5	0.2
Transportation												
Superintendent and staff							0.3	1.0	0.3	0.3	1.0	0.3
Central control							0.4	1.3	0.4	0.4	1.3	0.4
Train operations							1.5	4.4	1.5	1.8	5.4	1.8
Station operations							2.4	6.9	2.4	2.9	8.5	2.9
Maintenance												
Superintendent and staff							0.1	0.2	0.1	0.1	0.2	0.1
Vehicle maintenance							6.7	19.6	4.8	9.2	27.2	6.6
Maintenance control							0.5	1.6	0.5	0.5	1.6	0.5
Maintenance-of-way							23.0	67.7	13.6	30.7	90.2	18.0
Safety and assurance							0.9	2.5	0.9	1.0	2.9	1.0
General operating costs							6.1	18.0	5.7	7.1	20.9	6.6
Rail operations total							42.1	123.6	30.4	54.3	159.6	38.5
Executive office	1.3	3.4	1.3	1.4	3.5	1.3	1.1	2.9	1.1	1.3	3.2	1.2
Finance	10.0	27.3	9.7	11.0	30.0	10.6	10.6	29.1	10.2	11.0	30.1	10.5
Transit systems development	3.7	10.9	3.6	3.7	10.9	3.6	3.7	10.9	3.6	3.7	10.9	3.6
Administrative services	20.5	58.8	19.6	21.3	61.0	20.4	20.8	60.0	20.0	20.8	59.7	19.9
Public services	4.1	10.4	3.9	4.2	10.7	4.0	3.5	9.0	3.4	3.9	9.9	3.7
General operating costs	0.3	0.9	0.3	0.3	0.9	0.3	0.3	0.8	0.3	0.3	0.8	0.3
Administration total	39.9	111.7	38.4	41.9	117.1	40.3	40.1	112.8	38.5	40.9	114.6	39.2
System total	362.5	1059.2	291.9	388.0	1134.0	309.7	343.4	1004.0	274.0	342.4	1000.4	272.5

tics (e.g., miles, hours, and vehicles) and physical characteristics (e.g., stations, yards, and garages). Based on these system characteristics, FY2000 operating costs were computed and are presented in Tables 3 and 4. Table 3 presents the costs organized by the MTA management hierarchy. This presentation is useful in comparing future costs with current budgeted MTA operating costs. Table 4 presents the costs arranged by category, or object class. This presentation is useful in exploring cost components of special interest to management (e.g., union wages, diesel fuel) or to investigate inflation effects in sensitivity analyses.

Three values of cost are presented for each alternative in Tables 3 and 4. They involve the following considerations of inflation:

1. 1982 dollars. Costs are computed by using the inflation rates defined above, and then they are deflated on the basis of the Houston CPI alone. These costs, therefore, include the incremental effect of those cost components that inflate at a rate

Table 4. FY2000 costs by category (millions of dollars).

	Base Bus			Busway			Base Line Rail			Rail to North Belt		
Costs by Category	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation	1982 Dollars	Inflated Dollars	1982 Dollars, No Incremental Inflation
Operator/station agent	89.2	262.3	89.2	90.8	266.8	90.8	72.7	213.5	72.7	70.6	207.6	70.6
Repairmen/maintainer/cleaner	26.7	78.3	26.7	29.1	85.4	29.1	27.8	81.6	27.8	28.7	84.5	28.7
Other	4.3	12.6	4.3	4.5	13.2	4.5	4.0	11.6	4.0	3.5	10.4	3.5
Total wages	120.2	353.2	120.2	124.4	365.4	124.4	104.4	306.8	104.4	102.9	302.5	102.9
Front-line supervisors	5.6	16.4	5.6	6.0	17.7	6.0	7.4	21.8	7.4	7.8	22.8	7.8
Other	27.2	77.1	27.1	28.3	80.0	28.2	29.0	82.6	28.9	29.2	83.2	29.2
Total salaries	32.8	93.5	32.7	34.3	97.7	34.2	36.4	104.5	36.3	37.0	106.0	36.9
Total wages and salaries	153.0	446.8	152.9	158.6	463.2	158.6	140.8	411.2	140.7	139.9	408.5	139.8
Fringe benefits	28.4	83.2	26.6	29.5	86.2	27.6	25.9	75.7	24.2	25.7	75.1	24.0
Diesel fuel	62.3	183.0	29.4	64.1	188.5	30.3	49.3	145.0	23.3	46.8	137.5	22.1
Bus utilities	1.7	5.0	1.6	1.7	5.0	1.6	1.2	3.5	1.1	1.2	3.6	1.1
Rail electricity	0.0	0.0	0.0	0.0	0.0	0.0	13.8	40.5	6.0	18.8	55.4	8.3
Total fuel and utilities	64.0	188.0	31.0	65.8	193.5	31.9	64.3	189.0	30.5	66.9	196.5	31.5
Vehicle maintenance Parts, supplies, and services	72.5	213.1	40.1	87.3	256.7	48.3	64.8	190.4	35.8	61.4	180.4	34.0
Other maintenance Parts, supplies, and services	3.2	9.4	3.0	3.3	9.7	3.1	6.1	17.8	4.3	7.0	20.7	4.8
Other direct expenses	17.7	49.5	16.3	18.3	51.3	17.0	17.3	48.8	16.0	17.6	49.5	16.3
Purchased transportation	3.6	10.5	3.3	3.6	10.5	3.3	3.6	10.5	3.3	3.6	10.5	3.3
Total parts, supplies, and services	96.9	282.5	62.7	112.5	328.2	71.6	91.8	267.5	59.4	89.6	261.0	58.3
Insurance	6.7	19.8	6.2	7.7	22.7	7.2	8.7	25.5	8.0	8.4	24.8	7.8
Taxes	13.3	39.0	12.5	13.8	40.3	12.9	12.0	35.0	11.2	11.8	34.6	11.1
Total operating costs	362.3	1059.2	291.9	388.0	1134.0	309.7	343.4	1004.0	274.0	342.4	1000.4	272.5

Table 5. FY2000 number of employees.

	Base Bus		Busway		Base Line Rail		Rail to North Belt	
Computed Employees	Union	Non-Union	Union	Non-Union	Union	Non-Union	Union	Non-Union
Bus operations								
Operators/supervisors	4040	80	4110	84	3167	82	3042	80
Vehicle mechanics, cleaners/foremen	1015	93	1115	103	815	75	770	71
Facilities repairmen, cleaners, foremen	133	9	133	9	99	7	101	7
Other	185	267	195	267	163	254	147	246
Subtotal	5374	450	5553	464	4244	418	4060	404
Rail operations								
Operators/supervisors	0	0	0	0	50	14	66	14
Station agents/supervisors	0	0	0	0	80	20	99	25
Vehicle repairmen, cleaners, foremen	0	0	0	0	87	9	117	12
Maintenance-of-way repairmen, foremen	0	0	0	0	195	34	253	44
Other	0	0	0	0	6	92	6	98
Subtotal	0	0	0	0	419	169	542	192
Executive office	0	8	0	8	0	8	0	8
Finance	0	214	0	236	0	217	0	226
Transit systems development	0	77	0	77	0	77	0	77
Administrative services								
Security	0	292	0	306	0	303	0	293
Other	27	169	27	170	24	159	21	166
Subtotal	27	461	27	481	24	461	21	459
Public services	0	68	0	70	0	59	0	64
Total employees	5401	1278	5580	1337	4687	1409	4623	1431

different from that of the Houston CPI.

