Technical Report Documentation Page

1. Report No. UMTA MA-06-0049-78-1

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle

TRANSIT SERVICE RELIABILITY

5. Report Date December 1978

6. Performing Organization Code

7. Author(s) Mark Abkowitz*, Howard Slavin*, Robert Waksman*, Larry Englisher**, Nigel Wilson**

8. Performing Organization Report No. DOT-TSC-UMTA-78-18

9. Performing Organization Name and Address

*U.S.Department of Transportation
Research and Special Programs Admin.
Transportation Systems Center
Cambridge MA 02142

**Multisystems Inc.+
1050 Mass. Ave.
Cambridge MA 02138

10. Work Unit No. (TRAIS) UM927/R9742

11. Contract or Grant No. DOT-TSC-1083

12. Sponsoring Agency Name and Address

U.S.Department of Transportation
Urban Mass Transportation Administration
Office of Planning,Management,and Demonstrations
Washington DC 20590

13. Type of Report and Period Covered

Final Report
February 1976 - December 1978

14. Sponsoring Agency Code

15. Supplementary Notes

\+ Under contract to:
U.S.Department of Transportation
Research and Special Programs Administration
Transportation Systems Center
Cambridge MA 02142

16. Abstract

This report presents a comprehensive overview of the subject of transit service reliability. The major subject areas include the impact of service reliability on travel behavior, service reliability from the operator's perspective, empirical measures of service reliability, causes of service reliability problems, techniques for improving service reliability, and recommendations for further research.

Several findings can be reported from this study. First, it appears that transit service reliability is likely to be a significant determinant of traveler mode and departure time choices. In most cases, travelers will trade off reliability with other service attributes in making their travel decisions. Secondly, it is likely that service reliability is crucial in influencing the costs of providing transit service. The transit operator, by improving service reliability, has an opportunity to reduce operating costs through reduced vehicle requirements while inducing additional transit ridership through improved service. Finally, current evaluation measures are not able to capture the variety of impacts of service reliability on travel behavior and operator costs. Several proposed measures are recommended for use in future evaluation studies.

Several causes of reliability problems are identified, some which appear to be inherent in the specific transit service concept, and others which are more environmental in nature. A review of previous and current analyses reveals inconclusive findings in determining the relative importance and magnitude of each cause.

A number of strategies are considered to improve service reliability of fixed route and demand responsive transit systems. These strategies are generally directed at alleviating the initial cause of unreliability or serve as a corrective measure when the reliability problem has already developed. Unfortunately, few of these strategies have been tested in an applied environment.

Given these existing gaps in our knowledge, demonstrations and a variety of further research studies are proposed to gain a better understanding of the impact of reliability on transit operations and travel behavior and to apply this knowledge to improve transit service quality and reduce costs.

17. Key Words

Transit Service Reliability, Traveler Behavior, Transit Operator, Service Attribute Variability Measures, Transit Innovations, Transit Demonstration Projects, Urban Transportation, Demand Responsive Transportation.

18. Distribution Statement

DOCUMENT IS AVAILABLE TO THE PUBLIC
THROUGH THE NATIONAL TECHNICAL
INFORMATION SERVICE, SPRINGFIELD,
VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	192	

Preface

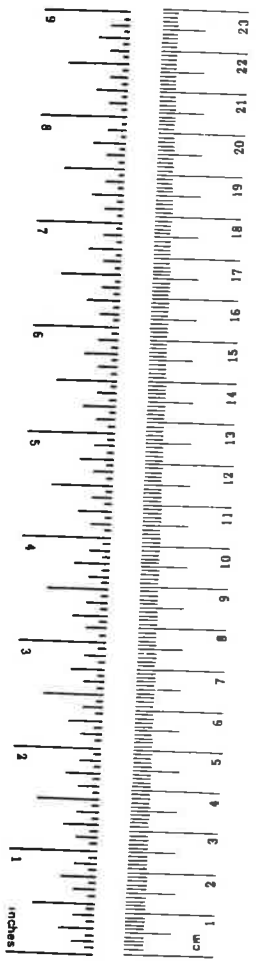
This report was prepared for the UMTA Service and Methods Demonstration Program by the Urban and Regional Research Division of the Office of Research and Analysis at the Transportation Systems Center, and by Multisystems, Inc., of Cambridge, Massachusetts. This study presents an overview of the subject of transit service reliability, emphasizing bus systems, and provides a conceptual approach to understanding and analyzing the traveler and operator impacts of service reliability. In addition, recommendations are made for future demonstrations and research studies. Chapters 1, 2, and 3 were authored primarily by TSC. Chapters 5 and 6 were authored primarily by Multisystems. Chapters 4 and 7 were jointly contributed by both TSC and Multisystems. TSC was responsible for the design and integration of the overall study.

The authors wish to acknowledge the assistance of a number of people who also contributed to this report. The authors are grateful for comments received from Professor Steven R. Lerman at the Massachusetts Institute of Technology and Professor Mark A. Turnquist at Northwestern University. Robin Blauer and Judith King of Raytheon Service Company contacted transit operators and obtained some of the information which forms the basis for the analysis in Chapter 3. Vera Ward at TSC, and Janet Burley at Raytheon are responsible for typing most of this report.

METRIC CONVERSION FACTORS

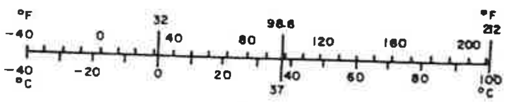
Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



CONTENTS

CONTENTS (cont'd)

CONTENTS (cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (cont'd)

LIST OF TABLES

READER'S GUIDE

This report presents an overview of the subject of transit service reliability (variability of transit service attributes), and provides a framework for a program of demonstrations and research studies which could be carried out under the Service and Methods Demonstration Program. Chapter 1 provides an Executive Summary of the report and contains an overview of the major issues and findings contained in the report.

Chapter 2 focuses on service reliability as viewed from the traveler's perspective. Attitudinal surveys of traveler behavior are reviewed to understand travelers' attitudes about reliability. A theoretical discussion on the effects of reliability on traveler behavior is also presented. Finally, the impacts of improved service reliability on travelers are examined.

