TRANSPORTATION RESEARCH RECORD 663

Recent Developments in Bus Transportation

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1978

Transportation Research Record 663
Price $4.80

subject areas
53 traffic control and operations
84 urban transportation systems

Transportation Research Board publications are available by ordering directly from the board. They may also be obtained on a regular basis through organizational or individual supporting membership in the board; members 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, D.C. 20418.

Notice
The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competence and with regard for appropriate balance.

This report has been reviewed by a group other than the authors 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 this report are those of the authors and do not necessarily reflect the view of the committee, the Transportation Research Board, the National Academy of Sciences, or the sponsors of the project.

This Transportation Research Record is one of a small number of Records in the 1978 series in which the papers have been reproduced from camera-ready copy furnished by the authors. This has been done to provide an opportunity to compare the quality and cost of publication of Records produced in this manner with the quality and cost of Records produced after editing and composition by the TRB staff. A decision as to whether camera-ready copy will be used in the future will be based on the results of this trial. The papers have been subject to the usual review for technical content.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Recent developments in bus transportation.

(Transportation research record; 663)
Includes bibliographical references.
1. Motor bus lines–United States–Management–Congresses.
2. Traffic engineering–United States–Congresses. I. Title. II. Series.
TE7.H5 no. 663 [HE5623] 380.5'08s [388.4'1322] 78-16696
ISBN 0-309-02692-X

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
E. Wilson Campbell, New York State Department of Transportation, chairman

Public Transportation Section
Kenneth W. Heathington, University of Tennessee, chairman

Committee on Bus Transit Systems
Herbert S. Levinson, Wilbur Smith and Associates, chairman
Richard J. Anderson, J. J. Bakker, Alvin G. Basmadjian, John W. Bates, Eugene T. Canty, Tapan K. Datta, Robert B. Deuser, J. R. Doughty, James C. Echols, Curt M. Elmberg, Edward R. Fleischman, Marvin C. Gersten, Ronald I. Hollis, Carol A. Keck, Henry M. Mayer, Neil Craig Miller, Donald A. Morin, Karen L. Paine, George W. Schoene, Franklin Spielberg, Vasant H. Surti, F. V. Webster, H. Donald White

Committee on Public Transportation Planning and Development
George E. Gray, California Department of Transportation, chairman
Joby H. Berman, Patricia Suanne Brooks, Daniel M. Brown, Robert C. Buchanan, John L. Crain, Russell Cummings, Frank W. Davis, Jr., John W. Dickey, James C. Echols, Jacqueline Gillan, David T. Hartgen, F. Norman Hill, William T. Howard, Eugene J. Lessieu, Lillian C. Liburdi, Ray A. Mundy, William T. Olsen, Philip J. Ringo, Gilbert T. Satterly, Jr., George M. Smerk, Donald R. Spivack, William L. Volk, Edward Weiner

Committee on Intermodal Transfer Facilities
John J. Fruin, Port Authority of New York and New Jersey, chairman
Walter H. Kraft, Edwards and Kelcey, secretary
Mark M. Akins, Colin H. Alter, Robert B. Anderson, David L. Andrus, Jr., Marjorie Brink, William J. Hayduk, Clark Henderson, Jeaninne M. Kahan, Hanan A. Kivett, George V. Marks, Jr., Neil Craig Miller, Satoshi Oishi, Robert A. Olmsted, Ira N. Pierce, David W. Randles, Wilfred Sergeant, Stanley A. Spinweber

GROUP 3–OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
Adolf D. May, University of California, Berkeley, chairman

Committee on Transit Service Characteristics
James E. Reading, Central Ohio Transit Authority, chairman
Nicholas E. Bade, Arline L. Bronzaft, Robert C. Buchanan, Howard B. Clarkson, Peter B. Everett, Gordon J. Fielding, Martin Flusberg, Harold H. Geissenheimer, Aaron Isaacs, Judith Kaplan, Karla H. Karash, Thomas Starr King, Brian C. Kullman, Barry D. Lundberg, Faye L. Mench, Ray A. Mundy, Ronald C. Pfefer, Robert W. Pully, Philip J. Ringo, Sherrill Swan, Peter Wood

W. Campbell Graeub, Transportation Research Board staff

The organizational units and officers and members are as of December 31, 1977.

Contents

THE SANTA MONICA FREEWAY DIAMOND LANES: FREEWAY ACCIDENT ANALYSIS

John W. Billheimer, Systan, Inc.

The Santa Monica Freeway Diamond Lanes, a pair of concurrent-flow preferential freeway lanes for buses and carpools linking the City of Santa Monica, California with the Los Angeles Central Business District (CBD), opened on March 16, 1976, and operated amid much controversy for 21 weeks until the U.S. District Court halted the project. One of the most disturbing aspects of the project was the high incidence of freeway accidents, which increased by a factor of 2.5 times pre-project levels when the barrier-free preferential lanes were operating. This paper tabulates accident levels before, during and after the project; postulates and analyses a number of hypotheses regarding potential accident causes; compares the Santa Monica Freeway accident history with that of other preferential lane projects; and identifies the most likely causes of the increased accident levels. This analysis is part of a broader study of the Diamond Lane Project sponsored by the Urban Mass Transportation Administration's Service & Methods Demonstration program. Factors contributing to the increased accident rate included the distracting effect of increased enforcement activities and the congestion resulting from the removal of freeway lanes from general use. However, it appears that the most significant factor was the pronounced speed differential between the free-flowing traffic in the sparsely-occupied preferential lane and the stop-and-go traffic in congested adjacent lanes, coupled with the frequent lane changes made by vehicles entering and leaving the freeway.The experiment in Santa Monica raises serious questions about the use of barrier-free preferential lanes.

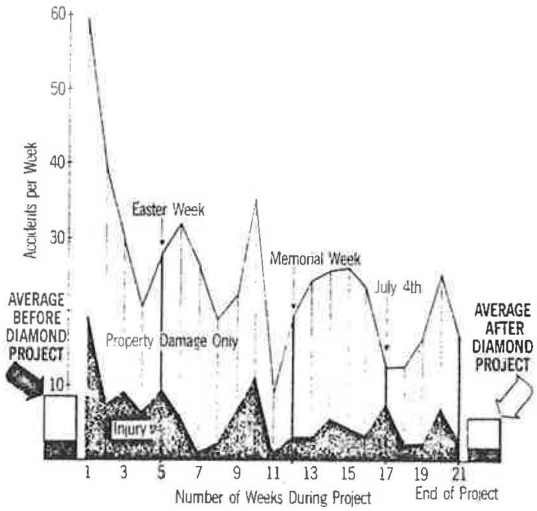
The Santa Monica Freeway, which connects the City of Santa Monica and downtown Los Angeles, is one of the most heavily traveled freeways in the world, and is served by a variety of sophisticated traffic devices, including metered on-ramps with preferential entry provisions at selected locations for two-person carpools, a computerized surveillance system, and centrally-controlled electronic displays. On March 15, 1976, the median lane in each direction of a twelve-mile, eight-lane segment of the Santa Monica Freeway was reserved for the exclusive use of buses and carpools carrying three or more occupants. The reserved lanes, known locally as the Diamond Lanes, operated in each direction during the peak hours of traffic flow. No barriers separated these lanes from the remaining flow of freeway traffic. Implementation of the Diamond Lanes was accompanied by the introduction of a variety of express bus services and the opening of three new Park-and-Ride lots in Western Los Angeles.

The Santa Monica Freeway project marked the first time preferential lanes had been created by taking busy freeway lanes out of existing service and dedicating them to the exclusive use of high-occupancy vehicles. Although the Diamond Lanes entailed no major physical modifications or construction on the freeway itself, they generated considerable physical and emotional dislocation among freeway drivers and other residents of Los Angeles. The first day of operations was disastrous, featuring bumper-to-bumper traffic, long queues at on-ramps, many accidents, outraged drivers, poor press notices and derisive news commentary. As the project progressed, freeway performance improved somewhat and both bus and carpool ridership increased, but accidents remained a problem and the climate of public opinion and media reaction grew more hostile. The preferential lanes operated amid much controversy for 21 weeks until August 9, 1976, when Judge Matthew Byrne of the U.S. District Court in Los Angeles halted the project and ordered additional environmental studies prior to its continuation.

Accident Statistics

Since the first week of operation, when fifty-nine accidents were reported during Diamond Lane operating hours, the total number of reported accidents dropped substantially, with an average of eighteen accidents per week during the last month of the project. Throughout the 21 weeks of the project, 527 accidents were reported during peak operating hours for an average of 25 accidents per week. This number is significantly higher than the average rate experienced prior to the project. Figure 1 (1) plots the average number of accidents occurring per week during the years 1972 through 1975, along with a week-by-week summary of accident levels during the first seven months of 1976. The level of accidents on the Santa Monica Freeway during the Diamond Lane operating hours was more than double the rate experienced during the period immediately preceding the

Figure 1. History of Freeway Accidents During Peak Operating Hours.



project, and more than two and one-half times the average rate experienced during the four years preceding 1976.

Trends Per Million Vehicle-Miles

Since accidents on the Santa Monica Freeway rose at the time that vehicle mileage decreased, the measured increases in accident levels are even more striking when considered in the light of the common index, accidents per million vehicle-miles (MVM). During the operation of the Diamond Lane project, the overall accident rate was 5.1 accidents per MVM. Accidents involving property damage only (PDO) averaged 3.9 accidents per MVM, and injury accidents averaged 1.16 accidents per MVM. The overall accident rate during the Diamond Lane project period was 3.64 times the rate recorded during the same period of the previous year (March 17 to August 12, 1975). This overall rate is biased by the influence of unusually high accidents and unusually low vehicle volumes during the early weeks of the project. As the project progressed, accident levels dropped while vehicle volumes increased, bringing about a steady decline in the accidents/MVM measure, a decline that was still continuing as the project ended. Following the close of the demonstration, accidents dropped below pre-project rates, emphasizing the role of the Diamond Lanes in generating the observed accident increases.

Breakdown By Severity, Time, Direction, Type, And Location

Table 1 (5) compares accident levels on the Santa Monica Freeway by severity, time, direction, type, and location for two time periods: the 21 weeks of the demonstration and a comparable 21-week period in 1975.

The three major categories of freeway accident severity are: (1) fatal; (2) injury; and (3) property damage only (PDO). Historically, few fatal accidents have occurred on the Santa Monica Freeway during peak operating hours--an average of one per year during the few years between 1972 and 1976--and no fatalities occurred during the Diamond Lane project. However, both injury and PDO accidents increased markedly with the implementation of the Diamond Lanes, with injury accidents increasing by a factor of 2.4 over a similar period in 1975 and by a factor of 2.5 over the average level recorded between 1972 and 1976. PDO accidents increased by a factor of 3.1 over the average recorded during the few preceding years.

Injury accidents may be further divided into three subcategories: (a) severe; (b) visible injuries; and (c) complaint of pain. Only one severe accident (slightly less than one percent of all injury accidents) occurred during the Diamond Lane project, while 38% of reported injury accidents entailed other visible injuries, and 61% of injury accidents resulted in complaint of pain. A sampling of injury accidents occurring during 1975 shows a slightly higher incidence of severe accidents (3% of all injury accidents), but statistical tests give no basis for concluding that the Diamond Lanes affected the relative severity of injury accidents on the Santa Monica Freeway.

Almost twice as many accidents occurred during the evening peak as during the morning peak. The dominance of the evening hours coincides with pre-project experience. The greatest relative increase in accidents by time and direction occurred in the eastbound lanes during the evening rush hours. For the corresponding period in 1975, 43 accidents occurred in these lanes during the evening peak. During the evening Diamond Lane operating hours, 178

Table 1. Accident Summary for Santa Monica Freeway Before and During Diamond Lane Operation.

	BEFORE		DURING		
	3/17/75 to 8/8/75 21 Weeks 8 Hours/Day		3/15/76 to 8/6/76 21 Weeks 8 Hours/Day Before 6/17 7 Hours/Day After 6/17		
	Total No.	% of Total	Total No.	% of Total	% of 1975
Total Accidents	180	100%	527	100%	293%
Accidents/Week	8.6		25.1		
Severity					
Fatality	0	0%	0	0%	0%
Injury	50	27.8%	120	22.8%	240%
a. Severe		3%	1	0.8%	
b. Other Visible		39%	46	38.3%	
c. Complaint of Pain		58%	73	60.8%	
Property Damage Only	130	72.2%	407	77.2%	313%
Time/Direction					
Eastbound AM	51	28.3%	108	20.5%	212%
Westbound AM	19	10.6%	65	12.3%	342%
Eastbound PM	43	23.9%	178	33.8%	414%
Westbound PM	67	37.2%	176	33.4%	263%
Eastbound	94	52.2%	286	54.3%	304%
Westbound	86	47.8%	241	45.7%	280%
Type of Collision					
Head-on	2	1.1%	2	0.4%	100%
Sideswipe	23	12.8%	47	8.9%	204%
Rear-end	122	67.8%	422	80.1%	346%
Broadside	12	6.7%	9	1.7%	75%
Hit Object	13	7.2%	42	8.0%	323%
Other	8	4.4%	5	0.9%	63%
Lane					
Median	14+	6.9%	24	4.6%	171%
Diamond Lane	55+	27.0%	27	5.1%	49%
No. 2 Lane	29++	14.2%	310	58.8%	534%
No. 3 Lane	29++	14.2%	51	9.7%	88%
No. 4 Lane	68+	33.3%	33	6.3%	49%
Other	9+	4.4%	82	15.6%	911%

\+ Prior to the project, more than one location was recorded for each accident when multiple collisions occurred. Hence, the total number of lane locations adds to a total greater than the number of accidents.

++ Prior to the project, accident data did not distinguish between the Number 2 lane and the Number 3 lane. A total of 58 accidents recorded in these median lanes have been split evenly between the two lanes.

accidents occurred in this off-peak direction, an increase of 314% over pre-project levels.

Rear-end collisions accounted for 80% of the accidents recorded during Diamond Lane operating hours. During a similar operating period in 1975, rear-end collisions accounted for only 68% of all freeway accidents. Thus, the relative incidence of rear-enders increased significantly during the project, reflecting an increase of stop-and-go conditions in the non-preferential lanes of the freeway.

In addition to the absolute increases in the number of accidents occurring during project implementation, certain changes occurred on the relative pattern of accidents. Perhaps the most notable was the marked increase of accidents in the Number 2 lane adjacent to the Diamond Lane. The number of accidents in the adjacent lane rose from under two accidents per week prior to the project to 14.8 accidents per week during the Diamond Lane operating hours, an increase of more than 13 accidents per week. The average increase in accidents on the entire freeway during project implementation was on the order of 15 accidents per week. Thus, a significant proportion of the overall increase in accidents was concentrated in the Number 2 lane.

Along the length of the project, most accidents occurred on those easternmost sections of the freeway near the CBD, where traffic volumes were highest. In the eastbound lanes, however, the greatest relative increases in accidents occurred farther from the CBD, at the point where cars from the San Diego Freeway entered the flow of traffic. The highest percentage of westbound accidents during both morning and evening hours, before and during the Diamond Lane project, occurred near the CBD, at a point where vehicles from a heavily used collector road and the Harbor Freeway entered the flow of traffic.

Probable Accident Causes

A number of potential causes were identified in an attempt to account for the observed increase in accident levels. These causes stemmed from a variety of factors, including increased CHP presence, increased congestion, the mechanics of Diamond Lane operation, the novelty of the Diamond Lane concept, and exogeneous events. The most prominent of these causes were listed in the form of hypotheses and examined in the light of available data. The remainder of this subsection discusses each of these hypotheses in the light of accident statistics reported during the period of project implementation. Hypotheses:

Accidents were related to increased CHP deployment and enforcement levels. The increased presence of the CHP may have led to increased accident rates for either of two reasons:

A. Minor accidents that previously would have gone unreported were more likely to be reported if more CHP units were present.

B. Increased ticketing rates led to gawking and unexpected slowdowns, causing accidents.

Details of CHP personnel deployment and enforcement activities before and after project implementation may be found in the official evaluation report (1 and 3). During the first week of the project, personnel deployment levels on the Santa Monica Freeway were approximately double pre-project levels. This level was reduced gradually over the demonstration period, so that the average d eployment level over the early weeks of the project was roughly 50% higher than normal. By the thirteenth week of the demonstration, the level of officer deployment approximated that experienced prior to the Diamond Lane project.

In an attempt to discover the extent to which increased CHP deployment and enforcement levels were related to observed accident increases, day-by-day correlations of accidents with both deployment and enforcement levels were undertaken for the period following the implementation of the Diamond Lane project. The results suggest that deployment and enforcement each had some small effect on accident levels, but are inconclusive for determining which effects were greater. The very strong correlation between deployment and enforcement levels makes it difficult to separate the effects mathematically. If overreporting were a significant factor in the accident increase, however, minor (PDO) accidents would have increased at a more rapid rate than more serious injury accidents. This did not occur. In the light of the proportional increases in both major and minor accidents, and the continued high level of accidents once deployment had returned to normal, it appears that any effect of increased CHP presence on Santa Monica Freeway accident levels was more likely to be a result of their ticketing activities than a result of any tendency to overreport minor accidents.

Prior to project implementation, accidents, deployment and enforcement levels were relatively low. Following implementation, accidents increased markedly, decreased, and settled at more than twice pre-project levels. Deployment increased by a factor of approximately 50% during the early weeks of the project and returned to pre-project levels early in the month of June. Enforcement activities, however, increased dramatically with project implementation and continued at levels well in excess of pre-project experience. The number of citations and warnings issued for Diamond Lane and entry ramp violations immediately following project implementation was more than four times the estimated number of citations issued for other traffic violations prior to the project. By the close of the demonstration, the total number of enforcement contacts stemming from illegal use of the Diamond Lane and Freeway on-ramps remained more than double the estimated pre-project level for all traffic violations. Thus, the general pattern followed by enforcement activities before and during project implementation parallels the pattern of accidents. These similar patterns, consisting of a marked increase followed by a decline to a level more than double pre-project levels reinforces the hypothesis that enforcement activities could have contributed to the increased accident level. It is clear from air surveillance traffic reports and observation of Freeway operations that Freeway traffic bunches up in areas in which tickets are being given. This bunching leads to stop-and-go conditions conducive to rear-end accidents.

Analysis suggests that it is unlikely that increased CHP presence on the Santa Monica Freeway led to any significant overreporting of accidents. It appears, however, that the distracting effect of the increased ticketing activities of the CHP may have accounted for some slight portion of the higher accident rate.

Increased accidents are a direct result of increased congestion resulting from the denial of a lane to non-carpoolers. A comparison of accident locations with vehicle volumes along the Freeway reveals the not-unexpected finding that, in general, the heaviest accident locations are found where vehicle volumes are heaviest. The type of accident most prevalent following project implementation--the rear-end collision in the Number 2 lane--would typically accompany increased congestion in that lane. Thus, the observed effects support the congestion hypothesis, although they also support many of the other proposed hypotheses.

No attempt was made to correlate measured congestion by time of day with accidents occurring on that day. The difficulty with this comparison is that accidents are all too frequently the cause of congestion, and hence will go hand-in-hand with measured congestion.

Although congestion undoubtedly contributed to the increased accident rate, three arguments make it seem unlikely that this factor is the primary cause of the marked increase in accidents:

1. Ramp meters were adjusted to minimize the effects of congestion and permit relatively unobstructed flow on the freeway. Speed runs made in the eastbound direction showed that the adjustments to the metered access ramps restored the non-preferential lanes to a condition of flow approximating that in existence prior to the initiation of ramp metering. Yet the average accident rate on the freeway did not exceed ten accidents per week during the two years prior to the introduction of ramp metering. Congestion increases severe enough to double the accident rate should have been reflected in slower operating speeds.

2. In the early months of 1967, the portion of the Santa Monica Freeway between Arlington and LaBrea Avenues was restriped to add a lane in each direction (4). The added capacity was accompanied by a reported accident drop of 10%, and a 15% decline in the accident rate per million vehicle-miles.

3. With the increase in carpool and bus ridership and the concurrent shifting of some drivers to city streets, the total number of vehicles per hour in each of the non-preferential lanes actually dropped slightly at several locations along the freeway.

Thus, if the demonstration project had simply taken one lane of the freeway out of general use, it is unlikely that the marked increase in accidents would have occurred.

Increased accidents may be traced to the barrier-free operation of the Diamond Lanes at speeds well in excess of the speeds in other lanes. The relative lack of vehicles in the Diamond Lane made it possible for vehicles using the lane to travel at speeds well in excess of the speeds in other, more congested lanes. On the average, Diamond Lane vehicles traveled 12 miles per hour faster than the general freeway traffic. This speed differential is considerably higher than that experienced on other preferential lane projects having no separation between reserved and non-reserved lanes. Observers have proposed that this condition may have led to increased accidents for a number of reasons:

1. The speed differential made safe lane changes more difficult to achieve. Motorists attempting to enter the Diamond Lane had to enter a faster traffic stream from a lower starting speed, while motorists attempting to leave the lane had to slow and attempt to find an opening in slower-moving traffic.

2. The ability to save time by using the Diamond Lanes attracted violators who dodged in and out of the lane unsafely, attempting to stay one jump ahead of the CHP.

3. Drivers in Lane 2 accustomed to the relative absence of vehicles on their left in the Diamond Lane, caused accidents by using the preferential lane as a safety valve to avoid rear-enders in their own lane.

4. The speed differential between the Number 2 lane and the faster adjacent lanes deluded the drivers in the Number 2 lane into believing they could travel faster than conditions in their lane allowed. Further, since traffic conditions were different in adjacent lanes, motorists received no cues from these lanes to indicate how conditions in their own lane were changing.

The most promising sources of information regarding the relative likelihood of the accident causes postulated above are the individual accident reports filed by CHP officers. Examination of these reports provides several insights into the relative incidence of these postulated causes.

The difficulty of changing lanes was often cited as a dangerous aspect of Diamond Lane operation. Analysis of vehicle movements prior to collision shows that the relative percentage of accidents in which at least one of the vehicles was changing lanes remained roughly the same before and during Diamond Lane operation. During Diamond Lane operation, 9% of all vehicles were changing lanes prior to the collision, while the corresponding percentage during a comparable period in 1975 was 9.3%.

Although the absolute number of lane-changing accidents increased markedly during Diamond Lane operation, the increase in other types of accidents was just as great or greater. Significant changes were noted in the relative percentage of accidents in which the vehicles involved were slowing, stopping or standing still prior to collision. These increases reflect the increased incidence of rear-enders in the Number 2 lane and the increased level of stop-and-go traffic in all non-preferential lanes.

Attempts to verify the relative importance of unsafe lane changes as a cause of accidents by tabulating the actions of colliding vehivles involved in the accident tend to be inconclusive. It is not uncommon for a vehicle changing lanes unsafely in congested conditions to escape unscathed while leaving a wave of braking vehicles in its wake that culminates in a rear-end collision well removed from the scene of the initial lane change. In such a case, the drivers involved in the collision are generally aware only of the proximate cause of their accident, and the accident report fails to record the lane change that initiated the chain reaction. Thus, although unsafe lane changes in and out of the Diamond Lane might seem to provide a plausible explanation for the observed increase in rear-end collisions in Lane 2, it is impossible to verify this explanation through a study of individual accident reports.

Early in the Diamond Lane demonstration, CHP officers noted that a few accidents were caused by violators dodging in and out of the preferential lane, attempting to stay one jump ahead of a ticket. Examination of the 51 accidents occurring in the Diamond Lane itself or on the median shows that at least five of these accidents were caused by vehicles carrying fewer than three passengers making unsafe lane changes. In three of these cases, the violators had been observed by the CHP prior to the accident.

One possible cause of accidents in the Diamond Lane itself was the sudden entry into the lane by motorists in Lane 2 trying to use the preferential lane as a safety valve to avoid rear-enders in their own lane. It has been proposed that motorists in Lane 2, used to comparative absence of vehicles on their left in the Diamond Lane, may have moved suddenly into that lane in emergencies, posing a hazard for faster moving traffic in the preferential lane. If this was a serious cause of accidents, it should show up in the reports of CHP officers.

A breakdown of 27 accidents occurring in the Diamond Lane itself during the project shows that 13 of these accidents, or 48%, were caused by vehicles swerving into the lane to avoid trouble in their own lane, and colliding with a Diamond Lane vehicle. An additional nine accidents, or 33%, were caused by unsafe lane changes made by vehicles facing no threat in their own lane. The remaining five Diamond Lane accidents, or 18%, were rear-end collisions between vehicles already in the Diamond Lane.

Vehicles swerving into the Diamond Lane to avoid rear-enders in their own lane were often out of control. In effect, they represented accidents about to happen, and the final nature of the accident depended only on whether there was an oncoming vehicle in the Diamond Lane. On at least 13 occasions, drivers originating in the Number 2 lane spun out to avoid rear-enders and collided with the highway median driver. On at least two other occasions, Diamond Lane drivers were forced into the median by automobiles bailing out of the Number 2 lane to avoid trouble. All of these accidents originated with stop-and-go conditions in the Number 2 lane, even though they were not reported as accidents in that lane.

It seems clear that the combination of high Diamond Lane speeds, when coupled with slow stop-and-go traffic in the non-preferential freeway lanes, contributed to the observed increase in accidents during the Diamond Lane demonstration. The exact extent of this contribution is impossible to determine. Under normal operating conditions, an accident-related slowdown in one lane generally results in a slowdown in all lanes. Given the reserved nature of the Diamond Lane, however, a slowdown in the remaining lanes usually just accentuated the speed differential between the Diamond Lane and the remainder of the freeway traffic. In recognition of the potential danger accompanying the juxtaposition of high Diamond Lane speeds and congested, stop-and-go traffic on the remainder of the freeway, Diamond Lane bus drivers were instructed not to exceed the speed of other freeway traffic by more than 30 miles per hour. There is some evidence, however, that the driver carpooling in the Diamond Lane failed to exercise such prudence when other lanes had slowed to a halt due to accidents or congestion.

The impact of the speed differential on accidents was exacerbated by the need for carpoolers to exit at many points along the freeway. The non-CBD orientation of Los Angeles traffic meant that carpool drivers had to slow and weave their way through stop-and-go traffic to exit at many points along the freeway. Carpoolers responding to the driver survey cited problems exiting from the Diamond Lanes as the greatest single difficulty encountered in using the lanes. Accidents might have been reduced somewhat if more drivers had followed the preferential lane to the end, where merging problems were minimized. The sprawling, multi-centered nature of the Los Angeles area, however, increased the need for carpoolers to merge with slower traffic all along the freeway, thereby increasing the safety hazard associated with the inter-lane speed differential.

Accidents are caused by the novelty of the Diamond Lane itself and the controversy surrounding it. The novelty of the Diamond Lane concept and the controversy surrounding it may have been a source of accidents for several reasons:

1. Driver confusion and experimentation in the early weeks of the project undoubtedly led to higher accident levels.
2. Faster movement in the preferential lanes tended to distract drivers in other lanes, making them more susceptible to accidents.
3. Driver aggravation with the concept may have led to reckless, aggressive driving.

There seems to be little doubt that the surge of accidents during the first two weeks of the project may be traced to the newness of the concept and driver uncertainty regarding the use of the lane. Accident increases have been experienced in the early weeks of other preferential lane projects in which no barrier separates buses and carpools from the remaining lanes of traffic. After initial increases, freeway accidents in both Portland and Miami, where preferential lanes were created by adding a lane to the existing traffic flow, dropped below pre-project levels by the second month of operation.

Drivers of one- and two-passenger automobiles reported a tendency to count the heads appearing in cars whizzing by in the preferential lane. To the extent that the single-occupancy automobile driver persisted in this headcounting, he was less likely to be able to control his own vehicle in an emergency. As drivers became used to the preferential lane, this tendency should have diminished. Although accident rates also diminished as the project continued, there is no sure way to determine the extent to which decreased headcounting accounted for decreased accidents. Specific sources of driver distraction were rarely noted in accident reports, and it is possible that the driver himself may not have been aware of the distraction or may not have wished to admit his inattention to the reporting officer.

Driver frustration and aggravation with the Diamond Lane itself may have contributed to the accident increase. One accident expert noted that the level of frustration would have been especially high among the aggressive drivers used to driving in the Number 1 lane. In testimony before the U.S. Superior Court, Paul O'Shea noted, "...you are taking the aggressive driver and the confident driver, and because he is not entitled to the Diamond Lane, putting him over into the slower traffic, which creates a tremendous frustration." Unquestionably, such frustration did exist, as manifested in the public outcry against the project. However, it is impossible to estimate the extent to which such private frustration may have increased the accident level. Examination of CHP accident reports shows that at least two accidents occurring during Diamond Lane hours may be traced to public frustration with the concept itself. On June 3, drivers opposed to the Diamond Lane concept staged a funeral procession in the lane to protest the lane's existence. The distraction resulting from this demonstration was listed as a contributing factor in two accidents occurring on that day.

Accident explanations not related to the Diamond Lane project. The overall accident level on all Los Angeles freeways has been increasing steadily since the early months of 1974, following the gasoline crisis and the introduction of the 55 m.p.h. speed limit. This increase has been less pronounced during peak operating hours, when there is less chance of a vehicle exceeding the 55 m.p.h. speed limit. On the Santa Monica Freeway during peak operating hours, a linear least-squares regression from January 1974 to the start of the Diamond Lane project shows that the accident level increased by .0215 accidents per week over this period. Extrapolation of this trend would lead one to predict an increase of 0.45 accidents per week during the 21 weeks of the Diamond Lane operation. This represents a small portion of the observed increase of 13.7 accidents per week accompanying the Diamond Lane project, indicating that general trends existing prior to the project had little effect on the accident situation. As noted, moreover, freeway accident rates dropped below pre-project levels following the close of the Diamond Lane demonstration, further discouraging any arguments that causes unrelated to the operation of the Diamond Lanes contributed to the pronounced increase in accidents during project operating hours.

Accident Levels of Other Preferential Lane Projects

In an effort to shed additional light on the causes behind the accident increase experienced on the Santa Monica Freeway during the Diamond Lane demonstration, accident levels on other barrier-free preferential treatment projects were investigated. Four other projects were selected for comparison: I-95 Freeway, Miami; South Dixie Highway, Miami; Banfield Freeway, Portland; and U.S. 101 Freeway, Marin County, California.

Physical Characteristics

The physical characteristics of each of these projects are summarized below.

I-95 Freeway, Miami. On this 7.5 mile freeway segment, a new barrier-free preferential lane for buses and carpools was created from the median shoulder and opened in December 1975. There are three and four lanes in each direction over the length of the project.

South Dixie Highway, Miami. This is a 5.5 mile segment of highway which runs north to the central business district of Miami. An existing concurrent-flow lane was reserved for carpools (two or more persons) and a contra-flow lane for buses was opened in July 1974. There are three lanes in each direction.

Banfield Freeway, Portland. On this freeway segment (3.2 miles inbound and 1.7 miles outbound), a new barrier-free preferential lane for buses and carpools was created from the median shoulder in December 1975. The roadway has three lanes in each direction overall, with one four-lane segment.

U.S. 101 Freeway, Marin. On U.S. 101 north of the Golden Gate Bridge, a 3.9 mile northbound contra-flow lane has been established. From the Richardson Bay Bridge north, new 3.8 mile northbound and southbound concurrent-flow lanes for buses and carpools were created from the median and shoulder.

On all of these projects except the South Dixie Highway, a new lane was created from the median and/or shoulders; on the South Dixie Highway, as on the Santa Monica Freeway Diamond Lane project, an existing lane was reserved. None of the projects have barriers between the preferential lane and remaining lanes. For the South Dixie Highway project, a carpool is defined as two or more persons; in January 1977, the Miami I-95 project also changed to two-person carpools. In the other three projects, a carpool is defined as three or more persons. All the preferential lanes operate only during peak hours. Speed statistics by lane are available for only three of the projects surveyed. The approximate average speed differentials are 12 miles per hour (mph) on the Santa Monica Freeway, 11 mph on the South Dixie Highway, and 6 mph on Prtland's Banfield Freeway.

Accident Levels

Table 2 summarizes both injury accidents and total accidents per million vehicle-miles for each of the four projects surveyed, in addition to the Santa Monica Freeway Diamond Lanes. Iniury accidents are plotted as a function of time in Figure 2. Because accidents involving property damage only (PDO) are not reported uniformly in all states, reports of injury accidents provide a sounder base for comparing the various projects.

Of the five projects compared, the two with the highest increases in total accidents per million vehicle miles over the "before" period are the Santa Monica Freeway and the South Dixie Highway, with accident rate increases of 264% and 89%, respectively. During both projects, accidents were initially high, but decreased during the succeeding months. Accidents also increased significantly in Marin, rising 82% above pre-project levels during the first years of the project. After initial increases on Portland's Banfield Freeway and Miami I-95, accident rates have varied above and below pre-project averages, but the increases have not been so marked as those of the Santa Monica Freeway, South Dixie Highway, and Marin projects. On Miami I-95, moreover, accident rates appear to have dropped slightly since project implementation. Both the Santa Monica Freeway and South Dixie Highway projects exhibit the highest accident increases and have certain physical and operating characteristics which the other three projects lack: Both were created by removing a lane from existing traffic, and both have a median pull-over area for enforcement use. Thus, enforcement levels on both projects have been high, and violation rates comparatively low. The congestion caused by lane removal and the distracting effects of violators help to account for the increased accident rates in both projects.

Implications of Accident Analysis

Summary of Findings

The creation of the Diamond Lanes on the Santa Monica Freeway through dedication of an existing lane to buses and carpools increased peak hour accidents on the freeway from approximately ten accidents per week to 25 accidents per week. The relative severity of accidents did not change significantly with the project. However, the relative percentage of rear-end accidents occurring in the Number 2 lane rose remarkably from under two accidents per week to 14.8 accidents per week.

A number of potential causes have been identified in an attempt to account for the observed increase in accident levels. These causes stem from a variety of factors, including increased CHP presence, increased congestion, the mechanics of Diamond Lane operation, the novelty of the Diamond Lane concept, and exogenous events. While it is likely that each of these were contributing factors in some instances, in light of the accumulated data, it seems unlikely that certain of the potential causes had a major influence on the accident picture. The increased CHP ticketing activities do not provide a direct explanation for the remarkable increase in accidents in the Number 2 lane, and a correlation of accident and enforcement levels during the demonstration period explains a relatively small proportion of the observed accident variation. Furthermore, equivalent congestion levels existed on portions of the freeway prior to both the 1967 lane enlargement and the introduction of ramp metering without causing pronounced accident levels.

The one potential cause which could not be discounted, and which does in fact appear to account for a large share of the accident increase is the pro-

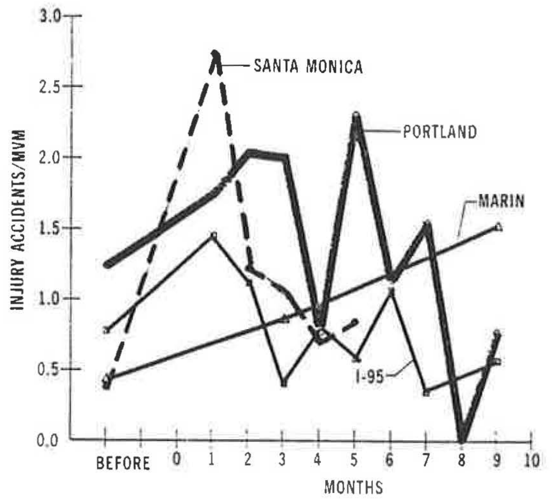
Table 2. Summary of Injury and Total Accidents Per Million Vehicle-Miles.

Project	Period Included in "After" Measurement (Months)	Injury and Fatal Accidents/MVM			Total Accidents/MVM		
		Before	After	% Increase (Decrease)	Before	After	% Increase (Decrease)
Miami I-95	9	0.78	.76	(-2.5%)	3.55	3.44	(-3.1%)
Miami South Dixie Highway	8	NA*	NA*	--	6.4	12.1	89%
Portland	9	1.24	1.37	10.5%	2.48	2.55	2.8%
Marin**	18	0.50	1.31	162%	2.33	4.23	82%
Santa Monica	4	0.40	1.16	190%	1.40	5.10	264%

*NA = data not available.

**Accident data based on the average for three non-continuous six-month periods.

Figure 2. Preferential Lane Project Injury Accidents Per Million Vehicle-Miles.



nounced speed differential resulting from the combination of unhindered traffic in the sparsely occupied preferential lane and congested conditions in the remaining lanes, coupled with the lack of barriers between lanes and the variety of possible destinations along the freeway.

Because of the shortened duration of the project, the effect of Diamond Lane novelty on accident levels can never be known with certainty. The tendency to gawk and count heads of passing carpoolers would certainly have diminished with time. It is not possible, however, to project with confidence the accident level that would have existed following a longer period of operation.

Implications for Planners

Given the nature of the most likely explanations for the increased accident rate, several occurrences could have brought about a decline in accident levels. To the extent that usage of the preferential lane increased with time, the speed differential would decrease as the preferential lane became more crowded and non-preferential lane congestion was reduced by the elimination of defecting carpoolers and bus riders. The reduction of CHP enforcement levels would also work in two ways to reduce the level of accidents: by eliminating the distraction of ticketing, and by permitting more violators to shift to the preferential lanes, thereby cutting the speed differential and easing congestion in the non-preferential lanes. Thus, to some extent, the elimination of either of these two accident sources tends to work against the presumed concept of the preferential lane. As the speed differential is reduced, so also is the inducement to use the lane. Moreover, any decision to relax enforcement must, by encouraging violators, run counter to the philosophy of a lane reserved for high-occupancy vehicles.

The apparent dilemma whereby reduced accidents might be achieved at the cost of lane operating efficiency highlights the delicacy of the control problem faced by planners attempting to design barricade-free preferential lanes for use in mixed traffic. On the one hand, if the preferential lane operates below capacity with a significant speed differential relative to adjacent congested lanes, accidents are almost certain to increase. If the lane is allowed to fill, however, either by allowing violators to infiltrate or by relaxing the requirements for the use of the lane(i.e., by allowing two-person carpools), much of the inducement for using the lane vanishes. In theory, the number of carpools should grow over time until the marginal amount of time saved by switching to a carpool exactly balances the perceived inconvenience of making the switch. In practice, the level of accidents occurring before this equilibrium point is reached may be unacceptable to society, or the equilibrium point itself may result in an unacceptable accident rate.

The specter of increased accidents raises serious questions regarding the feasibility of the barrier-free preferential lane concept. These questions appear to exist whether the lane is created by reserving an existing lane, as was done on the Santa Monica Freeway, or by creating an entirely new lane, as has been done in Portland, Miami, and Marin County, and was originally contemplated for the San Diego Freeway in Los Angeles. The extent of the problem is difficult to assess at present. Although accidents have risen markedly in Marin, neither Portland nor Miami has experienced significant accident increases to date. In both Portland and Miami, however, enforcement activities are reduced, the influx of violators is relatively heavy, and the speed differential is not so great as in the Santa Monica project. Conceivably, the addition of a new preferential lane to an existing freeway could also result in increased accidents if conditions similar to those on the Santa Monica freeway exist. Further investigations of the relationship between accident levels and the operation of barrier-free preferential lanes should be undertaken as soon as possible so that the risks attending these operations may be more clearly defined.

1. John W. Billheimer, Robert Bullemer and Carolyn Fratessa. The Santa Monica Freeway Diamond Lanes: An Evaluation.SYSTAN, Inc. for the Transportation Systems Center, U.S. Department of Transportation, Contract DOT-TSC-1084, April 1977.
2. California Department of Transportation. 21-Week Report on the Diamond Lanes, August 1976.
3. California Highway Patrol. Operation 500: A Study of the Effect of Increased Road Patrol, Final Report. Sacramento, California, April 1972.
4. California Transportation Agency, Department of Public Works, Division of Highways. Effect on Traffic Operation of Use of Shoulder as Traveled Way on Portions of the Santa Monica Freeway. Internal Report No. 68-1 of the Freeway Operations Department, District 7, Los Angeles, January 1968.
5. Memorandum from C.P. Sweet to C.E. Forbes, Caltrans. January 11, 1977.

THE SANTA MONICA FREEWAY DIAMOND LANES: EVALUATION OVERVIEW

John W. Billheimer, Systan, Inc.

The Santa Monica Freeway Diamond Lanes, a pair of concurrent-flow preferential lanes for buses and carpools linking the City of Santa Monica, California with the Los Angeles Central Business District (CBD), opened on March 16, 1976 and operated amid much controversy for 21 weeks until the U.S. District Court halted the project. The Diamond Lane project marked the first time preferential lanes had been created by taking busy freeway lanes out of existing service and dedicating them to the exclusive use of high-occupancy vehicles. Although the Diamond Lanes entailed no major physical modifications or construction on the freeway itself, they caused significant physical and emotional dislocation among freeway drivers, public officials and other residents of Los Angeles, and generated considerable controversy regarding the reported and actual impacts of the project. This paper summarizes the findings of the official, objective, independent evaluation of the project sponsored by the U.S. Department of Transportation as part of the UMTA Service and Methods Demonstration Program. The paper addresses a broad range of project impacts in the following major areas: traffic speeds and travel times; traffic volumes and carpool formation; bus operations and ridership; safety and enforcement; energy and air quality; and public attitudes and response. Analysis shows that the project succeeded in increasing carpool ridership by 65% and the increased bus service accompanying the Diamond Lanes caused bus ridership to more than triple. Nonetheless, energy savings and air quality improvements were insignificant, freeway accidents increased significantly, non-carpoolers lost far more time than carpoolers gained, and a heated public outcry developed which has delayed the implementation of other preferential treatment projects in Southern California and given planners and public officials in other areas ample cause for reflection before attempting to implement similar projects.