2. Inflated dollars. Costs are computed by using the inflation rates defined in Table 1. These costs, deflated by the Houston CPI, equal the 1982 dollar costs.

3. 1982 dollars, no incremental inflation. Costs include no inflation whatsoever and are based solely on the FY1982 base year unit costs, wages, and salaries. These costs can be considered as the costs to operate the FY2000 systems in 1982. Thus, these costs can be useful in comparing the model results with current transit industry experience (7).

Table 5 presents FY2000 employees for each alternative. These values are determined during the course of the cost model computations and are useful in explaining some of the differences in costs. In addition, they can provide guidance to management in the consideration of service expansion plans.

CONCLUSION

The transit operating cost model presented in this paper has several important features that make it a useful analytical tool for transit management. First, the model is rich in detail, capturing the cost effects of staffing levels, labor productivity standards, unit prices, and inflation for different cost components. Second, the model is user-oriented. It is formulated on the basis of data commonly developed in the budgeting process. Its responsibility center-based organization provides for both ease in comparing projections with current conditions and ease in updating various data values. Finally, the model can be applied either manually or on a computer. Simplified worksheets allow for organized computation. Both mainframe and microcomputer applications have been successfully performed.

There are fundamentally two potential applications of the cost model. For short-range planning, the model can be used in the budgeting process for quick-response sketch planning. It could be used in many of the what-if questions typically asked by management regarding the cost effects of alternative service changes or potential labor productivity changes. It can also be useful in the context of sensitivity analyses concerning rates of inflation or other unknowns.

In long-range planning, the model can apply current and anticipated cost experience to project operating costs in the financial and cost-benefit analysis of major capital investments. The cost model described in this paper provides a strong analytical foundation for multiyear analysis of transit investment in Houston, Texas. Other such applications should certainly be possible.

REFERENCES

1. J.M. Holec, Jr., and R.L. Peskin. Use of Productivity Measures in Projecting Bus and Rail Transit Operating Expenditures. TRB, Transportation Research Record 797, 1981, pp. 40-49.
2. Peat, Marwick, Mitchell and Co. FY80 Net Income Analysis. Washington (D.C.) Metropolitan Area Transit Authority, Draft Final Rept., June 1981.
3. Peat, Marwick, Mitchell and Co.; Kaiser Engineers. CSTS and Bus Operating Costs Models. New Orleans Office of Transit Administration, Working Paper, Aug. 1981.
4. Peat, Marwick, Mitchell and Co.; Barton-Aschman Associates, Inc. Bus and Light Rail Transit Operating Cost Models. East-West Gateway Coordinating Council and Bi-State Development Agency, St. Louis, Working Paper, Dec. 1981.
5. Peat, Marwick, Mitchell and Co. Feeder Bus Cost Model. Southeastern Michigan Transportation Authority, Detroit, Working Paper, April 1982.
6. Peat, Marwick, Mitchell and Co. Revised Transit Operating Cost Model. Metropolitan Transit Authority of Harris County, Houston, TX, Working Paper, May 1982.
7. Peat, Marwick, Mitchell and Co. Reasonability Test of Operating Cost Model Results. Metropolitan Transit Authority of Harris County, Houston, TX, Working Paper, June 1982.

Tri-Met Bus Operator Costing Methodology

JANET JONES

Traditional financial planning techniques are rapidly becoming inadequate as public mass transit confronts an environment characterized by limited and fluctuating revenues, funding shortfalls, and rising costs. The Tri-Met operator costing model is a part of a financial forecasting system approach toward the planning process in which short- and long-term consequences of alternative operating policies and performance can be determined. Tri-Met has drawn on past experience, research and review of existing methodologies, and future needs assessments to develop a costing methodology that combines the positive features of cost build-up and historical cost approaches and represents a sensitivity to the causal relationships underlying fixed and variable cost items at a marginal cost level. Bus operator costs are projected on a monthly basis over a six-year time frame as a function of service levels, service characteristics, work rules, productivity, and economic conditions. Common applications of the model range from service and scheduling changes to union labor contract provisions, assessments of part-time drivers, benefits, productivity, and absenteeism. The forecast technique has proved to be an invaluable tool of cost management and control, minimizing the risks involved in critical policy decisions.

Traditional financial planning techniques were sufficient tools of cost-revenue management when costs remained relatively stable and revenues were predictable and even sufficiently available. But growing complexities that characterize today's financial policy decisions require sophistication in planning, anticipating, and coping with financial uncertainties. Transit planning is increasingly complex due to demands to apply new and better tools for handling the dynamics of limited and fluctuating revenues, funding shortfalls, and rising costs. As a result, transit operators are directing greater attention toward cost effectiveness, efficiency and control, productivity, and performance analysis. It is fundamental to the responsibilities of transit operators to not only manage existing revenues and

Figure 1. Financial forecasting system.

REVENUE MODELS		COST MODELS	
INPUT	Revenue Line Items	Cost Line Items	INPUT
NON-CAPITAL REVENUES	Fare Revenues Tax Base Revenues Federal Operating Assistance (Section 5) Federal Technical/Demonstration Grants Miscellaneous (interest on investments, etc.) State Operating Assistance Other	Bus Operator Costs Other Transportation and Operations Costs Fuel Maintenance General and Administrative including pension cost and insurance and claims	OPERATING COSTS
CAPITAL REVENUES	Federal Capital State Capital Local Capital Other Local Assistance	Capital Costs: Vehicles, facilities and equipment Vehicle replacement Debt Service Project Scheduling Life-cycle Costing	CAPITAL COSTS
OUTPUT	SUMMARY STATISTICAL REPORTS SUMMARY FINANCIAL FORECAST REPORTS		OUTPUT

cultivate new resources but also to better anticipate, monitor, and control costs.

Although there are many types of approaches applied to address these needs, financial forecasting provides a forward-looking economic planning tool. It steps beyond the customary budgeting process--financial forecasting in its most basic form--to assess the often profound financial implications of certain courses of action. And, unlike budgeting, it captures a consideration of the real causes and consequences underlying many revenues and costs. This is especially useful when incorporated into the planning process to assist in shaping and evaluating alternative plans and policies. However, because costs and revenues are controllable only within certain limits, financial forecasting can bring about a greater awareness of marginal cost-revenue impacts of alternative policies and the extent to which they are within management control.

It is difficult to measure the usefulness of financial forecasting. Recognizing, however, that transit reaps the result of decisions and not the result of plans, financial forecasting is effective in increasing the opportunities for making better (or at least better-informed) decisions and minimizing the risk of making a poor decision. It can be an invaluable interface between the planning process and the decisionmaking process.

FINANCIAL FORECASTING SYSTEM

Tri-Met faces a continuing need for accurate, timely financial projections that are readily responsive to policy issues. Answering this need, Tri-Met has made strides toward the development, improvement, and application of forecasting for use in the financial planning process and in program planning, including an automated version of a bus operator costing technique, which is discussed in this report.