Chapter 3 investigates service reliability from the transit operator's perspective. Based upon discussions with transit operators in several metropolitan areas, operators' views of transit reliability and transit reliability problems are examined. A theoretical discussion of operators' tradeoffs between average travel time and travel time variability is presented. The potential impacts of improved reliability on transit operators are discussed as well as their implications for travelers.

Chapter 4 addresses the selection of appropriate measures of reliability based on their conformance to the decision variables and impacts relevant to travelers and operators. This includes a review of measures that have been utilized in SMD demonstration projects. A set of measures is recommended which should prove useful in analyzing the effects of changes in service reliability.

Chapter 5 examines the basic causes of unreliability in transit service. This includes a discussion of both environmental causes and those inherent to basic transit service concepts.

Chapter 6 investigates strategies to overcome unreliability resulting from sources identified in Chapter 5. The effectiveness of each strategy is examined through available empirical data.

The findings of this study suggest that transit service reliability improvement strategies can have significant beneficial impacts on both travelers and transit operators. However, it must be recognized that there are major gaps in our knowledge of transit service reliability problems, improvements, and impacts. The concluding chapter of this report, Chapter 7, proposes a coordinated program of experimental demonstrations and research studies with the ultimate aim of developing and demonstrating transit service improvement strategies of widespread applicability.

1. EXECUTIVE SUMMARY

This report presents an overview of the subject of transit service reliability and a framework for a program of demonstrations and research studies which could be carried out under the Service and Methods Demonstration (SMD) Program. The reasons for the initiation of this program are several-fold. First, the improvement of transit reliability is a major objective of the SMD Program and is also intimately related to other SMD objectives such as improving travel time and increasing productivity. Second, there is cause to believe that service reliability may be crucial in influencing the demand for and the costs of transit and thus of major impact on ridership and revenues. Third, reliability and perceived reliability are considered to be of particular importance in the design of innovative transit services and methods.

The above rationale notwithstanding, there is little but the most fragmentary evidence on the role and importance of transit reliability in the decisions made by travelers and operators. In this report an attempt is made to provide an overview of what is known about transit reliability, develop a conceptual approach for thinking about service reliability, and propose a framework for future research studies to fill the gaps in our knowledge and hasten its application to the SMD Program. As the authors forsee them at this time, major applications to the SMD Program include: (1) the fostering of transit reliability improvements through the dissemination of new awareness of its role and importance; (2) the identification and development of strategies for the improvement of reliability that can be operationally demonstrated; and (3) the improvement of the measurement and evaluation of reliability in complex demonstration settings.

For the purpose of this report, transit reliability will be viewed in terms of the consistency of transit system performance, excluding from emphasis the reliability of vehicles and other transit equipment. Thus the authors propose a formal definition of service reliability as the invariability of service attributes which influence the decisions of travelers and transportation providers. Equipment reliability is incorporated into the assumptions that will be made about the availability of vehicles to operators and of transit services to potential travelers.

RELIABILITY AND TRAVEL BEHAVIOR

Studies of the preferences of actual and potential transit users have been conducted by transportation planners in efforts to improve transit service, to evaluate demonstrations, and to formulate models useful for forecasting travel demand. These studies point to the importance that travelers place on reliable transportation services.

Study findings indicate that travelers consider reliability to be one of the most important service attributes for a variety of trip purposes and modes. Both work trip and non-work trip survey respondents in Baltimore and Philadelphia (Paine, et al., 1) placed a higher importance on reliability-related service attributes than average travel time and travel cost attributes, generally considered to be the predominant determinants of travel choice decisions. Results of attitudinal surveys on specific modes point out the importance users place on travel reliability regardless of the mode concerned. Patrons of dial-a-bus, park-and-ride service, fixed route bus, and taxis rated reliability as an extremely important attribute in determining travel choice (Keck, 2; Winchester, 3). One can interpret the findings of the attitudinal studies as an indication that reliability is an important attribute in the travel decision process, and that direct reliability measures and inclusion of specific reliability variables in quantitative demand models are needed to more accurately assess the importance of reliability in the travel decision-making process.

With this in mind, a theoretical basis was established for understanding the relationship between service variability (variability of travel times) and travel behavior, particularly mode choice and departure time behavior were developed. In making both mode choice and trip departure time decisions, travelers appear to make tradeoffs between alternatives with shorter travel times and those with improved reliability (among other factors), based on their relative preferences.

Furthermore, in making work trip departure time decisions, travelers are hypothesized to account for the probability and degree of being late (or early) in addition to the mean and variability of travel times. The magnitude of the expected loss depends on the degree of risk aversion associated with the individual (Marfisi, et al., 4). For the work trip, it is likely that individuals are risk prone for low variabilities about the official work start time and risk averse for higher variabilities.

To date, there has been a paucity of empirical work directed at understanding the effects of reliability on mode choice and departure time decisions. There has been little success in modeling reliability using objective measures. Much of this problem is due to the difficulty associated with obtaining disaggregate data on the distribution of travel times and the correlation between the mean and variability of travel times (Lerman, et al., 5). Gardner and Reid (6) have developed a network reliability supply model, a more disaggregate approach which should provide better reliability data for input to disaggregate demand models. Attitudinal reliability variables have also been included in mode choice models (Prashker, 7; Spear, 8). Inclusion of the attitudinal reliability variable added statistical significance to the model. However, because the attitudinal variable was scaled and not metric, the explanatory power of this variable is difficult to determine. The only empirical trip departure time research was conducted by Cosslett (9), who found that reliability appears to affect the departure time decision slightly, with travelers tending to schedule departure times away from the peak, whenever possible. One of the drawbacks of this study was that transit was not considered as one of the modes under consideration by the traveler.

When a transit service reliability improvement takes place, the traveler's behavioral response depends on his/her preference for reductions of the mean and variability of travel time. A reliability improvement will decrease a traveler's disutility for transit trips. The reduced disutility for transit trips may lead the traveler to make more frequent trips at more convenient times, to more desirable locations, and on different modes, or may even enable the traveler to make entirely new trips, in all ways increasing his/her mobility.