The Santa Monica Freeway, which connects the City of Santa Monica and downtown Los Angeles, is one of the most heavily traveled freeways in the world, and is served by a variety of sophisticated traffic control devices, including metered on-ramps with preferential entry provisions at selected locations, a computerized surveillance system, and centrally-controlled electronic displays. On March 15, 1976, the California Department of Transportation (CALTRANS), acting in conjunction with the California Highway Patrol (CHP) and local bus operators, reserved the median lane in each direction of a 12-mile, eight-lane segment of the Santa Monica Freeway for the exclusive use of buses and carpools carrying three or more occupants. The reserved lanes, known locally as the Diamond Lanes, operated in each direction during the peak hours of traffic flow. No barriers separated these lanes from the remaining flow of freeway traffic. Implementation of the Diamond Lanes was accompanied by the introduction of a variety of express bus services and the opening of three new Park-and-Ride lots in Western Los Angeles.

The project neither started nor ended as scheduled. The original starting date was delayed by a combination of concerns including operational readiness, financial problems, a local dispute over the implications of nationwide labor protective agreements, and the Southern California rainy season. When the Diamond Lanes finally opened, the first day of operations was disastrous, featuring bumper-to-bumper traffic, long queues at on-ramps, a malfunctioning ramp meter, many accidents, outraged drivers, poor press notices, and derisive news commentary. As the project progressed, freeway performance improved somewhat and both bus and carpool ridership increased, but accidents remained a serious problem as the climate of public opinion and the media reaction grew more hostile. The preferential lanes operated amid much controversy for 21 weeks until August 9, 1976, when Judge Matthew Byrne of the U.S. District Court in Los Angeles halted the project and ordered additional environmental studies prior to its continuation.

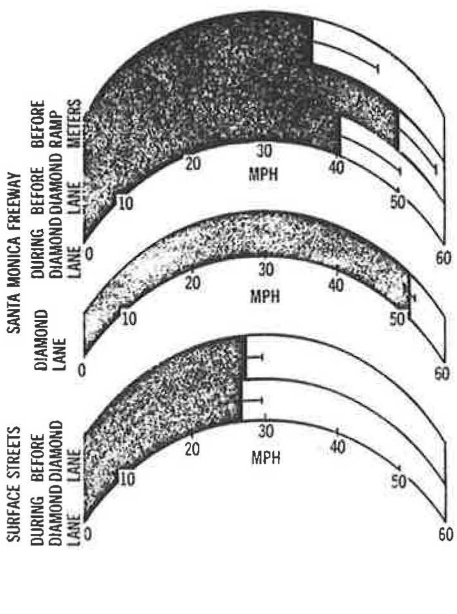
Much of the controversy surrounding the Diamond Lanes consisted of conflicting claims regarding the ability of the project to accomplish its stated objectives of conserving energy, improving air quality, and expanding effective freeway capacity by increasing the occupancy of buses and automobiles using the freeway. An independent analysis of the vast quantities of data assembled by both friends and foes of the project reveals that, although some of the stated objectives had been attained by the close of the demonstration, the cost in accidents, driver delay, and public outrage was far greater than anyone had anticipated. Major findings of the analysis are summarized below.

Traffic Speeds and Travel Times

Vehicle Speeds

The dedication of the Diamond Lanes to the exclusive use of buses and high-occupancy vehicles, and the accompanying changes in ramp metering rates, had a marked impact on vehicle speeds on the Santa Monica Freeway. *The speeds of vehicles using the Diamond Lanes were significantly faster and steadier than the speeds of vehicles in the remaining non-preferential lanes, which were generally slower and less predictable than pre-project speeds.* Although the speeds of non-carpoolers improved as the demonstration progressed, they never returned to meter-controlled, pre-project levels. Average freeway driving times for non-carpoolers traveling the full length of the project over the last seven weeks of the demonstration were slightly more than one minute longer than pre-project levels in the westbound direction during the evening and more than four minutes longer in the eastbound direction during the morning.

Figure 1. A.M. Eastbound Travel Speeds



Average statistics do not provide a complete picture of travel times in adjacent lanes, since these times varied considerably during the morning and evening hours of operation. The non-carpooler entering the eastbound freeway at 6:30 A.M. found his travel time to Los Angeles increased by roughly one minute during the demonstration. By 8:00 A.M., however, the additional delays in freeway travel time approached nine minutes per trip. The difference between the average travel time measured over the full span of Diamond Lane operating hours and the actual travel times experienced by motorists during specific peak travel periods, coupled with the greater uncertainty associated with travel during Diamond Lane operations, helps to explain some of the skepticism reported in the press while the project was still in progress.

Speeds recorded by carpoolers in the Diamond Lanes were both faster and more consistent than pre-demonstration speeds. Carpoolers traveling the full length of the Diamond Lanes were able to save 2 to 3 minutes over pre-project travel times in other lanes.

Aggregate travel speeds on the surface streets paralleling the freeway slowed slightly during the demonstration, dropping by about 4.5% as former freeway users transferred to surface streets.

Entry Ramp Conditions

Over a period of two years prior to the Diamond Lane demonstration, traffic signals were installed on the Santa Monica Freeway on-ramps to control the number and spacing of cars entering the freeway during the peak hours. Before these ramp meters were installed, general vehicle speeds on the freeway were slightly slower than the speeds attained by non-carpoolers during the Diamond Lane demonstration. The installation of these ramp meters greatly improved traffic speeds on the freeway by limiting entering vehicles to a fixed rate of flow. Although vehicles entering the freeway spent an average of two minutes waiting at the ramp meters, this delay was more than offset by the time saved in traveling in the improved traffic conditions on the freeway itself.

Metering rates on most freeway access ramps were adjusted during the week preceding the opening of the Diamond Lanes. In some instances, these adjustments represented severe departures from pre-project conditions. The adjustments were designed to alleviate anticipated freeway congestion and, in most cases, increased the length of time motorists were required to wait in queues before entering the freeway. As the project continued, metering rates were readjusted in response to actual traffic conditions, but these attempts to fine-tune the system did not match the sweeping changes made before opening day in either the magnitude of the adjustments or the number of ramps affected.

Once the confusion and adjustments of the first week were past, few changes in ramp delays were observed during the peak hours of travel. *Average delays at the metered ramps carrying the bulk of entering traffic increased between one and five minutes per car during the project.*

At 12 of the 30 metered entry ramps, preferential access lanes permitted buses and vehicles with two or more occupants to bypass the metered system. The bypass lanes at these selected ramps saved buses and two-person carpools between two and seven minutes per trip during the Diamond Lane demonstration.

The average increase in queue lengths at freeway on-ramps were not so pronounced as the increases in ramp waiting times. There were relatively few instances in which the Diamond Lane metering changes caused vehicle queues to extend dramatically beyond the ramp storage capacity, and speed measurements showed that the queue increases did not appear to cause additional interference with traffic on north-south feeder roads.

Total Trip Times

Measured Freeway Trip Times. Considering both ramp delays and slower freeway speeds, measured increases in average trip times for non-carpoolers traveling eastbound on the freeway in the morning were as high as six to seven minutes per trip for those drivers starting at one end of the project and traveling to the other. Drivers entering the freeway about midway along the length of the project experienced negligible increases in total travel times.

At each of the entry ramps with a bypass lane for buses and two-person carpools, the amount of time saved by using the ramp bypass exceeded the amount of time saved by traveling in the Diamond Lane to the lane's end. *That is, the relative delays imposed on single-occupant automobiles at preferential on-ramps were greater than those imposed by the Diamond Lane itself.*

Perceived Trip Times. The changes on freeway travel times encountered during the Diamond Lane demonstration may also be viewed in the light of the total door-to-door commuting times perceived by drivers in the freeway corridor. The average door-to-door trip reported by a sampling of 2,800 corridor drivers was 21 minutes long, and took 37.4 minutes in the morning and 43.2 minutes in the evening. Diamond Lane carpoolers responding to a survey questionnaire reported an average savings of 1.5 minutes over pre-project travel times. Non-carpoolers reported an increase in trip times of 8.3 minutes in the morning and 9.4 minutes in the evening. These perceived increases are slightly higher than freeway measurements indicate are likely, and include a number of impossibly high reports (greater than 30 minutes) of average trip delays. Not unexpectedly, non-carpoolers appear to have overestimated the average delays accompanying the Diamond Lane demonstration. Given the increased uncertainty accompanying travel in the non-preferential lanes, however, delays on any single day could have been much higher than the average measured increase of six to seven minutes.

Traffic Volumes

Freeway Traffic Volumes

The changes in travel speeds experienced during the demonstration were accompanied by significant shifts in traffic patterns. The total number of vehicles and people using the Santa Monica Freeway dropped markedly during the early weeks of the demonstration, and then rose steadily. The early decline in freeway traffic reflected a combination of carpool formation, growing bus ridership, and defection to surface streets by non-carpoolers. *By the close of the demonstration, the number of people using the easternmost segments of the freeway was within 2% of pre-project levels, while vehicle volumes had declined by 10%.* Summary Table 1 provides more detail on changes before, during and after the project as measured at observation points near the Los Angeles CBD.

Table 1. Average Daily Vehicle and Passenger Statistics, Santa Monica Freeway at Crenshaw Boulevard.

(Seven-Hour Peak Periods, Both Directions of Travel)

Statistic	Before Project	During Diamond Lane Project: First Seven Weeks	Second Seven Weeks	Final Seven Weeks	After Project
Total Vehicles	113,135	76,738	97,197	101,678	112,059
	--- *	(-32%)*	(-14%) *	(-10%) *	(-1%)*
Total People	138,873	101,643	128,180	136,421	140,507
	--- *	(-27%)*	(-8%) *	(-2%) *	1% *
Bus Ridership	1,171	3,092	3,569	3,810	2,916
	--- *	164% *	205% *	225% *	149% *
Passengers/ Vehicle	1.23	1.32	1.32	1.34	1.25
	--- *	8% *	7% *	9% *	2% *
Three-Person Carpools	3,479	4,345	4,923	5,749	3,652
	--- *	25% *	42% *	65% *	5% *

* Number
% Increase
(Decrease)

Measurements made at different points along the freeway reflect the same general pattern of usage depicted in Summary Table 1, although shifts in vehicle and passenger movement were less pronounced at locations farther removed from the CBD. Although directional trends on the Santa Monica Freeway are less pronounced than on most major freeways, the greatest changes in vehicle and passenger movement during the demonstration occurred in the peak directions of travel (eastbound in the morning and westbound in the evening), where congestion was greatest in the non-preferential lanes. By the last seven weeks of the demonstration, the freeway carried an average of 9% fewer people in 17% fewer vehicles in the peak directions of travel.

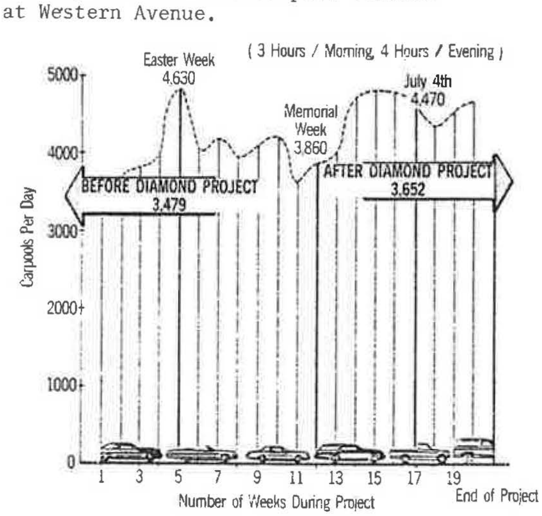
Vehicle volumes at all measuring points increased over pre-project volumes during the midday hours when the Diamond Lanes were not operational. The extent of the increase on vehicle volumes between the hours of 10:00 A.M. and 3:00 P.M. ranged between 2% and 6 % over pre-project levels, indicating that drivers who had some flexibility in their choice of travel times elected to travel during the midday lull rather than face the much-publicized freeway congestion during Diamond Lane operating hours.

Prior to the project, each lane of the Santa Monica Freeway carried approximately 1,800 vehicles per hour during peak periods of flow. During the project, the Diamond Lanes carried an average of 300 vehicles per hour in the peak eastbound direction and 500 vehicles per hour in the peak westbound direction. Thus, the preferential lanes operated at between 20 and 30% of their vehicular capacity, and appeared relatively empty when compared with the heavily congested adjacent lanes. Even so, the number of people carried by the Diamond Lanes approached the number carried by the remaining lanes by the end of the project, and the unused capacity in each preferential lane supplied the Santa Monica Freeway with at least as much reserve capacity as two additional lanes operating at pre-project occupancy rates.

Carpool Formation

The number of carpools carrying three or more people on the Santa Monica Freeway increased significantly during the demonstration, rising 65% above pre-project levels by the last seven weeks of the project. The growth of carpool usage was relatively steady throughout the project, with pronounced peaks

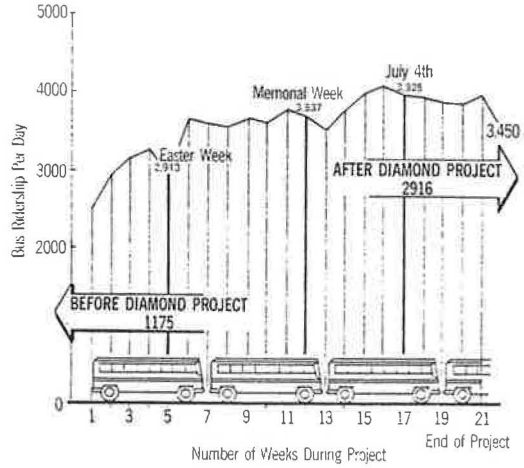
Figure 2. Daily Diamond Lane Carpool Volumes Counted at Western Avenue.



during vacation periods. Afternoon traffic in the eastbound Diamond Lane increased markedly during the Easter holiday week, and rose steadily following Memorial Day, suggesting that much of the increased Diamond Lane usage during these vacation periods may be attributed to groups of vacationing beachgoers returning from the ocean. Although no formal data were assembled to support this observation, Diamond Lane observers noted a number of surfboard sightings during the periods in question.

The average size of the carpools using the Diamond Lanes was 3.4 people. The primary incentive for forming a carpool mentioned by most of the carpoolers surveyed (63%) was to save money. Only 25% of the carpools responding to the survey were initially formed during the Diamond Lane demonstration period, and only 30% of these carpoolers identified the Lanes as the primary incentive behind their decision to carpool. With the disappearance of the Diamond Lanes, the number of cars on the freeway dropped to within 5% of pre-project levels, suggesting that the Lanes themselves were more of an incentive to those carpools formed during the demonstration than the survey responses indicated.

Figure 3. Average Daily Peak Period Bus Ridership on all Santa Monica Freeway Project Routes in the Peak Direction of Travel only.



Surface Street Volumes

Traffic volumes on surface streets parallel to the freeway rose between 10% and 15% shortly after the demonstration began, then appeared to subside somewhat during the summer months. Surface street vehicle occupancy rates did not change significantly during the demonstration.

By taking into account Santa Monica Freeway users, surface street travelers, and former Santa Monica Freeway users traveling on different freeways or during less congested time periods, a rough comparison of vehicle and passenger movement across the entire Santa Monica Freeway corridor near the CBD indicates that *by the last seven weeks of the project, 1% more people were traveling in 5% fewer vehicles than were being used prior to the demonstration.*

Bus Operation and Ridership

Operations

Two bus operators in the Los Angeles area participated directly in the Santa Monica Freeway Preferential Lane project by offering new services in conjunction with the opening of the Diamond Lanes: The Southern California Rapid Transit District (SCRTD), which operates 2,400 buses in the four-county Los Angeles area, and the Santa Monica Municipal Bus Lines (SMMBL), which operates about 100 buses in the Santa Monica area.

Bus service linking the Westside study area to the Los Angeles CBD improved significantly with the implementation of the demonstration. The addition of four new Diamond Lane feeder/express routes to the four already serving the Westside area more than doubled the number of Westside CBD workers living within walking distance of the express bus service. In addition, three new Park-and-Ride routes were introduced to serve those Westside residents who were not within walking distance of a feeder/express route.

On the first day of the demonstration, 74 express bus trips were offered from the Westside area to the Los Angeles CBD during the morning peak, an increase of more than four times pre-project levels. Even without the Diamond Lanes, therefore, the marked improvement in service significantly improved the travel time by bus from most sections of the study area to the CBD. In the initial stages of the demonstration, service headways on new routes were generally set so that buses were no more than 15 minutes apart. As the project progressed, headways were adjusted to reflect ridership.

The introduction of the Diamond Lanes significantly improved the on-time performance of those SCRTD routes in existence prior to project implementation, cutting two minutes off the freeway travel time of the busiest line. Diamond Lane buses also generally exhibited better on-time performance than buses using other freeways without preferential treatment and buses using surface streets.

Ridership

Daily bus ridership between the Westside study area and the Los Angeles CBD increased from 1,171 riders per day prior to the project to 3,793 riders per day during the last week of Diamond Lane operation. Bus ridership rose rapidly during the first month following implementation, and continued to grow throughout the project. While the growth patterns were essentially the same for both SCRTD and SMMBL, SMMBL carried 26% of the combined average daily ridership with only 15% of the total daily bus trips. By the close of the project, most of SMMBL vehicles were fully occupied, and the average occupancy during the project was 41.1 riders per trip, an occupancy rate of 82 percent.

In the case of SCRTD, the average occupancy during the project was 19.2 riders per trip, or 38% of the available seating capacity. This figure was well below pre-project levels, and stemmed from the policy decision to provide as much service as possible early in the project to maximize the possibility of attracting ridership. While the policy appears to have had the desired effect, it also put a large number of near-empty buses in public view during the early stages of the project. As the project continued, unprofitable runs were eliminated and SCRTD occupancy rates improved markedly.

In general, both the new feeder/express routes and those routes existing prior to the demonstration succeeded in attracting patronage from the ranks of automobile drivers during the project, and an overwhelming majority of the bus riders surveyed expressed sat sfaction with the service. *By the close of the project, the eight feeder/express routes had come* *close to meeting the aggregate long-term demand predictions for patronage on these routes, carrying nearly 30% of the CBD-destined trips projected to be within walking distance of a bus line.* The three new Park-and-Ride routes, however, fell far short of expectations and were all discontinued by September 1.

After the close of the demonstration and a five-week SCRTD bus strike, ridership on those freeway routes remaining in service was 17% below the peak attained during Diamond Lane operations. Ridership declines were greatest on those routes reporting the longest door-to-door travel times. Ridership drops were lowest on the one SMMBL route which continued operating through the strike. By early 1977, however, none of the routes had succeeded in attaining the peak ridership levels attained during the demonstration.

Attempts to isolate the impact of the Diamond Lanes themselves on bus ridership are frustrated by the short, uncertain life of the project, seasonal patronage variations, the media blitz, frequent and major changes in bus service frequency, fare increases, and the five-week strike of SCRTD workers which followed the closing of the demonstration. Recognizing these uncertainties, it can be argued that the extent of the Diamond Lanes' influence can at least be bounded by surviving ridership levels. If, in the light of service cutbacks, fare increases and a five-week strike, subsequent ridership levels still managed to rise to within 17% of their peak during Diamond Lane operations, it would seem that this 17% figure represents a fair estimate of the maximum drawing power of the Diamond Lanes alone. This aggregate figure varies from line to line, and might have been greater had the life of the lanes not been continually threatened. *Nonetheless, although the Diamond Lanes and the attendant publicity helped increase bus ridership, it appears that improvements in bus system coverage and service frequency were responsible for the bulk of the observed patronage increases.*

Police Deployment, Enforcement and Violations

Police Deployment

Highway patrol deployment doubled during the first weeks of the project, and gradually returned to normal (76 man-hours daily during the project operating hours) by the thirteenth project week. For the most part, the additional manpower used early in the project consisted in motorcycle units diverted from other freeways.

Enforcement

Although levels of police deployment returned to normal midway through the demonstration, enforcement activities remained considerably higher than normal throughout the life of the project. An average of 151 warnings and citations were issued daily, more than four times the estimated pre-project levels.

Enforcement of the Diamond Lane provisions was facilitated by the existence of a median strip where violators could be cited without being escorted across three or four lanes of traffic to the right shoulder of the roadway. Helicopter and roadside observers soon noted, however, that the use of the median for enforcement also interfered with the flow of traffic in other lanes. The use of the median for enforcement led to gawking and traffic slowdowns, particularly in the Number 2 lane adjacent to the Diamond Lane.

Violations

The Diamond Lane violation rate, defined as the ratio of vehicles with fewer than three occupants to the total number of vehicles in the lane, was high on the first day of the project and dropped immediately thereafter. On the opening day, 40% of all vehicles using the preferential lanes did so illegally. The violation rate then dropped rapidly, and fluctuated between 10% and 20% for the duration of the project. Most of the observed violations occurred at the fringes of the Diamond Lane operating hours.

Safety

Freeway Accident Patterns

One of the most disturbing aspects of the Diamond Lane project was the high incidence of freeway accidents accompanying the operation of the preferential lanes. Accidents increased markedly in the first week of the project, when 59 accidents were reported during Diamond Lane operating hours. Accident levels subsequently declined, dropping to an average of 18 accidents per week during the last month of the project, but they remained substantially higher than pre-project levels throughout the demonstration. During the 21 weeks of the demonstration, 527 accidents were reported during peak operating hours, an average of 25 accidents per week and roughly 2.5 times the pre-project average. A significant proportion of the overall increase in accidents was concentrated in the Number 2 lane adjacent to the Diamond Lane, reflecting an increase in stop-and-go conditions in this lane. Accidents in this lane increased from under 2 accidents per week to 14.8 accidents per week during the project.

Probable Causes of Freeway Accidents

A number of potential accident causes were postulated and analyzed in an attempt to account for the observed increase in accident levels. Results of accumulated accident data identify a number of these causes as minor contributing factors, including increased CHP ticketing; increased congestion due to the removal of the Diamond Lanes fron general use; and the confusion, distraction and aggravation accompanying the novelty of the Diamond Lane concept. A more detailed analysis of potential accident causes is discussed in an accompanying TRB paper. (3.)

The single factor that appears to account for the largest share of the accident increase is: *The pronounced speed differential between the free-flowing traffic in the sparsely occupied preferential lane and the stop-and-go traffic in congested adjacent lanes, coupled with the frequent lans changes resulting from the variety of possible origins and destinations along the length of the project.* Motorists attempting to enter the Diamond Lane had to enter a faster traffic stream from a slower starting speed, while motorists attempting to leave the lane had to slow and attempt to find an opening in stop-and-go traffic. This problem was exacerbated by the large variety of trip origins and destinations in the Los Angeles area, which led carpoolers to enter and leave the Diamond Lanes at many points along the freeway. Regular and occasional carpoolers responding to the driver survey cited problems merging with slower traffic in leaving the Diamond Lanes as the greatest single difficulty encountered in using the lanes, and regular carpoolers felt that the discomfort of traveling faster than vehicles in the other lanes was just as disturbing as the difficulty of merging with these vehicles. As the speed differential increased, moreover, the ability to save time by using the Diamond Lanes attracted a few violators who dodged in and out of the Lane unsafely, attempting to stay one jump ahead of the CHP.

Implications of the Accident Picture

Since the ability to travel faster in a preferential lane is the chief inducement for attracting carpoolers and bus users to that lane, the fact that this ability increased accident levels significantly on the Santa Monica Freeway raises serious questions regarding the feasibility of the barrier-free preferential lane in certain settings. These questions appear to exist whether the lane is created by reserving an existing lane, as was done on the Santa Monica Freeway, or by creating an entirely new lane, as was originally contemplated on the San Diego Freeway in Los Angeles. Conceivably, the addition of a new barrier-free preferential lane to an existing freeway could also result in increased accidents if stop-and-go traffic conditions exist in the non-preferential lane, a significant speed differential is maintained between these lanes and an underutilized preferential lane, and destinations are scattered so that carpoolers enter and exit at many points along the lane.

If the usage of a preferential lane increases with time, either because more carpools are formed or because enforcement is relaxed, the speed differential will decrease and accident levels can be expected to drop. As the speed differential drops, however, the inducement to use a preferential lane drops as well. In theory, the number of carpools should grow over time until the marginal amount of time saved by switching to a carpool exactly balances the perceived inconvenience of making the switch. In practice, the level of accidents occurring before this equilibrium point is reached may be unacceptable to society, or the equilibrium point itself may result in an unacceptable accident rate.

Surface Street Accidents

One of the potential side effects of the Diamond Lane project was the possibility that traffic diverted from the Santa Monica Freeway to surface streets might increase the number of accidents on those streets in the corridor surrounding the freeway. *Although surface street accident levels increased slightly (between 5% and 10%) during the demonstration, statistical evidence linking these increases with the Diamond Lane project is inconclusive.*

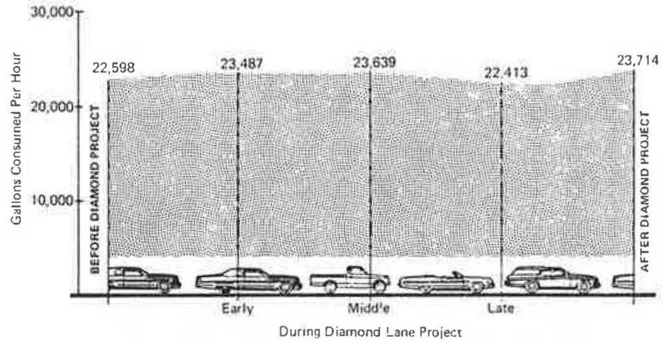
Energy and Air Quality

Fuel Consumption

Fuel consumption estimates based on vehicle mileage records indicate that, even allowing for increased idling time at on-ramps, gasoline consumption declined on the Santa Monica Freeway during the Diamond Lane demonstration. At the same time, fuel consumption actually increased on all parallel surface routes that were sampled. The net effect for the entire east-west corridor was a slight increase in fuel consumption of approximately 500 gallons per hour during the first fourteen weeks of the project. *By the last seven weeks of the project, the total energy consumption was 185 gallon per hour lower than the pre-project level of 22,958 gallons per hour, a savings of 0.8% over pre-project levels.*

Because of increased congestion and idling time, fuel consumption rates for non-carpoolers had increased by 6% by the close of the project. These increases were offset by the savings accompanying increases in carpool and bus usage. Each solo driver switching to a carpool or bus was estimated to save roughly eleven gallons of gasoline per week.

Figure 4. Average Energy Consumption per hour for the Santa Monica Freeway Corridor (gallons/hr).



Air Quality

On the basis of vehicle mileage computations, corridor vehicle emissions rose early in the project and dropped to pre-project levels by the time the project closed. Measured air samples showed a decrease in carbon monoxide concentrations during the project. In view of the small sample sizes, seasonal changes, meteorological variations, and analytic uncertainties, however, it is impossible to make conclusive statements regarding the precise impact of the Diamond Lanes on air quality.

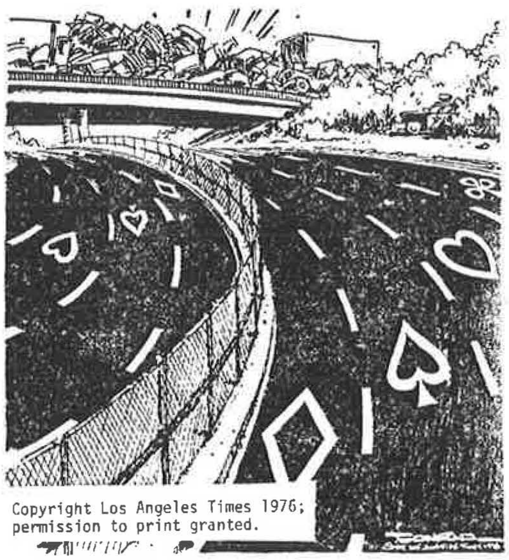
What Happened Off The Freeway

Statistical summaries quoting freeway speeds, vehicle volumes, bus ridership and accident rates do not begin to convey the full picture of the Santa Monica Freeway Diamond Lane demonstration. The Diamond Lane experience was not confined between the guardrails of the Santa Monica Freeway. The demonstration quickly became a media event, generating reams of newsprint, radio and television coverage, vocal public reactions, political debate, lawsuits, banners, slogans, badges, cartoons, and at least one song. *From their implementation to their dissolution, the Diamond Lanes were never far from public view and, when in view, they were treated as an eyesore.*

Media Coverage

During the 21 weeks of Diamond Lane operation, the three major daily newspapers covering the project--the morning Los Angeles Times, the afternoon Herald-Examiner, and the Santa Monica Evening Outlook--produced an average of nine articles and editorials per week on the Diamond Lanes. The predominant tone of the articles was negative, and the editorials were solidly against the project. Although the operations on the freeway improved following the disastrous opening day, when the newspapers carried banner headlines proclaiming "Freeway Chaos," newspaper coverage grew steadily more hostile as the demonstration progressed.

The Diamond Lanes were also a popular subject for radio and television coverage, and provided a platform for many public figures seeking public exposure. As in the case of the press, the general tenor of the coverage provided by local and national radio and television stations was hostile to the project. Perhaps the most hostile and least balanced of all media coverage was provided by the radio disk jock-



Copyright Los Angeles Times 1976; permission to print granted.

eys, whose jibes ("you'll get home tonight if it takes all year") reached motorists while they were in the middle of their congested commuting period.

Project Promotion

Although the full extent of the public and media outcry was not anticipated by the project participants, it was recognized in advance that the Diamond Lane project was likely to generate adverse public reaction, and an extensive marketing campaign was developed with the joint aims of promoting buses and carpools and encouraging public acceptance through a program of information and education. Given the extent of the pre-project advertising campaign, which included television and radio announcements, newspaper advertisements, the use of the changeable message signs on the freeway itself, and brochures distributed at freeway on-ramps, it is unlikely that many regular users of the Santa Monica Freeway were unaware that March 15, 1976 marked the opening of the Diamond Lanes. Although the appearance of the lanes themselves should have come as no surprise, opening day commuters did have reason to be surprised by several of the unannounced last-minute adjustments, including the tightening of ramp meter rates and the barricading of a slip ramp at the interchange of the Harbor and Santa Monica Freeways near the CBD. These unannounced adjustments undoubtedly contributed to the opening day confusion, and helped make March 15 "Mad Monday."

Following Mad Monday, the advertising campaign was drowned out by the media outcry and the project's sponsors, placed on the defensive, were able to do little to counter the tide of adverse public reaction.

Public Response

Surveys, interviews, telephone calls, newspaper polls, public hearings, and letters to newspaper editors generated during and after the project all revealed an overwhelmingly negative public response to the Diamond Lanes. *In the most extensive survey undertaken, 86% of the corridor drivers surveyed--including the majority of carpoolers--felt that the Diamond Lanes were either harmful or of no benefit whatsoever.* But public response to the Diamond Lane project was not limited to such formal avenues as survey responses and letters to editors. Residents of Los Angeles managed to find unique ways of expressing their general distaste for the Diamond Lanes. On opening days, nails were spilled in the lane by a disconsolate motorist, and a "baggy bomber" used paint-filled balloons to obliterate several of the painted diamonds in the lane. On June 3, the "Citizens Against the Diamond Lane" slowed Diamond Lane users by staging a mock funeral procession in the lanes, and they later attempted to hang anti-project signs from a freeway overpass. A smaller, less vocal group of "Citizens for the Diamond Lanes" was organized and developed a newsletter to champion their cause. Entrepreneurs sold bumper stickers and badges carrying comments on the lanes, while college students offered their services as riders for a fee to drivers wishing to qualify as carpoolers, and the media reported a brisk sale of mannequins designed to gull observers into believing one driver and two dummies constituted a three-person carpool.

Any attempt to lay the full blame for the hostile climate of public opinion on the media both oversimplifies and overstates the case. It is unlikely that the negative media reports alone could have generated such a hostile response if the reports were not reinforced by a negative impact on the lives of the public. In Los Angeles, the negative media image of the Diamond Lanes was reinforced daily for over 100,000 freeway users who found their daily commute trip lengthened by a project designed to benefit a perceptibly smaller proportion of the traveling public.

Institutional and Political Climate

Several factors contributed to the stormy political weather encountered during the Diamond Lane demonstration. These included:

1. The complexity of transportation planning, financing, and decision making in the Los Angeles area.
2. The changing philosophy, policies and personnel in the state transportation agency; and
3. The scheduling of the demonstration in an election year.

All of these factors combined in a setting where everyone talks about transportation conditions but few are able to do anything about them. Los Angeles's fragmentation of public power and authority meant that a large number of government agencies and elected officials had some purview over the Diamond Lane project. Each decisionmaker had his own concept of project goals, and the degree of involvement and commitment to the Diamond Lanes varied greatly from agency to agency. When the media spotlight turned on the project, the public saw not a united front but a number of public agencies and elected officials pointing accusing fingers at the lead agencies, while other officials remained prudently silent. The adversary role adopted by several public transportation agencies hindered both the free flow of project information and the coordination of project decisions.

CALTRANS, the lead agency responsible for project implementation, went from a state of flux immediately prior to the project to a state of seige during the demonstration. In the period immediately preceding the project, the agency was in a state of transition that included shifts in executive responsibility at the state level as well as sweeping layoffs locally. The shuffling of responsibilities, layoffs, and changes in management caused problems in both planning continuity and pre-project data collection. Once the project began, the new faces at CALTRANS were confronted with a new set of problems. Whereas the agency had become accustomed to public pressure over the building of freeways, the Diamond Lanes represented a new concept with a new set of aims and enough adverse side effects to lead some within the agency to question whether CALTRANS was justified in defending the project. As CALTRANS struggled to assess the operations on the freeway, deal with the hostile press, and evaluate a number of complex issues involving the project's future, an impatient press and public blistered the agency for its apparent intransigence and insensitivity to the needs of the citizens.

Public reaction and the media din were exacerbated by the frequent and public opposition of several elected and appointed City and County officials. The level of opposition ranged from responsible criticism on the part of some officials who had worked with project personnel in an attempt to make the Diamond Lanes more acceptable to their constituents, to simple attempts on the part of other officials to align themselves publicly with the opposition to a clearly unpopular project. Responsible opposition and objective analysis had to clamor for a hearing alongside of simplistic arguments, emotional appeals, and self-serving electioneering. The Diamond Lanes even became a pawn in the election-year battle for the approval of funds for a rapid rail system in Los Angeles (STAMP OUT DIAMOND LANES: VOTE YES FOR RAPID TRANSIT). In the face of the opposing clamor from the media, public, and elected and appointed officials, those officials who might have favored the project found it prudent to remain silent, and little in the way of a constructive public dialogue emerged. There is little doubt that the continual public threats to the lanes' existence led many potential carpoolers to deter any commitments to shared riding until the opposition was silenced and the project achieved a more permanent status.

The Legal End

Although the life of the Diamond Lanes was continually being threatened by the media and the public, and State and local officials had drafted terminating legislation, the demonstration was eventually done in by what at the time seemed the least likely suspect, a lawsuit in the U.S. District Court of Appeals. The lawsuit only indirectly addressed the merits of the project, focusing instead on the alleged failure of CALTRANS and UMTA to comply with the requirements of the California Environmental Quality Act (CEQA) and the National Environmental Policy Act (NEPA) by filing an Environmental Impact Report on the project. On Monday, August 9, 1976, Judge Matthew Byrne ruled that Environmental Impact Reports should have been filed under both state and national environmental laws, and ordered that the freeway be returned to pre-project status by Friday, August 13, 1976.

Observations and Implications

The Santa Monica Freeway preferential lane project succeeded to some degree in attracting riders to carpools and transit, and increased the person-moving capacity of the freeway without requiring additional levels of police deployment. However, the project brought about a significant increase in freeway accidents, non-carpoolers lost far more time than carpoolers gained, and a heated public outcry developed which has halted the implementation of other preferential treatment projects in Southern California, giving planners and public officials in other areas ample cause for reflection before attempting to implement similar projects.

The Negative Impact of Lane Removal

Whereas other preferential lane projects have constructed additional lanes or converted lanes in off-peak directions to preferential use, the Santa Monica Freeway Diamond Lane project marked the first time preferential lanes were created by taking busy freeway lanes out of existing service and dedicating them to the exclusive use of high-occupancy vehicles. This aspect of the project contributed to most of the negative impacts recorded during the demonstration. The removal of two lanes from general use contributed heavily to the congestion and confusion on opening day, was a slight but important factor in the increased accident rate, and appears to have been one of the chief sources of public dissatisfaction with the project. Many freeway users felt strongly that they had paid for the lanes with their gasoline taxes and were entitled to go on using them. The lane preemption and the resulting slowdown were viewed with hostility by most corridor drivers as a plot designed by meddling bureaucrats to inconvenience many for the sake of a few. Moreover, the number of project beneficiaries were perceived to be even fewer than their numbers indicated because they traveled three-to-a-car, or rode in buses that were often half-empty, and did not fill the Diamond Lanes to capacity.

The Effect of Geographic Sprawl

Because of the scattering of trip origins and destinations throughout Los Angeles, relatively few users of the Santa Monica Freeway are destined for the CBD. The lack of a focal point for trip destinations made carpool formation relatively difficult and decreased the pool of potential riders of the CBD-directed bus service. In practice, the scattering of origins and destinations also meant that drivers were likely to want to enter and leave the Diamond Lanes at points all along their 12.5-mile length, greatly increasing the possibility of accidents.

Accidents and the Absence of Barriers

Another factor in understanding the project's disappointing performance was the absence of barriers between the preferential lane and the congested adjacent lanes, Frequent vehicle shifts in and out of lanes operating at markedly different speeds contributed heavily to the increase in accidents. The problem of accidents in barrier-free operation is a serious one, and deserves further study. In other areas, and in Los Angeles itself, preferential treatment lanes separated from the general flow of traffic have been successful in improving carpool and bus ridership without increasing either accident rates or public acrimony.

The Success of Ramp Metering

One positive aspect of the Santa Monica Freeway experience which has been largely ignored was the performance of the ramp meters in alleviating freeway congestion and smoothing traffic flow before, during, and after the project. Prior to the project, the meters alone had so improved freeway traffic speeds that the Diamond Lanes suffered by comparison. The Diamond Lanes offered only a marginal one- or two-minute improvement over the meter-controlled speeds generally available to all traffic prior to the project. Conditions in the non-preferential lanes did not approach metered pre-project levels, although freeway speeds with both ramp meters and Diamond Lanes operating were faster than speeds where neither the meters nor the lanes were operational. Where available, moreover, carpool bypass lanes on the on-ramps offered more of a time savings to carpoolers than the Diamond Lanes themselves. *Thus, the ramp meter bypasses, which were safer and--surveys showed--less objectionable to the public than the Diamond Lanes, actually offered a greater time savings to carpoolers than the preferential freeway lanes, while the meters themselves improved freeway traffic flow.*

The Question of Credibility

One of the most serious controversies which emerged during the demonstration involved the question of data credibility. The sponsoring agencies were collecting data as the project progressed, and CALTRANS became the source for disseminating project statistics. As "CALTRANS' project" came under attack, so did the data it issued. Other agencies began drawing different conclusions from the CALTRANS data, and some local groups--including the press itself--began collecting and issuing their own data. The free-form use of different numbers and different reference bases during the demonstration made it difficult for the public to know who or what to believe, and led the press to question the credibility of project participants. The credibility of project foes was rarely questioned by the media.

Under the best of circumstances, there will always be some degree of ambiguity associated with traffic data. In many instances, statistics concerning the Diamond Lane project were produced under the worst of circumstances, having been hurriedly processed under rigid deadlines in the glare of publicity, and interpreted by agencies with a vested interest in attacking or defending the project. Problems encountered in the data collection and evaluation phases of the project ranged from simple human miscalculations to complex computer failures. In retrospect, the picture of the project that emerges from a more thorough examination of the data is somewhat different from that presented by both proponents and opponents of the project in the midst of the "battle of numbers" waged during the demonstration itself.

References

1. Billheimer, John W., Robert Bullemer and Carolyn Fratessa, The Santa Monica Freeway Diamond Lanes: An Evaluation, SYSTAN, Inc. for the Transportation Systems Center, U.S. Department of Transportation, Contract DOT-TSC-1084, April 1977.
2. California Department of Transportation, 21-Week Report on the Diamond Lanes, August 1976.
3. Billheimer, John W., The Santa Monica Diamond Lanes: Freeway Accident Analysis, submitted to TRB for publication, January 1978.