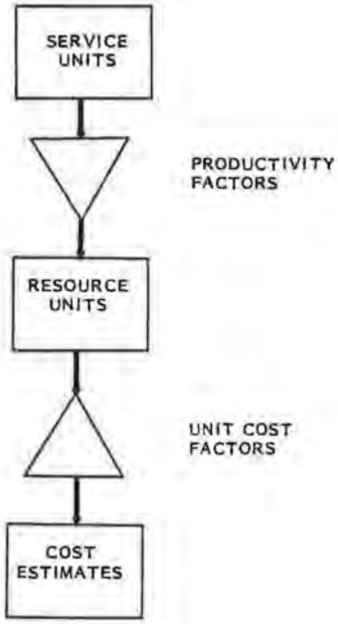
The financial forecasting system is composed of a set of models and a planning structure. The planning structure serves as an analytical framework within which the models reside. Financial forecasting is the process in which a number of techniques, or models, which share a common data base, calculate future costs and revenues in terms of cash flow. Each model, supported by one or more subprograms, forecasts a distinct segment of costs or revenues within the comprehensive system structure. (See Figure 1.)

The concept behind this modular approach is to achieve a great deal of flexibility for testing what-if kinds of questions that require varying appropriate levels of detail. For example, a six-year cash flow annual summary report may be desired, requiring application of all of the cost and revenue models, or any model can be run individually if a monthly breakdown of detailed departmental line item costs is required.

The models represent mathematical relationships between cost-revenue items and a simulation of cause and effect. The causal factors are input data (independent variables) and the effects are the cost and revenue output (dependent variables). Results of the models are calibrated to match observable data and validated against available historical data.

The planning structure provides an organized method to input, access and analyze data and control parameters, and specify output reports. The structure contains a multioption variable processor, coded in FORTRAN, which incorporates such features as parameter-driven inputs and built-in default values. It allows flexible interpolation of missing values and extrapolation of input data on growth-inflation factors. It facilitates input data file editing and labeling capabilities. Reporting is

Figure 2. Financial forecast model structure.



allowed at various levels of detail and aggregations may be made on a quarterly or annual basis from monthly projections.

These features were developed in light of several characteristics that were considered desirable attributes in a financial forecast system. Flexibility was a high priority because in setting up a system, one cannot hope to predetermine all requirements. Therefore, the system was structured to allow for model changes and enhancements as development proceeded. It was also recognized that the system should be relatively simple to work with from a user standpoint. The models were built within a framework designed to accommodate a variety of input options, built-in default values, and convenient data interpolation-extrapolation features with a modular format that will permit independent subroutines for testing or future enhancement. Another requirement was sensitivity within the models to small changes at a marginal cost level. This is especially useful where a change in assumptions might make a difference, but it is not clear how much difference it might make. Perhaps the most important attribute to be considered was the value of the system in application. The system has been successfully applied in policy alternatives analysis for major decisions at Tri-Met. It has provided quality information that has minimized the risks involved in critical policy issues.

TRI-MET APPROACH TO BUS OPERATOR COSTING

Organization of Bus Operator Model

The bus operator costing model was designed to provide detailed, accurate financial information that would reflect sensitivity to operational policy and performance changes at a marginal cost level and would also capture the primary interrelationships among fixed and variable costs and their causal factors. With expanding applications ranging from service and scheduling changes to union labor contract provisions, assessments of part-time drivers, benefits, productivity, and absenteeism, it became clear that overly simplified techniques were not only inflexible, but inadequate as well. Tri-Met has drawn on past experience, research and review of existing methodologies, and future needs assessments to develop costing techniques.

The methodology of the model combines the positive features of cost build-up and historical cost approaches. The unit cost build-up approach is used to develop labor requirements based on service characteristics (e.g., peak to base ratio, service hours, and miles). The historical cost approach is used to develop labor cost factors per productivity unit (e.g., extraboard, work rule constraints, supervision, fringe, etc.) based on historical data. The historical cost approach captures the inefficiencies of exception pay, replacement labor costs as a function of absenteeism, and some work rules under existing conditions. However, it does not achieve the sensitivity to costs incurred due to major changes in service and operating characteristics, which is better accomplished by the cost build-up approach.

The bus operator model represents a dynamic costing technique designed to project future costs as a function of service level, service characteristics, work rules, productivity, and economic conditions. The model is initially driven by daily-level service hour input and further responds to any alterations in the type of service such as changes in the peak to base ratio, adjustments in weekday versus Saturday or Sunday service, or a shift from urban radial to time transfer service. Productivity factors reflect the efficiency of service provided in terms of ratios between service hours, platform hours required to support a particular service level, and operator pay hours required to assure to certain reliability at that service level.

Underlying the operator cost component calculations are assumptions that reflect work assignment provisions such as extraboard rules, constraints on the use of part-time drivers, and specifications for wage rates, cost-of-living adjustments, and benefits. Productivity assumptions take into account absenteeism, extraboard requirements, and unscheduled overtime. Based on the number of scheduled operators, additional operators are figured in to cover absence exceptions such as sickness, vacations, holidays, and other miscellaneous absences. This is achieved by applying various productivity factors, efficiency ratios, and unit costs to derive total operator requirements and direct labor costs. Finally, economic assumptions impact variable overhead costs including benefits and pension and payroll taxes.

Structure of Model

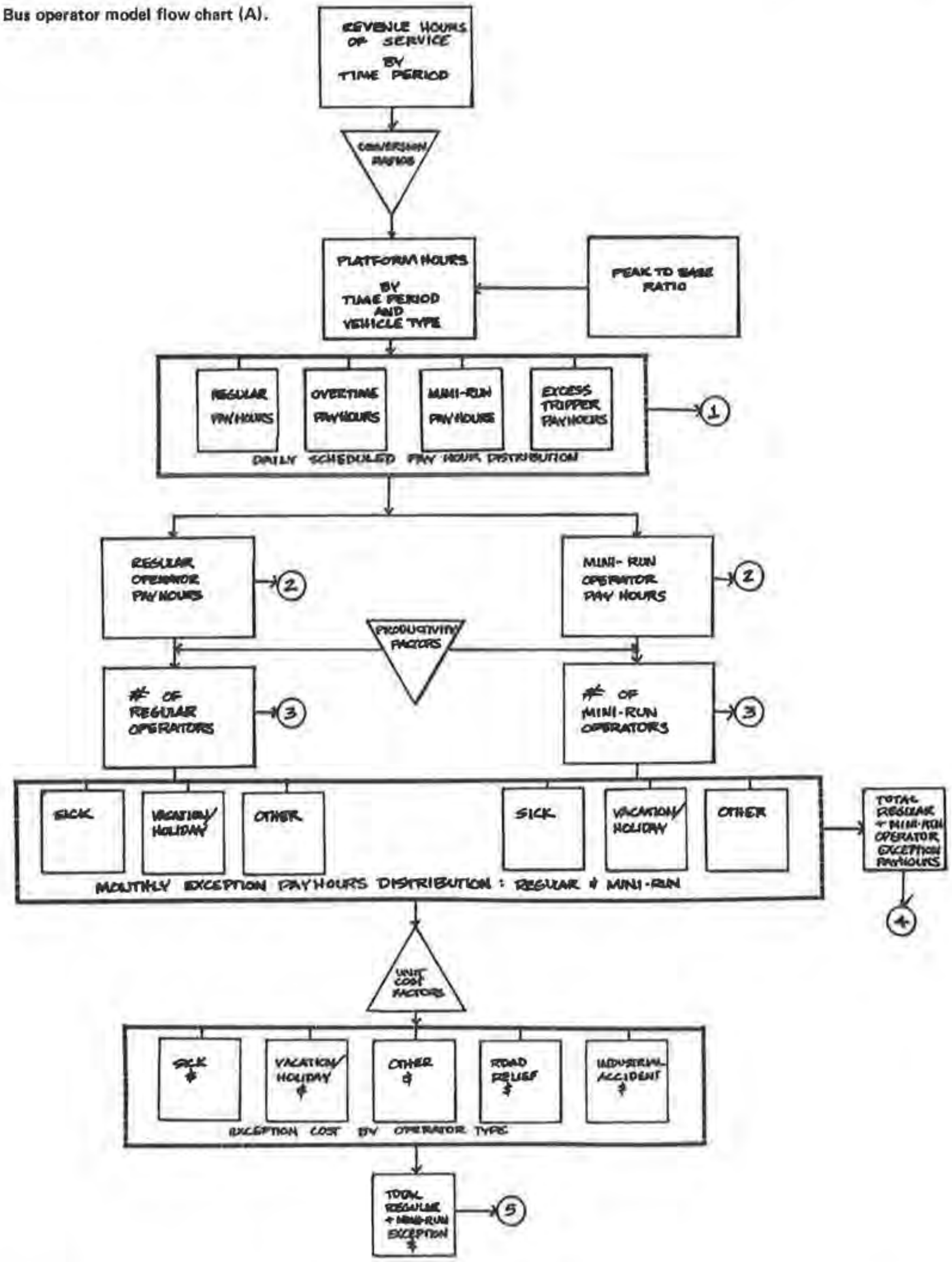
The bus operator costing process is structured in a hierarchical manner, starting from basic service units (service hours, miles, and vehicles) and working back to resource units (consumable items for which the transit operator must pay directly such as labor hours). The conversion from service units to resource units is accomplished through a series of productivity factors (such as pay hours per platform hour) that are based on primarily historical experience of Tri-Met and other transit operators. Once resource units are derived, they are converted to expenditures by applying (unit) cost factors. This general process is illustrated in Figure 2.