The departure time response to a reliability improvement depends on the traveler's sensitivity to on-time arrival in addition to the mean and variability of travel times. Travelers who place a higher value on the variability of travel times and on-time arrival than mean travel time are likely to take the improvement by departing from their origin at their usual time and increasing the probability of on-time arrival. Travelers who place a higher value on mean travel time will depart later for their trip, and maintain the same probability of on-time arrival as they had before the improvement was made. Most travelers are likely to realize part of the improvement in later departure and part in increased probability of on-time arrival.

RELIABILITY AND OPERATOR BEHAVIOR (FIXED ROUTE BUS SYSTEMS)

The variability of transportation service attributes has a significant impact on transit operators as well as on travelers. To account for these effects, bus operators must allocate additional drivers and vehicles to maintain schedules and scheduled headways. Otherwise, the service provided may be considerably less predictable, user level of service may deteriorate, and the operator may lose riders and revenues.

Discussions with bus operators in ten United States cities revealed that reliability from the bus operator's perspective is treated primarily as an operational consideration (i.e., adherence to schedules). They consider the maintenance of stable route operating characteristics to be one of their major concerns, and they often take specific actions to achieve this end. It is within this framework that the operators attempt to provide reliable service to their patrons. Unfortunately, improvements in schedule adherence do not necessarily yield improved reliability to travelers. (First of all, regular travelers are often more concerned about predictability of arrival times at their bus stop and their destination rather than schedule adherence at terminals and maximum load points. Secondly, when the improvements are made on faulty schedules, they may do little to improve the overall predictability of service to users). In addition, current monitoring efforts are not able to capture the large spatial and temporal variability in wait and travel times, and in demand levels. Thus aggregate improvements in schedule adherence may not be evenly, efficiently, or profitably distributed to travelers. It should be noted, however, that budget and resource constraints restrict the degree to which the operator can treat reliability problems and make improvements.

When schedule adherence problems have been found to be chronic, and neither rerouting nor bus driver diligence can improve the situation, the operator typically adjusts the schedule to prevent the propagation of schedule adherence problems from run to run. Operators tend to avoid taking action when buses arrive at their destination on time on the average, but with a great degree of day-to-day variability. Unfortunately, bus users view this situation as being very unreliable because of the high degree of variability in day-to-day arrivals.

When chronic overcrowding conditions exist, the operator may schedule additional buses on a route if they are available. However, if the overcrowding is caused by unreliable service (e.g., bus bunching), the addition of buses may do little to improve the situation.

In day-to-day situations where bus users are experiencing extreme delays and overcrowding, only half of the bus operators contacted indicated that they would take immediate corrective action. The general solutions suggested (sending out an extra bus over either a portion of or the entire length of the problematic route, rescheduling future runs, and detouring buses) are aimed mostly at restoring schedule adherence on the system and relieving overcrowding. While generally such strategies will improve reliability experienced by users, in some instances these modifications could actually worsen users' reliability.

It is evident that a variety of operator decisions influence the reliability of bus service. Although many of these decisions implicitly involve tradeoffs between travel time and travel time components (such as wait times and ride times), and component variabilities experienced by bus users, the decisions are based mainly on maximizing schedule adherence and, as a result, often ignore travelers' preferences. Conceptually, improved operating strategies may result from considering the preferences of existing and potential fixed route bus travelers.

Improved bus reliability can impact the economics of bus operations by affecting both costs and revenues. The extent to which each of these is affected depends largely on operator behavior. An improvement in bus reliability provides the operator with the opportunity to increase the availability of drivers and vehicles. The operator can respond in the following ways:

(1) Do nothing at all.

(2) Take advantage of the increased availability of drivers and vehicles by reducing the driver and vehicle fleet to provide the existing service (or one with comparable level of service) to the existing service area.

(3) Take advantage of the increased availability of drivers and vehicles by increasing frequency and/or coverage of service with the existing driver and vehicle fleet.

The operator who does nothing at all automatically realizes improvements in schedule adherence which will, in turn, result in improvements in level of service for travelers and the possibility of longer actual layovers for drivers. By choosing either the second or third option, the operator is converting improved reliability benefits into totally different benefits (i.e., cost savings with option 2 and improvements in other service components with option 3). This may result in some travelers experiencing less reliability than before the reliability improvement was made. All three options should improve the operator's net revenue, but in different ways. With options 1 and 3, improved service should increase ridership, and therefore revenues, without increases in cost. With option 2, cost savings are realized without loss in ridership.

It is apparent that operator policy decisions are extremely complex in their dimensions and consequences. In designing strategies for improving bus reliability it may be particularly important to ensure that an appropriate balance is struck between quantifiable benefits to the operator and less quantifiable benefits to travelers.

MEASURES OF RELIABILITY

A review of measures used in previous Service and Methods evaluations indicates several weaknesses in the current approaches to measuring transit reliability. The wide variety of measures of lateness and deviations from published schedules, which were often not adjusted properly to reflect the impacts of the project, have led to inconsistencies which make comparison of reliability levels between projects virtually impossible. In addition, few of the current measures have captured the projects' effects on transit reliability as it is experienced by travelers. Most measures have focused on bus travel time and its relationship to scheduled travel and arrival times. Furthermore, the collection of disaggregate data across several time periods to measure the variability of transit service has been totally lacking in current and previous demonstrations.

In light of the limitations of many of these reliability measures, it is desirable to develop criteria for selecting a set of summary statistics which accurately measure reliability, and which are useful for comparison in before/after evaluations at a site or for comparison between sites. To this end, a set of relevant traveler and operator service attributes was selected and alternate ways to measure their variability were examined. Some of the attributes whose variability are of concern to travelers include total travel time, wait time, in-vehicle time, and seat availability. Some of the attributes whose variability are important to operators include schedule adherence, headways, and seat availability. The distributions of travel times, headways, and schedule adherence are likely to be significantly skewed to the right.

In defining summary statistics which accurately assess the impacts of the variability of these distributions, three separate criteria were identified (Martland, 10). First, a measure of the compactness of the distribution is desirable for this is an indication of the predictability of service. In addition to a measure reflecting compactness of the distribution, it is also useful to have a separate measurement of the likelihood of extremely long delays. For users and operators, such measures have significance in identifying the likelihood of a real system failure. A final criteria is the selection of a reliability measure which can be used in comparative analyses at a particular site or between sites. Both the user and operator will experience changes in the mean and variability of a service attribute when a service improvement is implemented. Thus, to evaluate the "true" impacts of a service change on reliability, it may be necessary to normalize compactness and skewness measures by including the mean value of the service attribute.