A COMPARATIVE ANALYSIS OF RESULTS FROM THREE RECENT NON-SEPARATED CONCURRENT-FLOW HIGH OCCUPANCY FREEWAY LANE PROJECTS: BOSTON, SANTA MONICA AND MIAMI

Howard J. Simkowitz, Transportation Systems Center

Results from three recent non-separated concurrent-flow high occupancy freeway lane projects, Boston's Southeast Expressway, the Santa Monica Freeway in Los Angeles, and Miami's I-95, are compared. The Los Angeles and Miami projects have been terminated, and, in Miami, the carpool definition has been decreased to two or more persons per car. While carpooling and bus ridership increased, other results point out the many generic weaknesses in the concept: the large number of violators and the difficulty of enforcement; the potential for accidents; the inability of the reserved lanes by themselves to attract large numbers of new bus riders and carpoolers; and the political problems associated with removing an already existing lane from general use. A comparison of the performance of these non-separated reserved lane projects with the Shirley Highway reversible lanes and the El Monte busway indicates that when concurrent flow lanes are separated from the general lanes by a concrete barrier or an empty safety lane, the accident and enforcement problems are virtually eliminated and the reserved lanes are better able to perform their function of attracting and carrying high occupancy vehicles.

In order to move more people in fewer vehicles, and with a limited capital investment, a set of priority techniques for high occupancy vehicles (HOV) has been developed and implemented over the past several years. These traffic management options include concurrent-flow, contra-flow, and reversible lanes on arterials and freeways, exclusive lanes that bypass congested areas such as freeway ramps and toll plazas, exclusive access ramps to freeways, bus pre-emption of traffic signals, congestion pricing, transit malls, and auto restricted zones.

This analysis focuses on recent experience with non-separated concurrent-flow high occupancy lanes on freeways. For the remainder of this paper, the term "reserved" will be used to denote these lanes. Reserved lanes exist or have existed on Routes 101 and Route 280 in San Francisco, on the Santa Monica Freeway in Los Angeles, on the Banfield Freeway in Portland, on the Southeast Expressway in Boston, on I-95 in Miami, and on the Moanalua Freeway in Honolulu.

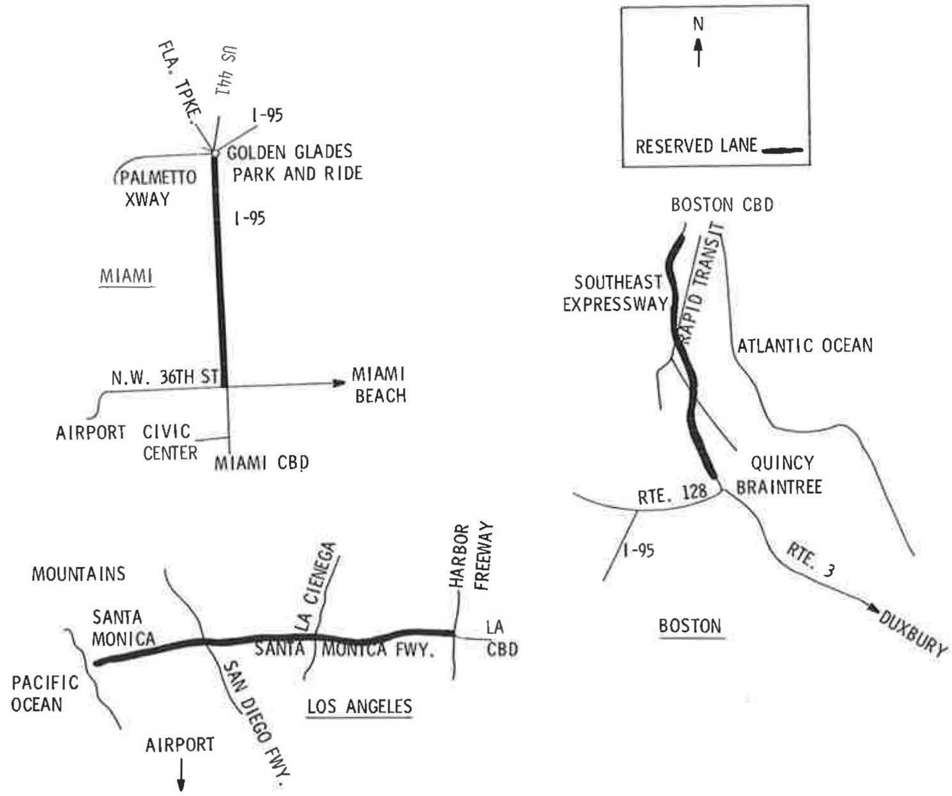
Through a comparative analysis of results of the three most recent concurrent-flow projects, Boston's Southeast Expressway, I-95 in Miami, and the Santa Monica Freeway, this paper attempts to develop a better understanding of the issues surrounding the reserved lane concept. Boston, Miami, and Santa Monica were chosen for comparative analysis for several reasons: all three represent recent experiments with the reserved lane concept; the three projects and project sites exhibit substantial differences; and evaluation efforts were conducted at each site.

Description of Reserved Lane Projects

The three reserved lane projects, even though each involved the concurrent-flow high occupancy vehicle lane concept, differed significantly from each other as to physical design of the freeways, hours of operation, entrance ramp treatment, transit characteristics, and other project related activities (Figure 1 and Table 1).

Boston's Southeast Expressway carries 121,000 vehicles per day, the Santa Monica Freeway carries 240,000 vehicles per day, and Miami's I-95 carries 170,000 vehicles per day. In Miami a lane for high occupancy vehicles was added to I-95 in both directions, completely eliminating the median area. In both Boston and Los Angeles existing lanes were taken away from normal use and dedicated to high occupancy vehicles. In Boston the left lane in the northbound (in-bound) direction only was

Figure 1. Drawings of the three projects.



reserved for buses and carpools of three or more occupants from 6:30 a.m. to 9:30 a.m. during weekdays. In Los Angeles the lanes were reserved for buses and carpools of three or more occupants in both directions from 6:30 a.m. to 9:30 a.m. and 3:00 p.m. to 7:00 p.m. In Miami the southbound (inbound) lane was restricted to buses and carpools of 3 or more occupants from 6:00 a.m. to 10:00 a.m. and the northbound (outbound) lane from 3:00 p.m. to 7:00 p.m. After a year of operations the times were changed to 7:00 a.m. to 9:00 a.m. and 4:00 p.m. to 6:00 p.m. and the restrictions changed to carpools with 2 or more occupants.

Access into and out of the lanes in Los Angeles and Miami was unrestricted. In Boston plastic inserts spaced at 20 or 40 foot intervals separated the lane from the rest of the roadway, and entry to or exit from the lane was allowed only at the beginning and the end. Weaving was prohibited but only sporadically enforced by the police.

Only Los Angeles employed ramp metering. Thirty on-ramps were equipped with meters (these existed before the project), and their timing was adjusted and pre-set to maintain free flow on the Freeway. Twelve of these ramps offered preferential access to buses and vehicles with two or more occupants. During the first three months of operation, the left lane on Boston's Southeast Expressway was blocked just before the beginning of the reserved lane, and all vehicles had to merge into the right lanes. This made it necessary for carpools and buses (and violators) to switch back into the reserved lane. The effect was similar to metering the Expressway. In Miami a flyover providing a direct connection between the major park and ride lot and the reserved lane was opened 12 months after the start of the project.

While all three sites stressed the need to use the existing freeways in a more efficient manner and to reduce energy consumption and air pollution by encouraging the use of high occupancy vehicles, the motivating force behind the reserved lane project in Boston was the need to reconstruct a portion of the roadway that would create a temporary decrease in capacity of up to 25 percent. The potential for serious congestion and the need for preferential treatment for high occupancy vehicles was clearly explained to the public.

The lane restrictions were heavily enforced in Los Angeles and only lightly enforced in Miami. The restrictions were voluntary in Boston during the first five months of operations, after which time enforcement was instituted by sending traffic citations through the mail.

In Boston few changes were made to the existing very extensive public transportation systems. One park and ride route was added, and back-up sections on existing bus and rapid rail routes were provided. Additional fringe parking spaces were made available.

In Los Angeles, up to twelve bus routes used the diamond lane. Five of the routes were new feeder express routes from the Westside area to the Los Angeles CBD. Three new routes provided service to the new park-and-ride lots. In all, the number of morning express bus runs was increased from 18 to 74. Headways on all the routes were 10 to 15 minutes.

Table 1. Comparison of the three preferential lane projects.

PROJECT	FACILITY	LENGTH (miles)	OPERATING DATES	LANE RESTRICTIONS	LANE ORIGIN	HOURS OF OPERATION	SPECIAL FACILITIES
Boston: Southeast Expressway	Freeway, 3 or 4 lanes each direction, including use of shoulder in peak direction during peak period	8	5/04/77-11/02/77	Buses and carpools (3 or more occupants)	1 existing lane reserved (inbound)	6:30-9:30 a.m. inbound only	Plastic inserts space 20-40 feet, freeway "metering" for 3 months
Miami: I-95	Freeway, 4 or 5 lanes each direction	7.5	3/15/76-present	Buses and carpools (3 or more occupants, changed to 2 or more)	2 lanes built in median area	6-10 a.m. (changed to 7-9 a.m.) inbound; 3-7 p.m. (changed to 4-6 p.m.) outbound	Flyover connecting major park and ride lot to I-95 after one year
Los Angeles: Santa Monica Freeway	Freeway, 4 or 5 lanes each direction	12.9	3/15/76-8/09/76	Buses and carpools (3 or more occupants)	2 existing lanes reserved	6-10 a.m. (changed to 6:30-9:30 a.m.) 3-7 p.m. inbound and outbound	Ramp metering, some with preferential bypass
Portland Oregon: Banfield Freeway	Freeway, 3 or 4 lanes each direction	3.3	12/15/75 present	Buses and carpools (3 or more occupants	resurfaced, removed shoulder, narrowed lanes	24 hours/day changed to 6:30-9:30 a.m. inbound and 3:30-6:30 p.m. outbound	

PROJECT	ACCESS/EGRESS	ENFORCEMENT	TRANSIT	EXPRESS BUS AVERAGE FARE
Boston: Southeast Expressway	Only at beginning and end	Voluntary for first 5 months, enforced last 2-1/2 weeks; increase in police	Minor changes to existing express and feeder bus, rapid rail, commuter rail, and commuter boat; new park and ride route	$1.25
Miami: I-95	Unlimited	Little enforcement; no increase in police	Park and ride and feeder/express bus service increased from 18 to 52 trips per day; new large park and ride lot	.60
Los Angeles: Santa Monica Freeway	Unlimited	Fifty percent increase in police, reduced to normal by 12th week	Four existing feeder/express bus routes increased to 9; 3 new park and ride routes and lots	.61

In Miami the express bus service was expanded in 1974. Not only was the express bus service increased to 55 trips per day, but also the size of the market area served was increased: at the northern end of the corridor, express buses provided increased residential coverage to the northwest and northeast of the Golden Glades interchange; at the southern portion of the corridor, the buses served two employment centers (Civic Center and Airport) formerly not served by express buses.

A parking lot with space for 1320 vehicles was constructed at the northern end of the reserved lanes at Golden Glades, the confluence of 5 major highways. The lot was fenced, well lit, and patrolled. Some bus runs originated at this parking lot, while others performed local collection service before converging at the lot to pick up park-and-ride, kiss-and-ride, and transfer passengers. The buses then traveled south along I-95 destined for one of four major employment centers.

The Golden Glades Parking Lot, by acting as a transfer point for the four feeder routes as well as a park-and-ride and kiss-and-ride facility, enabled travel between any point in the residential market area and any employment destination, whereas the former express bus service only operated between selected origins and destinations, with no transfer capability. Furthermore, the four new feeder routes provide far more efficient and direct service in the residential area than the three express bus routes that they replaced.

In Boston and Los Angeles computer carpool matching, a marketing campaign, and a telephone center were provided to assist and encourage travelers to use the reserved lanes. In Miami only a marketing effort was undertaken.

In Miami the lanes are still in operation although the definition of a carpool has been changed from three to two occupants. In Boston the police began enforcing the lane restriction 5 months after the project began. After two and one half weeks of significant political pressure and unfavorable articles in one of the daily newspapers, the Commissioner of the Massachusetts Department of Public Works suspended the project. In Los Angeles a federal judge ruled that an environmental impact report should have been filed under both federal and state environmental laws. This ended the Santa Monica project after 21 weeks of operation.

Project Results

The three reserved lane projects have met with differing degrees of success and failure. The reserved lane on the Southeast Expressway survived for 6 months only to be cancelled suddenly two and a half weeks after the lane restrictions became mandatory. A federal judge shut down the Santa Monica project after 21 weeks of operation because an environmental impact report had not been filed. In Miami, the inability to enforce the lane restrictions led to a lowering of the lane qualification to two or more persons per car.

The three projects resulted in an increase in the occupancy rate of those vehicles using the facility (Figure 2). However, in both Boston and Los Angeles person throughput on the freeways decreased (Figure 3). A promising trend had developed in Los Angeles, and when the project was terminated the Freeway was carrying only 1.8 percent fewer persons in 9.4 percent fewer vehicles (Figure 4). In Boston, the corresponding figures were 8 percent and 21 percent. In Miami, a rapidly growing area and where new lanes were constructed, person throughput increased by 28 percent while vehicle throughput increased by 20 percent.

Figure 2. Freeway auto occupancy.

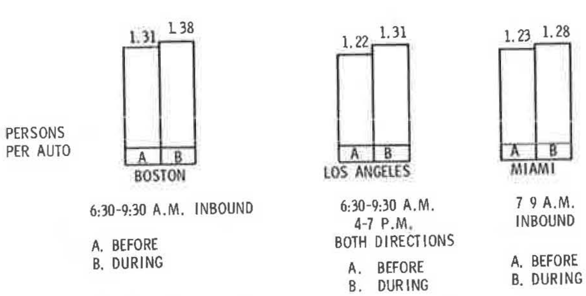


Figure 3. Freeway person throughput.

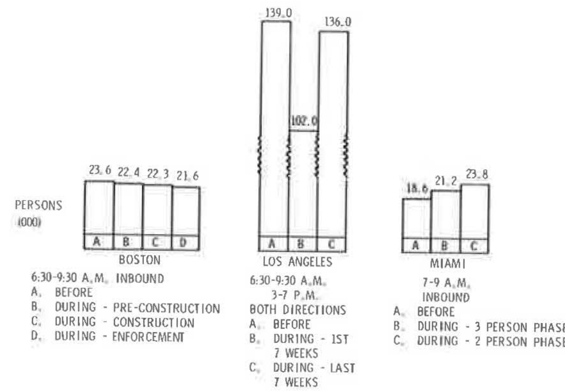
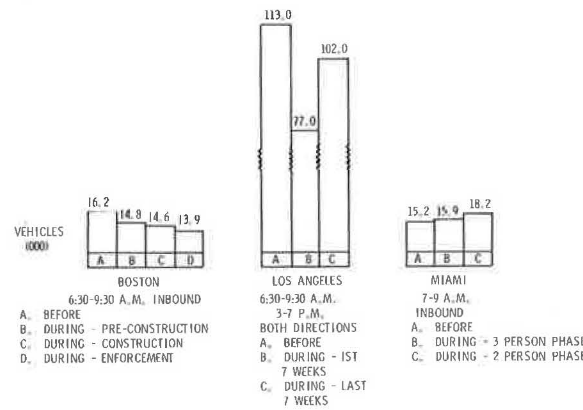


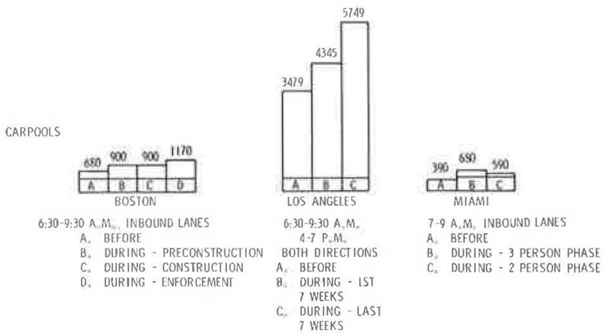
Figure 4. Freeway vehicle throughput.



In Boston, after the reserved lane was instituted but before construction began, the total number of persons carried by the Expressway during the peak period was 22,400, 5 percent less than during the March pre-project period. In June, person throughput declined to 22,300, a decrease of 6 percent from March. This additonal one percent decrease was probably the result of the combination of the construction further north on the Expressway and seasonal factors. During the enforcement period, the total number of persons carried was 21,600, a decrease of 8 percent from March. Since the dominance of Boston's core area as an attraction zone indicated a much greater potential for carpooling and bus ridership than in Los Angeles, it was possible that an increase in person throughput similar to that experienced in Los Angeles would have developed had the enforcement period continued. In fact, it is reasonable to assume that all three projects suffered from the public's perception that the lanes were not permanent. It was less likely for a person to form a carpool or learn about a convenient bus route if he believed that the reserved lane project was to be terminated when construction was completed or if political pressure became too great to maintain it.

At all three sites carpooling increased by about 70 percent (Figure 5). In both Los Angeles and Miami the primary reason given

Figure 5. Freeway carpools.



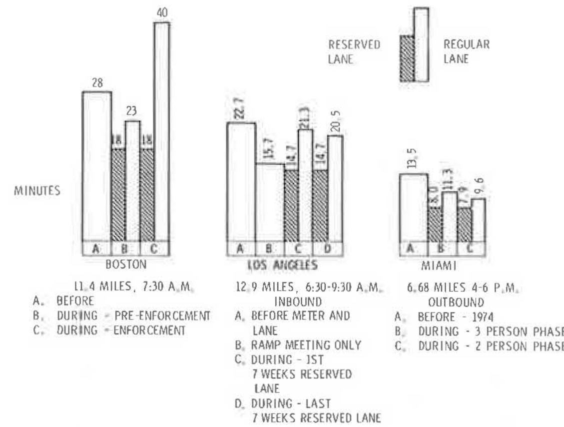
for carpooling was cost and not the time savings from using the lanes. While it was true that the majority of the carpoolers surveyed at each site had formed carpools before the reserved lanes were instituted, and therefore, their primary incentive would have been expected to be time rather than monetary savings, in Los Angeles 35 percent of members of carpools that were formed during the reserved lanes gave cost as the main reason for carpooling while only 30 percent gave time savings. However, the number of carpools fell to within 5 percent of pre-project levels after the project was terminated. It could be that time savings from using the reserved lanes were balanced by the additonal time it took for the collection and distribution portions of the trips.

Not everyone who was eligible for the reserved lanes used them. In Miami less than one-third of the eligible carpools used the reserved lanes. In Santa Monica 22 percent of eligible carpools were in regular lanes. For persons not making long trips it was probably not worth the effort to access the reserved lanes.

At all three sites the greatest benefits accrued to users of the lanes, carpoolers and bus riders, who experienced decreases in travel times and increases in arrival time reliability (Figure 6). In Los Angeles and Boston, these benefits needed to be weighed against any decreases in level of service experienced by non-users of the reserved lanes. In Los Angeles travel times increased for non-diamond lane users. In Boston, users of the regular lanes experienced a decrease in travel times during the pre-enforcement period. This was due to people shifting out of their cars and into carpools and buses on the Expressway and to other modes and routes which resulted in a 5 to 6 percent decrease in vehicles on the southern portion of the Expressway. It was also due to the "metering" of the Expressway just before the start of the express lane. As with ramp metering on the Santa Monica Freeway, this screenline metering worked well in creating free-flow conditions on the roadway. In Miami all users of the facility benefited, but this was a result of the opening of the two additional lanes, at a cost of $19 million, and had little to do with the lane restrictions.

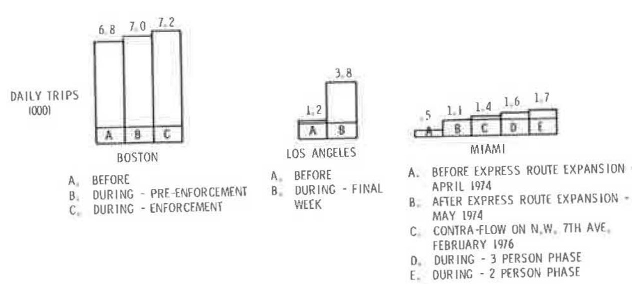
A disappointment with the reserved lane projects was their inability in and of themselves to attract large numbers of new

Figure 6. Freeway travel times.



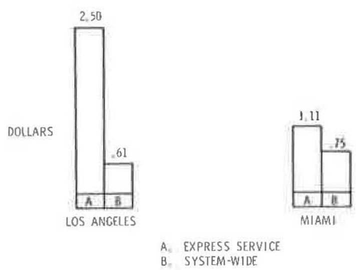
bus riders (Figure 7). In Los Angeles and Miami a large portion of the ridership increases appeared to have been the result of the increase in coverage and schedule frequency and not the travel time savings and increased reliability resulting from the reserved lanes. For most runs, the time spent in the reserved lanes did not represent a major portion of total in-vehicle travel time. However, the reserved lanes were useful in providing a focal point for the transit marketing campaigns and in creating a perceived, as well as a real, time advantage in the minds of the bus passengers. In Boston, where there were almost no transit level of service changes except decreased bus line-haul travel times, express bus ridership increased by only 3 percent. It was interesting to note that ridership on rapid rail and commuter rail increased by about 7 percent, possibly due to the higher visibility and public awareness of these modes.

Figure 7. Daily express bus ridership.



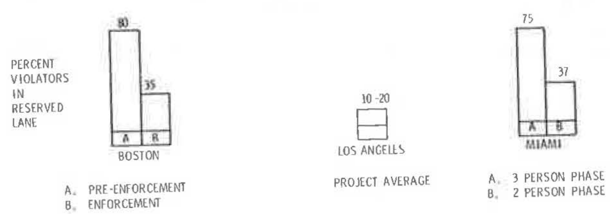
While the feeder/express routes in Miami and Los Angeles proved to be very popular, they also proved to be very costly since few buses could make more than one run during each peak period (Figure 8). Park-and-ride lots at the three sites met with mixed success, and this was a function of where they were situated and the frequency of the bus service. In Miami, the success of the park-and-ride service was due, in part, to the placement of a large parking lot 11 miles from the CBD at the confluence of 5 major highways. Buses travelled to four destinations, and headways were low.

Figure 8. Cost per bus passenger.



Another disappointment with the reserved lane concept was the number of lane violations that occurred and the difficulty of enforcing the lane restrictions (Figure 9). In Boston the plastic inserts did not prevent drivers from weaving in and out of the lanes. A median strip, where police could station themselves and stop violators, helped keep the violation rate in Los Angeles between 10 and 20 percent. Stiffer fines might have proven to be a deterrent, but the probability of being caught was not that great, especially if upon seeing an officer, the illegal driver was able to weave into the adjoining lane. In Boston and Miami a median area was not available. When Boston began enforcing the lane restrictions by sending tickets through the mail, the violation rate fell from 80 to 35 percent.

Figure 9. Violation rate.



One of the most serious problems with the reserved lane projects was the potential for accidents. Accidents were caused by the large speed differential between the reserved lanes and the normal-flow lanes and people making unsafe lane changes, weaving by violators to avoid detection, and by distressed motorists mistaking the reserved lane for a breakdown lane during non-operating hours. Lane changes could be limited by closely spaced plastic inserts, and reserved lane access and egress could be restricted to coincide with major entrances and exits. Boston did this to the extreme by permitting only one entrance and one exit, but motorists still managed to violate the no-weaving restrictions.

Carpool matching programs did not meet with great success. In Miami no carpool matching program was attempted since such a program had been tried on another project and failed. In Los Angeles commuter computer estimated that it was responsible for the formation of only 193 carpools. In Boston about 400 persons filled out carpool matching questionnaries. It was not known how many of these persons actually formed carpools. Most carpools in Los Angeles were formed among co-workers.

Due to the differing nature of the projects, the costs varied significantly (see Figure 2). For example, in Miami almost $19 million was spent just for construction of the two reserved lanes, a parking lot, and a flyover. The entire Santa Monica project cost just over $3 million, with $1.2 million being spent for data collection and evaluation and $886 thousand for bus operations. Boston spent only $245,000 for their entire project.

Recommendations for Future High Occupancy Vehicle Priority Projects

The results of the three non-separated concurrent-flow projects described in this paper point out the many generic weaknesses in this concept: the large number of violators and the difficulty of enforcement; the potential for accidents; the inability of the reserved lanes by themselves to attract large numbers of new bus riders and carpoolers; and the political problems associated with removing an already existing lane from general use.

Based on the Boston and Santa Monica results, it is not recommended that an existing lane be re-dedicated for preferential use unless there is a pressing need such as a reduction in capacity due to freeway reconstruction. If there is to be a decrease in freeway supply available to non-high occupancy vehicles, this decrease should be phased in order to cushion its effects and to encourage single occupant auto drivers to switch early to other modes or routes. A corridor whose transportation facilities are not already saturated will cushion the transition from pre-project to post-project equilibrium by allowing former users of the freeway the option to switch to alternate routes or other modes of transit if these are preferable to carpooling, taking an express bus, or staying on the freeway's normal lanes. These concepts were well-illustrated in Boston.

A comparison of the performance of these non-separated reserved lane projects with the Shirley Highway reversible lanes and the El Monte busway indicates that when concurrent flow lanes are separated from the general lanes by a concrete barrier or an empty safety lane, the accident and enforcement problems are virtually eliminated and the reserved lanes are better able to perform their function of attracting and carrying high occupancy vehicles. The appearance of permanence seems to contribute a great deal to convincing people to switch to HOV's.

Quite often these permanently or semi-permanetly separated configurations are not feasible for economic and/or engineering reasons. Boston attempted the minimum in physical lane separation by installing plastic inserts every 20 or 40 feet between the reserved and regular lanes. Unfortunately, these inserts did not prevent a large amount of illegal weaving between

Table 2. Costs of the three preferential lane projects.

	BOSTON		LOS ANGELES		MIAMI	
	UNIT COST	PROJECT COST ($000)	UNIT COST	PROJECT COST ($000)	UNIT COST	PROJECT COST ($000)
INVESTMENT COSTS						
LANE CONSTRUCTION		-		-		11,656
PARKING LOT(S)		-		199		1,711
FLYOVER TO LOT		-		-		2,981
PLANNING, DESIGN, AND SUPERVISION OF CONSTRUCTION		-		-		2,372
SIGNING		8		163		1,627
BUSES		-		-	20 @ 51,500	1,030
MARKETING		40		358		84
EVALUATION		55		1,232		973
PLASTIC INSERTS	3500 @ $11.	39		-		-
DRILLING HOLES	1500 @ $ 4.	6		-		-
OPERATING COSTS						
BUS OPERATIONS		-	2,588[1]	886[2]	461[1]	211[3]
ROADWAY & SIGNING MAINTENANCE		-		-		88
PARK AND RIDE LOT MAINTENANCE AND SECURITY		-		-		18
INSTALLATION AND REMOVAL OF INSERTS	26 WKS @ $3750	97		-		-
LOCAL AGENCY ADMINISTRATION		-		193		-
COURT COSTS		-		77		-
TOTAL PROJECT COSTS		245		3,108		22,751

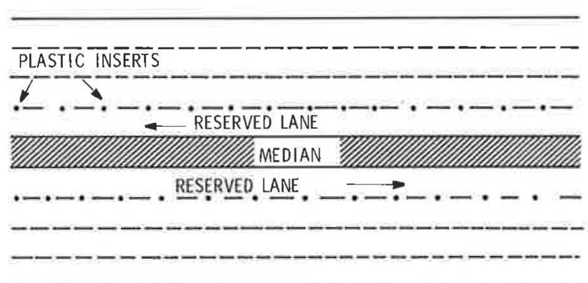
[1] OPERATING COST PER YEAR ($000)
[2] OPERATING DEFICIT FOR 22 WEEKS
[3] OPERATING DEFICIT FOR 1 YEAR

the two lanes. No only did non-carpoolers switch into the reserved lane, but carpoolers illegally left the lane to exit the Expressway.

The evidence indicates that there should be a median strip between the two directions of flow to provide both an area for motorcyle police to station themselves to control the violation rate and a safe area for distressed motorists to stop (Figure 10). To reduce the dangers of lane changing between two lanes travelling at significantly different speeds, the reserved lane entry and exit points should be limited to the beginning and end of the reserved segment and to a few intermediate points. The potentially large speed differential between the reserved lane and the regular lanes could possibly be reduced by electronic signs on the freeway that would limit the speed in the reserved lanes to some amount greater than in the regular lanes. This speed limit could be enforced if bus drivers were instructed to adhere to it. This concept has never been tested.

If the reserved lane configuration calls for inserts and a median, then it must be determined whether or not to leave the inserts in place on a 24 hour basis. It is costly to install and remove the inserts, the operation tends to confuse motorists, and it cannot be performed in the snow or dark. If the inserts were permanent, the lane restrictons would not necessarily have to be in effect or enforced on a 24-hour basis. However, this arrangement could be confusing to motorists as was the case in Miami where the solid striping used to separate the lanes during the early months of the project resulted in the reserved

Figure 10. Concurrent flow reserved lane with inserts and median.

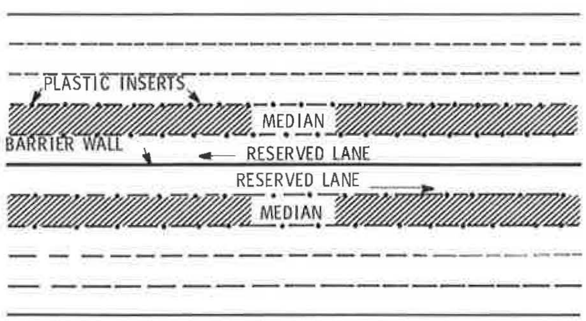


- inserts separate reserved lane from regular lanes
- median for police and distressed motorists
- entry and exit limited to beginning and end and a few intermediate points

lanes being mistaken for breakdown lanes during the non-restricted hours. Other drawbacks are that the inserts could create a safety hazard at night or during slippery conditions and plowing would be extremely difficult.

If space permits, the median could be shifted to the area between the reserved lane and the normal lanes as is the case of the El Monte Busway (Figure 11). Permanent plastic inserts would separate this safety lane from the rest of the roadway. The inserts would be spaced far enough apart so that this empty lane could be accessed by slow moving police and distressed motorists. Carefully designed slip-ramps would provide entry to and exit from the lanes at a few intermediate points. These ramps would be denoted by inserts, striping, and special pavement treatment so as not to be confusing to motorists.

Figure 11. Concurrent flow reserved lane with safety lane and inserts.



- safety lane between reserved and regular lanes
- safety lane used by police and distressed motorists
- inserts separate safety lanes
- entry and exit limited to beginning and end and a few intermediate points via carefully designed slip-ramps
- barrier wall between two directions of flow

Concurrent-flow lanes are applicable when the flow is balanced in each direction. When there is a large imbalance in peak directional flows, and if sufficient capacity exists in the off-peak direction then contra-flow or reversible lanes would be more appropriate.

In addition to the careful selection of the most appropriate form the HOV lanes will assume, this analysis has revealed factors related to site characteristics, implementation procedures, transit operations, and media treatment that must be considered.

The primary characteristic of the site that defines the market potential for the reserved lanes is a CBD that is the focal point for regional employment. This ensures a ready market for express bus patrons and facilitates the formation of carpools. In order to avoid citizen protest, it is important that the reserved lanes appear to be well-utilized to those travelling in the regular lanes and appear to be permanent.

Any increase in express bus operations should focus on the development of new feeder/express routes with the feeder component used to expand transit coverage, preferably serving more densly populated neighborhoods that currently have poor access to transit. Free and efficient transfer capabilities should be provided at park and ride lots if the buses go to different destinations. However, demand for priority facility bus services has proven to be inelastic with respect to fare; therefore, the fare should reflect the quality of the service being provided.

Park and ride service should be provided only from lots that are distant from the CBD and have good transit and highway access. The lots should be adjacent to the freeway and be large enough to support low headway service to several major destinations. Lots should be guarded, well lit, highly visible to the motorist, and contain amenities such as sheltered waiting areas, telephones, and toilets. The lots should have a convenient and adequate waiting area for afternoon kiss and ride automobiles. The transit operator should be aware of the high cost of operating this express bus service. High occupancy vehicles, such as double deck and articulated buses, could be used on these routes to minimize driver costs.

The public should be made aware of all aspects of the reserved lane project as early as possible. Commuter Computer estimated that carpool formation took an average of one month following a request. All travel options should be clearly described including estimates of level-of-service for each one.

Ramp metering, freeway metering, and pricing can be used along with, or in lieu of, reserved lanes. Ramp metering is relatively inexpensive, easy to install, and acceptable to the public. It worked well on the Santa Monica Freeway, making the average trip time both shorter and less variable. Many of the ramps provided preferential treatment, and the violation rates were low. A form of freeway metering was attempted in Boston and resulted in a decrease in travel time. However, freeway metering does not afford high occupancy vehicles preferential treatment.

The majority of carpoolers in Miami and Los Angeles indicated that their primary reason for carpooling was to save money. Thirty-five percent of members of carpools formed during the Santa Monica project reported cost incentives as the primary reason for carpooling while 30 percent listed the diamond lane. These results indicate that parking or toll policies favorable to carpools, in addition to preferential lanes, would do much to increase carpooling. The revenues generated could be used to expand the express bus service, which would further increase the use of high occupancy vehicles.

Acknowledgements

This work was sponsored by the Office of Service and Methods Demonstrations of the Urban Mass Transportation Administration.

References

Howard J. Simkowitz, Southeast Expressway High Occupancy Vehicle Lane Evaluation Report, U.S. Department of Transportation, Transportation Systems Center, January 1978.

John W. Billheimer, Robert Bullemer, and Carolyn Fratessa, The Santa Monica Freeway Diamond Lanes: An Evaluation, submitted to the U.S. Department of Transportation/Transportation Systems Center by SYSTAN, Inc., Los Angeles, California, April 1977.

Daniel Brand, John Attanucci, Howard Morris, and Charles Kalauskas, Southeast Expressway Reserved Lane for Buses and Carpools, presented at the 57th Annual Meeting of the Transportation Research Board, Washington, DC, January 1978.

Evaluation Study for the I-95/N.W. 7th Avenue Bus/Carpool Systems Demonstration Project in Miami, Florida, Three Quarterly Research Project Progress Reports, July 1, 1975 - March 31, 1976, Appendix D, University of Florida Transportation Research Center, June 1976.

Interstate 95 Exclusive Lanes Demonstration Project, Project Review (Draft), Florida Department of Transportation, Division of Mass Transit Operations District IV, December 1976.

Evaluation Study for the I-95/N.W. 7th Avenue Bus/Carpool Systems Demonstration Project in Miami, Florida, Quarterly Research Progress Report, April 1 - June 30, 1976, University of Florida Transportation Research Center.

I-95/N.W. 7th Avenue Bus/Carpool Systems Monthly Progress Report Summary, Florida Department of Transportation, May 1977.

Evaluation Study for the I-95/N.W. 7th Avenue Bus/Carpool Systems Demonstration Project in Miami, Florida, Quarterly Research Progress Report, October 1 - December 31, 1976, University of Florida Transportation Research Center.

Evaluation Study for the I-95/N.W. 7th Avenue Bus/Carpool Systems Demonstration Project in Miami Florida, Quarterly Research Progress Report, July 1 - September 30, 1976, University of Florida Transportation Research Center.

C.E. Wallace, An Investigation of the Economic Performance of an Express Bus Operation, Transportation Research Center, University of Florida, October, 1976.

Traffic Control of Carpools and Buses on Priority Lanes on Interstate 95 in Miami, final report, University of Florida Transportation Research Center, August 1977.

Abridgment

BUS PRIORITY SIGNAL CONTROL: SIMULATION ANALYSIS OF TWO STRATEGIES

Edward B. Lieberman, KLD Associates, Inc.
Ann Muzyka and David Schmeider, Transportation Systems Center

A study to evaluate the effectiveness of two bus priority strategies using a validated traffic simulation program is described. The test network, which is located in the Central Business District of Minneapolis, Minnesota includes two major arterials, each with a contraflow bus lane. The first control strategy consists of a fixed-time traffic signal pattern generated by the SIGOP-II model, which is designed to minimize passenger-delay, rather than vehicle-delay. The second control strategy is a real-time policy which preempts the fixed-time control to provide preferential treatment for approaching bus vehicles. The simulated results for each strategy were compared with those reflecting the existing fixed-time signal control. This study indicated that, for this application, the reduction in delay for bus passengers as predicted by the simulation program provided by both strategies outweighted the additional delay experienced by passengers in private vehicles; the preemption strategy provided greater improvement in performance than did the other. The study also demonstrates that a validated simulation model is an effective tool for evaluating alternate design configurations prior to field demonstration.

With attention focusing on techniques for encouraging the general public to increase its usage of mass transit facilities, several experiments have been undertaken to improve the performance of bus operations. One approach for evaluating strategies is to develop a preliminary design which is subsequently implemented in the field. This approach is both costly and time-consuming (1). Furthermore, should the preliminary design produce a degradation in traffic performance, there exists the prospect that the experiment will be terminated before refinements can be implemented.

Another approach is to develop the "best" preliminary design possible prior to field implementation by exploring different candidate designs. Each is then tested in a manner which replicates the proposed traffic environment. An effective methodology for such experiments, which has received increasing usage in recent years, is the application of traffic simulation techniques.

This paper describes the application of a microscopic simulation model of urban traffic, named SCOT(2,4) to a network in the central business district in Minneapolis. On each of two adjoining parallel, one-way arterials, a contraflow bus lane has been implemented. The purpose of this study was to identify the "best" preliminary design. Specifically, the following tests were conducted:

1. Evaluation of traffic operations on the network with the existing fixed-time control timing plan.
2. Evaluation of traffic operations with a fixed-time signal timing plan specifically designed to minimize person-delay.
3. Evaluation of traffic operations with a real-time bus preemption control.

The operational characteristics of general traffic and of bus traffic are specified as input to the simulation program separately. For general traffic, queue discharge headways, free-flow speed and turn movement percentages are specified for each link. Bus traffic is specified in terms of their respective route structures and the bus stations serviced. All bus stops are located appropriately and their respective [curb] capacities and observed bus dwell times are specified.

The traffic control is specified in terms of signal interval durations and signal offsets, at each node (intersection) of the network. For the on-line bus preemption control, detectors were specified in the locations where they would be installed in the pavement, and the control was specified in terms of minimum phase duration for the cross streets. The actual real-time pattern of signal indications was determined by internal logic.

The physical street system is represented as a network, as shown in Figure 1. Each north-south arterial services general traffic in one direction with a single bus contraflow lane. The cross-flow streets all service one-way flow, as indicated.

The urban portion of the SCOT simulation model moves individual vehicles along the network links (streets) and through the nodes (intersections) in response to the signal control. This portion is essentially synonomous with the UTCS-1 model (3) which was validated on a network servicing bus traffic. Statistics describing traffic operations are accumulated and listed for each link in the network; bus statistics are maintained separately.

Control Plans

The existing fixed-time control system exhibits a common cycle length of 90 seconds, relative offsets are zero for all links, and the signal split at each intersection is set at a G/C of 0.5, approximately. Right-turn-on-red is permitted for most approaches. General traffic may turn left across the bus contraflow lane.

The SIGOP-II model (5) was employed to obtain new signal timing plans. A small modification was introduced into this model specifically to replicate bus traffic operations and the dwell time experienced while servicing passengers at bus stations. The difference in passenger occupancy between buses and general traffic vehicles (40 vs. 1.3, respectively) was represented. This effectively transformed the objective function in the SIGOP-II model from "vehicle-delay" to "person-delay." The signal cycle length was retained at the current value of 90 seconds (4).

The bus preemption strategy is designed to alter the fixed-time sequence of signal phasing so as to provide preferential service for bus traffic. Briefly, the algorithm is based on a design where a bus station on a street is always located upstream of the detector which issues a "call" for signal preemption at the downstream signal. That is, there is no bus station between the detector and the stop line. Based on the projected arrival of a detected bus at the stop line, the algorithm determines whether to truncate the RED phase or extend the GREEN phase, or cycle rapidly to reinstate GREEN phase, or cycle rapidly to reinstate the GREEN phase subject to minimum phase duration constraints. The objective is to minimize bus delay. When competing buses vie for the GREEN phase, the algorithm resolves the conflict by implementing that strategy which minimizes total bus delay, subject to certain constraints.

The preemption strategy was programmed and integrated into the SCOT simulation model.

Experimental Results

The existing control was treated as the "base case." A total of 15 bus routes traversed the test network; each route exhibited an average headway of about 2 minutes. Hence, a bus entered the network every 8 seconds. Results describing the overall performance of bus traffic for the new signal timing plan are presented in Table 1, while those for general traffic appear in Table 2.

As indicated, the buses along the major arterials benefit significantly, while those along the cross streets experience sharp degradation in performance. General traffic experiences a moderate fall-off in operational performance. The overall bus performance experiences improved service as measured by a 12 percent reduction in the total delay relative to the base system. On the basis of the observed occupancies of 40 passengers per bus and 1.3 per auto, the net effect over a 15-minute period is a decrease of 395 passenger-minutes. Extrapolating this figure over the peak hour yields a net reduction in delay of 26.3 passenger-hours per hour.