The bus operator costing model is divided into eight principal sections. These are discussed in sequence in this section, supported by flow chart diagrams (Figures 3-5).

Variable Declaration and Identification

Variables are processed primarily through the main program. All variable arrays are identified, in-

Figure 3. Bus operator model flow chart (A).



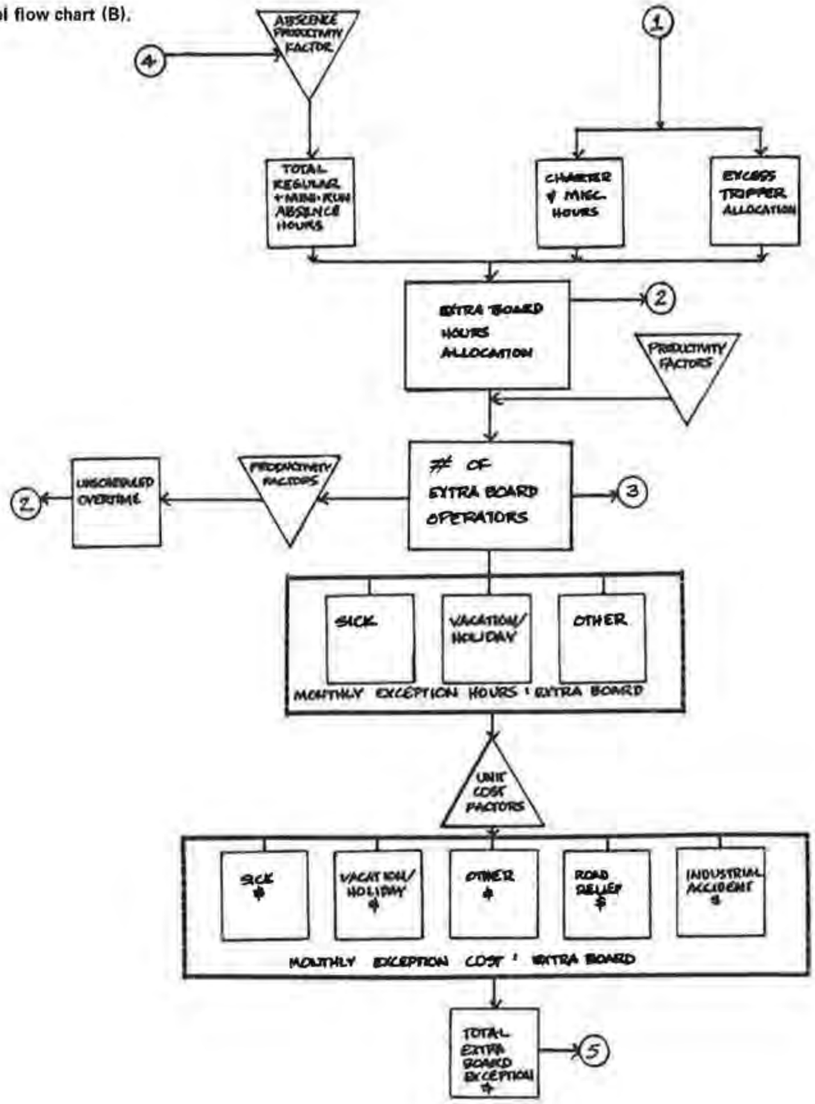
dexed, and declared as integer or real and are passed through the bus operator model using COMMON statements. The variables are input through a separate data file.

Program control parameters are input through the FORTRAN NAMELIST function. "&PARAM" is used to set the base year of the projections and "&SELECT" is used as a report parameter and contains arrays to control the output reports. The values assigned to these arrays call the appropriate subroutines, such as the bus operator model, and specify the desired output report.

Input/Computational/Output Variables are separated into two types of variables. The first type of variable includes time-based monthly data such as growth rates, unit cost factors, and operations costs. Each has a capacity for 72 time periods and is categorized by variable function. Each variable also is associated with a 20-character variable label and a 12-character units label, and it can be tied to a 12-digit accounting code. The second type of variable is characterized by changes that do not occur on a monthly basis, such as service levels and productivity factors. These values may be input with up to only nine changes, although they are calculated on a monthly basis.

Both types of variables require a label card that performs three functions. First, it simply defines the variable. Second, it designates an appropriate interpolation-extrapolation code to be performed on the variable, and third, it indexes the variable labels to correspond to the data cards.

Figure 4. Bus operator model flow chart (B).



Initial Input Data

The basis of all future cost estimates is the description of the future alternative transit service networks. These are quantified through the computerized simulation urban transportation planning system (UTPS) program, INET. This process converts the basic network description (route alignments, headways by time interval, running times, and layovers) into estimates of service units (revenue service miles, hours, and vehicles by time period by mode). The output must be refined to reflect actual conditions through several steps. Since INET is not a perfect simulation tool, evening peak and daybase statistics require adjustment by using calibrated factors: pure revenue hours (no layover)--peak, 0.80 and daybase, 0.99.