Based on these criteria, the following traveler time-related reliability measures are recommended:

1. mean time
2. coefficient of variation of time (for distributions which are significantly skewed, use standard deviation excluding values exceeding N minutes longer than the mean)*
3. % of observations taking N minutes longer than the mean.

For the operator-oriented measure of schedule adherence, the following parameters are recommended:

1. average difference between mean and scheduled arrival time
2. standard deviation (about mean arrival time and <u>not</u> about printed schedule times)
3. % of arrivals N minutes later than mean arrival time.

---

*It is suggested that N take on the value 2.32 standard deviations, as discussed further in Section 4.3.

Measures for the variability of headways and seat availability are also suggested (See Section 4.3).

The operator measures are defined in a manner which separates the effects of faulty schedules from the predictability of arrivals. Note that for both traveler and operator measures, it is also important to retain the mean as a separate measure. Otherwise, cases where the coefficient of variation increased, only because the percent reduction in the standard deviation was less than the percent reduction in the mean, might be misinterpreted as an increase in unreliability.

Both traveler and operator oriented measures should be taken across several days to uncover the day-to-day variability of service attributes. Moreover, both disaggregate demand and supply data are needed to assess the impact of reliability on travel. Thus, it is recommended that traveler data be collected and segmented by time-of-day, trip purpose, bus route, and the traveler's origin and destination bus stops. Operator data should be collected and segmented by time of day, route, and bus stop.

It seems logical that a series of mathematical relationships can be developed to relate measures of traveler and operator reliability. It is known that wait times are a function of the headway distribution, and that schedule adherence can be related to user travel time. Also, by measuring schedule adherence data with respect to mean arrival time, the effects of faulty schedules can be filtered out, and a more direct relationship between user travel time and service variability can be established.

Unfortunately, disaggregate demand modelling has only begun to address reliability as a component of travel mode choice, and the supply side implications of reliability for the transit operator have not been addressed. The measures selected here will hopefully contain enough information to continue to be useful at such future time when progress is made in these related areas.

CAUSES OF UNRELIABLE SERVICE

A number of factors have been shown to cause unreliable service in transit systems. They include "environmental" factors, related to the setting in which the system concept is applied and all exogenous random influences, and "inherent" factors specific to the service concept.

Environmental factors are common to all transit operations. Foremost among these factors are general traffic conditions, controlled intersections, variation in demand, and availability of drivers and vehicles.

General traffic conditions can cause reliability problems due to accidents, breakdowns, turning movements, lane speed changes, merging, stopping, parking, illegally parked cars, cross street traffic, weather, and driver characteristics. In addition, variation in traffic volume will cause variation in travel time, and therefore, unreliability. It is hypothesized that generally the variance of travel time increases with higher traffic densities.

One cause of variation in travel time is the interrupted flow produced by the presence of intersections. Signals can introduce significant variability by delaying a vehicle if it should happen to arrive at the intersection during the red phase.

Hourly, daily, and seasonal variations in transit user demand can also cause significant variation in level of service. Time spent boarding and alighting will produce fluctuations over time, resulting in service variability.

Finally, service quality may be affected by vehicle and driver availability. The impacts of vehicles and drivers on reliability are determined by vehicle design, maintenance policies, employee policies, benefits and morale, and the extent of reserve vehicles and drivers.

The basic inherent factor in causing fixed route bus unreliability is the instability of the headway distribution. The instability of headways in fixed route bus systems tends to become more pronounced further downstream. Exogenous factors, such as delays between stops due to traffic conditions, or delays at stops due to unusual loading conditions, trigger an initial deviation from schedule or scheduled headways. The inherent instability lies in the fact that any delay in arrivals results in an increased dwell time at that stop due to the increased passenger load that has accumulated over the previous long headway. Thus, late buses get later, and early buses (because of reduced demand) get earlier, eventually resulting in bunching and imbalanced loading (an inefficient utilization of resources).

An important step in determining improvement strategies is identifying the primary causes of unreliable service. A number of studies have addressed this issue. It appears that variation in run times is a major cause of unreliable service, while variation in boarding times does not significantly affect the reliability of service (Bly and Jackson, 11; Shields, 12). Variation in run times are directly related to the number of signalized intersections, route length (Sterman, 13), and traffic conditions. Variability in wait times (also significant in traveler trip time variability) is closely tied to travel time variation and headway irregularity (Kulash, 14). Because of the difference in study approaches and classification of causes, findings regarding relative importance of causes have been somewhat inconclusive.

Dial-a-ride operations also introduce several unique and inherent sources of variance that can cause unreliability. Among these are dynamic routing, the sensitivity of dial-a-ride service quality to demand, passenger cancellations, and passenger no-shows.

Vehicle operations in dial-a-ride are determined by a formal or informal set of rules which are applied to assign riders to vehicles. For a specific request, the actual wait time and ride time will depend on the current positions and activities of all vehicles, the set of competing users currently active on the system and the set of new users who will request service in the immediate future. Clearly all of these elements are highly stochastic, as are the resulting wait and ride times provided. Thus level of service will vary for similar trips, over the day, and between days even if the overall level of demand and the number of vehicles are constant.

While variation in demand has already been identified as an environmental factor affecting level of service in all transit operations, for dial-a-ride this sensitivity merits special attention. In fixed route scheduled service, the addition of a passenger affects overall service only through extra dwell time, which is virtually insignificant. In dial-a-ride, a new passenger will typically involve added vehicle-miles as well as a much longer dwell time. The net result is that dial-a-ride service quality is extremely sensitive to level of demand.

Because of the sensitivity of dial-a-ride service to demand levels, the cancellation of a trip which has been scheduled can create significant service quality changes for other passengers. In general, a cancellation will result in early arrival at subsequent stops on the affected vehicle's tour. These arrivals may be before the advertised time for some passengers, thus increasing the probability that they will not be ready and hence become no-shows. No-shows can have two effects downstream: (1) a longer than usual dwell time at the no-show stop can delay the vehicle at subsequent stops and, (2) the resulting cancellation of the associated delivery stop can result in the vehicle being early beyond this point in its tour. It is possible to imagine instabilities being created because of these no-show properties, particularly where some large initial perturbation is introduced such as a vehicle breakdown or major traffic delay.