Results describing the overall performance of bus operations for the bus preemption strategy are presented in Table 3, while those for general traffic appear in Table 4. The pattern of these results is similar to those described above for the new signal timing plan. As expected, buses along the main arterials benefit significantly while other components of the traffic stream experienced increased delay.

For this 15-minute time period, delay experienced by general traffic increased by 498 vehicle-minutes, while delay for buses decreased by 42 vehicle-minutes. Employing the same occupancy figures as previously, the net effect is a decrease of 1032 passenger-minutes in the test period or 68.8 passenger-hours per hour.

Conclusions

This paper has described a study employing traffic simulation to evaluate design alternatives to improve urban bus operations. The major conclusions are:

1. Strategies designed to improve bus operations involve a blend of several types of improvements. For each facility, it is advisable to explore several candidate strategies prior to the demonstration phase.
2. Simulation has been demonstrated as a viable tool for conducting such evaluations to identify that strategy (or limited set of candidate strategies) which exhibits the highest potential for a successful demonstration project.
3. On the basis of this study, it appears clear that any preferential bus strategy in an urban environ must include consideration of signal control. Furthermore, such consideration should be based upon people-movement measures as opposed to vehicle-movement measures exclusively.

Acknowledgments

This paper was based upon a more detailed report (6). Dr. James Woo applied the SIGOP-II model to generate the new fixed-time signal patterns for this study.

Table 1. Simulation results for bus system
Bus progression strategy
4:30 - 4:45 P.M.

Performance Measure	Base Case	Priority Case	% Change
Number of Buses	116	117	+1
Total Delay (Bus·Minutes) (Dwell Time Excluded)	232.7	204.6	-12
Mean Trip Time (Minutes) (Dwell Time Excluded)	3.19	2.94	-8
Mean Speed (MPH)	6.8	7.3	+8
Number of Intersection Stops	265	217	-18
Total Duration of Intersection Stops (Minutes)	108.8	88.5	-19

Table 2. Simulation results for network general traffic. Bus progression strategy
4:30 - 4:45 P.M.

Performance Measure	Base Case	Priority Case	% Change
Vehicle Miles	1035	1028	-1
Vehicle Trips	3537	3522	0
Vehicle Minutes	5325	5873	+10
Average Speed(MPH)	11.7	10.5	-10
Stops per Vehicle	1.56	1.77	+13
Delay per Vehicle (Sec)	47.9	57.6	+20

Table 3. Simulation results for bus system.
Bus preemption strategy.
4:30-4:45 P.M.

Performance Measure	Base Case	Priority Case	% Change
Number of Buses	116	117	+1
Total Delay (Bus·Minutes) (Dwell time excluded)	232.7	190.7	-18
Mean Trip Time (Minutes) (Dwell Time Excluded)	3.19	2.98	-7
Mean Speed (MPH)	6.8	7.3	+7
Number of Intersection Stops	265	235	-11
Total Duration of Intersection Stops (Minutes)	108.8	90.9	-16

Figure 1. Study network

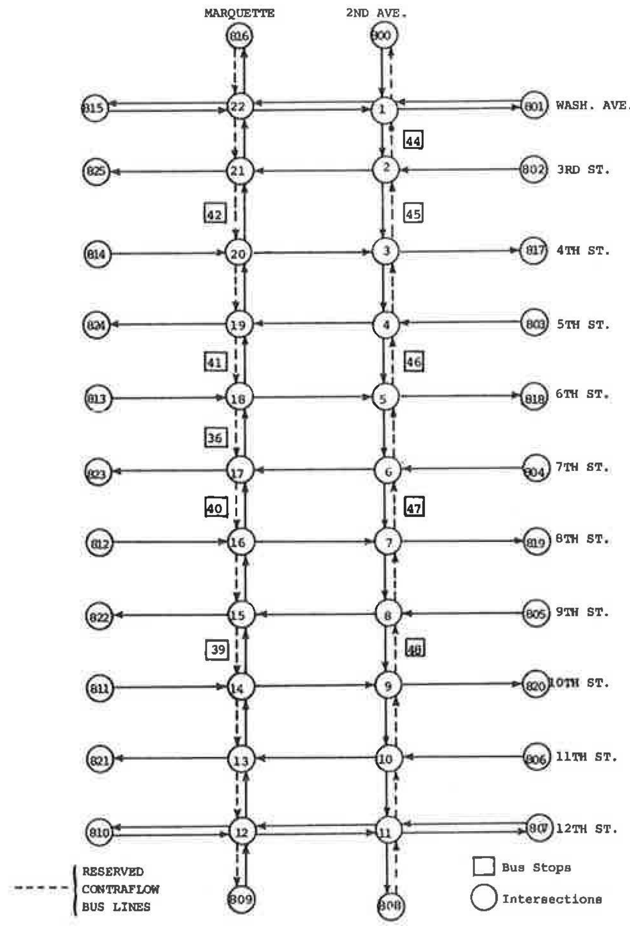


Table 4. Simulation results for network general traffic. Bus Preemption Strategy
4:30 - 4:45 P.M.

Performance Measure	Base Case	Priority Case	% Change
Vehicle Miles	1035	1022	-1
Vehicle Trips	3537	3507	-1
Vehicle Minutes	5325	5808	+9
Average Speed (MPH)	11.7	10.6	-9
Stops Per Vehicle	1.56	1.78	+14
Delay per Vehicle (Sec)	47.9	56.8	+19

References

1. Erdman, J. W. and Panuska, E. J., Jr., "Exclusive Bus Lane Experiment," Traffic Engineering, Vol.46, No.7, July 1976.
2. Lieberman, E. B., "Simulation of Corridor Traffic: The SCOT Model," Highway Research Record 409, 1972.
3. Lieberman, E. B., Worrall, R. D. and Bruggeman, J. M., "Logical Design and Demonstration of UTCS-1 Network Simulation Model," Highway Research Record 409, 1972.
4. Woo, J., "Simulation of Traffic for Analysis and Planning (STRAP), An Extension of the SCOT Model," KLD Associates, Inc., Huntington, N.Y., March 1976.
5. Lieberman, E. B. and Woo, J. L., "SIGOP-II: A New Computer Program to Calibrate Optimal Signal Timing Patterns," KLD Associates, Inc., presented at the 55th Annual TRB Meeting, January 1976.
6. Muzyka, A., et al., "Traffic Signal Strategies for Bus Priority - A Simulation Study," DOT/TSC, January 1977.

SOUTHEAST EXPRESSWAY RESERVED LANE FOR BUSES AND CARPOOLS

Daniel Brand, Charles River Associates Incorporated
John Attanucci and Howard Morris, Massachusetts Executive Office of Transportation and Construction
Charles Kalauskas, (Boston) Central Transportation Planning Staff

On May 4, 1977, the existing northbound left lane of Boston's 8-mile, 8-lane, heavily congested Southeast Expressway was reserved on a voluntary, unenforced basis, for buses and 3-or-more-occupant carpools during the morning peak period 6:30-9:30 A.M. This started Phase 1 of an effort to raise the vehicle occupancy of the highest volume roadway in Massachusetts in anticipation of several years of reconstruction of all bridge decks on and over the Expressway. Phase 2 of Reserved Lane operation began June 2, 1977 by carrying the Reserved Lane through a three-lane construction bottleneck and detour at its northern end. Phase 3, in which the 3-or-more-occupant per vehicle requirement was enforced, commenced the morning of October 18, 1977, and continued until the termination of the Lane on November 2, 1977. The operation of the voluntary lane in Phase 1 increased carpooling on the Expressway by 38 and 72 percent in the 3 hour A.M. peak period and peak hour respectively. In the peak hour, 184 more people were carried in 429 fewer vehicles. Fifty percent of the peak hour persons using the Expressway during Phase 2 were carried in the free flowing Reserved Lane. The entire Expressway operated in Phases 1 and 2 with less congestion than before, no increase in accidents and no measurable impact on alternate surface street traffic attributable to the Lane itself. During the Phase 1 peak period, over 50 percent of the reduction in autos on the Expressway was accounted for by increased vehicle occupancy on the Expressway itself. Rail transit ridership in the corridor increased, accounting for 25 percent of the peak period reduction in autos during Phase 1, indicating the complementarity of alternative high occupancy modes in a high volume corridor. Express bus ridership increased only slightly during all phases of operation. During the only two weeks of operation of Phase 3, travel times in the general-purpose lanes increased and varied from day to day. There was not a significant increase in the number of accidents on the Expressway. Despite continually increasing shifts to alternate modes (a greater than 10 percent shift of all autos on the expressway to carpools and public transportation), the public outcry and concern of public officials regarding the deteriorated travel conditions in the general-purpose travel lanes led to a decision to terminate the project after two weeks of enforced operation. Phase 3 results are presented in the paper, but are not felt to represent equilibrium results.

On May 4, 1977, the Massachusetts Department of Public Works (MDPW) reserved the existing left lane of Boston's Southeast Expressway for buses and three-or-more-occupant carpools. This started Phase 1 of an effort to increase the vehicle occupancy of the highest volume roadway in Massachusetts in anticipation of several years of reconstruction of bridge decks along the entire eight-mile length of the Expressway. Phase 2 of Reserved Lane operation began the morning of June 2, 1977 by carrying the Reserved Lane through a three-lane construction bottleneck at the northerly end of the eight-mile Reserved Lane. Phase 3, in which the three-or-more-occupant per vehicle requirement was enforced, commenced the morning of October 18, 1977 and continued until the termination of the Lane on November 2, 1977.

The Reserved Lane, called the Downtown Express Lane locally and in this paper, was an experimental cooperative effort between several Massachusetts transportation agencies and was the key element in a program for traffic maintenance during the Southeast Expressway reconstruction. Two aspects of the Lane are vitally important to consider when comparing this project to other preferential lane demonstrations:

1. The Lane removed an existing general-purpose traffic lane.
2. The Lane through Phases 1 and 2 reported on in this paper was a *voluntary* lane. Violators of the Lane were *not* ticketed.

Project Description

The Southeast Expressway is one of only three radial limited access highways penetrating all the way to the Boston CBD from the Route 128 limited-access circumferential highway about twelve miles out from Boston (with a metropolitan population of

3.5 million). Average daily traffic on this heavily congested Expressway has experienced only slow growth over the last several years and was 126,000 vehicles a day just south of the Massachusetts Avenue interchange at Southampton Street in 1976, making it the most heavily traveled highway in the State.

The Reserved Lane for buses and 3-or-more-occupant carpools extended over an 8-mile section of the Southeast Expressway between a point 1500 feet north of its intersection with Route 128 in Quincy to the Massachusetts Avenue interchange in Boston.

The Reserved Lane was the far left lane in the northbound direction and operated only on weekdays between 6:30 and 9:30 A.M. By reserving the Express Lane for high-occupancy vehicles only, these vehicles would be the recipients of substantially reduced travel times. These reduced travel times were intended to encourage the use of express buses and the formation of carpools to improve the "people-moving" capacity of the Southeast Expressway during the morning peak period.

Daily Operation of Express Lane

During the morning peak period, the Downtown Express Lane was separated from the three general lanes of traffic by yellow 19" high plastic posts which were inserted into 8" metal sleeves embedded in the roadway. The posts were spaced 20 feet apart in some heavily congested areas of the Expressway and 40 feet apart along the remaining length. They were inserted daily beginning at approximately 5 A.M. and removed after 9:30 A.M. (Setup and pickup times were each about 75 minutes using two truck crews.)

Phases of Operation

The results of three distinct phases of Downtown Express Lane operation are described in this paper. Phase 1 began May 4, 1977 and provided four weeks of Express Lane operation prior to actual Southeast Expressway reconstruction. These four weeks were intended to allow time for carpool formation, a massive publicity campaign (which actually started one month prior to Phase 1), and buildup of express bus ridership prior to the bottleneck caused by reconstruction of the bridge decks at the Massachusetts Avenue interchange.

At midday June 1, 1977, the four travel lanes in the northbound direction in the vicinity of the Massachusetts Avenue interchange at the northern end of the Express Lane were shifted over to a temporary three-lane detour roadway. The far left Reserved Lane was carried all the way through the three-lane bottleneck section. Phase 2 of the Lane's operation (with the detour through the three-lane bottleneck section) commenced on the morning of June 2, 1977.

Phase 3, the enforcement phase, began on October 18, 1977. Vehicles with less than three occupants traveling in the Lane were subjected to a fine of $20.00.

Enforcement

As noted in the introduction, the Downtown Express Lane was a voluntary lane for Phases 1 and 2. This means one- and two-occupant vehicles using the Lane were not ticketed.

The decision to operate the Express Lane with compliance on a voluntary basis during Phases 1 and 2 was made for several reasons.

1. The voluntary approach simplified the legal requirements for implementation and enforcement of the Express Lane project. (No federal money was involved in the operation of the Lane which removed the NEPA EIS requirements.)

2. Public acceptance of a voluntary lane would be greater and the concept could be proven without alienating those opposed to the project at the start. The responsibility for the success or failure of the project was, therefore, shifted to the general public (and each commuter then using the Expressway) and away from a focus on the police's ability or right to enforce the three-occupant carpool requirement.

The Phase 3 enforcement of the three-person-per-vehicle minimum requirement for use of the Express Lane began on October 18, 1977, after more than five months of voluntary operation. A regulation to enforce the Lane was issued by the Massachusetts Department of Public Works (MDPW). The MDPW regulation provided for a maximum fine of $20 for the owner of a vehicle cited for traveling in the Downtown Express Lane in violation of the posted signs. The police assured the MDPW that they would enforce such a regulation and, therefore, would issue citations by mail. The basic premise of enforcement was that the owner of the vehicle would be liable for the use of a vehicle in violation of the three-or-more-person requirement. The officer was not required to stop and cite the violator on the spot because of safety considerations, but simply noted the vehicle's license plate number. This procedure is similar to that for a parking ticket and allowed the direct mailing of the citation to the vehicle owner. A similar regulation, which provides for mailing of citations by the State Police upon observation of a toll evader on the Massachusetts Turnpike, has been in effect for many years and was upheld by the Massachusetts Supreme Judicial Court as within the normal police powers of the Commonwealth. (See *Commonwealth v. Pauley*, 331 NE 2d 901, 1975).

Cost of the Express Lane Project

The cost of implementing the Downtown Express Lane project consisted of a minor capital expenditure and a regular operating expense. Approximately $40,000 was expended for the publicity campaign including the special carpool matching effort. The 1500 plastic post inserts needed to separate the Express Lane from general traffic cost $11 each, totaling about $16,000. Two thousand replacement posts (approximately 15-18 posts needed replacement daily) cost $22,000. Signing and pavement markings for the Lane cost approximately $7,500. Approximately $5,500 was expended for labor and equipment to install the post sleeves, erect the signs and paint the pavement markings in anticipation of the start of the Lane. The total capital cost of the project therefore was approximately $92,000.

Operating cost of the Lane includes expenditures for MDPW crews to set down and pick up the plastic post inserts daily, and police protection for the crews performing these tasks. Weekly MDPW crew costs averaged $3,200 and weekly State Police overtime costs were $540. If the Lane were to operate year-round in this manner, yearly operating costs for the project would total approximately $195,000.

Phase 3 enforcement costs were not estimated due to the short duration of the Phase 3 operation. About five police officers in vehicles were assigned over the three-hour period per day. The fact that tickets did not have to be issued on the spot allowed increased productivity from a minimum number of officers. If necessary, hundreds of citations could have been issued.

Alternative Commuting Facilities to the Southeast Expressway

Alternatives Prior to the Reserved Lane Project

Before detailing the results of the Reserved Lane project, it is important to describe the travel choices available to South Shore commuters before and during the project.

The extensive public transportation system has four components: MBTA rail rapid transit with feeder bus, private carrier express buses, commuter rail trains and commuter boat. The importance of this network is illustrated by the high percentage of peak period trips from the South Shore to the CBD (approximately 60 percent) which are made by transit.

Rail rapid transit service between Boston and the South Shore is provided by two branches of the Red Line operating at 5-minute headways during peak periods on each line. Each weekday morning, 8,750 riders boarded at the three Quincy stations on the Quincy branch of the Red Line during the 6:30-9:30 A.M. peak period during March and April 1977 before Phase 1. Extensive feeder bus service is provided from many South Shore communities to Red Line stations in Quincy and Dorchester (Boston).

Express bus service direct to Boston's CBD is provided from a large number of communities south of Boston by four private carriers: Plymouth and Brockton (P&B), Almeida, Hudson and Bonanza. All express buses use the entire 8-mile length of the Southeast Expressway on which the Lane is located. Each weekday during the 6:30-9:30 A.M. peak period, approximately 100 bus-runs are made in the northbound direction on the Southeast Expressway carrying a total of 3,400 passengers.

Commuter rail service is provided to an adjacent (southwest) corridor as far south as Providence, R.I. Frequent service is provided only during peak periods, with weekday peak period inbound ridership totaling approximately 2,600 passengers.

Commuter boat service prior to the start of the Downtown Express Lane consisted of one trip each way each day from Hull, Massachusetts to Rowe's Wharf in downtown Boston. This service accommodates approximately 125 riders each way during the summer months.

In addition, the dense highway network in Boston and the South Shore provides many alternative surface street and arterial routes to the Boston CBD. Morrissey Boulevard is a six-lane arterial running approximately four miles between the Neponset River and the Massachusetts Avenue interchange. Other routes consist of many lesser streets and roadways which drivers connect up in almost infinite variety. Some Southeast Expressway users have origins and destinations far enough to the west of Boston and the South Shore that the large limited access circumferential Route 128 and radial highways to the southwest and west of Boston, including the limited access Massachusetts Turnpike are convenient alternate routes.

Transportation Services Provided for the Reserved Lane Project

A number of transportation improvements were provided for Southeast Expressway commuters in addition to the Downtown Express Lane as part of the traffic maintenance plan for Southeast Expressway bridge deck reconstruction. These transportation services were all aimed at using high occupancy vehicles, on or off the Expressway, and were aggressively promoted during the publicity campaign which preceded Phase 1. The service improvements included:

- Providing maximum service levels on the MBTA rapid transit Red Line serving the South Shore by increasing the number of transit cars available for service from 88 to 104, and providing some additional feeder bus service to Red Line stations.
- A completely new express bus route from two major commuter parking areas on Route 128 began May 9, 1977 providing service at 20-minute headways for the peak period to Boston's Government Center in the northern part of the CBD.
- The major private bus company promised to provide up to a 30 percent increase in numbers of bus runs consisting principally of extra sections on high-density portions of their extensive route system. The other smaller private carriers generally felt they had sufficient empty seats to serve up to a 50 percent increase in ridership.
- Seven existing and two new fringe parking lots were expanded and upgraded by the MDPW, MBTA and the Metropolitan District Commission (MDC) prior to, or during, Phase 1 of the Lane.
- Carpool matching assistance was provided to South Shore commuters through a variety of high visibility mechanisms.
- Additional one-round-trip-each-day commuter boat services were initiated in May 1977 from two South Shore locations to downtown Boston.

In addition to the new South Shore transportation services listed above, substantial increase in police patrols on the Expressway and additional emergency highway equipment including tow trucks and push bar equipped police vehicles were provided for quick removal of disabled vehicles in order to keep traffic flowing smoothly.

Results of Phase 1 and Phase 2 Operations

Introduction

The results of the Southeast Expressway Reserved Lane for buses and carpools are organized by four major categories of information:

1. General public acceptance.
2. Impact on travel and travel conditions on the Southeast Expressway.
3. Impact on travel and travel conditions on other modes and highways (i.e., off the Expressway).
4. Summary of where the cars "went."

These results are presented in this section for Phases 1 and 2, the five month voluntary period of Lane operation. The next section presents these results for the brief two-week period of enforced Phase 3 operation. A more detailed account of the methodology employed in the monitoring and evaluation program for the Downtown Express Lane is given in the first part of that section. (Central Transportation Planning Staff, *Southeast Expressway Evaluation of Downtown Express Lane*, December 1977.)

Public Acceptance

The response of the general public to the Downtown Express Lane before and during implementation of (voluntary) Phases 1 and 2 was generally positive and without major controversy. The justification of the project as the critical element in the traffic maintenance plan during Southeast Expressway reconstruction seemed to diffuse opposition to the concept. It is clear that the fact that the Lane was voluntary and not mandatory quieted an important segment of the population who otherwise would have vehemently objected to the Lane.

Editorial comments in the newspapers generally were favorable and expanded on the news reporting theme of "it's for everyone's good." Public meetings in the affected communities produced only a few interested citizens whose ideas would generally make the Lane's operation more complex. Legislators from the South Shore area were vocal that complementary actions such as securing additional fringe parking sites and providing additional public transportation service had not gone far enough.

Public response to a newly established C-A-R-P-O-O-L phone number to obtain matching information and a temporary information booth at a Howard Johnson's Restaurant on the Expressway was mixed. Large numbers of commuters requested information on the Express Lane project and related construction activity, but few of these commuters requested carpool matching assistance. In May and June, the first two months of operation of the Lane, a maximum of 120 calls per day was received at the C-A-R-P-O-O-L number, with average daily calls being far less than this number. At the Howard Johnson's information booth on the southbound side of the Expressway, a total of 640 inquiries were made by commuters during the eight-week period (April 11, 1977 through June 2, 1977) the booth was open. Out of all these requests for information over a two and one-half month period, only about 430 were requests for carpool matching information. About a third of these resulted in a match with at least one other person and the mailing of a carpool matching list.

Changes in Travel and Travel Conditions on the Expressway

The results of the Downtown Express Lane on travel at Southampton Street on the Expressway for the three reporting periods ("Before," Phase 1, and Phase 2) are presented in Table 1. Southampton Street is a cross street near the northern end of the Lane, a point where the highest volumes on the Expressway are generally observed.

There is no clear and consistent monthly variation in Southeast Expressway travel volumes between March, April, May and June. Therefore, the data are not "seasonally adjusted" for month or year. However, morning peak period travel for the 6:30 to 9:30 A.M. period that the Lane wasin operation declined substantially in July and August. Also, travel on Mondays and Fridays in this corridor is distinctly different from Tuesday, Wednesday and Thursday travel due to a carry-over of weekend travel to Mondays and Fridays in this corridor leading to the South Shore and Cape Cod. For these reasons, very little data were collected on Mondays and Fridays and during the months of July and August.

Travel on the Entire Expressway. As may be seen in Table 1, the number of carpools on the Expressway during the peak period (6:30-9:30 A.M.) grew by 38 percent or 331 carpools during Phase 1. This increase dipped to 15 percent or 133 carpools in Phase 2 relative to the "Before" condition. For the peak hour, the corresponding growths in carpooling were 72 percent or 268 vehicles for Phase 1 and 34 percent or 129 vehicles for Phase 2 relative to the "Before" condition. The table shows corresponding growths in the percent of total persons in cars carried in carpools and buses. Although the number of carpools declines in absolute terms between Phase 1 and Phase 2 of the Lane's operation, the percent of persons carried in carpools declined far less, and the percent of persons carried in carpools and buses increases because of the drop in numbers of persons and vehicles carried on the Expressway during the Phase 2 bottleneck. That is, carpooling in relative terms dips only slightly between Phase 1 and Phase 2. During the latter part of June 1977, vacations started and this caused more difficulty in carpooling which was reflected in the data. The Phase 2 data are average for the entire month.

Express bus ridership increased by only approximately 100 riders during the peak period and 65 riders during the peak hour, or about a 3 percent increase in both cases. The increases appear to have been solely due to the reduction in travel time in the reserved Lane, and not due to the new service provided during the Lane's operation. Ridership did not change between Phase 1 and Phase 2. These somewhat disappointing increases for the first two months of the Lane's operation match the experience of the contraflow lane provided on the Southeast Expressway during morning peak periods of daylight savings time months for the previous six years. It is also consistent with work purpose direct elasticities for line haul transit travel time of -.3 to -.4 (i.e., the approximate 10 percent decrease in line haul travel time has produced a 3 percent increase in ridership).

Table 1 shows that during Phase 1, 2,010 fewer autos used the Expressway, but only 1,130 fewer persons were accommodated on the Expressway during the peak period, and for the peak hour, 184 *more* persons were accommodated and there were *fewer* autos. This shows the significant effect of the Lane itself in increasing the average occupancy of autos (from 1.31 to 1.40 during the peak period, and from 1.34 to 1.49 during the peak hour for Phase 1), and in preparing the Expressway to accommodate passengers in higher occupancy vehicles during the Phase 2 Expressway reconstruction.

Shift in Time of Travel. Between 6:00 and 6:30 A.M., the number of vehicles using the Southeast Expressway during Phase 1 decreased in the same proportion as the reductions in 6:30 to 9:30 A.M. volumes shown in Table 1. During Phase 2, the 6:00 to 6:30 A.M. decrease was only one-third the 6:30 to 9:30 A.M. decrease, while during Phase 2, the post-peak decrease was one-half the peak period decrease. Person travel shifted by similar amounts due to the similar auto occupancy results for the peak and post-peak periods. There did not appear to be any shifting of travel within the three-hour peak period. The range of autos shifting to the post-peak period was 0-250 for Phase 1 and 100-500 for Phase 2.

Express Lane Utilization and Compliance. Table 1 shows the percent of total persons, vehicles and persons carried in cars in the Express Lane at Southampton Street for both the peak period (6:30 to 9:30 A.M.) and the peak hour (7:00 to 8:00 A.M.) for Phases 1 and 2. For the peak period, the Lane carried 37 percent and 46 percent of the total persons on the Expressway during Phase 1 and Phase 2 respectively. For the peak hour, the figures increased to 43 percent and 50 percent during Phase 1 and 2 respectively. These figures include persons carried in one- and two-occupant vehicles which "violated" the Lane restriction. Nevertheless, it is impressive that during the peak hour up to half the persons on the Expressway experienced a smooth and congestion-free ride (as will be shown in the next section) on the one Reserved Lane.

A significant result is that during the peak hour, the Lane carried a proportionate number of *vehicles* to the number of lanes available in each phase (i.e., 25.1 percent of the vehicles in 1 of 4 lanes and 31.2 percent of the vehicles in 1 of 3 lanes available). That the Express Lane moved freely is due to the lack of weaving in the Lane, and the fact that the right lane carries relatively few vehicles due to its high number of weaving movements near the frequent ramps.

Phases 1 and 2 of the Downtown Express Lane, as noted often above, were a voluntary lane which gave preference to high-occupancy vehicles. The most disappointing aspect of the Lane, therefore, in view of its high people-carrying capacity, and what will be shown below to be its safe operation and lessening of congestion for all Expressway users, was the high violation rate during Phase 1 and Phase 2 operation. The compliance rate (percent of total vehicles in the Lane which are buses and 3-or-more-occupant autos) at the beginning of the Lane for both Phase 1 and Phase 2 ranged between 23 percent and 53 percent and averaged about 36 percent. The compliance rate is highest during the peak hour when there are more carpools available to fill the Lane. The compliance rate at the northern end (Southampton Street) ranged between 16 percent and 24 percent during both Phase 1 and Phase 2. On the average, however, the Phase 1 compliance rate was 21 percent and for Phase 2, it dropped to 19 percent. These statistics indicate there are substantial numbers of violators weaving into the Downtown Express Lane along its length. However, police cruisers located at the beginning of the Lane did have an effect in dissuading non-carpool vehicles from entering the Lane at its beginning.

Travel Conditions on the Expressway and in the Lane.

Travel Times. Table 2 shows travel times for the entire length of the Southeast Expressway for the "Before," Phase 1, and Phase 2 periods at half-hourly intervals from 6:30 A.M. to 9:00 A.M. through June 1977.

In general, and in particular during the time of peak congestion between 7:30 and 8:00 A.M., it can be seen that the travel times during Phase 1 and Phase 2 were shorter for all lanes than before the implementation of the Downtown Express Lane. The fears of tie-ups from "taking away a lane" were unfounded. During the times of greatest congestion before the Lane, namely between 7:30 and 8:00 A.M., users of the Express Lane experienced travel time savings of 9 minutes, while general purpose lane users had time savings of between 4 and 8 minutes. These time savings for all lanes even increased slightly during Phase 2.

The possible slight increase in travel times at 6:30 A.M. during Phases 1 and 2 was not due to additional congestion in the normal sense. The time increases were caused by the dampening effect on speed of the presence of the lines of posts delineating the Lane and the barrier at the beginning of the Lane. This apparently had a positive safety effect.

An important probable cause for the decreased travel times on all lanes of the Expressway during Phases 1 and 2 was the metering effect of reducing the Expressway from four to three lanes at the start of the Lane (the Lane "started empty"). Also, the presence of the reserved Lane reduced weaving movements on the entire length of the Expressway. This resulted in smoother traffic flow downstream. (The travel times in Table 2 include the time to pass through the often congested area at the start of the Lane where the metering took place.)

Waiting Times at On-Ramps. Most morning peak period volumes on the four major on-ramps to the northbound roadway of the Southeast Expressway decreased in proportion to the decreased traffic on the Expressway itself during Phases 1 and 2. More importantly, and in line with the decreased congestion on the "main line," the average and maximum waiting times for these on-ramps decreased between the "Before" condition and Phases 1 and 2. For example, at the high-volume Neponset Avenue on-ramp, which has about 43 percent of the total on-ramp traffic of the four ramps combined, average and maximum wait times were reduced to about one-half their "Before" values during Phase 1, and to about one-quarter of their "Before" values during Phase 2.

Safety. Personal injury accidents on the Southeast Expressway for the months of May and June from 1970 through 1976 ranged from 0-9 with a 3.0 average for May, and 1-4 with a 2.3 average for June. Property damage accidents ranged between 2-8 with a 4.7 average for May, and 4-12 with a 6.7 average for June over the same seven years.

For better or worse, more careful accident reporting characterized the first two months of operation of the Downtown Express Lane than previous Mays and Junes. As noted before, police patrols were greatly increased on the Expressway which substantially improved the detection and reporting of accidents during the Lane's operation. It must be assumed that "fender-benders" and similar property damage accidents are included in the accident statistics for Phase 1 and Phase 2 in addition to the more major rear-end, head-on, and other accidents included in the standard reporting. The reporting of personal injury accidents would be less affected by the increased police patroling during 1977.

During the entire month of May 1977, including the Lane's Phase 1 operation from May 4 on, there were 6 personal injury accidents and 6 property damage accidents. During all of June 1977 (Phase 2), there were 3 personal injury accidents and 10 property damage accidents. Both months' figures fall within the range of accidents reported by the normal police patrols between the years 1970 and 1976. In addition, only two of the May 1977 accidents occurred in or could be associated with the Express Lane. The corresponding figure for June was 1 of the 13 accidents. There were no fatalities on the Expressway in May or June 1977.

Changes in Travel and Travel Conditions Off the Expressway

Rail Rapid Transit (Red Line). Seasonally adjusted ridership counts during the 6:30 to 9:30 A.M. peak period at the three Quincy stations on the Quincy Branch of the Red Line showed an increase in weekday peak period boarding of 600 persons for both Phase 1 and Phase 2 over the "Before" number of 8,750 boarders.

Ridership on the Ashmont Branch of the Red Line was not perceptibly affected. This branch, further from the impacted area, would have had any increases in ridership distributed over many stations.

Commuter Rail. Commuter rail from the Southwest Corridor experienced no significant increase in ridership during Phase 1. However, in June 1977 (Phase 2), seasonally adjusted commuter rail ridership increased by approximately 100 riders to 2,650 inbound boardings during the 6:30 to 9:30 A.M. peak period.

Commuter Boat. During Phases 1 and 2, in May and June 1977, total ridership on all three commuter boats

Table 1. Vehicles and Persons Traveling on the Southeast Expressway and in the Downtown Express Lane at Southampton Street

	Before March	Phase 1 May Number	Phase 1 May Change	Phase 1 May % Change from Before	Phase 2 June Number	Phase 2 June Change	Phase 2 June % Change from Before
Peak Period (6:30-9:30 A.M.)							
• All Lanes							
No. of Carpools (3 or more occupants)	877	1208	331	37.7%	1010	133	15.2%
% of Persons (in Cars) in Carpools	12.8%	18.3%	5.5%	43.0%	16.7%	3.9%	30.5%
% of Total Persons in Carpools & Buses	23.4%	28.9%	5.8%	22.2%	32.6%	9.2%	39.3%
No. of Bus Passengers	3400	3500	100	2.9%	3500	100	2.9%
Total No. of Persons in Autos & Buses	27916	26780	-1136	-4.0%	21800	-6116	-21.9%
Total No. of Vehicles	19429	17537	-1892	-9.7%	13740	-5689	-29.3%
Total No. of Autos	18677	16668	-2009	-10.8%	13010	-5667	-30.4%
Total No. of Single-Occupant Autos	14223	12018	-2205	-15.5%	9334	-4889	-34.4%
Total No. of Two-Occupant Autos	3577	3442	-135	-3.8%	2666	-911	-25.5%
Average Auto Occupancy	1.31	1.40	0.09	6.9%	1.41	0.1	7.6%
• Traveling in Downtown Express Lane							
% of Total Persons on Expressway	---	37.0%	---	---	46.0%	---	---
% of Autos on Expressway	---	28.6%	---	---	36.6%	---	---
% of Vehicles on Expressway	---	22.2%	---	---	30.7%	---	---
Peak Hour (7:00-8:00 A.M.)							
• All Lanes							
No. of Carpools	373	641	268	71.8%	502	129	34.6%
% of Persons (in Cars) in Carpools	14.5%	24.4%	9.9%	68.3%	21.2%	6.7%	46.2%
% of Total Persons in Carpools & Buses	23.4%	38.3%	14.9%	63.7%	45.3%	21.9%	93.6%
No. of Bus Passengers	2000	2065	65	3.3%	2065	65	3.3%
Total No. of Persons in Autos & Buses	11008	11257	249	2.3%	8542	-2466	-22.4%
Total No. of Vehicles	6902	6473	-429	-6.2%	4698	-2204	-31.9%
Total No. of Autos	6704	6185	-519	-7.7%	4490	-2214	-33.0%
Total No. of Single-Occupant Autos	4960	4140	-820	-16.5%	3048	-1912	-38.5%
Total No. of Two-Occupant Autos	1371	1404	33	2.4%	941	-430	-31.4%
Average Auto Occupancy	1.34	1.49	0.15	11.2%	1.50	0.16	11.9%
• Traveling in Downtown Express Lane							
% of Total Persons on Expressway	---	42.6%	---	---	50.3%	---	---
% of Autos on Expressway	---	25.5%	---	---	31.7%	---	---
% of Vehicles on Expressway	---	25.1%	---	---	31.2%	---	---

Table 2. Travel Times (in Minutes) on the Southeast Expressway Northbound from Union Street (Braintree) to Kneeland Street (Boston)

	Before (March)	Voluntary Phase 1 (May)		Voluntary Phase 2 (June)		Enforced Phase 3 (October)	
Start Time	All Lanes	General Lanes	Express Lane	General Lanes	Express Lane	General Lanes Average 3 Days	Express Lane
6:30 A.M.	16 (14-17)*	17 (16-18)	18	17 (15-20)	17	22 (17-25)	15
7:00 A.M.	20 (18-22)	20 (19-22)	20	18 (16-19)	15	28 (22-32)	14
7:30 A.M.	28 (25-31)	24 (23-25)	23	24 (19-27)	17	40 (35-43)	18
8:00 A.M.	28 (25-30)	22 (21-22)	21	21 (17-25)	19	36 (30-42)	18
8:30 A.M.	23 (20-26)	17 (16-18)	16	17 (15-19)	14	30 (25-35)	16
9:00 A.M.	17 (14-20)	16 (15-16)	16	17 (14-20)	14	21 (17-26)	14

*Numbers in parentheses are absolute ranges with the exception of the "Before" numbers, which denote the likely range based on a 95 percent confidence interval and a t-distribution.

including the two new services accompanying the start of the Lane, was 295 persons inbound to Boston, a seasonally adjusted increase in riders of 170 persons. A special survey indicated approximately one-third of the new commuter boat users were former auto drivers. This accounts for a seasonally adjusted removal of an estimated 50 automobiles from the Southeast Expressway as a result of improved commuter boat service.

Fringe Parking. No change in the utilization of the major fringe parking lots was observed during Phases 1 and 2, with the exception of one new surface lot serving the MBTA rapid transit Red Line in Quincy which opened during Phase 2. This lot was utilized by 224 cars during Phase 2, but served to relieve the capacity constraint of the three Red Line parking lots in Quincy. This indicates that new carpoolers found it more convenient to collect their friends and neighbors at their homes or at small widely scattered parking places. The presence of additional fringe parking appears also not to have significantly affected express bus use.

Alternative Highway Routes. Diversion of automobiles to alternative routes made up of local and major streets and arterials is difficult to measure in the South Shore corridor because of the presence of the dense road network. In addition, traffic from the south and southwest headed to points west and north of downtown Boston can use Route 128 and radial arterials from 128 to these destinations as an alternative to the Southeast Expressway. Numerous peak period volume counts during Phase 1 and 2 were made on 14 alternative major streets and highways including Route 128, and travel time runs were made on 10 different alternate routes. It is clear from the volume counts in Table 1 that traffic was substantially reduced on the Expressway, particularly during Phase 2. The next section will attempt to summarize what happened to the cars that "disappeared." Meanwhile, the complexity of the network and the highly variable traffic volumes in this corridor made it very difficult to measure the exact number of automobiles diverted to surface streets or even to detect any locations where statistically significant increases in traffic volumes occurred, particularly in Phase 1.

Travel time studies on the alternative routes showed no deterioration in service. This indicates that the shifted traffic did not concentrate on a small number of streets and that the volume increases were small compared to the available capacity. This was the case even for Phase 2 for which substantial reductions in automobile volumes on the Expressway were observed.

Summary of Where the Cars "Went"

As has been repeatedly stated, the overall purpose of the Downtown Express Lane was to minimize the impact of Southeast Expressway reconstruction on both Expressway travelers and highway travel in general in the corridor. Table 3 summarizes the estimates of the reduction in auto travel on the Southeast Expressway between 6:30 and 9:30 A.M. accounted for by diversions to the transportation alternatives described above.

The peak period increase in numbers of persons carpooling on the Expressway of 1200 and 500 during Phases 1 and 2 respectively, divided by the "Before" condition auto occupancy of 1.31 at Southampton Street, yields the diversion of 920 and 380 autos to carpools on the Expressway during Phases 1 and 2 respectively. Alternatively, the increase in peak period car occupancy at Southampton Street during Phase 1 of 1.40 represents a 6.9 percent increase over the "Before" occupancy of 1.31. This means that the same number of people could be carried in approximately 6.9 percent fewer autos or approximately 1,140 fewer autos. The similar result for Phase 2 is 970 fewer autos. The entries in Table 3 reflect the decrease in carpooling from Phase 1 to Phase 2 based on a combination of the two methods.

The increased weekday peak period Red Line boardings of 600 persons, divided by the "Before" auto occupancy of 1.31, yields the 460 auto diversion to the Red Line shown in Table 3. The 50 automobile diversion to commuter boats was described above. The express bus ridership increase on the Southeast Expressway of 100 persons during both phases is divided by 1.31. to obtain the 75-car figure shown in the table. Similarly, the 100-passenger commuter rail increase during Phase 2 is noted in the table as diverting 75 cars.

Three important conclusions can be drawn from Table 3. First, the results for Phase 1 show that, by itself, the reservation of an existing Expressway Lane on a voluntary basis for buses and carpools did not increase traffic on alternative surface streets and highways, much less affect congestion on these alternative roads. Between 75 percent and 90 percent of the reduction in automobiles on the Expressway is accounted for by modal shifts, with over 50 percent of the auto reduction accounted for by increases in vehicle occupancy on the Expressway itself. The usefulness and complementarity of the parallel public transportation service on its own right-of-way in the same corridor (the Red Line), which accounted for about 25 percent of the reduced number of cars in Phase 1, should also be highlighted in planning for similar reserved lanes.

Second, the results for Phase 2 shown in Table 3 must be viewed in the context of the substantial capacity constraint imposed by the construction detour just north of Southampton Street which narrowed the Expressway from four to three lanes. It seems clear that when the bottleneck occurred, the persons who perceived sufficient reason to change their travel behavior simply shifted their travel routes or their time of travel (minor) or decided not to make the trip. The publicity campaign preceding Phase 1 of the Lane and the travel time advantage of the Lane appear to have stimulated all who would carpool to shift modes to have done so during Phase 1.

Finally, it must be concluded that even though Phase 2 did not produce additional carpooling or express bus usage, there were fewer low-occupancy autos "available" to shift to alternate routes. In this sense, the Lane was successful in reducing the highway travel impacts of Expressway reconstruction in the South Shore corridor during Phase 2.

Results of Phase 3 (Enforcement Phase) Operations

Phase 3 of the Lane's operation began on October 18, 1977 with enforcement of the 3-or-more-person per vehicle requirement for use of the Lane and continued until termination of the project on November 2, 1977. The decision to enforce the Lane was based on several factors:

- The fact that violators were being rewarded with a congestion-free ride generated significant public and media demands for enforcement of the carpool requirement.
- The completion of construction at the northern end of the Lane in early October returned the northbound roadway to four lanes. This made it feasible

to enforce the Lane since general lane users would have three lanes along the entire length of the Expressway.