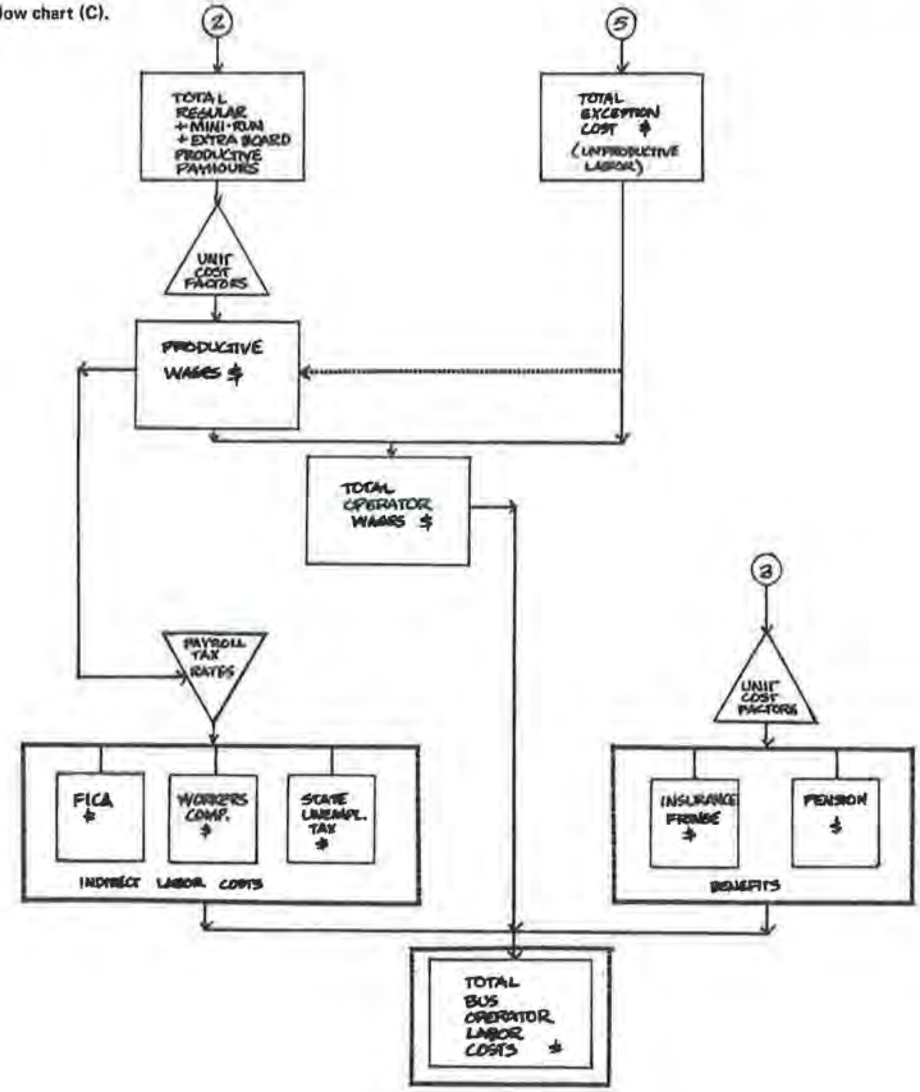
To derive "pure" revenue hours for all service periods in agreement with an actual run cut, as performed by RUCUS, INET/RUCUS conversion factors are applied. This is done in order to eliminate layover time included in the INET revenue hours summaries and to adjust systemwide statistics to realistic figures. (Assumed service period factors are shown in Figure 6). RUCUS simulations would yield more precise figures than INET, but RUCUS is data-intensive and quite difficult to use as a forecasting tool. The need for INET can be substituted by a SAS simulation, used to determine vehicle and service hour estimates of service changes.

Data availability can be a problem when estimating potential changes in service levels or service characteristics, so the program is constructed to handle several input options. Flags are used for indication of the desired input entry level. For example, if service hour data are unavailable for the evening peak and daybase, total weekday hours may be input instead. Input can be any of the following combinations: (a) weekday evening peak and daybase revenue hours; (b) weekday revenue hours; (c) weekday, Saturday and Sunday revenue hours; (d) weekday, Saturday and Sunday revenue hours plus weekday, Saturday and Sunday articulated bus revenue hours; or (e) weekday, Saturday and Sunday platform hours.

INET produces revenue hour figures in terms of the evening peak and daybase service levels. After these figures are factored to reflect actual conditions, the assumed service factors are employed to develop total weekday service levels from evening peak and daybase figures.

Non-peak weekday revenue hours are derived as a function of the daybase. Peak weekday revenue hours are derived as a function of the evening peak. The sum of weekday peak and non-peak revenue hours represents total weekday revenue (in service) hours.

Figure 5. Bus operator model flow chart (C).



Saturday and Sunday revenue hours are derived as functions of total weekday revenue hours.

Platform time, which includes hours of scheduled service operated (revenue hours) plus deadhead and layover time, is calculated in the next step of the process. No overtime, guarantee, or report-clear time is included. The conversion from revenue hours to platform hours requires ratios applied on the basis of the weekday, Saturday and Sunday relationships between total daily platform hours and daily revenue hours. The ratios, which account for deadhead and layover time combined, average about 1.32 under present service conditions:

Item	Weekday	Saturday	Sunday
Layover	0.20	0.26	0.27
Deadhead	0.12	0.04	0.07
Revenue	1.00	1.00	1.00
Total	1.32	1.30	1.34

The composition of these ratios in terms of proportion of layover and deadhead varies by time period as well as by weekday, Saturday and Sunday. This relationship for weekday service is shown in Figure 7.

Layover is a function of running time and remains fairly constant throughout the day. Deadhead is a function of peak service during the day and amount of tripper service. The proportion of deadhead time increases at the beginning and the end of each concentration of service hours as drivers begin and end their runs, and remains at a minimum during peaks of service. Whereas non-peak revenue hours include all service outside the morning and evening peaks, platform hours falling in the non-peak time periods include not only deadhead and layover time for off-peak service but also deadhead time for service that is provided during the peak. Consequently, in order to accurately evaluate the cost of peak service (and assess the portion of platform hours falling in the off-peak associated with service during the peak), it is necessary to allocate a cumulative portion of the off-peak deadhead to the peak revenue hours.

In comparing weekday, Saturday and Sunday hours, there is a higher proportion of deadhead time on weekdays than on Saturdays and Sundays because of the nature of tripper service that provides service exclusively to the peaks. Trippers, operating only on weekdays, have a high ratio of deadhead to layover time. Straight shifts, which generally involve a much greater number of repeated runs, have a larger proportion of layover time and a smaller proportion of deadhead time as compared in Figure 8. Additional comparisons of platform hour components revealed that although a low peak to base ratio is

Figure 6. Service period factors.

Period	Time Interval	# Hours		Factor	1980 Service Hours	1980 Service Hours/Hours
Wee Hours	1:00 - 5:00 A.M.	4.0	X	.05 X Midday	39.3	9.8
A.M. Shoulder #1	5:00 - 7:00 A.M.	2.0	X	.42 X P.M. Peak	301.9	151.2
A.M. Peak	7:00 - 8:00 A.M.	1.0	X	1.00 X P.M. Peak	359.4	363.8
A.M. Shoulder #2	8:00 - 10:00 A.M.	2.0	X	.71 X P.M. Peak	510.3	254.8
Midday (Base)	10:00 - 2:30 A.M./P.M.	4.5	X	1.00 X Midday	883.3	196.3
P.M. Shoulder #1	2:30 - 4:30 P.M.	2.0	X	.75 X P.M. Peak	539.1	271.1
P.M. Peak	4:30 - 5:30 P.M.	1.0	X	1.00 X P.M. Peak	359.4	359.4
P.M. Shoulder #2	5:30 - 7:30 P.M.	2.0	X	.70 X P.M. Peak	502.7	250.6
Evening	7:30 - 1:00 P.M./A.M.	5.5	X	.41 X Midday	442.6	80.1
		24.0			3,938.0	