For dial-a-ride, very little empirical work on causes of service unreliability has been performed. Further research is necessary to determine effects of no-shows, cancellations, demand variation, and dispatching procedures and algorithms on reliability.

STRATEGIES TO IMPROVE RELIABILITY

There are three basic techniques for improving reliability: priority, control, and operational. Priority methods involve special treatment of transit vehicles apart from other motor vehicles; control strategies concern dynamic changes in bus operations from the original schedule; operational methods deal with route, schedule, and resource allocation changes. Reliability improvement strategies can also be classified by the stage in the process of unreliability development during which they are applied. The two general categories are preventive and corrective/restorative. Preventive strategies are directed at reducing the probability of initial deviations; corrective/restorative strategies attempt to correct deviations once they have occurred, and restore operations to the initial schedule.

There are a few reliability strategies which are applicable to transit operations in general. The use of reserve vehicles and drivers either set aside for emergency situations or employed in day-to-day rescheduling, can be directed at reducing delays caused by vehicle breakdowns. Modification of schedules, routes, layover time, vehicle allocation and promised times of pick-up and drop-off can be made to counteract the effects that predictable variations in demand and traffic congestion over time will have on the variability of service components.

Several strategies can be applied specifically to fixed route bus operations. One set of strategies directed at controlling the effects of bus demand variation includes express service, exact fare/passes, and improved access onto the vehicle. Express service can reduce the magnification of deviations (caused by stopping) and decrease wait times downstream. Exact fare/passes and wider, more accessible vehicle entrances can reduce the per passenger boarding time and decrease the overall effects of increased boarding volumes.

Another set of fixed route bus strategies is directed at controlling unreliability caused by variation in traffic volumes and congestion. Providing exclusive lanes for buses may reduce the variance of link travel times as well as the mean. However, unless priority for buses is provided for entering the exclusive lane and for passing through intersections, a potentially large part of travel time variance will remain. The strategy of providing priority to buses at signals (pre-emption) may reduce the mean and variance of stopped time at signals, and hopefully alleviate the major effects of traffic on bus travel times and headway variation.

Once deviations and disruptions have occurred on a bus route, there are a number of corrective strategies directed at restoring the system to regular service. They are speed modification, control point holding strategies, adding a reserve vehicle, multiple route vehicle pool, turnback, skip stops, and pass first bus. Instructing drivers to change speeds according to deviations from the desired headway or schedule corrects for any initial deviation. Control point holding strategies are designed to stabilize the variation in headways by holding vehicles in mid-route at specified control points. In employing a strategy of adding a reserve vehicle, a trade-off exists between maintaining a regular lower frequency service (with reserve vehicles) and providing a less regular higher frequency service (utilizing all available vehicles). A multiple route vehicle pool would involve orienting all radial routes to a central point and exchanging vehicles between routes to allow shorter layover times. Turnback, turning a vehicle around before it reaches the terminus, may correct headway deviations upstream at some inconvenience to passengers who must deboard before reaching their destination in order for the vehicles to turn back. Skip stop, passing those stops where no one on-board wishes to alight, might be a difficult strategy to implement operationally. Pass first bus, changing the order of two buses which are bunched, can restore balance to the system by allowing the inherent instability to work in a correcting direction. The selective use of providing priority to buses at signals could be viewed as a corrective measure as well as a preventive strategy. It is also possible to reduce the overall effect of the reliability problem by scheduling imbalanced slack in schedules, where the more unreliable routes (or segments of routes) have added slack time built into the schedule.

Another set of strategies directed at improving fixed route service reliability includes major routing and scheduling changes. Because reliability tends to deteriorate as a bus proceeds further away from the dispatching point, the implementation of shorter routes may improve reliability. However, the effects of extra transfers may counteract the potential benefits of more reliable service from the user perspective. An alternate strategy may be to examine the tradeoffs of fewer routes (lower route density) and higher frequency service as opposed to more routes (higher route density) and lower frequency service (Turnquist and Bowman, 15).

A number of analyses have been conducted which examine the effectiveness of priority, control, and operational strategies. To date, the impacts of many bus service improvement strategies have generally been measured in terms of mean travel time changes; variability of service components have been largely ignored.

Bus priority strategies have been demonstrated in the United Kingdom and Ireland (Chapman, 16), resulting in a reduction in mean travel time; the evaluation of their effects on reliability was inconclusive.

Control strategies have been analyzed in empirical and simulation studies in London (Jackson and Stone, 17), Dublin, and a hypothetical crosstown bus route (Bly and Jackson, 11). Roadside control (London) and radio-telephone control (London, Dublin) offered reductions in mean wait time, although analyses of reductions in travel time variance and of the costs of implementation were not always conducted. Automatic vehicle monitoring may also be a viable strategy for improving reliability (London); however, the cost-effectiveness of this strategy is questionable. The studies also noted the importance of accurate and efficient schedules. Much predictable variation in bus travel time and travel time variability over time can be built into schedules by varying scheduled travel times and layover times to reflect these predictable variations. Thus, layover times and scheduling are important strategies for counteracting the effects of the inherent instability of fixed route bus systems.

Holding strategies at a few control points (bus stops) proved to be particularly useful in controlling headway variation. Various procedures for determining control policies have been developed. These include an approximate-optimal dispatching strategy for the control point (Barnett, 18) and a heuristic algorithm developed by Giraud and Henry (Reference 19) for application in Toulouse, France. Barnett's procedure was applied to a model of the northbound Red Line in Boston at the busiest stop on the line, resulting in a mean wait time reduction of nearly 10%. Not enough data was collected to draw definite conclusions using the Toulouse algorithm.