• The concept of the reserved express lane had been demonstrated to be operationally feasible (i.e., the Lane alone during Phase 1 had achieved the goal of increased carpooling and a decreased number of vehicles traveling on the Expressway with reduced travel times of non-Lane users).

• The continuing downward trend in the compliance rate over the summer was jeopardizing the success of the Lane (i.e., as the Lane became filled with autos with less than three occupants, the relative travel time advantage of the Lane decreased).

It must be stressed that the results for the two weeks of Phase 3 operation presented here are an attempt to represent dynamic phenomena. There are indications of favorable trends towards equilibrium, particularly during the Monday to Wednesday (October 31 to November 2), which were the last three days of the Lane's operation. However, great caution should be exercised by anyone seeking to use the results presented below as representative of equilibrium conditions for an enforced reserved lane that "takes away" an existing general purpose lane.

Public Response (Phase 3)

Only modest opposition was voiced when MDPW officials announced their intention to enfore the Lane. The most vehement opposition did not develop until the actual enforcement began.

With the introduction of the police officers on the roadway recording license plate numbers of violators on October 18, 1977, and the resultant traffic delays discussed elsewhere in this section, the opposition to the concept became more vocal and pervasive. The more conservative major Boston daily newspaper began running front-page columns that included a high degree of negative editorializing on the subject, after providing exceptionally objective and complete reporting of the summer's successful Phase 1 and 2 experience. Within a week, the paper called the test "a flop" in their lead editorial and began to run an array of letters in opposition to the enforced lane. The *Boston Globe* and the *Christian Science Monitor* remained editorially neutral and reported only the enforcement statistics of the first few days. The electronic media increased their coverage at the onset of the enforcement period, with television reports generally providing a gloomy picture of the Expressway experiment. Radio reports at first concentrated on warning commuters of the new fine being imposed on violators, but quickly began to take the editorial slant of the station or particular announcer.

Because first-day operations were typically confusing and traffic was snarled badly, commuters began a fairly steady flow of angry phone calls and letters to the MDPW, EOTC, and police agencies, legislative representatives and the Governor. Phone calls to the MDPW totaled between 200 and 300 during the two weeks of enforcement, almost all of which vigorously opposed the Lane. An emergency bill was filed to change the requirement for the Lane to two-person carpools, and a long-dormant bill was reactivated to abandon the Lane entirely. A well-publicized hearing was scheduled for Wednesday evening, November 2, 1977, at 5:00 P.M. to discuss the two bills. During the entire two week enforcement period, hardly a word was heard from new and old carpoolers, bus users and other supporters of the Express Lane concept.

Changes in Travel and Travel Conditions on the Expressway (Phase 3)

The impact of the Downtown Express Lane on travel during Phase 3 as compared with the "Before" condition (March 1977) is presented in Table 4 at a point near the southern end (or beginning) of the Lane (Furnace Brook Parkway). Since November, travel is very similar to March travel on the Expressway, no seasonal adjustment factors were applied.

Travel on the Entire Expressway. As may be seen from Table 4, the number of carpools at Furnace Brook during the peak period increased by 71 percent, from 681 in the "Before" condition to 1,166 during Phase 3. Approximately 225 of these 485 additional carpools were newly formed during the two-week enforcement period. During the peak hour (7:00-8:00 A.M.) at the same location, the number of carpools increased from 388 in the "Before" condition to 641 in Phase 3 which represents a 65 percent increase. Increases in carpooling at the northern end (Southampton Street) were not quite as large, probably because the impacts of construction activity at the Massachusetts Avenue interchange continued during Phase 3 even though the northbound detour was removed in early October 1977.

The percentage of persons traveling in carpools on the Expressway increased during the peak hour and peak period. At Furnace Brook Parkway in the peak hour, the percentage of persons traveling in carpools increased from 17.3 percent in the "Before" condition to 29.6 percent in Phase 3. During the peak period, the percentage of persons in cars that traveled by carpool more than doubled at Furnace Brook Parkway, from 10.5 percent in the "Before" condition to 22 percent during Phase 3.

Express bus ridership increased by only about 200 riders during the peak period to approximately 3,600 riders. The 200 new riders represent an increase of about 6 percent from the "Before" condition. Bus ridership was showing an upward trend during the two weeks of enforcement, with the largest increases of the entire Downtown Express Lane project occurring during the final week of the enforced operation.

Table 4 shows that at Furnace Brook Parkway, there was a 14.4 percent decrease in travel volume in the peak period (2,333 fewer vehicles), but the number of persons decreased by only 8.2 percent (1, 937). This represents an increase in auto occupancy from 1.30 to 1.39 in the peak period. During the peak hour (7:00-8:00 A.M.) at Furnace Brook Parkway, the corresponding percent reductions in total vehicles and total persons were 15.2 percent and 5.8 percent, representing an even greater increase in auto occupancy (from 1.37 to 1.49).

Shift in Time of Travel. As a result of increased travel times for general-purpose lane users on the Expressway during Phase 3, some travelers shifted their time of travel. From counts taken during the hour immediately following the operation of the Lane (9:30-10:30 A.M.), it has been estimated that up to 250 vehicles shifted their travel to this post-peak hour. Because of a lack of data for the half hour preceding the Lane's operation (6:00-6:30 A.M.), a similar analysis could not be completed for that period.

Express Lane Utilization and Compliance. During the Phase 3 peak period, the Downtown Express Lane carried one-third of all commuters on the Expressway in approximately 15 percent of the vehicles, and during the peak hour it carried over 40 percent of

Table 3. Summary of Estimated Changes in Travel Behavior on Southeast Expressway at Southampton Street, 6:30 - 9:30 A.M.

	AUTOS Phase 1 (May 1977)	Phase 2 (June 1977)	Phase 3[a] (October 18–(November 2, 1977)
Reduction in Number of Cars on Expressway (from "Before" Condition)	2010	5670	2600-3900[b]
Where they went:			
• Shifted Mode			
Carpooling (Increased auto occupancy)	920-1140	500-700	900
Red Line (Quincy Stations)	460	460	425[b]1000
Commuter Rail	0	75	155
Commuter Boat	50	50	0
Express Bus	75	75	155
Sub-Total	1505-1725	1160-1360	1535[b]2210
• Shifted Time (Made trip after 9:30 A.M.)	0-250	100-500	0-250
• Shifted to Alternate Route	125-250	3250-3500	900[b]2400
• Did Not Make Trip	-	500-700	-
Total Accounted for (By Estimation)	1730-2225	5010-5760	2435[b]4860

[a]These do not represent results at equilibrium.

[b]The bottom of the range is at Furnace Brook Parkway (the Southern end).

Table 4. Vehicles & Persons Traveling Northbound on the Southeast Expressway in the Downtown Express Lane at Furnace Brook Parkway During Phase 3

	Before March	Phase 3 October		
	Number	Number	Change	% Change from Before
Peak Period (6:30-9:30 A.M.)				
• All Lanes				
No. of Carpools (3 or more occupants)	681	1,166	485	71.2%
% of Persons (in cars) in Carpools	10.5%	22.0%	11.5%	109.5%
% of Total Persons in Carpools & Buses	24.9%	35.0%	10.1%	40.6%
No. of Bus Passengers	3,400	3,600	200	5.9%
Total No. of Persons in Autos & Buses	23,580	21,643	-1937	-8.2%
Total No. of Vehicles	16,218	13,885	-2333	-14.4%
Total No. of Autos	15,548	13,021	-2527	-16.3%
Total No. of Single-Occupant Autos	12,026	9,631	-2395	-19.9%
Total No. of Two-Occupant Autos	2,841	2,224	-617	-21.7%
Average Auto Occupancy	1.30	1.39	.09	6.9%
• Traveling in Downtown Express Lane				
% of Total Persons	---	31.7%	---	---
% of Autos on Expressway	---	11.8%	---	---
% of Vehicles on Expressway	---	11.8%	---	---
Peak Hour (7:00-8:00 A.M.)				
• All Lanes				
No. of Carpools	388	641	253	65.2%
% of Persons (in cars) in Carpools	17.3%	29.6%	12.3%	71.1%
% of Total Persons in Carpools & Buses	33.6%	45.4%	11.8%	35.1%
No. of Bus Passengers	2,000	2,124	124	6.2%
Total No. of Persons in Autos & Buses	10,080	9,491	-589	-5.8%
Total No. of Vehicles	6,098	5,171	-927	-15.2%
Total No. of Autos	5,892	4,947	-945	-16.0%
Total No. of Single-Occupant Autos	4,325	3,422	-903	-20.9%
Total No. of Two-Occupant Autos	1,179	884	-295	-25.0%
Average Auto Occupancy	1.37	1.49	.12	8.8%
• Traveling in Downtown Express Lane				
% of Total Persons	---	42.3%	---	---
% of Autos on Expressway	---	9.9%	---	---
% of Vehicles on Expressway	---	10.1%	---	---

the commuters in less than 20 percent of the vehicles. These figures include the relatively few persons in one- and two-occupant vehicles which violated the Lane restriction. Nevertheless, this indicates the efficiency of the Lane in that it moved many more people in fewer vehicles than any of the general lanes of traffic. Despite the high volumes of persons traveling in the Express Lane, these commuters experienced smooth and congestion-free travel over the entire length of the Expressway, as will be shown elsewhere in this section.

Although police enforced the MDPW regulation prohibiting low-occupancy vehicles from entering the Lane during Phase 3, a number of violators chose to travel in the Lane because of the travel time savings. A total of 1,583 citations were mailed out during the two weeks of Phase 3, which averaged 132 citations per day. As a result of the enforcement effort, compliance rates in the Lane improved significantly. At the start of the Lane (near Furnace Brook Parkway), on November 2, the compliance rate was 65 percent during the peak period and over 77 percent during the peak hour, as compared to only a 15 percent peak period complaince rate before Phase 3 commenced.

Travel Conditions on the Expressway and in the Lane

Travel Times. The key to the relative success or failure of the Downtown Express Lane project was the travel times on the Expressway. Table 2 summarizes average travel times for the "Before" condition of March 1977 (presenting the likely range based on a 95 percent confidence interval assuming a t-distribution), and for Phases 1, 2 and 3. As shown in the table, the travel times in the Express Lane in Phase 3 were consistently lower than the lower ranges of the "Before" condition for the 6:30-9:30 A.M. period. However, Phase 3 travel times in the three general-purpose lanes exceeded the upper ranges of the "Before" condition during the same time period. The three-day average in Table 2 should, of course, not be construed to be a reliable estimate of an equilibrium condition on the roadway. Also, the travel times were taken on (only) the last three days of operation of the lane (Monday through Wednesday, Oc ober 31-November 2, 1977). As noted earlier, data collection was almost always avoided during Mondays and Fridays on this (in part) recreational route leading to the South Shore and Cape Cod, and "Before" data do not reflect the usually higher than average Monday travel times. As might be expected, travel times on the unreserved lanes decreased from Monday to Wednesday. This trend, and the trends showing increased express bus and rapid transit ridership indicate that equilibrium conditions had not been reached by the end of two weeks of operation of this lane.

Safety. A history of traffic accidents for the two-week period of October 18 through November 2 for the years 1970 through 1976 shows the number of injury accidents ranged from one to three, with a two week average of 1.6. In 1977, during Phase 3 operations, only one accident involving an injury of any sort was reported. For the same time period, from 1970 through 1976, the number of property damage accidents ranged from two to five, with an average of three. A total of eight property damage accidents occurred between October 18 and November 2 in 1977. As noted before, the high number of property damage accidents reported during the two week "enforcement" period in 1977 can be in part attributed to a high rate of reporting resulting from the greatly increased number of police on the Expressway. Prior to the start of the Lane in Spring 1977, the police avoided cruising the Expressway because their presence tended to cause shock waves to form on the saturated facility. Of the nine accidents that occurred during Phase 3, four involved cars traveling in the Express Lane. Each of these four accidents was caused by an auto traveling in the far left general-purpose lane crossing illegally into the Express Lane.

Changes in Travel and Travel Conditions Off the Expressway

- Rail Rapid Transit.

Manual counts at the three stations on the Quincy branch of the Red Line showed an increase of 1300 boardings during the morning peak period for Phase 3, compared to counts during the "Before" condition. These counts, of course, reflect the impact of slight natural ridership growth and somewhat unpredictable seasonal variation, from March through October 1977. An increase of 550 Red Line riders was observed in the two weeks immediately following the start of the enforcement. It appears that Red Line ridership was increasing during Phase 3, with over 10,000 peak period boardings counted on the morning of November 2, 1977. Unfortunately, due to limited resources, similar boarding counts were not taken on the Ashmont branch, which may have experienced similar increases in ridership.

- Commuter Rail.

Inbound boardings on the two affected commuter rail lines for the month of October showed a substantial increase of 200 boardings once enforcement of the Lane began.

- Commuter Boat.

Commuter boat ridership typically declines in the autumn because of cooler weather, and 1977 was no exception. Commuter boat ridership did not increase significantly during Phase 3.

- Fringe Parking.

As in Phases 1 and 2, the only fringe parking lot to show increased use during Phase 3 was the facility at the North Quincy Red Line station, which experienced an increase of 100 parkers per day.

- Alternate Highway Routes.

Alternate highway routes were not as closely monitored during Phase 3 as they were during Phaes 1 and 2, because of a lack of resources. However, an estimate of Southeast Expressway users who diverted to local streets can be made from Expressway volume counts. In Phase 3, between 900 and 2,000 vehicles were shifted to alternate routes during the 3-hour peak period instead of traveling on the Expressway at the Furnace Brook Parkway location. Most of these vehicles used Route 128, Route 3A, or Routes 28 and 138 through Dorchester. In addition, between 1,000 and 1,500 vehicles either exited or did not enter the Expressway north of Furnace Brook, largely due to the lingering construction near the Massachusetts Avenue interchange and the relative attractiveness of alternate routes with respect to travel time. Despite the diversion of traffic from the Expressway to various alternate routes, travel times on alternate routes generally were unaffected by the implementation of the Express Lane, voluntary or enforced, with the exception of isolated small increases in travel time (5-10 minutes) during the peak half-hour (7:30-8:00 A.M.) of the two week enforcement period. These delays were encountered in a street parallel to the section of the Expressway which also experienced the greatest congestion.

Summary of Where the Cars "Went" During Phase 3

Table 3 again contains the estimates of the reduction in auto travel on the Southeast Expressway between 6:30 and 9:30 A.M. accounted for by the diversions to transportation alternatives described above for Phase 3. These may be compared with the results of Phases 1 and 2. However, it must be stressed again that these do not represent results at some equilibrium set of conditions.

In summary, the results of the experiment (with the exception of the general-purpose lane travel-time delays) were quite positive. A greater than 10 percent 3-hour peak period mode-choice shift (up to 2210 cars "shifted mode" out of 18,680 total cars at Southampton Street), and a 70 percent increase in the number of three-or-more-occupant carpools could not have been generated by any other type of action in such a short period of time at almost no cost. The generalization that it is impossible to change individual travel behavior through short-term policies clearly was proven wrong, although we are unfortunately unable to say whether the experiment would have provided in the long run an acceptable level of service for general-purpose lane users.

Termination of Phase 3 and Lane Operation

On November 2, 1977, at the Joint Transportation Committee's legislative hearing on the two bills restricting the operation of the Express Lane, MDPW Commissioner John C. Carroll announced the immediate termination of the Downtown Express Lane. He cited the overwhelming public opposition, the travel-time delays for general-purpose lane users, and his own feeling that (despite some significant commuter mode shift), "it just isn't working," as the reasons for his decision.

At the time of the termination, there was little, if any, visible political support for the project from either the commuting public or any elected public officials. Following the announcement and the Lane's demise, however, the MDPW received approximately 50 phone calls and numerous letters expressing great displeasure at discontinuing the Lane. Once again, those who were negatively impacted only reacted when the results of the decision were physically implemented. However, no significant comment was made by any public official in response to this Express Lane user backlash. It can be stated, however, that the feasibility of providing permanently separated reversible lanes on the Expressway, within the existing right-of-way, which could facilitate various vehicle-management options, is now being analyzed.

Conclusions

Conclusions concerning the overall impact of the Downtown Express Lane are as follows:

1. The Downtown Express Lane allowed the Expressway to operate at higher vehicle occupancies and lower total volumes, which accomplished its primary purpose in the view of the responsible public officials.
2. The massive publicity campaign and the coverage and editorializing by the media which preceded the implementation of the voluntary Lane was vital in explaining the purpose of the Lane and obtaining the public's cooperation during Phases 1 and 2. The negative media reaction during Phase 3 (enforcement) contributed to the level and intensity of public opposition to continuation of the project.
3. The results of the Express Lane for Phase 1 (voluntary operation) presented in this paper are more representative than Phase 2 results for evaluating the impacts of the Lane, because of the construction bottleneck at the northern end of the Expressway during Phase 2.
4. The results of the Express Lane for Phase 3 (enforced operation) are clearly those of a dynamic system in which travel conditions apparently had not yet reached equilibrium.
5. The compliance rate was inversely proportional to the number of carpools available to fill the Lane during the voluntary operation (Phases 1 and 2). However, it appears that compliance with a voluntary reserved lane always will tend to be low. Compliance rates proved much higher (and certainly acceptable) during the two-week enforcement period.
6. During Phase 1 (before the construction bottleneck) over 50 percent of the reduction in autos was accounted for by increased vehicle occupancy on the Expressway itself. Only 10 percent of the reduction in autos was accounted for by shifts to alternate routes or by drivers not making the trip at all. The Phase 3 enforcement results, while not as encouraging in this regard for the short period of operation of the Lane, did have 60 percent of the reduction in autos accounted for by a mode shift (to carpools and transit). This accounted for more than 10 percent of all autos using the expressway during the three-hour peak period.
7. Reserving a lane for buses and carpools did not hurt rapid transit ridership on parallel routes. In fact, rail transit ridership increased substantially, accounting for 25 percent of the reduction in autos during Phase 1 and Phase 3, and reflecting the complementarity of alternative high-occupancy modes in a high-volume corridor.
8. Once the Lane was enforced, travel times decreased in the Express Lane, but increased significantly and were unpredictable in the remaining three lanes, encouraging commuters to use alternate local street routes. However, whether or not these times would have remained unreliable if the enforced Lane operation had continued is unknown.
9. The absence of breakdown lanes and adequate acceleration lanes on the Southeast Expressway during the peak period increased the average time required for recovery from accidents and breakdowns during Phase 3 (the enforcement operation).
10. Finally, operational changes of this kind are difficult to implement. Since travelers who are inconvenienced by such a change (even for a short time) are more vocal than those who benefit by the change, public officials have difficulty responding to the resulting political pressures. The Downtown Express Lane experience shows once again that government is often unable to resolve issues in which short term private interests appear to conflict with the overall public good.

Acknowledgements

The authors of this paper wish to acknowledge the very able help and cooperation of many persons and agencies in carrying out such a large and complicated project as the Southeast Expressway Reserved Lane Project. Commissioner John J. Carroll of the Massachusetts Department of Public Works is particularly thanked for his personal leadership in implementing the Project.

Robert Tierney, Kay Krekorian and Thomas F. Humphrey of the MDPW played important roles in overall project management. Also, James DeAngelo, Joseph Schraffa and Charles Sterling of the MDPW were instrumental in the traffic engineering phases of the Project. Dominick Spinale (MDPW) supervised the actual operation of the Lane, and James O'Halloran (MDPW) led the traffic measurement phases of the Project in the South Shore Corridor. The Transportation Systems Center (TSC) of the U.S. Department of Transportation (USDOT) and its Urban Mass Transportation Administration (UMTA) Service and Methods Contractors (Multisystems, Inc., and CACI, Inc.), under the direction of Howard Simkowitz (TSC), made traffic measurements outside the South Shore Corridor. The carpool matching prior to and during Phase 1 was carried out by AMV, Inc. and SOM, Inc. The marketing staff of the Massachusetts Bay Transportation Authority (MBTA) helped design the publicity campaign that was integral to the success of the Project, and Karen Epstein of the Public Information Office of the MDPW contributed greatly to the public relations aspect of the Project.

The authors also recognize the invaluable assistance of many individuals in assembling and processing the massive amount of data collected as part of the evaluation program. The assistance of Adrian Marvin (EOTC), Carla Schrage (CTPS), and Lee Morrison (CTPS) was also invaluable in the preparation of this report.

The authors of this paper extend their sincere thanks to all these people and their apologies to other whose contributions remain unacknowledged.

References

1. Central Transportation Planning Staff (Boston), *Southeast Expressway Evaluation of the Downtown Express Lane*, report prepared December 1977.
2. Thomas A. Domencich, Gerald Kraft and Jean-Paul Valette, "Estimation of Urban Passenger Travel Behavior: An Economic Demand Model," *Highway Research Board Record 238*, 1968, pp. 64-78.
3. John W. Billheimer, "Santa Monica Freeway Diamond Lanes: Evaluation and Overview," prepared for January 1978 TRB meeting.
4. Howard J. Simkowitz, *A Comparative Analysis of Results from Three Recent Non-Separated Concurrent-Flow High Occupancy Freeway Lane Projects: Boston, Santa Monica and Miami*, prepared for January 1978 TRB meeting.

A SIMULATION STUDY OF ALTERNATIVE REAL-TIME BUS HEADWAY CONTROL STRATEGIES

David Koffman, Crain & Associates, Menlo Park, California

Bunching of buses, due to variations in passenger loading, degrades service and operating efficiency. A simulation model is developed of a single-direction bus route, including explicit simulation of traffic signals. Using a hypothetical situation, several strategies are tested for the control of bus headways in real-time. These are: 1) holding points, 2) skipping stops, 3) selective application of bus priority signalization, 4) reducing dispatch uncertainty. The last two prove most promising. It may be possible to field test selective application of bus priority signalization using existing computerized traffic signal systems.

Background

There is growing interest in improving urban bus transportation through the application of new technology, particularly in the areas of computer-aided planning and real-time control. The purpose of the work described is to evaluate the relative merits of several such real-time control concepts.

Two forms of real-time control are currently being tested in various places. Automatic vehicle monitoring (AVM) systems have been tried in Hamburg (11), Chicago (3,4,8), and London (10). Aside from allowing management response to exceptional circumstances, such as breakdowns and robberies, the primary benefits of such systems are supposed to be in controlling deviations from schedule or scheduled headway. If the position of all buses in a system or route is known at a central point, then deviations from desired positions can be noted and attempts made to correct them through some kind of control strategy. In the systems referred to above, the available means of control is verbal instructions to drivers sent by radio.

The impetus for development of such systems comes from observations on the sources of deviations from schedule in bus operations. The primary sources are:

1. Random variations in traffic, road conditions, and driver behavior;
2. Traffic control signals;
3. Variations in loading and unloading times due to uneven demand.

Source (3) implies the phenomenon which has been called the "dynamic instability" of a bus route (4). Longer loading times due to more waiting passengers at some stop, initially due simply to random fluctuations, will make a bus late, which in turn will cause it to find still more passengers than average at the next stop. The bus following will find ever fewer than average waiting passengers and hence gain on the bus in front. Thus the familiar phenomenon of bunching occurs. The equilibrium is unstable; small variations get amplified and propagated back through the chain of buses on the route. In general the deviations from regularity will get worse as one moves away from the start of the route, assuming that the buses start there on time with even headways. Since dispatch headways cannot, in fact, be exactly even, there is an additional source of built-in instability.

These observations are generally held to imply that effective control must be real-time and able to be applied at frequent intervals along a route. The AVM systems are designed to meet these criteria. Unfortunately, these systems lack any really effective means of control. Instructions to drivers by radio are advisory only and very limited in the range of action that can be suggested. The possibilities are:

1. Tell a lightly loaded bus that is gaining on its leader to slow down or wait at the next stop.
2. Tell an overloaded bus that is falling behind to skip picking up passengers at one or more stops.
3. Direct a bus to turn around to fill a gap in the opposite direction.

All these actions, although possibly effective in correcting schedule or headway deviations, also degrade service, at least to some passengers -- in particular, those most in a position to observe the operation. A simulation study of a London subway route that used only strategy (1) to correct deviations showed that it did not significantly increase the overall level of service (7). A possible justification for such a strategy, however, would be the econometric data which show that transit demand is much more sensitive to waiting times than to line-haul travel times (5). Thus some trade-off of the latter in favor of the former would be justified. Strategy (1) would have this effect if it evened out the headways. A second possible justification is based on cost considerations. This follows from the more even passenger loading that would result from more regular headways. Thus it would be possible to employ smaller buses without increasing the likelihood that passengers would have to stand or wait for more than one bus. Moreover, regularity itself should have economic value to passengers, as well as permitting the operator to draw up schedules to closer tolerances and hence make more efficient use of men and machines.

Strategy (2), skipping stops, would tend to increase line-haul speeds at the cost of increased waiting time for those passengers passed up. On the other hand, passengers at the next stop would wait less time than otherwise. Presumably, if headways were in fact evened out, average wait times would go down, as would measured time-regularity of service. But passengers watching a bus go past without stopping might be hard to convince on this point.

A second form of real-time control now being experimented with is bus priority signalization. This assumes some form of instrumentation such that intersection control equipment can recognize the arrival of a bus on one approach and allow it to pass through sooner than it might, either by advancing or extending the green indication for that approach when necessary. Such a system operated in Washington, D.C. as part of the UTC system of centralized digital control of traffic signals (9) and in Kent, Ohio (2). A similar system is operating on an experimental basis in Boston.

Bus priority signalization attacks one of the sources of running variance directly. However, the usual justification for giving buses priority at intersections has been to reduce line-haul travel time. To be sure, it has been shown that by reducing total time spent waiting for traffic lights, bus priority signalization will also partially eliminate one source of random delays and so reduce schedule variance (1).

There is no technical reason, however, why traffic signals cannot be used more imaginatively to reduce waiting time and wasted seat capacity by selective application of priority signalization according to headway criteria. Buses would be granted priority at an intersection only if the time since the previous bus was greater than some threshold headway. Thus no bus would experience a delay greater than it would normally, but would sometimes experience less delay. This strategy is similar to strategy (2) above (skipping stops), except that no passenger would be passed over. If feasible, such a strategy might therefore increase both average bus speed and headway regularity, and hence also reduce average passenger wait times. Of course automobile traffic might suffer, and it would have to be clear that the trade-off was worth it. Using traffic lights enjoys a significant advantage over other control methods, in that bus drivers can be expected to obey traffic signals, whereas other methods will encounter problems with enforcement, improper execution of instructions, and resentment between drivers and supervisors.

The Model

A computer simulation model was constructed to explore the relative merits of several headway control strategies and the current practice, whereby control is exercised, if at all, only at start-up and turnaround points, and occasional checkpoints by roving supervisors. The model compares the wait and travel time elements of level of service produced by the different control concepts as perceived by passengers. Since these level of service parameters are primary determinants of the short run demand for travel by transit, it was decided to include some demand equilibrium in the model. The model may be used to show the scheduled operating frequency and bus size required to service various levels of patronage under the different control strategies, as well as the resulting efficiencies of equipment use. In order to be as realistic as possible and to permit modelling of the bus priority signalization strategy, the model makes explicit allowance for the effects of traffic signals along the route.

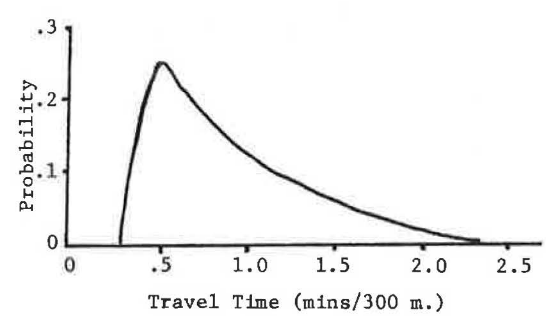
In its basic form, the model simulates the following major events for all control strategies:

1. Bus starts route.
2. Bus travels to bus stop or traffic light.
3. Bus waits at traffic light.
4. Bus takes on passengers.
5. Bus discharges passengers.
6. Bus finishes route.
7. Passenger arrives at bus stop.
8. Passenger boards bus.
9. Passenger arrives at destination.

The simulation is of a one-way bus route. This one-way assumption introduces no additional unreality to the model except insofar as it affects the bus priority logic at intersections. The model was implemented on an IBM 360 computer using the IBM General Purpose Simulation System (GPSS).

The model contains two transaction streams. Buses are created at intervals specified by the user, with dispatch uncertainty also specified by the user. Buses then go through a cycle of traveling followed by waiting, either at a traffic light or a bus stop. The user has complete flexibility in specifying the locations of any number of bus stops and traffic lights. In the hypothetical case used, a 2700 meter (9000 foot) route has stops every 150 meters (500 feet) for 2250 meters (7500 feet), plus one stop at the end of the route. Two different arrangements of traffic lights and light timings were used, as will be described later. Bus travel times are drawn from a distribution of the approximate shape shown in Figure 1.

FIGURE 1. BUS TRAVEL TIME DISTRIBUTION



Waiting time at traffic lights is computed according to traffic light timings specified by the user. At each bus stop, waiting passengers are picked up and passengers on the bus with that stop as their destination are discharged. The bus is delayed by 4.3 seconds for each boarding passenger and by 2 seconds for each leaving passenger; these values were estimated by Kulash (6) for a bus route in

Cambridge, Mass. The bus experiences an acceleration/deceleration delay of 10 seconds at every traffic light and every bus stop at which passengers actually board or leave. After the last bus stop, buses vanish from the system.

The movements of passengers are the second transaction stream. Passengers are created at a rate, specified by the user, representing the overall demand for travel by bus along the route under ideal service conditions. In the runs conducted, passengers were created at a rate of 360 per hour. Once created, each passenger is assigned an origin stop and a destination stop, using user-specified probabilities representing the relative densities of trip-creating and trip-attracting activities in the neighborhood of each stop. Only passengers whose assigned origin and destination imply a trip in the direction of the modeled route go to their origin bus stop and continue through the system. In the hypothetical situation modeled, all 17 bus stops were equally likely as origins and destinations, except that the first stop and middle stop were three times as likely as origins as all others, no passengers were assigned the first three stops as destinations, and the middle and last stops were three times as likely as all others as destinations.

In order to introduce some shortrun demand equilibrium into the model, each passenger is assigned a maximum acceptable wait time, drawn from a probability distribution derived from a very simplified assumed demand curve. Passengers with longer desired trips are given longer acceptable wait times. Only those passengers with maximum acceptable wait times less than or equal to a moving average of current actual passenger wait times continue through the system. The moving average is recomputed every time a passenger boards a bus, using a time constant of 300 passengers (50 minutes). Since wait time appears to be the most important perceived cost of transit, this seems a minimum-acceptable representation of demand equilibrium. It is admittedly incomplete, since it includes no sensitivity to travel time, or even to any other measure of wait time except the mean, e.g., reliability. This may be significant, since one assumption which provides motivation for controlling bus headways is that time-reliability is important to passengers. Moreover, the output of the model must be examined to establish that the moving average of wait time behaves reasonably, that is, does not oscillate too much.

All passengers at a bus stop board the first bus that arrives. Each passenger stays on the bus until it arrives at the pre-assigned destination stop, and then leaves the system.

Modelling the Control Strategies

Base Case

This is the basic model form as just described. Buses were dispatched every 10 minutes with a dispatch uncertainty of ±5 minutes.

Holding at Bus Stops

If a bus stops to pick up or let off passengers, it will be delayed if the time since the last bus at that stop is less than some tolerable level. The delay is the difference between the desired mean headway and the observed headway, rounded to the nearest whole minute. Tolerance levels of 65%, 75% and 85% of desired headway were tested.

Bus Skips Loading

When a bus arrives at a stop, no passengers are loaded if the observed time since the last bus at that stop is greater than some tolerable level. Tolerances of 115%, 125% and 135% of desired headway were tested. Passengers can always get off.

Bus Priority Signalization

When a bus arrives at a traffic light, it is "granted priority" if the observed time since the last bus at that light is greater than or equal to 75% of the desired headway. This bus experiences no delay at the light if this can be accomplished by extending or advancing the green by no more than a set maximum deviation; otherwise the duration of the delay is reduced by the same maximum amount. Values of 15 and 30 seconds for the maximum change to the green were tested.

Dispatch Control

This is the same as the basic form of the model, except the dispatch uncertainty is assumed to be reduced to ±1 minute, as a result of some network operational procedures.

Test Situations

The model was run with two different arrangements of traffic lights and numerous variations of the control strategies. In the "uncoordinated, wide tolerance" test situation, traffic signals were set up as follows on the hypothetical 2700 meter route:

Light #	Distance	Green	Cycle
1	600 m.	40 sec.	90 sec.
2	1500 m.	60 sec.	120 sec.
3	2100 m.	45 sec.	90 sec.

All lights operated with zero offset. In this test situation, the following control strategies were tested:

1. Base case
2. Holding at bus stops if headway under 75% of desired.
3. Holding at stop if headway under 65% of desired.
4. Skip loading if headway over 125% of desired.
5. Skip loading if headway over 135% of desired, except at first stop.
6. Priority signalization; green may be changed by up to 30 seconds.
7. Dispatch control.

In the "coordinated, narrow tolerance" test situation, traffic signals were set up as follows:

Light #	Distance	Green	Offset
1	300 m.	50 sec.	20 sec.
2	600 m.	60 sec.	40 sec.
3	900 m.	55 sec.	60 sec.
4	1200 m.	50 sec.	80 sec.
5	1500 m.	60 sec.	10 sec.
6	2100 m.	55 sec.	50 sec.

All signals operated on a 90 second cycle length.

In this test situation, the following control strategies were tested:

1. Base case.
2. Holding at stops if headway under 85% of desired.
3. Skip loading if headway over 115% of desired.
4. Priority signalization; green may be changed by up to 15 seconds.

The two situations represent, in effect, different levels of investment in control equipment. The control strategies in the first test situation are relatively insensitive and the traffic lights are very sparse and operate without any coordination. The control strategies in the second test situation are more sensitive, the traffic lights are more frequent, and they operate in a coordinated fashion, on a common cycle, presenting a "green wave" moving at 54 km/hr (34 mph). The green wave is, of course, for the benefit of automobile traffic; a maximum change in green time of 15 seconds is chosen so as not to disrupt this pattern too much and to represent probable limitations in bus detection hardware (the longer the green is to be extended, the further downstream the bus must be detected and its arrival time at the light predicted). Neither situation represents any actual existing bus route, but rather an artificial "typical" situation.

Results

Tables 1 and 2 summarize the results of the model runs. Statistics were collected after two and four hours of simulated operation. To conserve space, only the four-hour statistics are shown, as these appear to represent the steady state. Means and standard deviations are shown, the latter as a measure of reliability. The "95% Level" shown is another measure of reliability; it is the least extreme, that is "least bad," value of the measure such that at least 95% of the observations do not exceed that value.

Base Case

The base case is shown in the first column of both tables. The base case is useful for comparison of other strategies and as a check on the validity of the model. The results agree well with the observed behavior of real bus routes. In particular the bus speed works out to be near 13 km/hr

TABLE 1. RESULTS FOR FIRST TEST SITUATION: Uncoordinated, Wide Tolerance

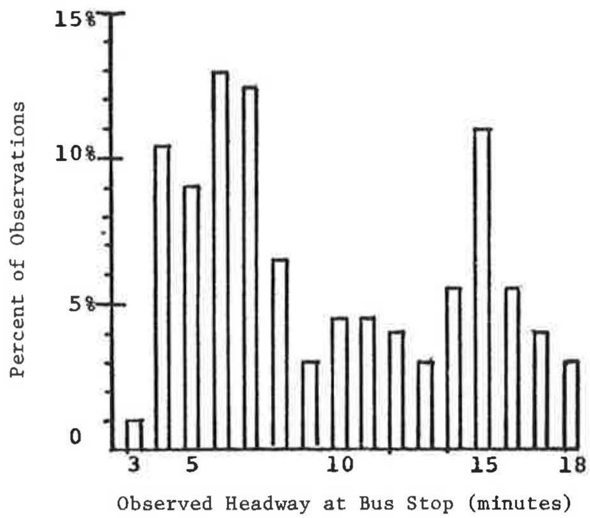
Strategy:	Base	Holding		Skips Stops		Priority	Dispatch
Tolerance:	--	75%	65%	125%	135%	--	--
Pass. wait time							
(mins.) Mean:	6.1	6.3	5.9	8.4	6.0	5.7	5.2
Std. Dev.:	3.9	4.3	4.1	7.6	4.4	3.7	3.0
95% Level:	13	14	13	24	14	12	10
Bus travel time							
(mins.) Mean:	12.1	17.0	22.4	10.8	11.3	10.9	12.0
Std. Dev.:	2.2	7.0	12.2	2.1	2.1	1.9	1.5
95% Level:	15	28	30+	14	14	13	14
Pass. speed							
(mph) Mean:	8.7	7.3	6.4	9.3	9.1	9.9	9.1
Std. Dev.:	2.2	3.6	3.8	2.4	2.4	2.3	2.3
95% Level:	5	1	1	5	5	6	6
Bus headway							
(mins.) Mean:	10.3	10.8	10.2	Not Available		9.6	9.9
Std. Dev.:	4.1	4.9	5.2			4.0	1.3
95% Level:	17	20	19			16	12

TABLE 2. RESULTS FOR SECOND TEST SITUATION: Coordinated, Narrow Tolerance

Strategy:	Base	Holding	Skip Stops	Priority
Tolerance:	--	85%	115%	--
Pass. wait time				
(mins.) Mean:	6.6	5.8	9.9	5.3
Std. Dev.:	3.9	4.1	9.0	3.5
95% Level:	13	13	30	12
Bus travel time				
(mins.) Mean:	12.0	31.6	10.6	11.0
Std. Dev.:	1.8	12.0	2.5	1.4
95% Level:	15	30+	14	13
Pass. speed				
(mph) Mean:	8.9	4.3	9.1	9.9
Std. Dev.:	2.3	3.4	2.8	2.6
95% Level:	5	1-	5	6
Bus headway				
(mins.) Mean:	11.4	10.1	Not Available	9.3
Std. Dev.:	3.5	4.5		3.8
95% Level:	16	20		15

(8 mph), which is realistic, as is the average passenger wait time at rather more than half the average headway. In addition, the distribution of headways is markedly bimodal (see Figure 2), which is indicative of the bunching effect discussed above. The moving average of wait time appears to stabilize quite soon after the start of simulation. The slightly poorer performance on the wait time criterion with coordinated traffic lights is probably due to the "green wave" being poorly timed for the bus, hence acting as a source of periodic instability. In the uncoordinated situation, 23% of potential passengers were lost due to the demand equilibrium, in the coordinated situation 25%.

FIGURE 2. DISTRIBUTION OF BASE CASE HEADWAYS



Holding at Stops

Selective delays produce very minor or no improvements in wait times at the expense of longer travel times, as expected. In the coordinated case, the improvement is more noticeable due to the relatively poor performance of the base case there. A more significant evening out of bus loadings occurs; these numbers are not shown, since there is some question about their interpretation. The particular headway tolerance chosen appears to have little effect on the wait time or bus loading improvement produced. However, a wider tolerance appears to minimize the bad side-effects of this strategy. This is fortunate from the point of view of technical ease of implementation, but unfortunate from the point of view of acceptability to public and drivers, since rather long, apparently senseless delays en route are implied. The worsening of travel times is so bad even in the best case, however, that implementation would seem unwise. Moreover, the model's demand equilibrium on the basis of wait time only is probably unrealistic in the face of such extreme travel time changes. In other words, revenues would probably suffer in reality. It is tempting to speculate on whether refining the control response (i.e., allowing a correction to an accuracy better than whole minutes) would improve performance. Probably it would; it seems unlikely, however, that it could ever be implemented given real-world constraints.

Skipping Stops

For headway tolerances of 115% and 125%, this strategy has exactly the opposite effect of that sought--it worsens wait times by more than it improves line-haul travel times and drives away potential passengers. Since a narrow tolerance appeared to aggravate matters, and even reduce the travel time benefits, it was natural to speculate on whether requiring a larger deviation from scheduled headway would work better. In retrospect, it seems obvious that we should require considerable deviations before applying such a brick-bat form of control. In addition, it is likely that this strategy was suffering since initial deviations due to dispatch uncertainty were probably causing the first stop (which has a high demand rate) to be skipped over often--clearly an unrealistic procedure. Therefore the deviation required to skip a stop was increased to 135% and the first stop was never skipped. As expected performance improved. Wait times were reduced, wait time reliability improved (though they were still no better than in the base case) and higher speeds than in the base case were retained. Unfortunately, the lower wait times appear to be partially a result of an equilibrium process that loses potential passengers, since wait time reliability is still somewhat worse than in the base case. Indeed an experiment with a "superwide tolerance" of 150% of desired headway produced a drastic worsening of results and increased instability.

Bus Priority

In both test situations, selective priority at traffic signals both decreases wait times and increases line-haul speeds. In addition, service reliability is improved, as shown by a decrease in wait time uncertainty. A minor potential improvement in productivity also occurs, as shown by decreased bus travel time uncertainty. The improvements are more dramatic in the coordinated situation, due partly to the poorer comparison base and partly to the more refined control employed. The increased number of traffic lights (i.e., possible control points) and the weaker control response (maximum change in green of 15 rather than 30 seconds) increase the stability of this strategy (as shown by inspection of the behavior of the moving average of wait time), maintain the improvement in travel time, and increase the improvement in wait time shown in the uncoordinated situation.