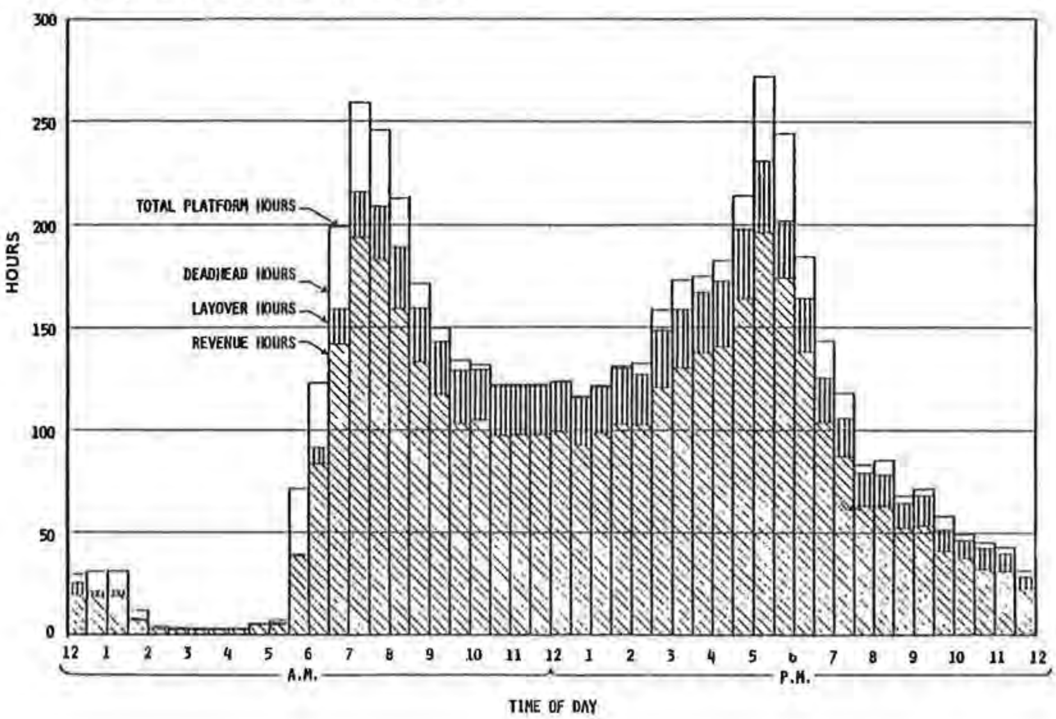
Total Daily Revenue Service Hours = 6.95 X Midday + 7.16 X Peak
= 6.95 X 196.3 + 7.16 X 359.4 = 3,937 hours

Total Saturday Revenue Service Hours = 9.62 X Midday = 9.62 X 196.3 = 1,888 hours

Total Sunday Revenue Service Hours = 5.10 X Midday = 5.10 X 196.3 = 1,001 hours

Total Annual Revenue Service Hours = 2,568 X Midday + 1,926 Peak

Figure 7. Comparison of platform hours to revenue hours.



associated with a large proportion of layover time, a higher ratio is characterized by a small proportion of layover and greater deadhead requirements as shown in Figure 9.

Scheduled Pay Hours

From daily platform hours, scheduled pay hours, which represent the hours required to operate the platform service, are derived. They range up to 110 percent of scheduled platform time, accounting for union work rules and provisions involving report-clear and overtime pay.

The reciprocal ratio of platform time to scheduled pay hours serves to identify the percentage of scheduled operators pay that applies to productive platform service. This efficiency factor generally ranges from 90 to 91 percent. The RUCUS programming system strives to optimize the various runs within the work rules for the best efficiency and thus the lowest system cost per platform hour.

In converting from platform hours to pay hours, the peak to base ratio is incorporated into the model in order to account for its relationship between regular scheduled pay hours, scheduled overtime pay hours, and minirun (tripper service) pay hours. As a function of the peak to base ratio, the allocation of these pay hours can be determined as shown in Figure 10. The peak to base ratio is found on the horizontal axis and the percentage of total pay hours along the vertical axis.

Currently at Tri-Met, 14 percent of the number of existing full-time operators are employed as minirun operators. At the current 2:1 peak to base ratio, there are 180 trippers worked by minirun operators, comprising 7 percent of total pay hours. Beyond a 2:1 ratio, minirun trippers remain constant at the maximum allowable level, and unassigned (open) trippers occur. For each 15 additional trippers, there is a corresponding 1 percent increase in overtime pay hours. As shown in Figure 10, as the peak to base ratio increases, especially beyond 2:1, the proportion of overtime and open trippers continues to increase and the proportion of regular pay hours declines. Weekday, Saturday and Sunday scheduled pay hours are interdependently assigned in the model based on the weekday peak to base ratio in order to reflect these relationships.

At this point in the model, daily platform hours are converted to total weekday, Saturday and Sunday pay hours per quarter. They are then aggregated to sum quarterly (a) regular operator pay hours, (b) scheduled overtime pay hours, and (c) scheduled minirun pay hours. These quarterly pay hours take

Figure 8. Straight shift and tripper platform time.

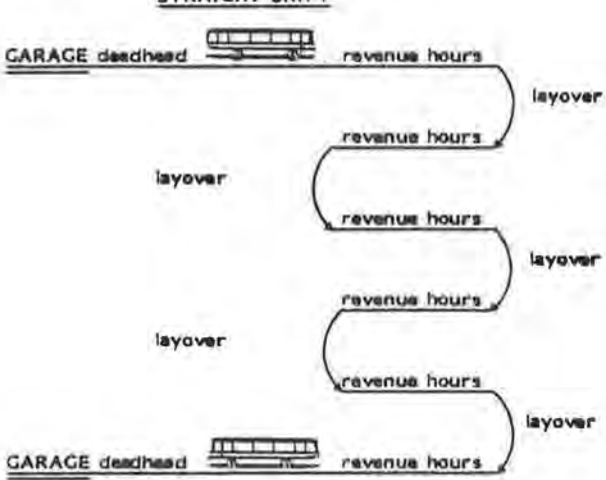


Figure 9. Effect of peak to base ratio on platform hour ratios.

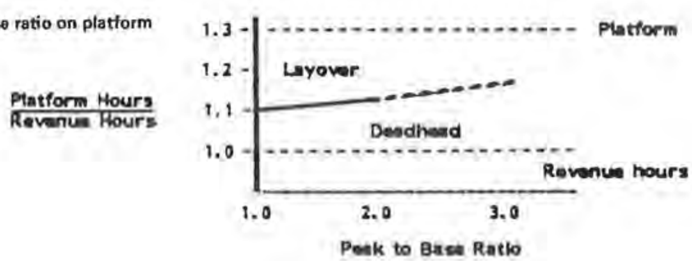
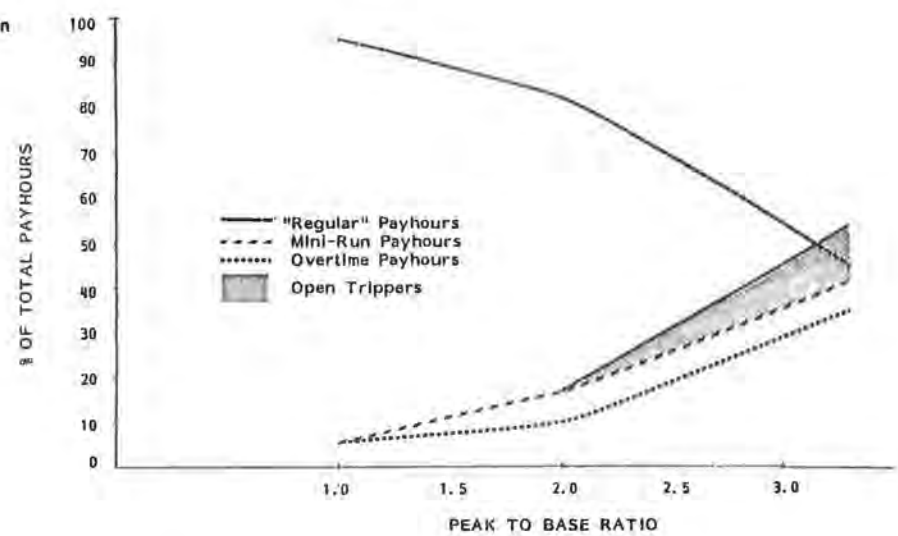


Figure 10. Proportion of payhours as a function of the peak to base ratio.



into account all of the scheduled pay hours required for scheduled service.