Little work has been performed analyzing the effectiveness of operational policy strategies. A Transportation and Road Research Laboratory (TRRL) study (Bly and Jackson, 20) utilizing analytic and simulation models investigated two strategies: (1) providing a reserve pool of drivers and buses to replace buses that are cancelled; and (2) rescheduling buses on a day-to-day basis, utilizing all available vehicles and providing equal headways in a schedule which would change daily. The analysis showed that both the strategy of rescheduling buses and the strategy of reserve bus pools provided similar reductions in passenger waiting time (compared to original schedule, 10% of buses missing) for service headways below about 12 minutes. At longer headways, reserve buses provided still larger waiting time reductions while rescheduling caused waiting times to increase. Due to the difficulty of constant rescheduling, the reserve bus pool strategy may be preferable at all headways.

Since dial-a-ride has no specified route or headway, it does not experience the same error magnification process as does fixed route service. However, dial-a-ride has its own inherent unreliability due to the effects of demand variation, dynamic routing and scheduling, and cancellations and no-shows. Strategies directed at improving dial-a-ride reliability include guaranteeing advance requests only, computer dispatching, restructuring routes and schedules for cancelled requests, setting maximum dwell time policies, transition from DRT to scheduled service, signal priority, speed modification, and reserve vehicles and drivers. Guaranteeing service to those who make advanced requests and accepting other requests only when conditions permit would allow for service norms to be established and maintained. Computer dispatching of dynamically-routed transit vehicles may reduce variance of waiting times through better assignment and estimation of arrival times. Restructuring routes and schedules due to a cancelled request would allow a vehicle to utilize the resulting extra time more productively. Setting maximum dwell time policies will alleviate long unexpected dwell times that are associated with an occurrence of a no-show. Transition from demand-responsive transit to more reliable forms of transit (e.g., subscription, route deviation) may be desirable where trip densities permit and when the market served is particularly sensitive to reliability. Signal priority, speed modification, and reserve vehicle and driver strategies are similar to the fixed-route application of these strategies.

Limited analyses performed to date suggest considerable promise for a variety of the dial-a-ride strategies. Computer dispatching is currently being used in Rochester, NY, and initial analysis has produced very tentative results which suggest that improved reliability may be achievable through computer dispatching. Also, surveys conducted in Rochester found advance request users placed different weights on service components than immediate request users. This finding would support application of a strategy in which a multi-class service is offered with different service norms (and fares) for each class. Maximum dwell time policies used in Gaithersburg, Maryland have resulted in a significant reduction in very long dwell times.

PROPOSED DEMONSTRATIONS AND FUTURE STUDIES

To supplement our knowledge about reliability and to hasten the implementation of strategies for improving transit service reliability, a set of research activities are proposed. They include: (1) analytical studies to examine, in more detail, the importance of reliability and its effects on operator and traveler behavior, and (2) implementation and evaluation of Service and Methods Demonstration projects explicitly directed at improving transit service reliability. It is likely that several of the analytical studies can be conducted using data collected from the proposed SMD projects.

Research projects and related studies include:

(1) Development of causal empirical models of the determinants of transit service reliability. Utilizing simulation models and the empirical data collected before and during the proposed demonstration projects, the following impacts on reliability could be explored and incorporated in the empirical models: (1) demand variation in time and space, and in mode and path, (2) traffic networks and volumes, (3) vehicle availability, (4) traffic management, and (5) weather.

(2) Development of reliability measures. The reliability measures developed in this report should be empirically tested and, where necessary,

modified as the demonstration projects are implemented and evaluated. Measures should be examined for appropriateness as tools for evaluating service attributes from both the traveler's and transit operator's perspective, and for feasibility and cost-effectiveness of collection.

(3) Development of travel behavior models. Empirical data collected during the demonstration projects and results obtained in deriving causal reliability relationships could be used to develop these models. The models should explain the role of reliability in influencing traveler decisions of travel frequency, destination choice, mode choice, choice of route, and departure time. In developing new travel behavior models, consideration must be given to identification of variables that can accurately measure the effects of transit reliability. The incorporation of reliability variables into travel behavior models also requires a more disaggregate view of transit supply than is generally considered in current models. It is hoped that an important result obtained from the models would be the determination of the traveler's value of the variability of time relative to his value of average travel time in different travel contexts. These relationships would explicitly define the basic traveler tradeoffs between travel time and its variability, and would provide planners with insight into the effectiveness of various improvement strategies.

(4) Development of the relationships between unreliability and the costs of transit operations. Using empirical data collected before and during the proposed demonstrations, the effects of unreliability on vehicle and driver productivity and costs of operation should be studied in detail. Operator behavior in response to changing levels of unreliability induced by the demonstration project could also be examined. A desired outcome of this study would be the development of improved operator policies.

In some or all of the demonstration projects, simulation models could be used to (1) assist in the final design of the innovative project features, (2) to fine tune these features during the project and (3) to assist in understanding the outcome of the project. As discussed earlier, simulation might also be used as a tool in developing empirical models of the determinants of transit unreliability.

Proposed demonstration projects include:

(1) Control point holding strategies. Control points would be located along two or three of the most unreliable routes in a fixed route bus system. Buses would be equipped with radio telephone equipment and all drivers would maintain communication with a bus dispatch center. Each bus would report to the dispatcher when reaching a control point, relaying information on time of arrival, volume on the bus and the number of people waiting to board at the control point. The dispatcher would make a decision of whether to hold the bus and the duration of time to hold the bus. The effects of this strategy would be measured in terms of benefits to the users and operators and the tradeoffs between benefits to people waiting for a vehicle and those delayed by a holding strategy.

Depending upon the success of employing a holding strategy, the demonstration could be extended to include other control strategies which would involve radio-telephone communication and a dispatching process. This would include coordination of the following strategies:

1. Transfering passengers and turning particular buses.
2. Injecting standby buses.
3. Running buses out of service until reaching particular points on the route.
4. Altering bus routes on a temporary basis.

(2) Priority Schemes. Priority schemes, particularly exclusive right-of-way for buses, have been significant elements of the SMD Program. These schemes have been predominantly directed at reducing mean transit travel times. While a significant benefit of all of the proposed priority scheme demonstrations should be a reduction in mean travel time, the evaluation of the new demonstration would be to also focus on improvements in reliability and their effect on

travelers and transit operators. The following demonstrations are proposed:

1. Signal pre-emption only (selected and general use)
2. Exclusive lanes only
3. Exclusive lanes with signal preemption (selected and general use)
4. Exclusive lanes with separate signal phases
5. Exclusive lanes with speed modification or other control strategies
6. Retiming traffic signals to facilitate bus movement (signal progression).