A possible criticism of the model for this strategy is that priority signalization would affect automobile traffic; hence different distributions of probable bus travel time between stops and signals ought to be used in the different test situations. If priority signalization worsened automobile congestion, the improvements shown might be cancelled out.

The improvements shown can certainly be questioned, given their not overwhelming magnitude and the model's untested nature. They are promising, however, especially considering that in a city which already has centralized computer control of traffic signals, this strategy might be implemented at reasonable additional cost, certainly compared to the expense involved in AVM systems. As over a hundred U.S. cities currently have such computer control systems in various stages of development (see The American City, August 1974), an experiment along these lines might be hoped for soon.

Dispatch Control

Reducing the dispatch uncertainty produces improvements in wait time as great or greater than any of the on-route control strategies. Travel time uncertainty is also noticeably improved, although travel time is not. The question is, of course, whether such a reduction in dispatch uncertainty is an achievable goal, institutionally or economically. Some method of reducing running time variation is probably a precondition of reducing dispatch uncertainty. The results of this test case also raise some doubts about the model itself, since the observed headways appear fairly stable and do not exhibit the bimodality of those in the original base case. It is possible that a wider variance in speeds between bus stops and lights is called for. In all fairness, however, at the demand rate given, and on such a short route as the one modeled, it is not clear that extreme bunching would really occur without the initial source of instability.

Acknowledgment

This work was done while the author was a graduate student in city planning at the Harvard Graduate School of Design. Special thanks to Professor Nigel Wilson for whose course at MIT this work was done as a term project.

References

1. Ir. B. von Bilderbeek, "Priority Treatment of Buses at Traffic Signals," in Organisation for Economic Cooperation and Development (OECD), Improvements and Innovations in Urban Bus Systems, Paris, October 1969 (NTIS PB 188 030).
2. H. Evans and G. Skiles, "Improving Public Transit through Bus Preemption of Traffic Signals," Traffic Quarterly, October 1970, p. 533.
3. Ansel V. Gould, "Automatic Vehicle Monitoring Applied to Bus Operations," IEEE Transactions on Vehicular Technology, Vol. VT-22, No. 2, May 1973.
4. Institute for Public Administration and Teknekron, Inc., "Public Urban Locator Service (PULSE) --Background and Conference Proceedings," Washington, D.C., October 1968 (NTIS PB 180 116).
5. Michael A. Kemp, "Some Evidence of Transit Demand Elasticities," Transportation, Vol. 2, No. 1, April 1973.
6. Damian Kulash, Routing and Scheduling in Public Transportation Systems, unpublished Ph.D. thesis (Civil Engr.), Mass. Inst. of Tech., Cambridge, November 1970.
7. S.M. Levene, J.G. Weston, and D. Williamson, "An Assessment of the Application of On-Line Computers to Control of an Underground Railway--a Simulation Study," Transportation Research, Vol. 8, No. 2, April 1974.
8. W.C. Scales, "Benefits and Costs of an Automatic Vehicle Monitoring System," MITRE Corp., McLean, VA, September 1971 (MTR-6064).
9. Sperry Rand Corp., Sperry Systems Management Div., "Advance Control Technology in Urban Traffic Systems," Vol. 1A: "Bus Priority System Description," Great Neck, NY, 1970 (NTIS PB 190 847).
10. D. Stone, "Systems for the Automatic Control and Monitoring of Bus Services," Traffic Engineering and Control, December 1970.
11. Hans Tappert, "Real-Time Bus Control," in OECD, op.cit.

INCENTIVE PROGRAM FOR BUS CARRIERS

Edward M. Abrams, Director of Transit Operations, Simpson & Curtin
Harold J. McLaughlin, Jr., Supervisor, Terminals Forecast and Analysis, The Port Authority of New York and New Jersey

One of the most elusive aspects of public assistance to independently operated transit services has been the development of incentives to the operator to provide quality service on a cost-effective basis. To deal with this, the New Jersey Department of Transportation intends to implement an incentive program which will result in monetary rewards or penalties to companies depending on the quality of service provided. This paper presents the essential elements of that program which were developed as part of an overall study to revamp the method that NJDOT follows in providing operating assistance to independent carriers in the State. The incentive system recommended is keyed to three principal areas that are perceived by transit users to be important - - 1) condition and cleanliness of the bus, 2) courtesy and driving skills exhibited by the bus operator, and 3) on-time performance. For each of these areas, surveillance procedures have been developed and rating forms designed to relate field observations to a numerical scoring system. The scoring system weights each survey element by its estimated relative importance as compared to all other elements. The program calls for each carrier to be rated quarterly in all three areas, and a method is described for translating a company's overall score for a quarter into a monetary reward or penalty.

One of the most elusive aspects of public support of independently operated transit services has been the development of incentives to the operator to provide quality service on a cost-effective basis. All too often, subsidies have supported, or even encouraged, the deterioration of both the quality of service and the efficiency and cost-effectiveness of management.

The quality of bus service has been of special concern in the State of New Jersey wherein about 28 bus operators presently receive operating assistance through the New Jersey Department of Transportation (NJDOT). To deal with this, NJDOT intends to abandon its current bus subsidy program, which is designed simply to pay out sums equal to a company's operating deficit, in favor of a system which will motivate carriers to be cost-effective, operate efficiently and provide high quality service to the public. It is fully expected that these objectives can be realized through the implementation of a program recently developed for the Department as part of an Urban Mass Transportation Administration (UMTA) sponsored study.(1)

The proposed program is comprised of two major elements. First, the State would make quarterly determinations of the subsidy due a carrier from the amount which an NJDOT projection of farebox receipts fails to cover an NJDOT estimate of expenses the carrier can reasonably be expected to incur in furnishing the scheduled quantity of service, as specified by the State. In this program (with one minor exception), every expense component is related to a service parameter - - either bus kilometers, bus hours or number of peak buses - - and an appropriate unit cost. Some of the unit costs developed (e.g., fuel cost per km.) are applicable to any carrier operating in the State, while other unit costs (e.g., drivers' wages per bus hour) are tailored to specific carriers. In either case, however, once a cost is defined, it will be revised quarterly solely in accordance with the movement of an appropriate BLS index without regard to actual costs incurred by the carrier.

The second element of the program, and the subject of this paper, deals with adjustments to those sums due based on the quantity of service furnished, as discussed above, to reflect the quality of the service rendered. The proposed incentive system will financially reward carriers providing high quality service and fine carriers rendering substandard service.

Proposed System

The aim of the incentive program is to improve the quality of bus service in New Jersey and consequently:

1. Increase present riders' satisfaction with the service, thereby assuring their continued patronage.
2. Attract additional riders to available bus service.
3. Increase farebox revenue which would benefit the carriers in the short run and the State over the

long term.

Consideration was given to development of incentives through a system of rewards and penalties on the basis of tangible results such as increases or decreases in riders. However, many of the factors affecting ridership are outside the control of the carrier; changes in competitive service, land use, and employment shifts all affect ridership levels. Therefore, it was determined that the incentive program should focus on characteristics of a transit system that relate to specific measures of quality rather than to apparent results. In other words, the recommended incentive program would monitor the causes of quality service, not the exhibited effects.

Quality service, as it is perceived by the patron, became the prerequisite for identifying characteristics considered significant for the incentive program. A passenger cares if his or her bus is clean, on time, free of defects and is driven by a competent courteous operator. These are the conditions that riders notice and care about; these are the conditions which, if improved, would have a positive affect on maintaining existing riders and even attracting new riders to the service. And finally, these are the characteristics that are within the control of the carrier to improve.

To determine how a carrier performs in relation to the above-noted characteristics, three types of field surveys are required:

1. Garage surveys to determine the condition and cleanliness of buses.
2. Riding surveys to observe driver performance with regard to courtesy and bus handling.
3. Street surveys to measure on-time performance.

A description of the field survey procedures, the elements to be checked and the relative weight given each element are described below:

Bus Garage Summary

It is proposed that inspectors from NJDOT compile the data on bus cleanliness and condition at the carrier's garage just prior to the bus departure for the first morning trip. This early morning surveillance will enable the check to be solely dependent on how well the carrier prepares its equipment for use in revenue service, independent of circumstances outside the control of the carrier, e.g., patrons' littering the bus.

The Bus Garage Summary involves 12 separate checks, as shown in Table 1, varying from verification of the proper operation of fare collection equipment to whether or not the seats are clean. Along with each check is an assigned set of numerical values for scoring each item relative to the actual condition of the bus for that particular check. For example, in the category of seat cleanliness, the score can be either a plus two, a zero, or a minus two, depending on the actual condition of the seats. The total score for an individual bus will be the algebraic sum of all 12 checks.

A bus carrier's overall score for this apsect of the incentive program will be the median total score recorded during the quarter. In other words, all total scores for the individual bus surveys of the company made in the quarter will be arrayed with the final score determined by the middle value. This method will minimize the effect on the final score of extreme individual values, both high and low, including any which may have been the result of biased ratings.

Riding Surveys

To properly obtain unbiased data in this survey element, NJDOT inspectors will be required to ride carriers' buses, concealing their identity and purpose. This aspect of the survey is important in order to obtain data which accurately reflect normal driver habits. Inspectors will be expected to ride the bus only long enough to obtain sufficient data to fill out the survey form.

As noted in Table 2, the riding survey involves five separate checks. Similar to the Garage Surveys, a set of numerical values is assigned to each. For example, if a driver is observed to be pleasant, helpful and considerate, a rating of plus 10 would be given. However, for a discourteous driver, a minus 10 would be scored.

The total score for an individual riding survey will be the algebraic sum of all five checks. As in the case of Garage Surveys, a carrier's overall score for a quarter will be the median total score for all riding surveys made during the quarter.

Street Surveys

This type of survey involves probably the most important surveillance element - - on-time performance. Poor on-time performance is the most often cited criticism voiced by transit users. Patrons expect, and rightly so, that bus services strictly adhere to the carrier's own published schedules. Of course, there are circumstances beyond the control of an operator which can affect on-time performance; therefore, a standard which calls for 100 percent schedule adherence is not realistic. In fact, there are situations when a bus should be deemed "on-time" even if it is five minutes late.

While service that is running late may be due to no fault of the carrier, there is no excuse for service that is early. Furthermore, standards for schedule adherence should vary over the length of the route - - being strictest at the origin where the departure time is within the carrier's control, and then gradually becoming more liberal as the bus approaches its final destination. Properly designed operating schedules should make allowance for delays by building in sufficient recovery time to insure that the next trip can depart on time.

A system was developed for assigning a numerical score to the on-time performance of a bus anywhere along its route. The array of possible scores is shown in Table 3. The highest score attainable for a single observation is plus 50, whereas a bus running sufficiently behind or ahead of schedule may rate a score of minus 200. Other features incorporated in the numerical scoring system as are follows:

1. If 80 percent of all measurements are determined to be on-time with the others either very early or very late, an overall rating of zero will result.
2. Minus scores are more heavily weighted toward early service than for service that is late.
3. The scoring system becomes progressively more liberal as the point of observation is moved further from the route's point of origin in recognition of the cumulative effect of circumstances beyond the driver's control which can affect on-time performance.

In conducting the on-time performance surveys, most data should be acquired from checks occurring within the first three columns (up to 50 minutes from routes' origins). Generally, service which is more than 50 minutes from a route's origin is in a

Table 1. Incentive Program, Garage Survey

Bus Co. ______ Date ______ Time ____ M
Location ______ Rated by ______

Survey Element	Value	Route ____ Bus ____
Fare Collection Equipment		
Inoperable	- 3	
Working Property	0	
Interior Lighting		
Operable, No Outages	+ 3	
One or Two Outages	+ 2	
Less than 20% Outages	0	
More than 20% Outages	- 2	
Inoperable	- 4	
Heating/Air Cond./Ventilation		
Normal Operation	+ 3	
Rough Sounding/Substandard Effectiveness	0	
Not Operable	- 3	
Condition of Seats		
One or More Seats Broken/Torn	- 2	
Seats Badly Worn or Marked	- 1	
Average/Acceptable	0	
New Appearing	+ 2	
Cleanliness of Seats		
Soil/Film	- 2	
Average/Acceptable	0	
"Just-Cleaned" Appearance	+ 2	
Condition of Windows		
One or More Broken	- 2	
One or More Cracked	- 1	
Some Scratches or Discoloration	0	
New Appearing	+ 2	
Cleanliness of Windows		
Soil/Film/Heavy Streaks	- 2	
Average/Acceptable	0	
"Just-Washed" Appearance	+ 2	
Condition of Floor		
Uneven or Mismatched	- 2	
Badly Worn	- 1	
Average/Acceptable	0	
New Appearing	+ 2	
Cleanliness of Floor		
Considerable Litter	- 2	
Some Litter	- 1	
Average/Acceptable	0	
Broom Clean, No Litter	+ 1	
Spotless	+ 2	
Interior Surfaces (other than floor, seats and windows)		
Damaged/Heavily Soiled/Marked	- 2	
Average/Acceptable	0	
Undamaged, Unmarked, and Clean	+ 2	
Condition of Exterior		
Noticeable Dents or Body Rot	- 2	
Appears Overdue for Repainting	- 1	
Average/Acceptable	0	
Slight Wearing Apparent	+ 1	
Excellent Condition	+ 2	
Cleanliness of Exterior		
Heavy Dust/Soil/Film	- 2	
Average/Acceptable	0	
"Just-Washed" Appearance	+ 2	
Score Totals		

Table 2. Incentive Program, Riding Survey

Bus Co. ______ Date ______ Time ____ M
Location ______ Rated by ______

Survey Element	Value	Route ____ Bus ____
Route/Destination Designation		
Route No., Routing (where applicable) and Destination Correctly Displayed	+ 3	
Either Route No. or Destination Displayed (Correctly)	+ 2	
Neither Route No. nor Destination Displayed	- 2	
Incorrect Route No. or Destination Displayed	- 5	
Driver Appearance		
Neat, Clean, Well-Groomed	+ 5	
Average/Acceptable	0	
Soiled Attire or Dishevelled	- 5	
Driver Courtesy		
Pleasant, Helpful, Considerate	+10	
Helpful and Informative	+ 5	
Average/Acceptable	0	
Gruff or Un-Communicative	- 5	
Discourteous, Inconsiderate	-10	
Driver Performance		
Secure Feeling	+ 5	
No Sudden Starts and Stops	+ 3	
Average/Acceptable	0	
Sudden Starts and Stops	- 3	
Insecure Feeling	- 5	
Ride Comfort		
Smooth, Quiet Ride	+ 3	
Average/Acceptable	0	
Unnecessarily Bumpy Ride	- 3	
Score Totals		

passenger discharge rather than a pick-up phase of service, and riders are much more sensitive to their departure times than to arrival times.

From the rating chart shown in Table 3, a numerical value will be assigned for each observation based on the exhibited on-time performance and the distance between the observation point and the route's point of origin as measured by scheduled running time in minutes. For example, a bus which is scheduled to arrive at a point which is 20 minutes' running time from the route's origin, at say 10:20 a.m., actually arrives at 10:25 a.m. or five minutes late, the score assigned to this trip would be minus 50 (the 11-30 minute column and the plus five minute row). The scores recorded for each carrier in each quarter will be algebraically summed and divided by the number of observations made to obtain an average overall score for on-time performance. However, before calculating a carrier's overall score, the five percent of the lateness scores with the highest number of penalty points will be deleted. The rating system will thereby compensate for the fact that unusual events or conditions which can cause a bus to be late are inevitable and impossible for a carrier to anticipate or a driver to avoid.

Monetary Considerations

To be effective, an incentive program must be based on a system of rewards and penalties. Privately

Table 3. On-time performance rating sheet.

Observed On-Time Performance (Minutes)		Scheduled Running Time in Minutes from Point of Origin 0-10	11-30	31-50	51-70
B	-8 or More	-200	-200	-200	-200
U	-7	-200	-200	-200	-150
S	-6	-200	-200	-150	-100
	-5	-200	-150	-100	- 50
E	-4	-150	-100	- 50	0
A	-3	-100	- 50	0	+ 25
R	-2	- 50	0	+ 25	+ 25
L	-1	0	+ 50	+ 50	+ 50
Y					
	On Time	+ 50	+ 50	+ 50	+ 50
B	+1	+ 50	+ 50	+ 50	+ 50
U	2	+ 50	+ 50	+ 50	+ 50
S	3	0	+ 50	+ 50	+ 50
	4	- 50	0	+ 50	+ 50
	5	-100	- 50	0	+ 50
L	6	-150	-100	- 50	0
A	7	-200	-150	-100	- 50
T	8	-200	-200	-150	-100
E	9	-200	-200	-200	-150
	10 or More	-200	-200	-200	-200

owned but publicly subsidized transit systems need external motivation to improve service quality, and a program that relates quality service to dollar signs can provide this motivation. Therefore, the incentive program, as recommended, would pay or fine a carrier a sum of money during a quarter depending on the carrier's performance during the prior quarter as determined in the three aspects of surveillance.

In order that the dimensions of the monetary motivation be equitable for all carriers, it is proposed that the amount of reward or penalty be related to a carrier's size. The most appropriate measure for this is standard cost; i.e., the NJDOT estimate of expenses the carrier can reasonably be expected to incur in furnishing regular route service. However, this standard cost amount would be adjusted to eliminate the effects of special circumstances that cause an unusually high or low expense projection.

It is recommended that the maximum reward or fine be six percent of the adjusted standard cost amount for that carrier. If a carrier had a total annual calculated cost of one million dollars, the reward or penalty limit would be $60,000 a year. If the aggregate of all fines levied was equal to the total of all rewards, the program could be self-supporting and result in a zero cash flow for the State. However, it is recommended that the State demonstrate good faith in the program and emphasize its positive aspect by paying out more rewards than it levies in fines. The appropriate net pay-out by the State has been estimated at three percent of the total calculated standard cost for all carriers. The three percent figure should be used by the State for budgetary purposes and would represent the maximum State liability. If payments during the first two or three quarters of a fiscal year, in the aggregate, exceed the three percent State commitment, then the individual payments would be proportionately scaled down in the latter part of the year to keep the total payments within the overall three percent limit.

Relation of Payments to Survey Results

Each subsidized carrier should be rated every quarter, with monetary results based on calculated standard costs and survey results for that quarter. For example, if the incentive program were initiated with the quarter beginning April 1978, the incentive dollar amount would be based on the percentage of calculated standard costs for April, May and June and the survey data collected in that quarter. Rewards in the form of cash payments and penalties in the form of sums withheld from subsidies would be made in monthly installments over the following quarter.

Results of the three types of surveys will be compiled into a single, composite value for each carrier each quarter by algebraically adding together the three overall scores. As previously noted, individual scores for poor on-time performance can range to minus 200. Therefore, if all the service provided by a carrier was observed to be very early or very late, an average overall score of minus 200 theoretically could result. Although it is unlikely that a carrier's overall score for on-time performance would be poorer than minus 50, if this improbable event occurred, it could overshadow the results of the other two surveys. To guard against such occurrence, it is proposed that for the purpose of calculating a composite value, the maximum negative overall score for on-time performance be limited to 50 points. Composite values, therefore, can range from a low of minus 106 to a maximum of plus 100, as shown in the table below:

Survey	Maximum Score Positive	Negative
Garage Survey	24	28
Riding Survey	26	28
On-Time Performance	50	50
Total	100	106

If a carrier achieved a 100 positive score or a 106 negative score, the reward or fine would be six percent of its calculated adjusted standard cost amount. Furthermore, each single positive point represents .06 percent (6 percent ÷ 100) and each single negative point represents .0566 percent (6 percent ÷ 106). Therefore, a carrier obtaining a positive score of 60 would receive a monetary payment equal to 60 x .06 percent or 3.6 percent of its adjusted standard cost. If the point total were a minus 60, the fine would be 60 x .0566 percent, or 3.4 percent of adjusted standard cost.

It is estimated at this time that the factors cited above for converting survey scores to monetary rewards and penalties will create a net cash out-flow equal to three percent of total standard cost; however, these conversion factors are subject to fine-tuning dependent upon actual program results. The necessary adjustments can be readily accomplished simply by changing the maximum percentage reward or penalty. For example, if it is found that the system as presently proposed, results in a pay-out which exceeds the three percent limit, the maximum reward could be cut back from six percent to five percent and the maximum fine increased from six percent to seven percent. Similar adjustments may have to be made from time to time (but not more often than annually) to compensate for changes in the mix of carriers under subsidy and improved performance levels brought about by the incentive program. And, of course, the three percent pay-out limit has been adopted only as a matter of policy and is, therefore, subject to revision each year by either the State Legislature or the Administration. Regulation of the cash pay-out caused by a change in the three percent limit can be accomplished in the same manner as described above.

In lieu of a fine for the first quarter that the incentive program is in effect, carriers will be given only a warning identifying the incentive program findings. Thereafter, any fine will be deducted from the following quarter's subsidy contract amount.

It is expected that all carriers, regardless of their standing in the incentive program, will be made aware of their individual performances as determined by the results of the three types of surveillance checks. In this way, the carriers will be given the opportunity to focus on areas where they perform below par.

References

1. Simpson & Curtin. Standard for Bus Service Contract Payments and a System of Incentives, prepared for the New Jersey Department of Transportation, December 1976.

PEAK-BASE COST ALLOCATION MODELS

Walter Cherwony and Subhash R. Mundle, Simpson & Curtin, Philadelphia

During the past several years, most transit agencies have been faced with the problem of rising deficits and limited tax resources to meet operating subsidies. For this reason, renewed emphasis has been placed on examination of the system's financial performance on a route-by-route basis. While route revenues can be determined by surveys and field counts, operating costs are more difficult to ascertain by route. Typically, the cost analysis has been conducted utilizing multivariable cost allocation models in which each expense account in a system is attributed to a particular resource (e.g., vehicle kilometers). This paper presents the cost analysis performed for the Metropolitan Transit Commission (Minneapolis-St. Paul) as part of the monitoring and evaluation program of the I-35W Urban Corridor Demonstration Project which tested the feasibility of express bus service on a metered freeway. The paper calls for the development of cost formulae that are sensitive to peak and base conditions rather than a single systemwide model. Also described in the paper is the development of labor productivity and service indices which can be used to compute both peak and base unit cost factors. The theoretical derivation of the relationship between the unit cost factors with systemwide costs and the indices, as well as the application of this theoretical concept, are presented.

With transit agencies beset by rising deficits and increased citizen opposition to higher taxes to support public transportation services, transit operators are more carefully examining their system's financial performance. Usually, this analysis involves an examination of each route or service type as an individual operating entity or "cost center." Route revenue, cost and margin are computed to determine the "profitability" of each transit line(1) and the extent of accommodation of service. Route revenues are relatively simple to determine by a variety of survey techniques - - passenger counts, origin-destination surveys or farebox checks. Route-by-route costs are more difficult to ascertain. The widely used method in the transit industry is the development of multivariable cost allocation models in which each system expense account can be attributed to one or more resources such as vehicle kilometers or vehicle hours. The cost allocation model is then applied to the resources required to provide service on each route to determine individual route cost. For the most part, cost formulae developed for transit properties throughout the nation represent systemwide averages which do not completely differentiate between cost associated with peak and off-peak transit services. This paper presents the development of a traditional cost allocation model, as well as the theoretical framework and computation of peak-base cost formulae.(2)

Traditional Cost Model

The first step in the development of peak-base cost formulae is the computation of a traditional cost allocation model.(3, 4) In this case, a "three-variable" model was computed rather than a more complex formula including numerous other variables such as passenger revenue.

Allocation of Expense Accounts

The Metropolitan Transit Commission's (MTC) monthly operating expense accounts were allocated to one of the three variables - - vehicle hours, vehicle kilometers or peak vehicles.

Vehicle Hours. Certain transit operating costs such as drivers' wages, which account for nearly half of the total operating costs, transportation supervision, etc., are directly related to number of vehicle hours. Therefore, these and some other expense categories which vary with the amount of service hours are appropriately allocated to vehicle hours. The use of vehicle hours which is a surrogate for pay-hours is preferred in cost allocation analysis since it is much easier to compute vehicle hours by line than payhours.

Vehicle Kilometers. Many operating costs are directly related to the vehicle kilometers of service provided. Expenses such as fuel, oil, tires and tubes, repairs to revenue equipment and servicing of revenue equipment are directly allocated to vehicle kilometers because they vary with kilometers of service operated.

Peak Vehicles. Many individual expense items do not vary as functions of either of the foregoing parameters - - vehicle hours or vehicle kilometers. For example, expenses for vehicle storage facilities are a function of the system's peak vehicle requirements rather than the number of kilometers or hours of service provided. Such peak vehicle-related expenses include supervision of shop and garage, maintenance of buildings, fixtures, grounds, service car equipment and other miscellaneous shop expenses. A number of broad overhead expenses also vary with the system's peak vehicle requirements including depreciation of revenue equipment, structures, service cars, and shop and garage equipment.

The Allocation Formula

The results of a traditional three-variable cost allocation model for a typical month (September 1974) are shown in Table 1. The three-variable formula resules in the apportionment of 59.3 percent of aggregate monthly cost on the basis of vehicle hours, 23.4 percent on a vehicle kilometer basis, and the remaining 17.3 percent as a function of the system's peak vehicle requirements. The costs attributable to vehicle hours result in a unit cost of $9.90 per hour and the costs attributable to vehicle kilometers of operation yield a unit cost of $0.19 per vehicle kilometer, while the costs allocated to peak vehicles produced a unit cost of $612.75 per peak vehicle per month.

Table 1. Development of three-variable cost allocation model, September 1974, Metropolitan Transit Commission.

Basis of Allocation	Total Allocated Cost	Percent of Total Cost	Total Operating Statistics	Unit Cost
Vehicle Hours(H)	1499400	59.3	151500	$ 9.90 per vehicle hour
Vehicle KM(K)	590800	23.4	3116000	$ 0.19 per vehicle km
Peak Vehicles (V)	437500	17.3	714	$612.75 per peak vehicle (per month)
Total(C)	2527700	100.0		

During the I-35W Urban Corridor Demonstration Project, a three-variable cost allocation model was prepared for each month from October 1972 to December 1974 - - the duration of the monitoring program.

Peak-Base Theoretical Framework

The traditional cost allocation model to some extent addresses the issue of different cost by time of day through the use of a peak vehicle unit cost factor. It does not account for the major cost differences between peak and base time periods in the labor-intensive transit industry in which drivers' wages represent the largest single expenditure. It is widely accepted that it costs more to operate a bus during the peak period than during off-peak hours because of the provisions in most labor agreements which require more payhours per vehicle hour for peak period service than base operations. Typical provisions which impact costs include:

1. Straight runs insure that at least some peak period drivers will have a continuous uninterrupted workday.
2. Combination time prescribes penalties for peak period only drivers to receive a full day's pay for less than eight hours of work.
3. Spread time provides premium pay for any work performed beyond a fixed daily time span (e.g., 10 hours).
4. Guarantee time sets minimum weekly pay regardless of hours worked (e.g., 40 hours pay per week).

While it is evident that these prohibitions and penalties associated with drivers' wages cause higher vehicle hour unit costs for peak period service, the quantification of these differences is yet another matter. The vehicle hour unit cost factor determined by the traditional cost allocation model represents a weighted average of both peak and base conditions. As noted previously, vehicle hours represent an easily quantified surrogate variable for payhours. Thus, it would be desirable to relate peak and base unit cost per vehicle hour factors to the systemwide unit cost (traditional model). Further, this relationship should include some measure of labor productivity (payhours/vehicle hours) and the service levels operated in each period (peak/base vehicle hours). These indices would be computed possibly one month a year and then used for model development in each of the 12 months of that year. The mathematical derivation of these desired relationships is presented below.

Consider the following definition of terms - -

VH_P = Peak period vehicle hours
VH_B = Base period vehicle hours
PH_P = Peak period payhours
PH_B = Base period payhours
TC = Total cost allocated to vehicle hours
UC_S = Vehicle hour unit cost (traditional cost model)
UC_P = Peak period vehicle hour unit cost
UC_B = Base period vehicle hour unit cost.

In a traditional model, UC_S is computed as shown in Equation 1.

$$UC_S = \frac{TC}{VH_B + VH_P} \tag{1}$$

Further, the relationship between payhours and vehicle hours can be established.

$$E_P = \frac{PH_P}{VH_P} = \text{Peak period labor productivity} \tag{2}$$

$$E_B = \frac{PH_B}{VH_B} = \text{Base period labor productivity} \tag{3}$$

The indices which should be related along with UC_S to peak and base unit cost factors are:

$$n = \frac{E_P}{E_B} = \text{Relative labor productivity} \tag{4}$$

$$s = \frac{VH_P}{VH_B} = \text{Service index} \quad (5)$$

It should be recognized that these values can be determined for each transit operator. Relative labor productivity (n) is a measure of the various features of the labor agreement while the service index (s) measures the relative amount of service offered in each time period. Mathematically, the desired relationship for peak and base vehicle hour unit costs are presented in Equations 6 and 7, respectively.

$$UC_P = f(UC_S, n, s) \quad (6)$$

$$UC_B = g(UC_S, n, s) \quad (7)$$

As noted previously, vehicle hours is an easily computed surrogate variable for payhours. Also, for derivation purposes, it is necessary to define payhour unit cost as follows:

$$UC_H = \frac{TC}{PH_B + PH_P} \quad (8)$$

By substituting Equations 2 and 3 in Equation 8, it can be shown:

$$TC = UC_H E_B VH_B + UC_H E_P VH_P \quad (9)$$

Since the sum of the unit costs multiplied by the appropriate quantities for each operating period must equal total cost, UC_P and UC_B can be defined as follows:

$$UC_P = UC_H E_P \quad (10)$$

$$UC_B = UC_H E_B \quad (11)$$

By various substitution of terms in Equation 10, it can be shown:

$$UC_P = \frac{n(1+s)}{1 + ns} UC_S \quad (12)$$

The term multiplied by UC_S can be thought of as an adjustment factor to compute UC_P. Similarly, by various substitution of terms in Equation 11, it can be shown:

$$UC_B = \frac{1 + s}{1 + ns} UC_S \quad (13)$$

The term multiplied by UC_S can be thought of as an adjustment factor to compute UC_B. Also, from Equations 12 and 13, it can be shown:

$$UC_P = n\, UC_B \quad (14)$$

Thus, it has been derived that the peak and base vehicle hour unit cost factors are a function of the systemwide unit cost, relative labor productivity, and service index. Because of space limitations, the complete derivations of Equations 12 and 13 are not presented in this paper. Traditional cost allocation model vehicle hour unit cost (systemwide) underestimates the cost of peak period service and overestimates the cost of base period service. The greater the values of either relative labor productivity (n) or service index (s), the greater the disparity in peak and base vehicle hour unit costs.

When the relative labor productivity equals one (no prohibitions or penalties in labor utilization), vehicle hour unit costs for the system base and peak periods are the same regardless of relative service levels (Figures 1 and 2). For any given value of service index, the peak adjustment factor is directly proportional to the relative labor productivity index which implies a widening disparity between system and peak period vehicle hour unit cost factors. Care should be exercised in interpreting the relationship portrayed in Figure 1. When relative labor productivity is greater than one (the typical situation), a greater value for the service index produces a lower value for the peak adjustment factor. For example, when the relative labor productivity equals two, the peak adjustment factor is larger when $s = 1 (VH_B = VH_P)$ than when $s = 2$ $(VH_B = \frac{1}{2}VH_P)$. At first glance, this may seem illogical; except, it should be noted that the systemwide vehicle hour unit cost factor is not fixed. As the service index increases, UC_S will also increase. Thus, with increasing values of service index, the systemwide unit cost factor becomes more similar to the peak unit cost factor with the peak adjustment factor approaching one. The overall result of greater value of service index is that both systemwide and peak unit cost factors would be greater.

As shown in Figure 2, the base adjustment factor also represents a family of curves which all intersect when the relative labor productivity equals one. For a given value of relative labor productivity, the base adjustment factor is inversely proportional to the service index. Thus, as the cost structure of the system more closely resembles the peak unit cost factor, the disparity between base and systemwide costs becomes greater.

The derived relationships between unit cost factors with relative labor productivity and service index are presented in Figures 3 and 4. As shown in Figure 3, for all values of relative labor productivity, the peak adjustment factors converge on the value of one with increasing values of service index. Conversely, the base adjustment factor diverges from one with increasing values of service index (Figure 4).

Model Application

The first step in applying the derived formulae is to select a month to compute the two indices -- relative labor productivity and service index. Since the tabulation of vehicle hours and payhours for both peak and base periods requires considerable data collection and manipulation, this effort was performed for only a single month. The indices computed from this data tabulation were then applied to subsequent months. It should be recognized that the indices would have to be recomputed when the labor contract changed affecting relative labor productivity (n) or when service levels were changed thereby affecting the service index (s). The results of this data tabulation for the "audit" month are presented in Table 2. Not surprising, the peak period requires 31 percent more payhours than vehicle hours, while the base period has only 14 percent more payhours than vehicle hours, which produced a relative labor productivity of 1.15. These results clearly indicate that provisions in labor contracts which restrict driver utilization and provide for penalty payments can affect costs as significantly as the drivers' hourly wage rates.

To compute the peak-base unit cost factors, the first step was the development of the traditional cost model as described previously. The next step was to apply the index values in Equation 12 and Equation 13 to determine the peak and base vehicle hour unit cost factors, respectively, for the month being

Figure 1. Peak adjustment factor vs. relative labor productivity.

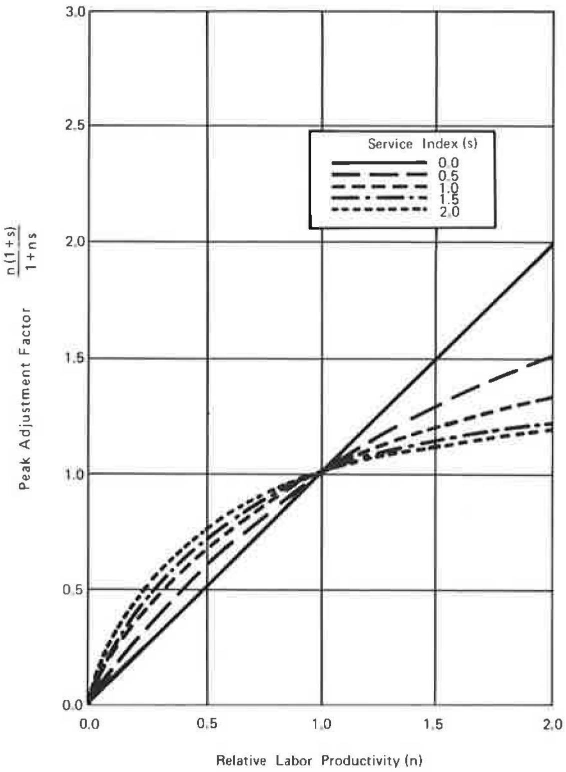


Figure 2. Base adjustment factor vs. relative labor productivity.

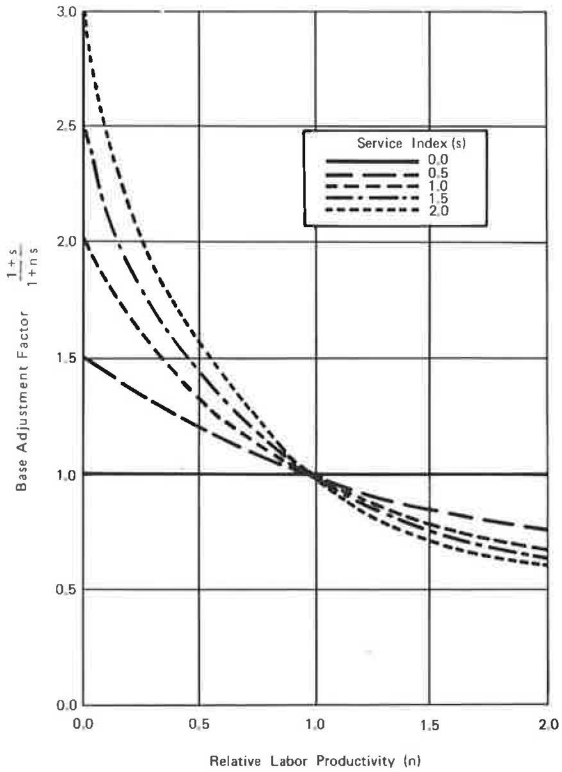


Figure 3. Peak adjustment factor vs. service index.

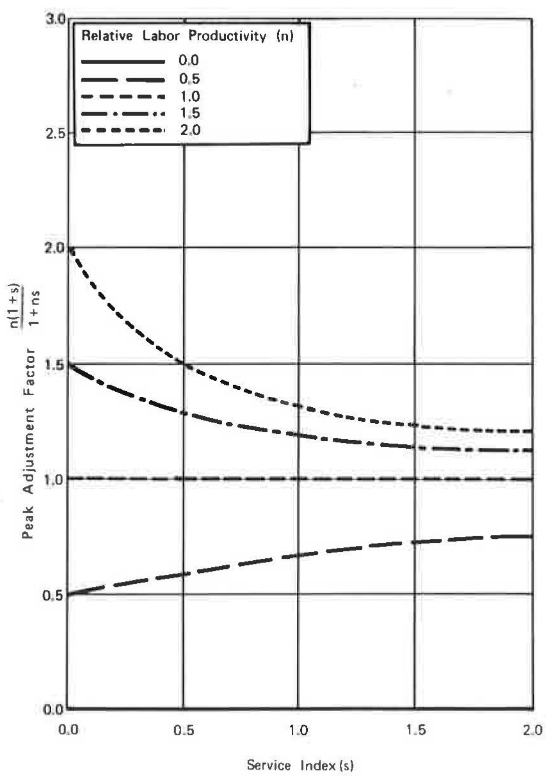
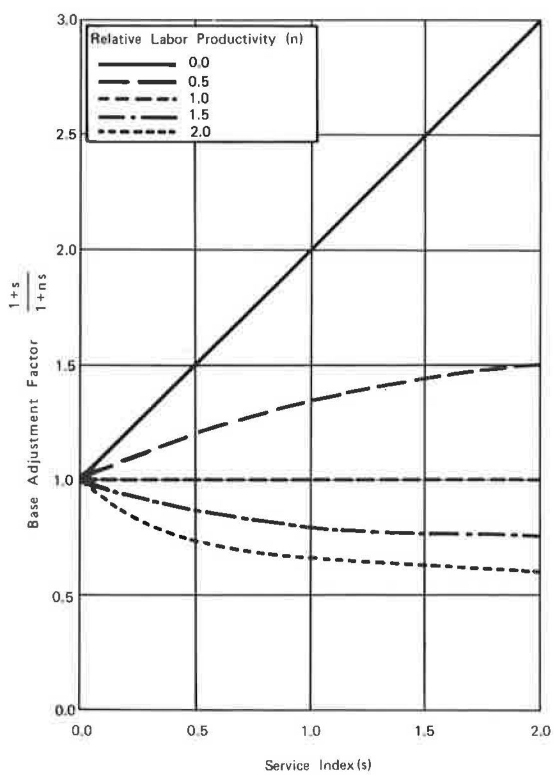


Figure 4. Base adjustment factor vs. service index.



analyzed. The resulting cost formulae for September 1974 are presented below:

$$\text{Peak:} \quad C = 10.57H + 0.19K + 612.75V \qquad (14)$$

$$\text{Base:} \quad C = 9.20H + 0.19K \qquad (15)$$

Table 2. Computation of indices, "Audit" month — March 1974.

	Peak	Base
Vehicle Hours	74967 (VH_P)	72947 (VH_B)
Payhours	98130 (PH_P)	83086 (PH_B)
Labor Productivity	1.31 (E_P)	1.14 (E_B)
Relative Labor Productivity	1.15(n)	
Service Index	1.03(s)	

The difference in cost estimates between the peak and base models can be illustrated by determining the cost of bus service in the I-35W Corridor by both formulae. For September 1974, the peak cost model would yield monthly operating costs of about $302,000, while the base cost model would estimate bus costs of $193,000 - - a difference of 56 percent. These results are not surprising since the base cost model has a lower vehicle hour unit cost and does not include a third variable to reflect peak vehicle requirements. The disparity between cost by the two models would confirm the need to develop separate cost formulae by time period.

By developing models requiring route statistics on vehicle hours rather than payhours, the cost model can be readily applied to individual line data to compute route-by-route cost.

Conclusions

The foregoing analysis permits the following conclusions to be drawn:

1. The use of traditional cost allocation formulae only partially explains the different cost structure of peak and base services. This traditional approach, by use of a peak vehicle unit cost, only accounts for the higher cost of providing service attributable to those cost items such as administrative and physical plant costs which are a function of the maximum number of vehicles in service at any one time.

2. In view of the labor-intensive nature of transit operations, the systemwide vehicle hour unit cost factor represents an average of differing costs by time of day. Clearly, there is a need to define peak and base vehicle hour unit cost factors. Also, the determination of these factors should reflect the consequences of various prohibitions and penalties in the utilization of drivers.