Operator Requirements

Regular and Minirun Operator Requirements

The number of regular operators required is based on a weekly aggregated number of scheduled regular operator pay hours plus overtime pay hours that exceed the 8-h limitation. Similarly, weekly scheduled minirun pay hours determine the number of minirun operators required. The number of minirun operators is limited to 14 percent of the number of full-time operators, however, by current union contract provisions. Should the number of minirun pay hours dictate a requirement of minirun operators in excess of this limitation, the "unassigned trippers" are allocated to regular operators.

Regular and minirun drivers perform work that is assigned to them, according to their preferences, exercised on the basis of seniority, through periodic "sign-ups". These individuals sign up for various runs as well as two days off per week and vacation or holiday time. A considerable portion of transit service cannot be assigned in this manner, however, and there is also service (although unassigned) that must be filled due to scheduled operator absence. It is the function of vacation relief and the extraboard to meet these needs on a day-to-day basis. Minirun drivers are never allowed to perform work on the extraboard.

Regular Operator Exception Pay Hours

Exception pay hours include all of the paid "nonproductive" hours of labor. Exceptions consist of three major categories: (a) sick, (b) vacation and holidays, and (c) other exceptions that include jury duty, accident reporting, funeral leave, student training, and operators in other positions. The sick exception calculation incorporates measurement of cost savings due to changes in the rate of absenteeism. Costs of road relief and industrial accidents are calculated using a unit cost-per-operator.

Minirun Operator Exception Pay Hours

Although minirun operators are now eligible for only a small portion of exception pay, the calculation procedure is the same for minirun operators as for regular operators. This enables testing of costs associated with potential proposals liberalizing these benefits and easily accommodates future changes.

Minirun operators are eligible only for student training, industrial accident, and accident reporting exception pay. They become eligible for holiday pay (on a pro rata basis based on hours worked) after one year and, with two years of service, become eligible for pro-rata vacation pay.

Extraboard Operator Requirements

The extraboard work assignment process is complex as it is responsible for meeting a variety of needs. There are runs, portions of runs, or pieces of work termed "open" work, which cannot be assigned because they do not fit within the work rule constraints. This occurs, for example, if trippers exceed the 60:40 ratio of "straight" to "split" runs, or if there are too many trippers to be worked exclusively by minirun operators. There are runs that become open between sign-ups due to promotion, retirement, termination, or long-term disability. There are reliefs to be provided for supervisors, dispatchers, operating clerks, and time for "breaking in" operators to be able to relieve these positions. There are special services that arise such as charters and overloads. There are regularly assigned runs that must be filled due to sickness and absence on the part of regular and minirun operators. The extraboard fills in for more than just their exception pay hours--in fact, there is up to 60 percent additional unpaid exception time beyond paid exception hours. There are exceptions on the part of the extraboard operators to be allowed for as well. There is the added premium of unscheduled overtime that varies according to the size of the extraboard.

In order to include these points in the calculation of extraboard sizing, the initial hours-per-operator procedure relies on the input of average quarterly charter hours, excess tripper hours (if not allocated to regular operators and cannot be met by minirun operators), and regular and minirun operator exception hours that are factored to reflect unpaid absences as well as paid exceptions.

Extraboard Operator Exception Pay Hours

In order to account for extraboard operators exception hours to be covered, the exception pay hour procedure is followed. It is assumed that extraboard operators have the same rate of exceptions (factors) as the regular operators. On deriving the total extraboard exception pay hours, this figure is factored to account for unpaid exception hours and then added to regular and minirun exception hours. The sum equals total paid and unpaid exception time for all part-time and full-time operators.

Extraboard Operator Unscheduled Overtime

Unlike regular and minirun operators, extraboard operators are assigned their work on a day-to-day basis. On the previous day, a list of the extraboard operators and their preassigned work for the next day is posted. Also, operators are assigned to report at various time intervals on the given work day as final adjustments are made to the work assignments. When the demands on the extraboard become particularly heavy, extraboard operators are first contacted to report back and stand by for additional work, then those on a regular day off are asked to come in, and, finally, regularly assigned operators who have either completed their regular work or are on a regular day off. When this occurs, costs expand to include a minimum guarantee of eight hours of pay, plus report time and any unscheduled overtime, at overtime rates for all operators called in on a regular day off. Alternatively, maintaining an abundance of extraboard operators requires a minimum guarantee of eight hours pay, five days a week, as well as fringe benefits, payroll taxes, and other variable overhead costs for each additional extraboard operator.

Enlarging the extraboard, while reducing unscheduled overtime and increasing guarantee time, assures an available supply of operators for work assignments that would otherwise be handled by unscheduled overtime or missed entirely. The benefits derived from increasing service reliability must be balanced against the high overhead costs of additional operators. To accurately simulate an ideal extraboard size that minimizes cost would require a submodel to address such issues as optimization, given the trade-offs between guarantee time and overtime; optimal size of the report crew, given probabilities of absence patterns on a daily basis; and sensitivity of extraboard costs to changes in absenteeism, work rules, and schedules. Such a submodel, currently under development, will serve to greatly enhance this costing technique and will be accommodated as more data and resources become available. The costing model bridges these considerations by using productivity factors based on historical data to derive total unscheduled overtime.

Wages

Total operator wages are separately calculated for productive labor that includes straight time, scheduled and unscheduled overtime and premium pay, and for unproductive labor that consists of the total of all exception pay. Regular straight time labor includes wages paid to regular operators, minirun operators, and extraboard operators. These "regular" extraboard hours consist of the paid and unpaid exception hours of regular and minirun operators. Note, however, that extraboard exception hours are not included as part of "regular" extraboard hours because extraboard operators do not have regularly scheduled pay hours requiring specific replacement by another operator. Because extraboard requirements are a function of exceptions and not exclusively a service level, it is clear that if the exception rate improved, fewer extraboard operators would be required. Overtime labor includes scheduled overtime worked by the extraboard. Premium pay includes an additional $0.50/h for time spent driving an articulated bus. The components of total driver pay hours follow: regular operator scheduled pay hours, regular operator exception pay hours, regular operator scheduled overtime pay hours, extraboard "regular" pay hours, extraboard exception pay hours, extraboard unscheduled overtime pay hours, minirun operator scheduled pay hours, minirun operator exception pay hours, articulated bus scheduled platform hours, exception wages--regular operators, exception wages--extraboard operators, and exception wages--minirun operators.