(3) Restructuring routes. Such a demonstration would involve dividing long routes into two shorter routes, or into express and local routes. While the objective of this demonstration would be to determine the effects of route redesign on reliability, an important element will be to examine the effects of redesign on other aspects of service. For example, the inconvenience of transfers and the predominance of particular travel patterns must be considered in route redesigns.

(4) Rescheduling. A rescheduling demonstration would involve increasing recovery time between runs and increasing the number of spare drivers and vehicles available for various contingencies. Such a demonstration would determine the actual effects on transit users and operators of making some simple changes on specific routes and systems. The key aspect of this demonstration would be to identify and measure all of the level of service tradeoffs experienced by transit users.

(5) Traffic engineering strategies. There are a wide range of general improvements of roadway facilities which would improve transit reliability. They include:

1. improving the physical condition of curb lanes,
2. providing bus stops of adequate length,
3. providing adequate turning radii at corners,
4. reorganizing parking restrictions,
5. enforcing parking restrictions,

6. channelization where bus turning movements are heavy (even if few other vehicles make the turn),
7. improving the location/phasing/control of pedestrian movements, and
8. improving bus stop location.

The effectiveness of each strategy has yet to be determined. All are included in the Transportation Systems Management (TSM) Element of Transportation Improvement Plans (TIP). Study of a sample corridor should help analyze the impact of TSM strategies.

(6) Reconfiguration of demand-responsive transit (DRT) service. The development of new hybrid service concepts can offer many of the advantages of dial-a-ride but which overcome some of its unreliability problems. Such modifications to the original many-to-many dial-a-ride service concept include scheduled fixed stops with deviation service, many-to-few service, integrated service (zonal dial-a-ride and fixed route), and restricting doorstep service to special needs groups/specified time periods. An existing dial-a-ride system could be selected to adopt these modified policies. The aim of the project would be to test the effectiveness of these polices in improving reliability and to determine if the resulting perceived level of service is acceptable to current and potential users.

(7) DRT multiple service options. Algorithms to assign passengers based on specific service level preferences should be developed and tested. A demonstration could assess the ability of the computer dispatch system to improve total passenger satisfaction with service and possibly expand the market to those who previously found the service too unreliable in one respect or another. As part of the demonstration, a set of pricing schemes could be implemented for the different service options, to determine their effects on ridership, reliability, other level of service components, and net revenues.

(8) Restricted DRT service. This demonstration would test the acceptability and efficiency of service which strives for reliable service levels for regular users and puts other users on a stand-by basis at a lower level of service.

(9) DRT dwell time/re-routing policies. This demonstration would experiment with alternate policies to deal with cancellations and no-shows on an existing dial-a-ride system. The aim is to develop policies which will minimize the effects of such perturbations on other users. The effects on service reliability could be analyzed for each policy including any significant change in the occurrence of no-shows and cancellations which can be related to the change in policy.

CONCLUDING REMARKS

The findings of this study suggest that transit service reliability improvement strategies can have significant beneficial impacts on both travelers and transit operators. The limited number of studies to date indicate that a variety of strategies may be capable of yielding significant reliability improvements.

However, major gaps in our knowledge of transit service reliability problems, improvements, and impacts still remain. The authors hope that the ideas and suggestions contained in this report will prove useful in enhancing our understanding of the effects of service reliability on various elements of travel, and will promote the design and implementation of more effective policies directed at improving transit travel and operator efficiency.

REFERENCES

1. Paine, F.T., A.N. Nash, S.J. Hille and G.A. Brunner. _Consumer Conceived Attributes of Transportation_. College Park, MD, Department of Business Administration, Univ. of Maryland, 1967.

2. Keck, Carol A. and G.S. Cohen. "Characteristics and Attitudes of Dial-A-Bus and Park-and-Ride Users in New York State." _Transportation Research Record 590_. (1976), pp. 21-25.

3. Winchester, S.A. "Haddonfield Dial-a-Ride Demonstration, Third Household Survey." Mitre Corp., McLean, VA, 1976.

4. Marfisi, et al. _Economic Criteria for Federal Aviation Agency Facility and Equipment Expenditures_. WP-210-U4-108, U.S. DOT, July 1975. *

5. Lerman, Steven, et al. _A Method for Estimating the Patronage On Demand-Responsive Transit Systems_. Cambridge Systematics, DOT-TJT-77, 1977.

6. Gardner, Wendall and F. Reid. _Transit Reliability Data for Disaggregated Travel Demand Models_. Working Paper No. 7603, Univ. of California, 1977.

7. Prashker, Joseph N., "Development of a Reliability of Travel Modes Variable for Mode-Choice Models." (PhD Dissertation, Northwestern Univ., Dept of Civil Engineering, Evanston, IL, June 1977).

8. Spear, Bruce D. "Generalized Attribute Variable for Models of Mode Choice Behavior." _Transportation Research Record 592_. (1976), pp. 6-11.

9. Cosslett, Stephen. "The Trip Timing Decision for Travel to Work by Automobile." _Demand Model Estimation and Validation_. The Urban Travel Demand Forecasting Project, Phase I, Final Report, Vol. V, Institute for Transportation Studies, Univ. of California, June 1977.

10. Martland, Carl D. _Rail Trip Time Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data_. Volume 2, MIT, R72-37, 1972.

*Material on file: U.S.Department of Transportation, Transportation Systems Center, Cambridge, MA.

11. Bly, P.H. and R.L. Jackson. *Evaluation of Bus Control Strategies by Simulation*. TRRL Laboratory Report 637, Transport and Road Research Laboratory, Dept. of Environment, Crowthorne (Great Britain), 1974.

12. Shields, L.R. *Reliability Study of a Central London Bus Route*. London Transport, Great Britain, 1974.

13. Sterman, Brian P. "Factors Affecting Reliability of Urban Bus Services." (M.S. Thesis, Northwestern University, IL, 1974).

15. Turnquist, Mark A., and L. A. Bowman. "The Effects of Network Structure on Reliability of Transit Service." Northwestern University, presented at International Symposium on Travel Supply Models, Montreal, 1977.