3. Traditional cost allocation models underestimate the cost of peak period service and overestimate the cost of base period service.

4. The peak and base vehicle hour unit cost factors are a function of the systemwide cost structure, as well as the relative labor productivity and service index. Further, these relationships can be mathematically derived.

5. Restrictions on driver utilization and penalty payments can affect transit operating cost as significantly as the drivers' wage. For example, a change in spread time from 10 to 9 hours may produce the same increase in cost as a 10-cent increase in the drivers' wage rate.

6. Expansion of service in peak periods at a relatively greater rate than base periods will also adversely affect transit operating costs. This conclusion has particular relevance to many transit properties that have embarked on ambitious programs to serve journey-to-work travel through express bus service and park-ride facilities - - peak period operations.

7. Although vehicle hours is really a surrogate for payhours, the ease of computing vehicle hours by route as opposed to computing payhours by route suggests its use in cost formulae.

8. The proposed peak-base approach described in this paper only requires data collection at infrequent intervals to compute the necessary indices (relative labor productivity and service index). By utilizing these indices, detailed peak and base cost formulae can be readily determined and applied each month to accurately assess the financial performance of each route.

References

1. Water Cherwony. A Cost Center Approach to Transit Performance. Transit Journal, American Public Transit Association, Fall 1977.
2. Bather, Ringrose, Wolsfeld, Inc. and Simpson & Curtin. I-35W Urban Corridor Demonstration Project. Final Report for the Metropolitan Council of Minneapolis/St. Paul. Minneapolis, MN: Bather, Ringrose, Wolsfeld, Inc., August 1975.
3. Michael G. Ferreri. Development of a Transit Cost Allocation Formula. HRB, Highway Research Record 285, 1969.
4. Walter Cherwony and Brian E. McCollom. Development of Multi-Model Cost Allocation Models. Proceedings of Fourth Annual Intersociety Conference on Transportation, American Society of Mechanical Engineers, 1976.

Abridgment

AN INTERACTIVE BUS TRANSIT MANAGEMENT INFORMATION SYSTEM USING CREDIT CARD FARE COLLECTION DATA

Frank DiCesare, Rensselaer Polytechnic Institute
Gary C. Sullo, General Motor's Technical Center

This research has shown that the development of a bus transit management information system that uses data from a bus credit card fare collection system is feasible from a hardware and software standpoint. An assessment of the data available from the credit card system shows that valuable and timely ridership and revenue information, which would not be readily available otherwise, can be provided to assist in management decisions regarding changes in service, i.e., adding or abandoning service, changing service modes, changing level of service and changing fare levels and/or structure. To match the large amount of available information to the transit manager's needs, a two level form of presentation is proposed with daily indicators for monitoring transit system performance and detailed reports available on demand.

Research in transit management information systems (MIS) is particularly pertinent at this time of scarcity when transportation policy is tending towards making better and more efficient use of the transportation facilities already existing in urban areas. Public officials [1] have been stating their concern for "finding better ways to manage transit" and citing "better management information systems" as a means for improving transit management.

The research reported here addresses the feasibility of using data obtained through credit card collection as a basis for a bus transit operations MIS.

The information system is built around a fare collection system having the following characteristics:

1. Each bus has a modified fare box which accepts credit card insertion.
2. Card insertions cause a record to be entered on a cassette tape (or other storage medium).
3. Data relating to the passenger is obtained from the card, and data relating to the vehicle is obtained from the fare box according to driver entries, an internal clock and an odometer input.
4. Credit card passengers are charged as a function of distance traveled and must insert their cards upon both boarding and deboarding.
5. These cassette tapes are then removed and processed at the end of each day.

The assumptions for this scenario of usage were based in part on an UMTA demonstration project in the Valley Transit District (VTD) of Connecticut [2], where the data was used only for fare calculation and billing [3].

Previous research at Stanford University [4] examined the use of bank credit cards for fare collection and recognized the potential usefulness of the data. In the near future, UMTA plans to further demonstrate the credit card concept, including use of the data for MIS.

Transit Management Information Systems

The information systems thus far developed or discussed in the literature can be categorized by their major emphasis: maintenance, planning, or management of operations [5,6,7,8,9].

The MIS developed at the Rensselaer Polytechnic Institute (R.P.I.) during this research is aimed at transit systems offering a full range of services from fixed route to demand responsive and whose principal means of fare collection is the credit card. Further, the system developed allows for data inputs other than those generated by credit card insertions such as the dispatcher file (where dispatching is computerized) and survey data files.

Research Objectives

The objectives of this work were: to analyze the software, hardware, and data requirements; design a data base which incorporates credit card trip records; implement an example of the software; and to illustrate, by example, typical usage of the system in transit decision making.

The methodology was to assess credit card data for its management information, determine those decisions which would likely be impacted by such information and then design a system which would make the data useful to the manager in considering these decisions.

Input Data

The bus cassette records the following entries for each passenger card insertion: Record type, Passenger ID number, Fare structure code, Time, Distance, Zone, Service type(route number or service region), and Number of persons riding on this fare account.

The dispatching file adds the following data for demand responsive services: Date, Intended time of pick-up, Bus number assigned, Service number of region, Origin zone, Destination zone, Passenger ID number, Fare structure code, Number of passengers to be served (at this pick-up), Maximum fare for this ride, Lead time (advance notice received by dispatcher), and Data code (unserved demand is recorded).

Assessment of the Data for Information Content

It is clear that this data can provide valuable statistics on ridership: volumes by zone of origin, zone of destination, origin-destination pair, time of day, week, month, etc., service mode, area or route, individual vehicle, and so on. The range of operating decisions this information can directly effect includes: adding service, abandoning service, changing service mode, increasing or decreasing level of service, and changing fare levels or structure.

A second set of statistics would be those concerning bus useage: vehicle occupancy, mileage, time, and travel speed; each by service mode and area or route. This information can be useful in the decisions cited previously as well as in the allocation, maintenance, and replacement of equipment.

System Design

The information system consists of three basic components:

1. A set of programs which read and edit the raw data and process it to create and update files for storage.
2. The online data base of files of which two, a passenger trip record file and an indicators file were implemented.
3. A set of programs which operate on the data base to generate information for the manager at two levels, aggregate performance indicators and sets of detailed statistics.

Output Information

A two level presentation hierarchy was used to avoid information overload of the transit manager. The first level is performance monitoring where the manager can compare system INDICATORS with expectations or standards in order to detect changes. The second level provides detailed systems statistics in the form of REPORTS so that a diagnosis of the problem can be made and solutions proposed.

Figure 1 gives the list of performance INDICATORS available in this system. To specify the information he wishes to see, the manager is guided interactively by the program as shown in Figure 1.

In Figure 2, a typical indicator display is shown. As a variation in this display one could also view all indicators for each period of the day, where the column headings would then read: AM peak-Base period-PM peak-Evening. This could be done for any one service type or for the system as a whole. Other variations include displaying a history of any one indicator.

The more detailed passenger movement reports produced by the information system are:

1. Origin-Destination - the number of passengers by O-D pair.
2. Ridership - a time profile of the number of passengers on board any service by hourly interval.
3. Bus Occupancy - ridership broken down further as a separate report for each vehicle, rather than by service type.
4. Passenger Boarding - ridership broken down to shown zone of boarding and deboarding within the specified service area.
5. Wait Time Distribution - distribution of time spent waiting for a late vehicle (demand responsive service).
6. Lead Time Distribution - distribution of advance notice times for demand responsive systems.

Figure 1. An example of the manager-computer interactions resulting in a particular indicator display.

```
ENTER THE SERVICE TYPE DESIRED (1-4).
ENTER 5 FOR THE CUMMULATIVE TOTAL OF ALL
SERVICES. ANY ENTRY GREATER THAN THIS CALLS FOR
ALL SERVICES TO BE DISPLAYED INDIVIDUALLY.
USE TWO DIGITS.
---> 02

USING ONE DIGIT, ENTER THE PERIOD OF THE
DAY TO BE ISOLATED. THE CODE IS:
     1   AM PEAK
     2   BASE PERIOD
     3   PM PEAK
     4   NIGHT PERIOD
     5   ALL DAY (SUM OF 1-4)
ANY ENTRY GREATER THAN 5 CALLS FOR ALL PERIODS
TO BE DISPLAYED CONCURRENTLY.
---> 9

ENTER TWO DIGIT CODE TO SPECIFY THE INDICATOR
TO BE PROCESSED. ANY ENTRY GREATER THAN 15 CALLS FOR
ALL INDICATORS TO BE DISPLAYED, IF POSSIBLE.
     01 - NUMBER OF BUSSES OPERATING
     02 - BUS REVENUE HOURS
     03 - RIDERSHIP
     04 - PERCENT COIN RIDERS
     05 - UNSERVED DEMAND
     06 - VEHICLE LOADING BY MILES
     07 - VEHICLE LOADING BY HOURS
     08 - PASSENGER-HOURS STANDING
     09 - AVERAGE DEVIATION FROM DIAL-A-RIDE SCHEDULE
     10 - AVERAGE DEVIATION FROM FIXED ROUTE SCHEDULE
     11 - AVERAGE TRIP TIME
     12 - AVERAGE TRAVEL SPEED
     13 - TOTAL REVENUE
     14 - REVENUE PER VEHICLE-MILE
     15 - REVENUE PER VEHICLE-HOUR
---> 99
```

Figure 2. A typical display of all indicators for the total system by periods of the day.

ALL INDICATORS
SYSTEM AS A WHOLE
FOR ONE PERIOD OF 2 DAYS
BEGINNING DAY 10

	AM PEAK	BASE PERIOD	PM PEAK	NIGHT PERIOD	ALL DAY
BUSES IN REVENUE SERVICE	20.00	20.00	20.00	20.00	80.00
BUS REVENUE HOURS	40.00	140.00	40.00	121.98	341.98
RIDERSHIP	63.00	278.00	75.00	178.00	594.00
PERCENT COIN RIDERS	0.0	0.0	0.0	0.0	0.0
UNSERVED DEMAND	0.0	0.0	0.0	0.0	0.0
VEHICLE LOADING BY MILES	0.02	0.02	0.02	0.02	0.02
VEHICLE LOADING BY HOURS	0.55	0.76	0.63	0.46	0.61
PASSENGER HOURS STANDING	0.0	0.0	0.0	0.0	0.0
AVERAGE DEVIATION FROM PROMISED PICKUP(MIN)	-0.76	-0.34	-0.11	-0.30	-0.34
AVERAGE DEVIATION FROM SCHEDULE(MIN)	0.0	0.0	0.0	0.0	0.0
AVERAGE TRIP TIME(HRS)	0.23	0.25	0.23	0.21	0.24
AVERAGE TRAVEL SPEED	16.16	18.35	19.94	18.68	18.39
TOTAL REVENUE($)	37.98	206.64	48.24	105.84	398.70
REVENUE/VEHICLE-MILE	0.11	0.18	0.15	0.12	0.15
REVENUE/VEHICLE-HOUR	0.95	1.48	1.21	0.87	1.17

These operational reports would be called for by interactive input similar to that for displaying the daily performance indicators. In this case the questions provide the user with a menu for selecting one or several of the reports, and allow for specification of which service type to be processed. Other input would relate to time of day and dates of data to be reported.

Conclusions

The major conclusion of this research is that a useful management information system can be readily achieved in situations employing credit card fare collection.

The programming requirements are moderate and a system of this type could be implemented on a mini-computer. The flowcharts, source code, and explanatory notes available from this project [10] provide sufficient detail such that they can provide assistance in writing of functional specifications for the software in a particular implementation of this type of management information system.

A logical extension to this information system would be to adapt it to a non-credit card transit operation by the use of rider tickets issued upon entry and retrieved upon exit. Since no billing will be done with this system, ridership data can be collected on a sampled basis therefore obviating the need for a box on all buses simultaneously.

Acknowledgements

The work reported in this study was performed at the Rensselaer Polytechnic Institute and was sponsored by the University Research Program of the Urban Mass Transportation Administration under grant number DOT-UMTA-NY-11-0001. The opinions and conclusions expressed are the independent results of the authors and do not necessarily reflect the views of the sponsor or of R.P.I. and neither represents or includes any research efforts or positions taken by General Motors.

References

1. "Patricelli Views Transit Applications and Grants", Metro, Vol. 72, No. 1, Jan.-Feb. 1976.
2. RRC International, Inc., "System Operation Aspects of the Valley Transit Demonstration Program", Troy, New York, 1974.
3. Nelson, Gary, "FAIRTRAN: Operation of a Credit Card Transit Fare System", presented at the 55th Annual Meeting of the Transportation Research Board, January 1976.
4. Buckley, R.F., et. al., "Design and Analysis of an Automatic Credit Card Fare Collection System", Report No. UMTA-CA-008-74-1, NTIS No. PB-237-092, Stanford University, Stanford, CA, 1974.
5. Thurlow, V.S., SIMS Overview Service: Inventory, and Maintenance System, MITRE Corporation, Mc-Clean, VA 22101, (NTIS: PB241-495/AS), 1973.
6. Worrall, R.D., Conceptual Design for an Urban Transportation Planning Information System, Peat, Marwick, Mitchell and Co., Washington, D.C., Federal Highway Administration, Washington, D.C., (NTIS: PB-240639/5ST), 1973.
7. Dial, R.B., "Demand Forecasting for New Options and Technology-Resource Paper", Highway Research Board Special Report No. 143, pp. 77-88.
8. Fielding, G.J. and Shilling, D.R., "Dial-a-Ride: Opportunity for Managerial Control", Transportation Research Board Special Report No. 147, 1973, pp. 69-77.
9. Hartzler, R.E., "System Performance Data Processing for a Demand-Responsive Public Transportation System", Report No. UMTA- A-06-0012-74-3, MTR-6794, MITRE Corp., November 1974.
10. DiCesare, F. and Sullo, G., "A Computerized Bus Transit Management Information System Using Credit Card Fare Collection Data", Final Report to UMTA, Rensselaer Polytechnic Institute, Troy, New York, 1976, available from NTIS.

Abridgment

COST ANALYSIS OF CURRENT U.S. SURFACE TRANSIT FARE COLLECTION SYSTEMS

James I. Scheiner and Subhash R. Mundle, Simpson & Curtin, Philadelphia

U.S. surface transit fares have typically been collected on-board the vehicle. The objective of this analysis is to measure capital and operating costs associated with collecting fares. A literature review indicated that transit systems do not record such directly related operating expenses.(1) Therefore, six transit systems were visited to obtain detailed data. The systems, selected on the basis of size and uniqueness of fare structure or collection equipment, were:

1. Westport Transit District (WTD), Westport, CT - Annual transit passes account for more than 70 percent of the system revenue, and about 90 percent of passenger boardings.
2. Red Rose Transit Authority (RRTA), Lancaster, PA - A system using conventional fare collection methods typical of small properties.
3. Capitol Area Transit (CAT), Harrisburg, PA - A medium-size system with the latest Duncan Industries fare collection equipment technology.
4. CNY Centro, Syracuse, NY - a medium-size system with the latest Keene Corporation fare collection equipment and vacuum revenue removal, and automatic data retrieval (ARCOM) technology.
5. METRO, Seattle, WA - A large transit system with conventional fare collection equipment but extensive "Magic Carpet" service - a downtown free ride zone.
6. MBTA Surface Transit, Boston, MA - A large multi-modal system with the largest utilization of employee payroll deduction passes in the country.

This discussion addresses two broad categories of costs - direct and indirect. Direct costs are the quantifiable capital and operating costs. Indirect costs, not easy to quantify, are related to system-wide revenue loss, operator involvement in fare-related activities, and impact of fare collection procedure on the system's insurance liabilities.

Direct Fare Collection Costs

Capital Costs

The capital cost estimates were based on an inventory of on-board fare collection equipment, off-board fare processing equipment, and other fare collection-related equipment. Annual depreciation cost for each system was derived from 1976 Reproduction Cost New and assumes an average 15-year life for fare collection equipment. As shown in Table 1, the annual depreciation cost per unit of operating parameter (peak vehicles and vehicle kilometer) is lower for the case-study systems with non-registering fareboxes than for those with registering fareboxes.

Table 1. Fare collection capital cost comparison

System Location	Total Annual Depreciation Cost[a]	Annual Depreciation Cost: Per Peak Vehicle	Annual Depreciation Cost: Per Vehicle km
Westport, CT[b]	$ 530	$ 53	0.074¢
Lancaster, PA[b]	1360	50	0.068
Harrisburg, PA	10250	160	0.329
Syracuse, NY	55300	·398	0.725
Seattle, WA[b]	18600	37	0.050
Boston, MA	$202100	205	0.372

[a]Based on 1976 Reproduction Cost New and an average 15-year life for fare collection equipment.

[b]Systems with non-registering fareboxes.

Not surprisingly, the CNY Centro registering fareboxes, vacuum revenue removal and ARCOM data processing system engender the highest annual depreciation cost per peak vehicle ($398), and per vehicle kilometer (0.725¢).

Operating Costs

The fare collection-related operating costs were grouped into four categories - Personnel, Materials and Supplies, Service Contracts, and Accidents and Insurance.

Personnel Costs. Typically, maintenance, transportation and accounting department personnel are involved in fare collection-related activities.

Table 2. Fare collection-related annual operating cost.

System Location	Cost Category: Personnel	Materials and Supplies	Service Contracts	Accidents and Insurance	Total
Westport, CT	$ 7700[a]	$ 2585	$ 895	$ 100	$ 11280
Lancaster, PA	19900[b]	825	0	2460[c]	23185
Harrisburg, PA	55000[d]	4325	2670	1990	63985
Syracuse, NY	70300[a]	17360	10610	30020	128290
Seattle, WA	334300[a]	104050[e]	43165	2155	483670
Boston, MA	1590500[a]	17240	24050	129400[c]	1751190
Total	$2077700	$146385	$71390	$166125	$2461600
Percent of Total	84.4%	6.0%	2.9%	6.7%	100.0%

[a]The fare collection-related activities of bus operators have been considered a collateral duty.
[b]One on-board fare collection activity of bus operators, making change, has been included in personnel cost.
[c]Includes an estimate of certain revenue losses.
[d]Ten minutes of bus operator turn-in time is included in personnel cost, but the operators' on-board fare collection activities, which include selling of tickets, are excluded.
[e]Includes annual cost of farebox-related road calls.

Materials and Supplies. Typical M&S items are spare parts inventory, transfer tickets, punch tickets, various types of adult passes, senior citizen identification cards, and student passes.

Service Contract. Services typically contracted include maintenance of special fare processing equipment, burglary protection, and armored car service.

Accident and Insurance. These costs include those of on-board accidents directly involving fare collection equipment and insurance premiums for employee bonding and felonious assults.

Table 3. Fare collection operating cost comparison.

System Location	As Percent of Total Operating Cost	As Percent of Total Operating Revenue	Per Peak Vehicle	Per Vehicle Km
Westport, CT	2.9%	10.3%	$1130	1.55¢
Lancaster, PA	1.7	4.6	860	1.18
Harrisburg, PA	2.3	3.5	1000	2.05
Syracuse, NY	2.0	3.5	925	1.67
Seattle, WA	1.3	4.3	970	1.24
Boston, MA	1.4%	8.3%	$1780	3.22¢

Operating Cost Comparison

Table 2 shows estimated annual fare collection-related operating cost of the six case-study systems. It is evident that current collection and processing activities are labor-intensive: more than 84 percent of the total cost is personnel cost. For MBTA Surface Transit, almost 91 percent of the fare collection operating cost is personnel-related.

Table 3 relates fare collection operating cost with total operating cost, total operating revenue, cost per peak vehicle and cost per vehicle kilometer. For the six case-study systems, the cost varies from a low of 1.3 percent to a high of 2.9 percent of the total operating cost. Costs of fare collection as a percent of total operating revenue ranges from 3.5 percent to 10.3 percent. This calculation is particularly appropriate since it compares the cost of collecting revenue with the amount of revenue collected. Both WTD and MBTA Surface Transit show fare collection costs as percent of total operating revenue much higher than those of other case-study systems. This is attributable to their charging of lower base fares.

Comparison on the basis of unit operating parameters, peak vehicles and vehicle kilometer, shows the fare collection cost of MBTA Surface Transit to be much higher than the other systems. For instance, the operating cost of fare collection for the Lancaster system is $860 per peak vehicle and 1.18¢ per vehicle kilometer. The corresponding cost factors for MBTA Surface Transit are $1,780 per peak vehicle and 3.22¢ per vehicle kilometer. Yet, MBTA Surface Transit fare collection costs are a typical proportion of total system operating costs (1.4 percent).

Indirect Fare Collection Costs

It has been said that the closer something can be measured, the less important it is likely to be. This may be the case with the costs of fare collection. The directly quantifiable costs of surface transit fare collection, which have been measured with some precision for six bus transit systems with widely divergent fare structures and revenue processing systems, are not a significant percentage of total system operating cost. Based on the previous analysis, measurable transit fare collection costs are in the range of two percent of total operating cost. This indicates that further cost-cutting in the area of revenue collection is unlikely.

Given the relatively small, directly measurable cost of transit fare collection, is there any purpose in the investigation of alternative mechanisms? Before answering this question, we should look beyond the measurable costs into those costs of fare collection that cannot be directly quantified. These are discussed below.

Revenue Losses

Only two of the six transit systems surveyed offered any information on revenue losses associated with fare collection. In no case was the survey able to determine with any precision the ratio of bank deposits to the aggregate fares that riders are supposed to pay under posted tariffs.

Newspaper reports suggest that in some systems revenue losses could be a serious problem. In Washington, D.C., published interviews with WMATA bus operators cited instances of fare underpayments.(2) In Philadelphia, SEPTA management reported to the Board that it feared the theft of up to $1.6 million in revenue annually - about two percent of the system's total.(3)

Revenue losses, like shoplifting, are a cost of doing business. To the extent that some patrons cannot pay their fair share, or that some employees misappropriate funds, the revenue loss must be made up by other patrons and by tax dollars.

Operators' Wage Costs

Only two of six transit systems had information on operators' wage costs associated with the fare collection system. In the other systems, all fare collection activities of drivers were reported as "collateral."

It is difficult to speculate on how organized labor in different transit systems might react to being taken entirely out of the fare collection cycle. Organized labor might also be interested in the job-creating aspects of an off-board fare collection system. If such a system induced patronage, it could lead to higher service levels and more jobs.

All of these factors would have to be examined in the context of collective bargaining at the individual transit system. In no event would off-board fare collection result in a lower wage rate for operators. It could be used in new contract negotiations, however, to mitigate wage increases.

Insurance Liability

Currently, transit systems pay anywhere from three percent to over ten percent of total operating costs for bodily injury and property damage liability insurance coverages. The cost of such insurance has increased rapidly in the past few years, and its availability is now severely limited. Off-board fare collection could impact insurance costs in two ways.

One possible impact is reduction of insurance premiums and claims resulting directly from the farebox. These costs have been discussed in a previous section.

A second possible impact on insurance costs is more subtle and legally more uncertain. About 20 to 30 percent of total liability insurance claims involve bus patrons (the remaining claims involve accidents external to the bus). When transit was a private industry, a legal contract was created when the patron paid a fare. Liability for injury to a patron who had paid the fare was adjudicated on the basis of simple negligence. With tax dollars now paying in excess of 50 percent of a typical transit system's operating cost, has the simple negligence standard changed? Would it change if transit service were provided free of direct user charges, much as streets and highways are provided? If the answer to this question is affirmative, significant liability insurance cost savings would result.

Discouragement of Revenue Patronage

The final category of indirect costs related to present fare collection methods is perhaps the most important. To imagine the possible revenue-generating implications of an off-board fare collection system, one only has to ask the question: What if the only kind of telephone were the pay phone, and the only way to call were to deposit exact change in the call box? How much of the current telephone revenue would be lost? Of course, there are significant differences in the relative market positions and feasible billing practices of a telephone company and a transit system. Yet, there can be no question that the current U.S. exact-fare, average-cost-pricing fare collection method is a significant deterrent to increasing revenue ridership.

Conclusions

While the transit industry's current fare collection costs are inexpensive relative to total operating costs (less than two percent), its fare collection methods may not be efficient. By requiring exact change in the farebox for each and every ride, the patron must bear the administrative burden of payment. This is contrary to the dominant trend in private industry toward credit cards and other "convenience" payment forms. The key question which remains to be investigated is: "Would a more convenient fare collection system generate enough additional revenue to offset higher transit agency fare collection costs?"

References

1. American Public Transit Association, "Transit Operating Report for Calendar/Fiscal Year 1975," Washington, D.C.
2. Washington Post, August 8, 1975, followed by Washington Star stories.
3. Philadelphia Inquirer, December 18, 1976.

BUS TERMINAL PERFORMANCE MEASURED WITH TIME STAMPING

John P. Braaksma, Associate Professor, Department of Civil Engineering, Carleton University, Ottawa, Canada
William P. Johnson, Transportation Analyst, Airport Services and Security Branch, Ministry of Transport, Ottawa, Canada

Previously, the terminal planner faced with the task of evaluating the overall performance of a facility, has had available only disjointed pieces of information concerning parts of the problem. This lack of comprehensive, quantitative data lead to the analysis of separate pieces of the terminal, under different loading conditions, resulting in a piecemeal assessment of the terminal's performance. The research described in this paper was aimed at correcting these problems, through the application of a new survey technique called "time-stamping". After its initial application in an air terminal, this study was undertaken to expand the technique to bus terminals. The time-stamping technique was applied to the intercity bus terminal in Ottawa, Canada. Analysis of the data showed that no capacity related problems existed at present. The future location, magnitude and approximate timing of capacity problems were determined. Specifically, the following information was provided: 1. The amount of area in the terminal used over the survey period (at a given level of service) and its distribution both temporally and spatially 2. average length of stay and its distribution 3. average distance walked 4. desire line mappings indicating layout problems 5. average occupancy curves of terminal facilities, by busload. The processing capacity of the facility was determined, plus the expected impacts of scheduling revisions on increasing the useful lifespan of the terminal. Also, impacts of terminal layout revisions on reducing walking distances was predicted.

In recent years, there has been increasing interest in terminal planning, design and analysis, owing to the realization among transport planners of the importance of this facet of the intercity trip. Experience has indicated that new designs have better served both the traveller and the terminal owner/manager, but quantification of terminal performance remains largely unexplored.

The terminal planner faced with the task of evaluating the overall performance of a facility, previously had only disjointed pieces of information concerning parts of the problem. This lack of comprehensive quantitative data lead to the analysis of separate pieces of the terminal, under different loading conditions: resulting in a piecemeal assessment of the terminal's performance. The need for further work in the area is summarized by Hoel and Rozner (1):

> Until very recently, little attention has been paid to methods for evaluating the performance of transition points between modes such as . . . transitions between intercity and interurban transportation networks . . . there is a pressing need for systematic procedures and analytical methods of facility design and evaluation.

Purpose and Scope of Study

Before attempting to optimize terminal design, the objectives of such a transfer facility should be delineated. The passenger is interested in effecting the modal transfer with a minimum amount of delay, a maximum degree of comfort and safety and at the preferred time. The owner/manager wishes to satisfy all demand, but must be concerned with return on investment through cost minimization and revenue maximization.

A search of the literature relevant to the state of the art in terminal evaluation/planning indicated that various parts of the terminal have been investigated separately. There remains the need for quantitatively examining the combination and interaction of these elements on a broader scale. This would be of more use to the practicing planner in ascertaining a facility's adequacy. Bits and pieces, which comprise the total, do not act in isolation; and so to analyse them as if they did, cannot be realistic.

Design standards and rules-of-thumb are available for rough "sizing out" of platform lengths, recommended area square footages etc. (2). These are usually based on gross yearly flow figures only and are derived from generalized past experience. Fruin (3) has defined pedestrian levels of service for walkways, stairways and queueing/waiting areas. These have been used extensively in the following analysis.

Computer simulations of terminal facilities have also been prepared (4). But none of the efforts surveyed have specifically addressed the problem of complete terminal performance evaluation. It is felt that one of the main reasons for this deficit has been the lack of data collection techniques suited to the task.

This paper presents the work of Johnson (5) who attempted to quantify as many of the stated design objectives as possible, measure these for a facility in operation and subsequently use these measures to evaluate the performance of a complete terminal.

This paper will therefore deal only with those design objectives which are readily quantifiable and measurable. Revenues and costs will not be dealt with directly. Table 1 specifies those areas that will be investigated and the measures to be used for each.

Physically the scope of the study is confined primarily to the terminal building itself, i.e. from the time the pedestrian enters until he leaves it.

Field Work

The intercity bus terminal in Ottawa, Canada, owned by Voyageur-Colonial Ltd. was selected as the example for analysis owing to its proximity, manageable size, relatively heavy usage (in peak periods), and newness of design.

The measures of performance sought were quantitative and detailed in nature and so the data base had to be the same. Basic elements such as occupancy counts, flow volumes, flow patterns and processing rates within the terminal had to be accurately provided. With these as a basis more complex measures could then be calculated.

The survey techniques considered were:

1. Personal interviewing.
2. Self administered questionnaires i) collected by survey personnel ii) mailed back.
3. Counts and observations.
4. Videotape and time lapse photography.
5. Inference from existing sources.
6. Tailing.
7. Time-stamping.

For detailed descriptions of the first 6 techniques see the Airport Travel Survey Manual (6). The last one, time-stamping, was chosen as the most appropriate in terms of data quality and content, total terminal coverage, cost and ease of execution and analysis.

Time-stamping is premised on the fact that virtually all quantitative data can be derived for a terminal, if a time-space trace of each of its occupants can be made. A simple objective, but previously not possible using data from other survey techniques.

In this method, each person upon entering the building is given a card to carry (illustrated in Figure 1) which contains a coded time and location stamp. As the pedestrian proceeds through the facility, his card is further stamped by surveyors at the entrance/exit (called checkpoints) of each part of the terminal to be examined. Examples of these checkpoints are entrances/exits to ticket lobbies and queues, waiting areas, baggage claim areas, restaurants etc. When the traveller prepares to leave the terminal, his card is stamped and collected. The completed card is, in fact, exactly the required time-space trace. For further details on this method, see Braaksma (7).

Figure 2 shows the layout of the bus terminal, along with the location of checkpoints and surveyors. The survey was carried out during the weekly peak period of Friday, 14:00-19:00 H., on January 23, 1976.

University students were hired as surveyors and required approximately one hour to be briefed and trained.

Entrants to the terminal were asked if they were passengers or visitors, with the cards of the former being marked with a large "P" by the surveyors.

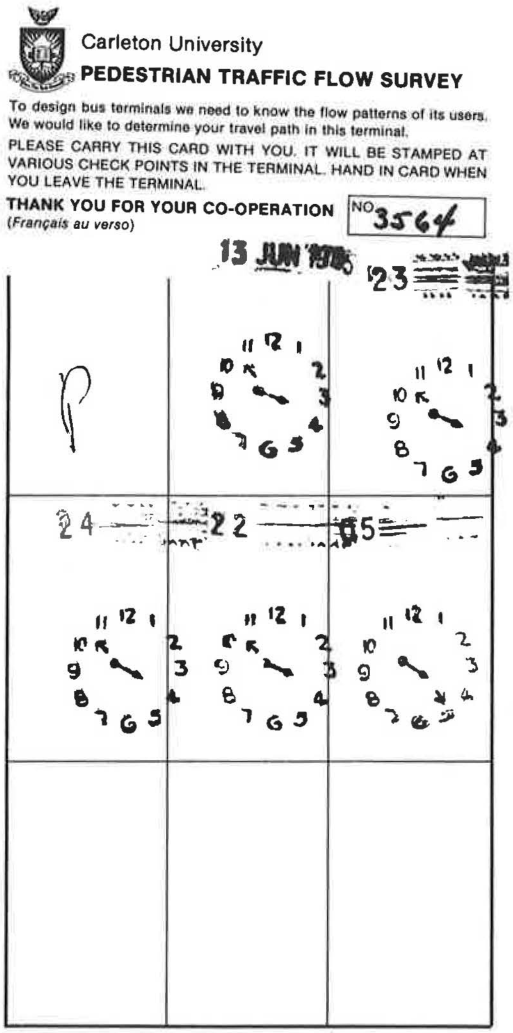
Problems in executing this survey were minimal with a few of the stamping clocks requiring replacing and there being insufficient surveyors available to permit coffee breaks. It was felt that surveyors were visible enough that participants could readily locate them, but did not interfere with or influence normal flows or activities.

The acceptance-completion rate was 45% of terminal users, giving a return of 1199 cards in 5 hours of survey time. It should be noted that determining the exact acceptance rate is vital to allow expansion of the data to simulate actual conditions.

Table 1. Terminal design objectives.

Terminal Design Objectives	Measure(s) Used
User-Related	
minimize time spent in terminal	. length of stay distribution . mean, median length of stay . length of stay variation with time
minimize distance walked	. average distance walked . desire line map (trip tables)
minimize crowding	. load to capacity ratios using appropriate levels of service
maximize available "extra" services	. remaining excess capacity
maximize convenience of arrival and departure times	. scheduling impact information; standard accupancy figures
Owner-Related	
maximize revenue	
. maximize use of existing facility	. area usage ratios . load to capacity ratios
. maximize scheduling efficiency	. scheduling impact information: standard occupancy figures
minimize service disruption	. problem prediction using expected growth and load to capacity ratios

Figure 1. Pedestrian traffic flow survey card.



Performance Measures

Data processing was done manually, owing to the relatively small number of cards. Since the completion of this work, the Airports Services and Security Branch of Transport Canada has developed computer programs for the calculation of most of the following data.

Occupancy Counts

The number of occupants in each part of the terminal over time was determined. This is shown in Figure 3. Peak occupancy for the whole terminal was 411 people at 5:10 P.M.. Figure 3 also illustrates relative values of occupancies in each part of the terminal.

Load to Capacity Ratios

More significant to the planner than simple occupancies is the ratio of what an area is holding to what it can/should hold, at a specified level of service, i.e. load to capacity ratios (L/C).

Many authors feel that 15 minute peaks should be planned for, and so, 15 minute average occupancies are used throughout. This means that if an area has a load/capacity ratio greater than 1.0, then, on the average, the desired level of service is not being met for this period.

To carry out a comprehensive analysis of the terminal, all of its elements must be included. This includes the static holding areas such as the restaurant, bar, botique, and waiting areas, the processors such as ticket wickets and lobby and bus bays; and links such as corridors.

Examples of load to capacity ratios over time are presented in Table 2 for the two processors in this terminal, the main waiting area and the ticket lobby as well as the "auxiliary" services. These have all made use of Fruin's (4) level of service in determining the capacity.

An off-shoot of the load to capacity ratios is the consideration of what percentage of available space is used for what part of the time over the survey period. This is expressed by the formula:

$$\frac{\text{space used} \times \text{time it is used} \times 100}{\text{total available space} \times \text{total time available}}$$

These area-temporal utilization ratios differ from simple L/C's, in that time is included, plus it covers the complete survey period. In fact, differing distributions of L/C's, over this period could result in the same area-temporal utilization. Examples of this are shown in Figure 4 with the distance between horizontal lines representing the facility's capacity. When these are expanded until one of the "bars" reaches the capacity line, the resulting utilization ratio will indicate how much of the area will be used when the facility reaches its capacity at the daily peak. Of course, a flatter curve is more desirable as the loads are then more evenly spread over time, and the effects of efforts to achieve this (such as differential fares) will be made evident here.

Table 2. Load to capacity ratios, %.

Time	Main Waiting Area	Ticket Lobby	Restaurant	Bar	Boutique
14:15-14:29	16	9	2	2	12
14:30-14:44	28	13	22	7	16
14:45-14:59	26	5	13	4	16
15:00-15:14	24	9	11	7	20
15:15-15:29	23	8	26	22	24
15:30-15:44	20	13	27	42	24
15:45-15:59	27	18	39	49	36
16:00-16:14	41	28	63	44	48
16:15-16:29	67	41	71	42	44
16:30-16:44	61	46	56	49	20
16:45-16:59	53	35	69	40	48
17:00-17:14	59	31	61	31	32
17:15-17:29	70	42	55	42	40
17:30-17:44	61	41	57	64	48
17:45-17:59	53	39	39	49	12
18:00-18:14	41	22	36	20	20
18:15-18:29	29	13	11	9	4
18:30-18:44	9	4	0	4	0

Occupancy Curves

To help determine the effects of each busload on the terminal's facilities, the occupancy in the terminal of outbound and inbound passengers was broken down into individual busloads. These curves were then superimposed and averaged for inbound, outbound "local" and outbound "express" bus passengers. An example is shown in Figure 5. These curves, in effect, indicate at any time relative to the scheduled departure or arrival, what portion of the busload can be expected to be in any part of the terminal. They also show graphically the impact scheduling can have on the terminal and surges that occur in loading are made evident. Similar curves have been developed for the waiting area, ticket lobby and combined auxiliary services. Through the superimposition of these curves, schedule revisions have been simulated. This indicated that such changes can have significant impact on extending the life of the facility.

Volume to Capacity Ratios

The processors and links are the dynamic elements of the terminal system. For these, the flow volume-to-capacity ratios have been calculated, again using Fruin's recommended levels of service (see Table 3).

Processing Capacity of Terminal

The facilities that are needed to complete the intermodal transfer, are the ticket wickets, ticket lobby, main waiting area and outbound bus bays. In order to determine the capacity of the terminal as a whole these are the parts that must be given priority. Table 4 contains the list of the daily peak L/C or V/C ratios:

Thus, the main waiting area should be the first essential part of the terminal to encounter capacity problems. When present loads are increased by roughly 50%, its capacity will be reached, at the specified level of service. For the hour surrounding this peak, there were 525 outbound passengers and 856 persons who moved through the terminal. Thus, one would expect that if more than approximately 800 outbound passengers or 1300 persons were to move through the terminal in one hour, there would be capacity problems. This is at a "peak hour factor" of .70 (see Highway Capacity Manual (8)). This has been defined as the terminal's processing capacity -- 800 outbound passengers per hour or 1300 persons per hour.

Table 4. Daily peak L/C or V/C ratios.

Terminal Element	Daily Peak L/C or V/C (%)
Restaurant	71
Waiting area	70
Bar	64
Ticket wickets	61
Boutique	48
Ticket lobby	46
Outbound bus bays	25

When combined with the expected growth rate in patronage, the capacity of the terminal can give the expected life of the facility. It was found here to be roughly 10 years. That is, the demand for use of the terminal will have increased by 50% in this time, and without modifications to the layout or operation, there will be capacity problems.

However, problems will not wait for 10 years to surface because the volume of demand and the resulting service provided are not linearly related in transportation facilities, as stated by de Neufville (9).

> When the arrival rate approaches the maximum rate of service, delays increase disproportionately faster than the rate of arrivals . . . Extraordinary delays result, therefore, from any service system operating near its capacity . . . This behaviour is characteristic of all manner of service systems: check-in counters for passengers: conveyors and sorters for baggage; corridors for pedestrians: runways serving arriving and departing aircraft: and so on.

Trip Tables

From the survey data trip tables were prepared between pairs of checkpoints. These were prepared for inbound and outbound passengers, visitors, and total persons.

When the O-D trip table is presented graphically, the result is a desire line mapping as shown in Figure 6. This can be used to reorganize the layout

Table 3. Volume to Capacity ratios, %.

Time	Door #16	Restaurant Doors	Bar Doors	Boutique Door	Inbound Bus Bays	Ticket Wickets	Outbound Bus Bays
14:15-14:29	2.1	4.1	0.1	4.0	11.1	21.0	0.8
14:30-14:44	1.6	3.1	0.2	5.1	3.6	40.0	8.0
14:45-14:59	1.5	1.6	0.4	3.3	2.8	28.0	16.0
15:00-15:14	2.8	1.3	0.4	2.6	15.0	12.0	3.0
15:15-15:29	2.1	3.1	0.7	5.6	4.0	16.0	5.0
15:30-15:44	1.8	3.0	0.6	3.5	2.8	22.0	4.0
15:45-15:59	2.2	5.3	0.7	6.7	13.8	33.0	5.0
16:00-16:14	3.1	6.9	1.2	10.0	2.8	34.0	3.0
16:15-16:29	3.4	8.4	1.7	9.2	6.7	51.0	7.0
16:30-16:44	3.5	9.1	2.1	7.2	13.8	61.0	25.0
16:45-16:59	3.7	5.8	2.0	7.9	7.1	48.0	17.0
17:00-17:14	3.0	5.4	1.0	8.9	4.4	43.0	7.0
17:15-17:29	3.8	8.1	1.6	10.2	12.3	54.0	16.0
17:30-17:44	3.5	5.3	2.3	7.9	5.2	57.0	13.0
17:45-17:59	2.0	2.8	0.6	5.9	2.0	43.0	12.0
18:00-18:14	1.3	2.8	0.6	5.6	5.2	28.0	13.0
18:15-18:29	1.4	1.8	0.1	3.0	3.2	17.0	11.0
18:30-18:44	0.9	0.0	0.0	0.2	0.0	6.0	4.0

Figure 2. Voyageur-Colonial bus terminal layout, Ottawa

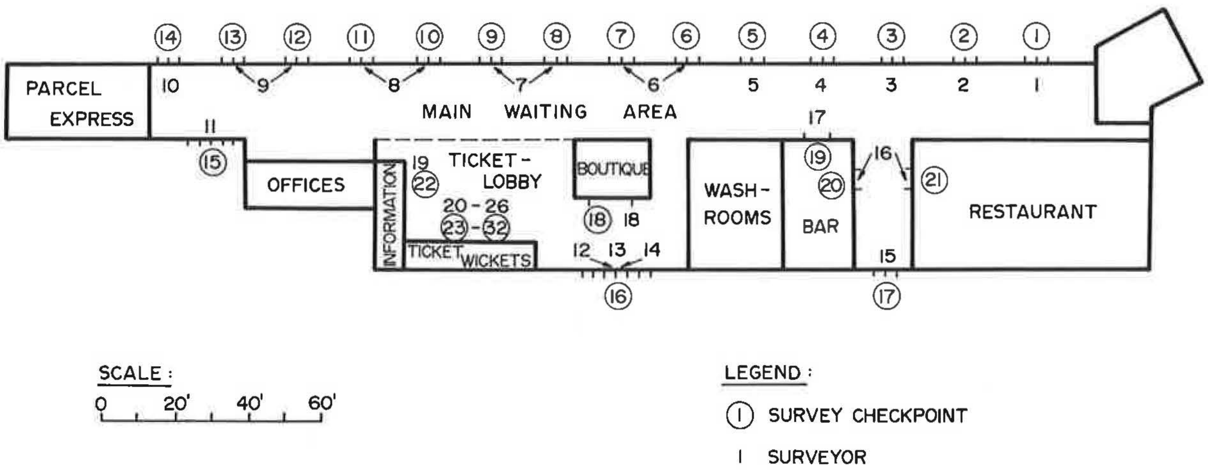


Figure 3. Area occupancies over time.