In order to derive total regular operator productive wages, the model first sums regular operator scheduled pay hours and scheduled overtime. This represents total regular operator scheduled pay hours. Regular operator exception pay hours must be subtracted from this to separate out the unproductive pay hours. The ratios of scheduled regular-scheduled overtime pay hours to total scheduled pay hours are then applied to split the correct proportions of regular and overtime pay hours. The scheduled overtime pay hours are then multiplied by 1.5 times the pay rate. Both regular pay hours and overtime pay hours are then multiplied by the appropriate wage rate to derive total regular operator productive wages.

Minirun total scheduled pay hours are separated according to the productive and nonproductive pay hours by subtracting the exception pay hours. Productive pay hours are then multiplied by the appropriate wage rate to derive total minirun productive wages.

Extraboard pay hours consist mostly of "regular" pay hours, comprised of the regular operators' and minirun operators' scheduled pay hours that are not worked due to a variety of exceptions. There also exist unscheduled overtime pay hours that result primarily due to variations in the sizing of the extraboard, depending on the number of extraboard operators available to perform the work at straight time. The number of unscheduled overtime pay hours is multiplied by 1.5 times the pay rate. Both the productive "regular" pay hours and overtime pay hours are then multiplied by the appropriate wage rate to derive total extraboard operator productive wages. Exception wages calculated for regular, extraboard, and minirun operators are combined to sum total exception (unproductive) wages.

All operators receive an additional $0.50/h for time spent driving an articulated bus. Total articulated bus platform hours are derived from articulated bus revenue hours, the multiplied by the premium pay rate.

Fringe Benefits

Fringe benefits include the cost to Tri-Met of providing life, medical, and dental insurance and a pension plan to its employees. Operator pension costs (which are expensed as defined by a standard actuarial study) are based on a cost-per-operator factor, multiplied by the number of regular and extraboard operators. Other bus operator fringe costs are determined by the total number of regular and extraboard operators and the small number of eligible minirun operators, multiplied by a cost-per-operator figure for life, medical, and dental insurance coverage. The sum of these costs equals total bus operator fringe cost, yielding nearly a 12 percent additive to the cost of operator wages.

Indirect Labor Costs

Indirect labor costs include employer-paid overhead costs of payroll taxes including Social Security, workers' compensation, and unemployment compensation. A tax rate (currently 0.0670 percent on a taxable base of $32 900 annual individual income) is applied to total labor costs to determine Social Security taxes. The rate of taxation is subject to escalating changes as shown in the current Social Security Administration's schedule as past, present, and proposed rates, shown below:

Year	Taxable Base ($)	Rate (%)
1980	25 900	6.13
1981	29 700	6.65
1982	32 900	6.70
1983	33 900	6.70
1985	38 100	7.05
1987	42 600	7.15

Maximum tax levels are established each year in association with the above bases and rates and are also subject to change. Workers' compensation is similarly calculated by applying a tax rate to total labor costs, excluding vacation and merit bonus pay. (Unemployment taxes are paid on an individual basis, rather than as a taxable rate paid to the State of Oregon. The cost model simplifies this calculation by applying an average cost-per-operator factor to the number of operators. Total indirect labor costs represent approximately an 11 percent additive to total wages.

Output

Report output is specified using the "&SELECT" option in the main program. Optional informational reports are available including a distribution of weekday, Saturday and Sunday pay hours, a breakdown of operator requirements by type, and exception pay hours and costs by operator type. Final reports can be printed showing monthly, quarterly, and/or yearly costs. Variable and data listings are also available. Output report options include operator statistical reports--distribution of weekday, Saturday and Sunday pay hours, operator requirements, distribution of exceptions by operator type; and bus operator cost reports--monthly, quarterly, and annual.

Figure 11. Sample financial forecast.

NON-CAPITAL
COST REDUCTION REDUCED/DEFERRED CETIP

	FY 81	FY 82	FY 83	FY 84	FY 85
NON-CAPITAL REVENUES					
Farebox Revenues	19029	19990	23338	26978	29098
Other Operating Revenue	1041	2025	1664	2031	1631
Payroll Tax	34965	37820	40701	45760	50950
State Operating Revenue	---	1000	1100	1150	1200
Federal Operating Assistance	5890	5890	3877	1994	---
Federal Tech/Demo Assistance	1566	1300	100	100	100
Miscellaneous	(12)	---	---	---	---
Interest	2319	2100	1900	1600	1200
New Revenue Source					
TOTAL NON-CAPITAL REVENUE	64798	70125	72680	79613	84179
NON-CAPITAL COSTS					
Bus Operators	25710	29569	31772	32854	35517
Fuel	4944	5415	6312	7157	8139
Maintenance	11178	12911	13904	15085	16489
Operations Adm/Support	5940	7278	9628	10694	12046
General & Administrative	11137	12903	12903	13935	15050
Banfield LRT Project	451	---	---	---	---
TOTAL OPERATING COSTS	59360	68076	74519	79725	87241
Debt Service	336	399	400	480	967
TOTAL NON-CAPITAL COSTS	59696	68475	74919	80205	88208
NET WORKING CAPITAL PROVIDED FROM OPERATIONS	5102	1650	(2239)	(592)	(4029)
CONTINGENCY			1000	1000	1000
BEGINNING WORKING CAPITAL		17181	18831	15592	14000
ENDING WORKING CAPITAL	17181	18831	15592	14000	8971

APPLICATIONS

Tri-Met Experience

Tri-Met commonly uses the bus operator costing model in assessing operating policy and performance, and short- and long-range financial impacts resulting from changes in economic conditions. Probably the most common application is the assessment of service improvements or expansion in the short range. This application also takes into account the associated marginal costs of new increments of service, with leading costs such as the hiring and training of new operators in advance of implementation. It is less effective, however, in the application of assessing reductions in service hours, necessitating layoffs on a low seniority basis, and changing the distribution of the run cuts.

The bus operator model, when applied in concert with all of the cost/revenue models, yields a cash flow status in terms of working capital after all the current period's revenues are summed, less costs, from the previous period's ending working capital figure. Through cash flow analysis, overall financial consequences can be quickly determined. This process provides a tool of cost management, and perhaps more important, an understanding of degrees of cost control. Through a process of sensitivity analysis, in which key variables are changed and the cost impacts are noted, comparisons can be made among alternatives to determine to what extent costs are controllable, and to what degree costs are impacted by policy and performance assumptions.

An example of the cash flow forecast is shown in Figure 11. An analysis was conducted to assess the impact of amounts of proposed new service and timing adjustments regarding implementation under a range of pessimistic, optimistic, and most-likely conditions. The result of these forecasts was the guidance to a series of decisions to pare the proposed amount of service by one-third and defer implementation by nine months. Six-year projections indicated that despite a financially healthy situation today, the expansion of a $7 million service improvement (roughly equivalent to 10 percent of Tri-Met's current annual operating budget) during a period of dwindling revenues threatened to produce a deficit within three years. Sometimes the role of the financial forecaster is to raise a red flag before the agency commits itself to a costly long-term improvement plan.

It is acknowledged that the forecasts are only one input among many other less-structured inputs within the institutional and political processes that simply depend on sound judgment. Adding forecasting to these processes does not make the financial uncertainties any less unpleasant, but by increasing the awareness of those who must make policy decisions, financial forecasting offers direction and depth of insight.