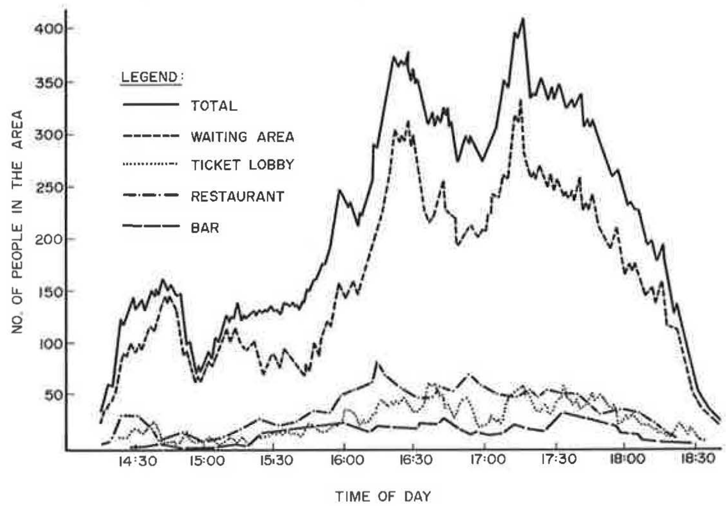


Figure 4. Area-temporal utilization.

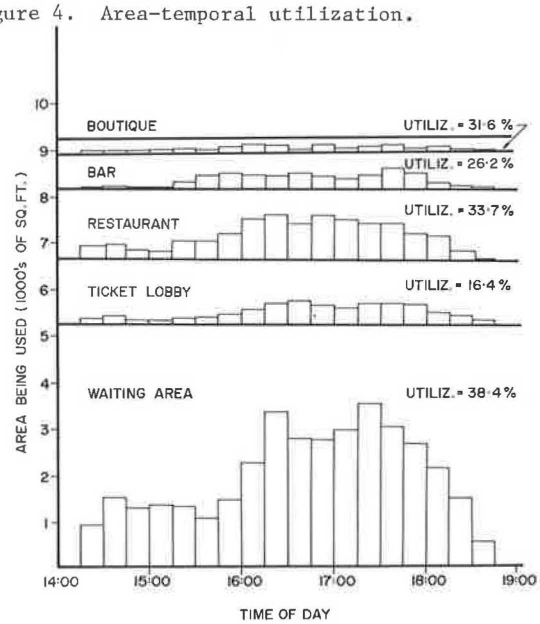
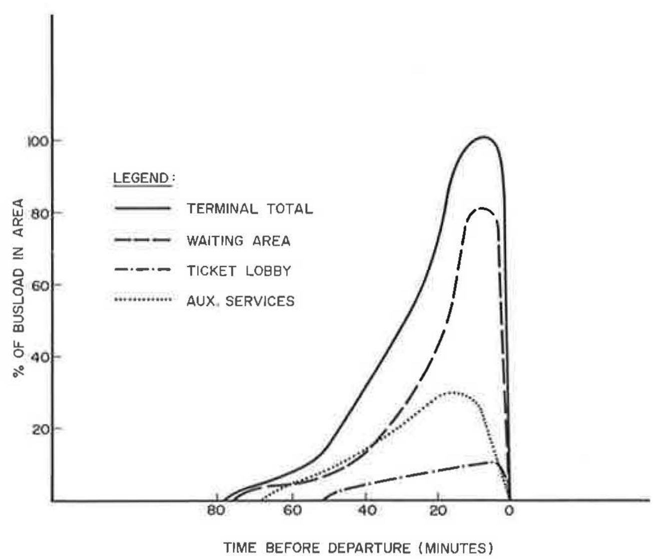


Figure 5. Outbound passenger area occupancy curves.



and to reduce distances walked. In this instance, exchanging bus bay #1 through #6 with bus bays #7 through #12 was found to reduce the average walking distance by almost 15%. This could also be an aid to designers when several layouts can be examined and compared.

Weighted Average Walking Distances

Weighted average walking distances have been calculated for the inbound and outbound passengers, visitors and the total. This was done by multiplying the elements of the origin-destination trip matrix with the corresponding elements of a matrix of walking distances between each O-D. The result is then a matrix of the number of person-m (person-ft) walked between each O-D pair. The sum of all elements, when divided by the number of people, produces this weighted average walking distance. Table 5 gives the results, with a "trip" being the movement between a pair of checkpoints. The relative proportion of terminal user types is also shown.

Length of Stay Distribution

The distribution of length of stay in the terminal was determined. For this particular setup, the outbound passenger had the longest average length of stay (mean = 22.79 minutes, standard deviation = 22.64 minutes), followed by visitors (mean = 19.31 minutes, standard deviation = 26.08 minutes) and inbound passengers (mean = 17.76 minutes, standard deviation = 22.46 minutes). The overall mean value was 20.77 minutes with a standard deviation of 22.18 minutes. The high degree of dispersion, as evident by the high standard deviation to mean ratio indicates a widely spread, skewed distribution of occupancy times as shown in Figure 7. Thus, the mean values alone are not truly indicative of the service being offered by the terminal. This distribution can be used in before and after type studies to determine the effects of terminal layout or operations changes on the length of stay.

Conclusions

1. At the present time, the Voyageur-Colonial terminal in Ottawa has no capacity related problems during a busy weekly rushhour. Examination of the holding area L/C ratios and processor/link V/C ratios showed the most heavily taxed essential element of the terminal was the main waiting area. It has a peak 15 minute average occupancy of 70% of its holding capacity.

According to expected traffic growth predictions, the terminal should experience capacity problems around the year 1985. This assumes all factors such as traffic mix, length of stay in the terminal etc. will remain relatively constant.

2. With scheduling revisions alone, it is estimated that the present levels of service can be preserved for roughly 10 years, thus extending the useful life of the terminal.

3. The time-stamping survey technique has proved to be applicable to a terminal of this size, layout and mode. The data base provided a wide variety of quantitative data concerning the terminal, in operation, with a minimum of disruption during its execution.

It is recommended that such surveys be incorporated in a continuing program to update the data base and to investigate the effects of seasonal variations on service provided to the user.

4. The planner can now produce a great deal of quantitative information specifically concerning the present performance of an entire terminal, in operation. Using this, future problem prediction and solution may be carried out, leading possibly to improved design techniques in the future.

Acknowledgments

The authors are grateful to Voyageur-Colonial for permission to use their terminal in this research.

Financial support for this research was derived from Grant A 8927, National Research Council of Canada.

References

1. L. Hoel and E. Roszner. Transit Station Planning and Design: State of the Art. Carnegie Mellon University, January, 1975.
2. The Ralph M. Parsons Co. The Apron and Terminal Building Planning Manual. U.S. Department of Transportation, Washington, D.C., July, 1975.
3. J. Fruin. Pedestrian Planning and Design. Metropolitan Association of Urban Designers and Environmental Planners Inc., New York, N.Y., 1971.
4. New York City Port Authority. Bus Bay Simulation Study. New York, N.Y., 1974.
5. William P. Johnson. Intercity Bus Terminal Performance Analysis. Master's Thesis, Department of Civil Engineering, Carleton University, Ottawa, Canada, March 1977, 179 p.
6. Barton-Aschman Associates Inc. Airport Travel Survey Manual. Washington, D.C., 1975.

Table 5. Weighted average walking distance.

Statistic	Terminal User			
	Outbound Passenger	Inbound Passenger	Visitor	Total
No. of card-carriers	616	290	293	1,199
Total distance walked (person-m)	45,894	16,504	15,862	78,260
Weighted ave. distance walked (m)	74.5	56.9	54.1	65.3
Standard dev. of distance walked (m)	15.1	19.6	17.5	17.0
Average No. of "trips" per person	2.4	1.6	1.7	2.1
Average "trip" length (m)	31.0	34.7	31.2	31.1

1 m = 3.28 ft.

Figure 6. Typical desire line map.

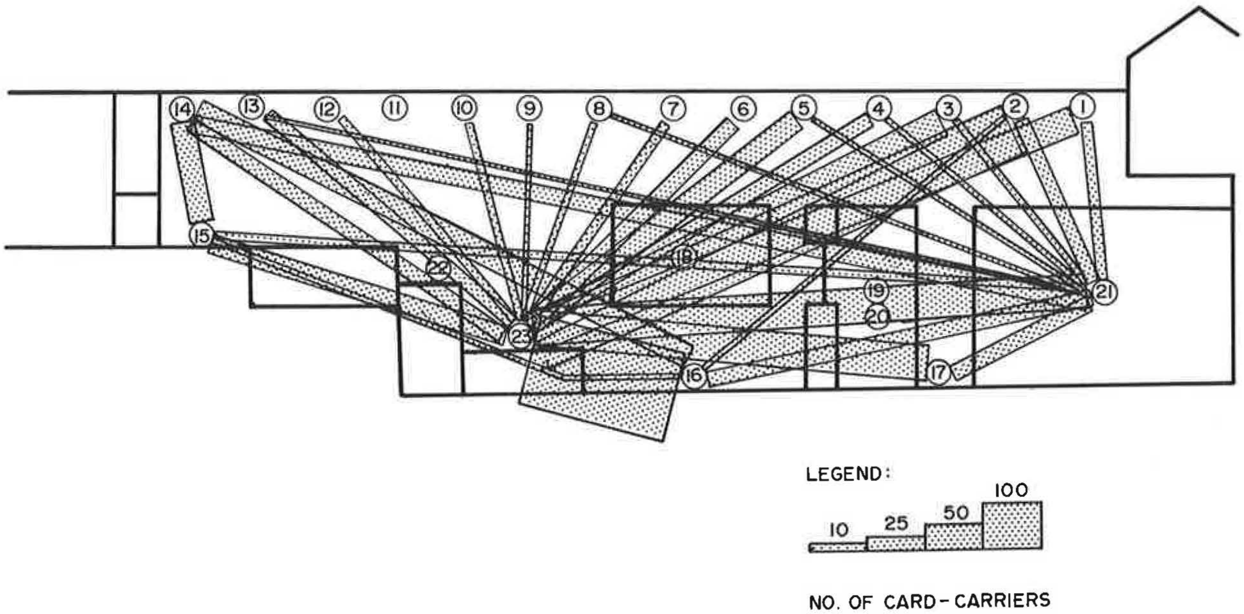
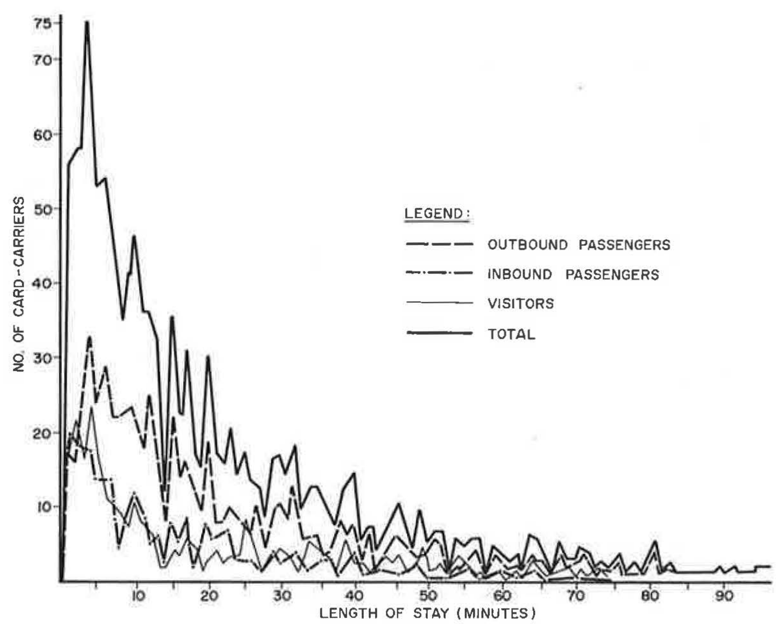


Figure 7. Terminal occupancy duration frequency distribution.



7. John P. Braaksma. "Time-Stamping: A New Way to Survey Pedestrian Traffic in Airport Terminals". Transportation Research Record 588, 1976.
8. Highway Research Board. Highway Capacity Manual. Special Report 87, National Academy of Sciences and National Research Council. Council Publication 1328, Washington, D.C., 1965.
9. R. De Neufville. Airport Systems Planning. The M.I.T. Press, Cambridge, Mass., 1976.

A MODEL FOR INVESTIGATING THE EFFECTS OF SERVICE FREQUENCY AND RELIABILITY ON BUS PASSENGER WAITING TIMES

Mark A. Turnquist, Northwestern University

A model of bus and passenger arrivals at a bus stop is proposed. Passengers are considered to be either random arrivals or non-random arrivals whose time of arrival is planned so as to insure a given probability of catching a selected bus. Buses are modeled as having a lognormal distribution of arrival times from day to day. The impacts on expected wait time of service frequency and reliability for both random and non-randomly arriving passengers are identified. The effects of frequency and reliability on the proportion of the user population who plan their arrival time are also explored through a small empirical study. The empirical results support the conceptual basis of the model, and indicate that it should be a useful tool for transit operators and planners.

The wait time experienced by transit users is one of the most important elements of the level of service provided by a transit system. For this reason, it is important to understand the effects on wait time of changes in basic service characteristics, such as frequency of service and schedule reliability. Such information is vital to the transit operator for evaluation of the costs and benefits of service changes.

The commonly used model which asserts that average wait is one-half the headway is true only under very special circumstances. This model requires that all passengers arrive at the bus stop at random and that headways be perfectly regular. These conditions are not generally met in the real world.

Several improvements to this simplistic model have been suggested by various authors (2,3,4,5,6,7), but with one exception have not formally treated non-random arrivals of passengers. The inclusion of this category of passenger arrivals is one major objective of the model presented here.

A second major objective of the current study is to incorporate the effect of service reliability on passenger wait time. By focusing clearly on this issue, a model can be formulated which will allow the transit operator to evaluate the impacts of operating changes designed to improve the reliability of service.

Previous Studies

Several previous studies have investigated the waiting times experienced by bus transit passengers. If passengers arrive at a bus stop at random, the average time they will have to wait before a bus comes is

$$E(w_r) = \frac{E(h)}{2}\left\{1 + \frac{V(h)}{[E(h)]^2}\right\} \quad (1)$$

where w_r = wait time for a randomly arriving passenger
h = headway between buses
$E(\cdot)$ = expected value of a random variable
$V(\cdot)$ = variance of a random variable

This expression has been derived by a number of authors including Welding (7), Holroyd and Scraggs (2), and Osuna and Newell (5). However, if buses tend to adhere to a fixed schedule and there are passengers who make the same trip frequently, it may be expected that some passengers will plan their arrival at the bus stop so as to be there just before the bus comes. In this case, we might expect average waiting time to be less than given by equation 1. O'Flaherty and Mangan (4) and Seddon and Day (6) have provided data and analysis which support this assertion. Both studies found average wait time to be considerably less than that predicted on the basis of randomly arriving passengers, and Seddon and Day used regression analysis to arrive at the following relationship:

$$E(w) = 1.71 + 0.57\, E(w_r) \quad (2)$$

where $E(w)$ = observed average waiting time.

More recently, Jolliffe and Hutchinson (3) proposed an improved model based upon considering passengers to be of three types: a proportion q, whose arrival time is causally coincident with the bus; a proportion (1-q)p, who arrive so as to minimize expected wait time; and a proportion (1-q)(1-p), who arrive at random. The proportion q, whose arrivals are coincident with the bus arrival, represent those people who run to the stop because they see the bus coming, and thus wait zero time.

The arrival time which minimizes expected waiting time is found in the following way. For times t (in one-minute increments) the waiting times to the next bus were found for each of several days of observations. These times were then averaged to obtain EW(t), the expected waiting time for a passenger arriving at t. By doing this for many values of t, a minimum of EW(t) was observed, and this waiting time was taken to be the average wait for the proportion (1-q)p who arrive so as to minimize expected wait. A simpler procedure based on a model of the distribution of bus arrival times was discussed in an appendix to the paper, but was not used in the empirical work.

The present paper represents a further modification of the Jolliffe and Hutchinson model which is different in three significant ways. First, arriving passengers are simply considered to be either random or non-random. This simplification is based on the premise that passengers who run and catch the bus because they see it coming are really either random arrivals or non-random ("planned") arrivals. These two groups make decisions as to arrival time which are clearly different. However, this is not true for passengers whose arrival is coincident with the bus. These people are really random or non-random arrivals whose original decision as to arrival time is modified slightly as a result of seeing the bus coming. They do not constitute a behaviorally distinct group, and thus should be included in the two larger groups which are distinguishable.

The second way in which this study differs from that of Jolliffe and Hutchinson is in the use of a theoretical model for the probability distribution of bus arrival times. The observed data on bus arrivals are used to estimate this distribution, and then the arrival times of non-random arrivals are derived from the estimated distribution. This eliminates considerable computation and also permits the exploration of decision rules other than the minimization of expected wait time.

The third major difference from the Jolliffe and Hutchinson work is that non-random, or planned, arrivals are assumed to minimize expected wait time subject to a constraint which results in a fixed (small) probability of missing the selected bus. The imposition of this constraint reflects more risk-averse behavior, and is more consistent with anticipated actions of people who must either reach their destination (e.g., work) on time, or make connections to other scheduled transit services.

The Model

Passengers are considered to be either random arrivals or non-random arrivals. The observed average wait time of all passengers as a whole is then

$$E(w) = aE(w_n) + (1-a)\, E(w_r) \tag{3}$$

where $E(w_n)$ = expected wait time for non-random arrivals
$E(w_r)$ = expected wait time for random arrivals
a = proportion of non-random arrivals, $0 \le a \le 1$

The expected wait time for random arrivals is given by equation 1. The expected wait time for non-random arrivals can be developed using the following model.

Suppose we observe bus arrival times at a given stop over several successive days. It is likely that the potential reduction in waiting time resulting from planning one's arrival time at the bus stop arises from the ability to predict the time of arrival of a given bus on different days rather than the regularity of headways on a single day. It is thus of interest to construct a probability distribution of arrival times of a given bus on different days.

Such a distribution should reflect several facts about the service. First, there is a definite earliest time of arrival, dictated by the distance from the terminal to the stop and the speed limit in effect; thus, the distribution should be truncated to the left. Second, there is a finite probability of the bus being very late, or even cancelled; so the distribution should have a long tail to the right. Finally, one of the major sources of lateness in the bus arrival times is increased dwell time at stops further up the line because of late arrival, and hence larger boarding volumes than expected. Thus, once a bus is initially delayed, subsequent delays become longer and longer. If we argue that delay at a stop is proportional to lateness arriving at that stop, we obtain a model of lateness as the result of a series of multiplicative effects.

A probability distribution consistent with all of these characteristics is the lognormal, and this distribution will be used here. If the arrival time of a bus, t, is distributed lognormal, its density function may be expressed in terms of two parameters, μ and σ, as follows:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[\frac{1}{\sigma}(\ell n\ t - u)\right]^2\right\} \tag{4}$$

The parameters μ and σ may be given intuitive meaning by noting that if t is lognormally distributed, $\ell n\ t$ is normally distributed. We then have

$$\mu = E(\ell n\ t) \tag{5a}$$

$$\sigma^2 = V(\ell n\ t) \tag{5b}$$

The mean and variance of t may be expressed in terms of μ and σ as

$$E(t) = e^{\mu + \sigma^2/2} \tag{6a}$$

$$V(t) = [E(t)]^2\ [e^{\sigma^2} - 1] \tag{6b}$$

It will be assumed that non-random arrivals choose their time of arrival so as to insure that the probability of missing the bus is no greater than some value X. This time is found by utilizing the relationship between the cumulative distribution function (CDF) for the lognormal and that of the standard normal random variable. It can be shown (see, for example, (1)) that the following relationship holds.

$$F(t) = G_z\left[\frac{1}{\sigma}(\ell n\ t - u)\right] \tag{7}$$

where $F(t)$ = CDF of a lognormal random variable with parameters μ and σ evaluated at t
$G_z(\cdot)$ = CDF of the standard normal random variable (i.e., N(0,1))

By setting the probability level, X, we can solve equation 7 to find the value of t, denoted t_a, at which the passenger should arrive. This is done by setting

$$\frac{1}{\sigma}(\ell n\ t_a - \mu) = G_z^{-1}(X)$$

or

$$t_a = \exp\left[\sigma G_z^{-1}(X) + \mu\right] \tag{8}$$

The exact probability level chosen is somewhat arbitrary, but an appropriate value is likely to be in the range .005 - .05. For the empirical analysis in this paper, the value X = .01 has been used. Should another value be deemed more appropriate by a user of this model, the substitution in equation 8 is straightforward.

Given t_a, we must compute the expected wait time for a passenger arriving at that time. The wait time is

$$w_n = \begin{cases} t_1 - t_a, & \text{if } t_a \le t_1 \\ t_2 - t_a, & \text{if } t_a > t_1 \end{cases} \qquad (9)$$

That is, if the passenger arrives before the desired bus arrives at t_1, he waits $t_1 - t_a$. If, however, he has missed the desired bus, he must wait for the succeeding bus which arrives at t_2. If the distributions of bus arrival times are assumed independent the expected wait is

$$E(w_n) = \int_0^{t_a}[E(t_2) - t_a]\, f(t_1)dt_1 + \int_{t_a}^{\infty}(t_1 - t_a)f(t_1)dt_1$$

$$= E(t_2)\ F(t_a) - t_a + \int_{t_a}^{\infty} t_1 f(t_1)dt_1$$

$$= E(t_2)\ F(t_a) - t_a$$

$$+ [\exp(\mu + \sigma^2/2)]\left[1 - G_z\left(\frac{\ln t_a - \mu}{\sigma} - \sigma\right)\right] \qquad (10)$$

We are thus able to predict the mean waiting time of non-random arrivals in terms of the bus service characteristics. One important element of the model expressed in equation 3 is still missing, however. It is to be expected that the proportion, a, of all passengers who are non-random arrivals will also be a function of the service characteristics. For example, as headways become longer the potential benefit from learning the schedule and planning one's arrival time becomes larger; hence, we would expect a to increase. Likewise, as service becomes more dependable, planning one's arrival time becomes easier, and again we would expect a to increase.

In order to explore the effects on a of changes in service characteristics, a small empirical study was undertaken of several services in Chicago with varying headways and reliability.

The Empirical Study

Four different services were observed for eight days each during the morning peak period in spring and summer of 1977. Data were collected on bus arrival times and passenger wait times. From observed bus arrival times, μ and σ were estimated for each service using equations 5a and 5b. Equation 8 then allowed the value of t_a to be determined, and the associated mean waiting time was found from equation 10. The expected wait for random arrivals was found by estimating the mean and variance of the headway distribution between successive buses, and applying equation 1.

Because passenger wait times were observed simultaneously with the bus arrivals, equation 3 may be rewritten to solve for a, the proportion of non-random arrivals:

$$a = \frac{E(w_r) - E(w)}{E(w_r) - E(w_n)} \qquad (11)$$

Summary statistics for the observed services are shown in Table 1. The values of a observed range from .49 to .74.

If the model suggested in this paper is to be a useful tool for predicting the effects on waiting time of changes in service characteristics, it is necessary that we be able to predict the value of a as a function of these service characteristics. As discussed above, it may be expected that increasing headways would lead to increasing values of a, as the potential gain from planning one's arrival time is larger. Also, as service becomes more reliable from day to day, the task of planning one's arrival to correspond to the arrival time of the bus becomes easier, and we may expect a to increase. One useful measure of the day-to-day reliability of the service is the coefficient of variation in the bus arrival time distribution. For the lognormal distribution, the coefficient of variation has a simple mathematical expression:

$$CV = \frac{\sqrt{V(t)}}{E(t)} = (e^{\sigma^2} - 1)^{\frac{1}{2}} \qquad (12)$$

The coefficient of variation is preferred to the standard deviation as a measure of reliability, as it also incorporates information regarding the mean of the distribution. In a skewed distribution like the lognormal, this is advantageous.

A simple model for a might be proposed as follows:

$$a = b_0 + b_1\ E(H) + b_2\ CV + e \qquad (13)$$

where $E(H)$ = expected headway
e = error term
b_0, b_1, b_2 = constants.

We would expect to find $b_1 > 0$ and $b_2 < 0$.

The data on E(H), CV and a from the four services observed are summarized in Table 2. These observations are clearly insufficient for reliable statistical inference; however, a regression using these four points results in the model

Table 1. Summary statistics for observed bus services.

Service Number	Mean Headway (minutes)	Headway Variance	μ	σ	E(w) (minutes)	$E(w_r)$ (minutes)	$E(w_n)$ (minutes)	a
1	10.09	.165	-.467	.834	3.02	5.06	0.90	.49
2	11.85	.478	-.267	.920	3.33	5.94	1.20	.55
3	7.75	2.68	.946	.387	2.38	4.05	1.79	.74
4	9.06	3.41	.041	.767	2.50	4.72	1.31	.65

$$\hat{a} = .602 + .023\ E(H) - .27\ CV \quad (14)$$

Thus, the data observed to date are at least consistent with a priori expectations about the model form. Additional data are being collected from other services in the Chicago area, but are not yet available at the time of this writing.

Table 2. Data on E(H), CV and a for services observed.

Service Number	E(H) (minutes)	CV	a
1	10.09	1.00	.49
2	11.85	1.33	.55
3	7.75	0.16	.74
4	9.06	0.80	.65

Application of the Model

The ability to predict the proportion, a, of non-random arrivals, as well as the expected waiting times for both random and non-random arrivals as functions of basic service characteristics allows the model to be easily applied in evaluation of proposed service changes. It has significant advantages over many previously available models as a result of incorporating non-randomly as well as randomly arriving passengers. As a result, the model proposed here is much more reflective of reality, and should provide a much better estimate of the impact of service changes. It also represents a significant improvement over previous models by providing a mechanism for evaluating the effects of reliability improvements.

Additional data are being collected to further verify the model empirically. These additional data will also provide an opportunity to test the sensitivity of the model results to the assumption of the probability level governing the arrival of non-random passengers, and to the assumption of a lognormal distribution for bus arrivals.

The major use of additional data, however, is to provide a firmer basis for estimating the proportion a, of non-random arrivals as a function of basic service parameters.

Conclusions

A model has been proposed to predict average passenger wait time at bus stops as a function of the headway distribution between successive buses and the arrival time distribution of a given bus from day to day. The model considers both random and non-random passenger arrivals. Its emphasis on the influence of service reliability on wait time sets it apart from previous models. For the first time, it provides the transit operator with the ability to predict the impact of changes in operations designed to improve reliability of service.

A very limited empirical study has produced results consistent with theoretical expectations, and has provided the motivation for further development and empirical verification of the model. This work is continuing, and further results are anticipated in the near future.

Acknowledgment

The work reported in this paper was supported by grant ENG76-09607 from the National Science Foundation. This support is gratefully acknowledged.

References

1. J. Benjamin and C.A. Cornell. Probability, Statistics and Decision for Civil Engineers. McGraw-Hill, New York, 1970.
2. E.M. Holroyd and D.A. Scraggs. Waiting Times for Buses in Central London. Traffic Engineering and Control, 8:3, 1966, pp. 158-160.
3. J.K. Jolliffe and T.P. Hutchinson. A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop. Transportation Science, 9:4, 1975, pp. 248-282.
4. C.A. O'Flaherty and D.O. Mangan. Bus Passenger Waiting Times in Central Areas. Traffic Engineering and Control, 11:9, 1970, pp. 419-421.
5. E.E. Osuna and G.F. Newell. Control Strategies for an Idealized Public Transportation System. Transportation Science, 6:1, 1972, pp. 52-72.
6. P.A. Seddon and M.P. Day. Bus Passenger Waiting Times in Greater Manchester. Traffic Engineering and Control, 15:9, 1974, pp. 442-445.
7. P.I. Welding. Time Series Analysis as Applied to Traffic Flow. Proceedings of the Second International Symposium on the Theory of Road Traffic Flow, OECD, London, 1963, pp. 60-72.

OPTIMAL URBAN BUS SIZE

Francis P.D. Navin, Assistant Professor of Civil Engineering, University of British Columbia

London's city transit has been shown by Webster and Oldfield to be operating at almost optimal occupancies of 38 and 18 passengers during the peak and off-peak respectively. The mathematical model developed duplicates these results as well as observations of conventional commuter routes and a long established crosstown line. The relationship combines the complex interaction of vehicle, patron and street into one general equation. The equation may be manipulated to give "mathematically optimal", passenger productivities, vehicle occupancy and fleet size.

London's famous double-decker buses have an average peak hour passenger occupancy of 38 and 18 in the off-peak. Optimal occupancies calculated by Webster and Oldfield (1) are 36 and 23 respectively. Many years of operating experience in North America has developed a balance between the transit service and demand, which in terms of average system productivity, is roughly 60 passengers per hour within cities such as Vancouver and Toronto. There is always the nagging thought that there ought to be more passengers in the vehicle to increase its productivity.

One of public transit's roles is to transport the most people with the fewest vehicles in a way that considers the unique combination of demand, street traffic, vehicle operation and vehicle occupancy. The variables included in this vehicle-patron-street system are listed and defined in Table 1. The transit operator directly controls the passengers carried by a bus P, through vehicle selection, total boarding-alighting time d and the acceleration a, deceleration b. Stop spacing is also at the operator's discretion and therefore the number boarding-alighting at any stop. The city traffic department sets the maximum travel speed V, and to a limited degree the amount of congestion. Finally the vageries of urban travel and city structure determine the number of travellers H and their average transit trip length.

The following develops a simple theoretical model to determine optimal bus size and vehicle productivity. The data also demonstrates that London's and indeed other large cities' buses do operate within the range of the optimal productivity.

Intuition suggests that if too few passengers are transported in each vehicle then many vehicles are needed. Similarly, if too many are carried travel times become extremely slow and again a large fleet is required.

Table 1. Bus-Patron-Road Variables

Variable	Units	Symbol	Purpose
Demand	$\frac{\text{trips}}{\text{h}}$	H	Maximum usage
Trip Length	km	L	Transit network & multiple rides
Dwell time	sec	d	Ease of boarding
Boardings	$\frac{\text{patron}}{\text{stop}}$	n	Stop collection efficiency
Accelerations	$\frac{\text{m}}{\text{sec}^2}$	a,b	Vehicle responsiveness
Passengers	$\frac{\text{patrons}}{\text{bus h}}$	P	Vehicle's income potential
Occupancy	$\frac{\text{patrons}}{\text{bus}}$	R	Vehicle's seat capacity
Congestion	% of h	C	Street efficiency
Speed	km/h	V	Street processing efficiency
Fleet Size	bus/h	N	Operator's cost

The number of vehicles required along a very long one-way transit route, to maintain an inter-vehicle time headway of h, is the total time to traverse the route, T, divided by the headway or

$$N = \frac{T}{h} \qquad (1)$$

If no delays exist along this very long route, then the minimum vehicle requirement is the driving time T_D divided by vehicle headway,

$$N_D = T_D/_h \qquad (2)$$

Equating the headways, and rearranging terms gives the estimate of the number of vehicles as,

$$N = \frac{T}{T_D} N_D \tag{3}$$

Assume that the demand for service over any segment of the long route is Ht person bus hours of travel and each bus may provide Pt person bus hours of travel. The minimum number of vehicles is then given by assuming that the number of vehicles required by the patrons demand for service (H/P) is the same as that given by no delays to travel. This idealized service permits no time to be lost due to boardings, traffic signals or road congestion.

The number of vehicles needed along a route, considering the road and passenger delays is then:

$$N = \left(\frac{T}{T-t}\right) \frac{H}{P} \tag{4}$$

where t is the time spent not moving and estimated by:

$$t = dP + \left(\frac{V}{a} + \frac{V}{b}\right) \frac{mP}{n} + C \tag{5}$$

where m = 1 if, at each stop the same number get on as get off and,
m = 2 if all ons and offs occur at different stops.

Combining the preceding equations the total fleet size is:

$$N = \frac{TH}{P\{T- \left[dP + \left(\frac{V}{a} + \frac{V}{b}\right) \frac{mP}{n} + C\right]\}} \tag{6}$$

Differentiated this equation with respect to P yields the "optimal" passengers per bus hour P* and fleet size N* as:

$$P^* = \frac{T-C}{2d + \frac{2m}{n}\left(\frac{V}{a} + \frac{V}{b}\right)} \tag{7}$$

$$N^* = \frac{4TH\left[d + \frac{m}{n}\left(\frac{V}{a} + \frac{V}{b}\right)\right]}{(T-C)^2} \tag{8}$$

The last equation implies the truly detrimental outcome of congestion on fleet size. Congestion represents a real loss to the transport system.

The determination of optimal size is possible for three transit operations. The first is downtown commuting. If all passengers must pass the maximum load point then H assumes this value and P* is the average passenger loading per bus and is the "optimal". The theory is for very long routes and does not consider buses making several trips past the maximum load point within the time T. The theory may be adjusted to accept multiple cycles, by dividing the resulting fleet size by a factor representing the number of times each vehicle passes the maximum load point for example. If P* sets the vehicle size then the operator must control variables such as loading time d, vehicle speed V, number boarding at each stop n, and time lost to other traffic C.

A more interesting bus route is the crosstown. Assume it transports a uniform number of passengers at all times over its entire length which is very long. The average loading, if the boarding and alighting is uniform, is the total driving time T_D proportioned by the actual driving time passengers spend on their journey, multiplied by the number of passengers carried during the time of interest. Mathematically it is:

$$R = \frac{P(L/V)}{T-t} \tag{9}$$

At any instant, for an optimal P*, the optimal uniform occupancy R* is:

$$R^* = \frac{L}{V} \left[\frac{1}{d + \frac{m}{n}\left(\frac{V}{a} + \frac{V}{b}\right)}\right] \tag{10}$$

Dial-a-bus may be considered as a set of vehicles in continuous service over a very long and erratic route. The average vehicle occupancies is the same as that for a crosstown bus.

The optimum equation of N* and P* is shown in Fig. 1 by the dashed line. The parameter characteristics are those for a Canadian city of one million people. The somewhat parabolic shape of the curves demonstrates the consequences of attempting to load too many persons into a vehicle. If the boarding and alighting time per passenger is, for example, 10 seconds, adding 10 more passengers over the maximum increases productivity by 20 percent while increasing the fleet size only slightly. Adding 10 more passengers increases the productivity by 15 percent but the fleet size by somewhat more than 12 percent. Any further increase in passenger productivity gives an even more rapidly increasing fleet size.

The most productive urban transit vehicles reported in Canada are those within the cities of Toronto, Montreal, and Vancouver where on an average they transport 60 passengers per hour. If the average total dwell time per passenger is 6 to 8 seconds and thirty percent of a vehicle's route service time is lost, from Fig. 1, then a value between 55 and 65 would appear optimal. Winnipeg, Edmonton and Regina have a passenger productivity of 50 and Calgary is down at 40. These values would reflect the vehicle size if all patrons behave as knowledgeable commuters and travel past a maximum load point. A lower limit on vehicle size is that of a crosstown bus and for the preceding example R* is 20 people in the vehicle.

Data from three very different transit operations are summarized in Table 2 together with estimates of the average vehicle occupancy and fleet size. The crosstown route in Vancouver and the assumed crosstown route in the City of London have an estimated occupancy, R*, well within ten percent of the observed value. London's "optimal" bus occupancy suggested by Webster and Oldfield is 36 and this is within fifteen percent of the result estimated from Equation 10. The vehicle occupancies for the two dial-a-bus experiments depart by twenty-five percent from the optimal riders. The fleet size in both cases represents the next largest whole number. The estimated fleet size is a very utopian view of transit service neglecting; mechanical breakdown, stocastic flows, and for dial-a-bus lost drivers, cancellations and no-show passengers.

The values of R* underestimates the number of seats needed in the vehicle since it represents an "average" occupancy or size and makes no allowance for irregular loading. Webster and Oldfield increased the average occupancy by 50 percent and equated this to the number of vehicle seats to accommodate all the irregularities of passenger loading. Navin (2) has shown that for many busy midday routes the mean occupancy (R*) plus two standard deviations (σ) of the occupancy can serve almost all routes without anyone being denied service. The observed coefficient of variation (σ/R*) has a typical value of 0.35 for crosstown type routes. If the optimal size of bus is set at R* + 20, this equals 1.7R* and for Vancouver during the off-peak is 36. During the off-peak most operating agencies have a policy of providing one seat per customer, therefore 36 seat buses are needed in Vancouver.

The peak hour vehicle productivity P* may represent the vehicle occupancy past the maximum load point. The Canadian transit operating practice is to provide seating for P*/1.5 people (3). The commuters in Vancouver travel approximately 7.7 km and for typical commute conditions the optimal productivity is 70 giving 47 seats per vehicle or thirty percent more than during the off peak. Buses in Vancouver have crush loads of 80 and 50 seats, a ratio of 1.6, slightly greater than the operating policy of vehicle loads being 1.5 times greater than seated capacity at the maximum load point. All P* patrons passing the maximum load represents the most extreme loading condition and gives the maximum vehicle size. If not all the patrons pass the maximum load point then the vehicle size may be reduced. The Vancouver experience indicates that sixty to seventy percent of the people do not pass the maximum load point. The average vehicle size may be reduced to serve 52 people, and seats may be provided for most.

Table 2. Observed and Estimated Transit Characteristics

	London U.K.	Vancouver Canada (3)	Dial-a-Bus (1)	Dial-a-Bus (2)
H	-	n/a	15	45
L	2.7	5.0	2.5	3.0
d	3	7	60	5
n	4	2	1	1
V	40	50	50	40
C	n/a	20	20	20
R	38	17.5	1.5	11
N	n/a	n/a	3	2
R*	41.7	17.0	2.0	9.8-13 (4)
N*	n/a	n/a	2.2	1.8

(1) Columbia Md., (2) Bay-Ridges Ontario,
(3) Crosstown route (4) Productivity estimated by P*

The mechanics of transit vehicles travelling along very long fixed routes can be manipulated to develop equations to give a theoretically "optimal" productivity and fleet size. The equations use variables that are both easy to estimate and representative of the observed vehicle-patron-street system. The theory also points out the detrimental impact of overloading vehicles or employing too large a vehicle that may become overloaded given the travel market characteristics and demand.

Comparisons of theory and experience lends credibility to the simple equations. Estimates for the average vehicle occupancy along fixed route transit are withing ten percent of the observations and even for small dial-a-bus services the estimates are within twenty percent. Webster and Oldfield's optimal vehicle occupancy for London buses comes within fifteen percent of those estimated by the simplified procedure presented. The optimal peak hour bus for Vancouver is estimated to have 47 seats and room for a total of 52 passengers. The most severe condition has all the commuters passing the maximum load point which suggests a total vehicle capacity of 70. The equations tend to overestimate as would be expected of theoretically "optimal" values.

"Optimal" productivity represents an idealized objective towards which transit operators may strive. The number of seats and passenger handling characteristics of the vehicle and operation can be designed to help accomplish the goal of transporting the most people in the fewest vehicles.

References

1. Webster, F.; Oldfield, R.; "A Theoretical Study of Bus and Car Travel in Central London", Traffic Flow and Transportation, Ed. Newell, G., American Elsevier Publishing Co. Inc. 1972.

2. Navin, F., et al; "Urban Bus Design", Transportation Research, Report No. 9. University of British Columbia, February 1975.

3. Archambault, C.; "Les objectives de service des transports collectifs dans les villes canadiennes de moyenne importance", Road and Transportation Association of Canada, Vancouver, 1977.

Figure 1. Transit bus productivity.