16. Chapman,R.A. "Bus Reliability-Definition and Measurement." (Working paper No. 18). Transport Operations and Research Group, Univ. of Newcastle Upon Tyne (Great Britain), 1976.

17. Jackson, R.L. and D. Stone. *Experiment in Bus Service Control Using an Interactive Model of a Typical Urban Bus Route*. Univ. of Newcastle Upon Tyne, (Great Britain), 1976.

18. Barnett, Arnold. "On Controlling Randomness in Transit Operations." *Transportation Science*. Vol. 8, No. 2, May 1974.

19. Giraud, Alain and J.J. Henry. *On-Line Control of a Bus Route-Studies and Experiments in Toulouse*. Third International Symposium on Control in Transportation Systems, Columbus OH, 1976.

20. Bly, P.H. and R.L. Jackson. "Scheduling for Depleted Services." Univ. of Newcastle Upon Tyne (Great Britain), 1975.

2. RELIABILITY AND TRAVEL BEHAVIOR

Improvements in the reliability of transit service or changes in the service reliability of other travel alternatives are likely to influence a variety of travel choices made by travelers. These choices may include those that govern trip frequency, destination and mode choice, routing, the time of day of travel, departure time, and trip chaining, and combinations of these choices.

The value travelers place on reliability relative to other service attributes (as well as the magnitude of changes in service reliability) will determine the degree to which their travel decisions are influenced. This chapter provides a framework for assessing the potential impacts of transit service reliability on travel behavior.

Section 2.1 provides a review of attitudinal studies which have been conducted concerning the determinants of traveler behavior. These studies suggest that attributes which are related to service reliability may be very important in the travel decision-making process.

Section 2.2 presents a theoretical discussion of traveler behavior as it is influenced by the reliability of transit service. The effect of reliability on travel choice behavior is examined, with a particular emphasis on the trip departure time decision (given that transit has been selected as the mode choice). Knowledge of transit schedules, schedule adherence, and the daily predictability of bus arrivals appear to have a considerable effect on the distribution of passenger arrivals at transit stops. This suggests that transit service reliability probably has an impact on trip departure time decisions and provides a basis for a theory that is proposed regarding the effects of service reliability on traveler departure times. Because of the inherent difference in service characteristics between dial-a-ride and more conventional forms of transit, a separate discussion of the effect of transit service reliability on departure time on dial-a-ride is presented.

Section 2.3 concludes the chapter with a discussion of the potential traveler benefits of improved transit reliability. This includes an examination of potential savings in mean travel times and variability in travel times stemming from strategies directed at improving transit reliability.

2.1 TRAVELERS' ATTITUDES ABOUT RELIABILITY

Studies of the preferences of actual and potential transit users have been conducted by transportation planners in efforts to improve transit service, to evaluate demonstrations, and to formulate models for the purpose of predicting patronage. These studies indicate the importance of reliable transportation services to the traveler. In this section the findings on travelers' values are discussed with particular emphasis on travelers' perspectives on system attributes related to reliability. Reliability is typically associated with the attribute "arrive at the intended time".

The importance of reliability to travelers relative to other service attributes is indicated by many studies which show reliability to be one of the most important service attributes. In a study conducted by the University of Maryland (Paine, et al., 1), each respondent in Philadelphia was asked to rate for both auto and transit 35 listed attributes on a 1 to 7 scale of importance for the work and non-work trips, with 7 being most important. In Baltimore, the same procedure was performed for 44 attributes on a 1 to 5 scale. The system attributes for work trips, their mean value, and rank by mean importance appear in Table 2.1. Note that commuters in Baltimore and Philadelphia rated "arrive at intended time" second in importance of all the attributes. Only "arrive without accident" was viewed as more important.

The importance commuters place on travel reliability as found by University of Maryland researchers is substantiated by surveys conducted in other metropolitan areas. A survey of Boston commuters (Spear, 2) found "arrive at the intended time" to be the most important level of service attribute ("arriving without accident" was not included in the survey). A similar survey conducted in Chicago (Stopher, et al., 3) found "arrive at the intended time" second in importance (following "easily accessible station").

Considerably less information is available with which to assess the importance of travel time reliability in non-work travel. The surveys in Baltimore and Philadelphia (Paine, et al., 1) provide data for analyzing this question. It is, perhaps, noteworthy that on-time arrival is still important relative to average travel time and cost factors, generally thought to be the predominant determinants of travel demand. The findings in Table 2.2 indicate that reliability attributes are of lesser, albeit still considerable, importance for non-work travel.

TABLE 2.1 RANK BY IMPORTANCE, WORK TRIP FOR COMBINED PHILADELPHIA, BALTIMORE SAMPLE

Description	Rank by Mean Importance	Mean Value
Arrive without accident	1	6.49
Arrive at intended time	2	6.40
Safest vehicle	3	6.39
Avoid stopping for repairs	4	6.30
Shortest distance	5	6.14
Fast as possible	6	6.13
Avoid changing vehicle	7	6.10
Vehicle unaffected by weather	8	6.01
Protected from weather while waiting	9	5.97
Shortest time	10	5.88
Avoid waiting more than 5 min.	11	5.83
One-way cost of 25¢ rather than 50¢	12	5.66
Comfortable	13	5.62
One-way cost of 25¢ rather than 35¢	14	5.59
Clean vehicle	15	5.59
Feel independent	16	5.52
Avoid walking more than a block	17	5.48
One way cost of 3¢ rather than 15¢	18	5.46
Cost	19	5.28
Avoid unfamiliar area	20	5.17
Travel when traffic is light	21	5.14
Uncrowded vehicle	22	5.10
Package and baggage space	23	4.70
Pride in vehicle	24	4.70
New modern vehicle	25	4.66
Friendly people	26	4.55
People you like	27	4.50
Need not pay daily	28	4.11
Avoid riding with strangers	29	4.07
Listen to radio	30	3.87
Ride with people who chat	31	3.87
Look at scenery	32	3.56
Take along family and friends	33	3.56

Source: Paine, et al (1)

TABLE 2.2 RANK BY IMPORTANCE, NON-WORK TRIP FOR COMBINED PHILADELPHIA, BALTIMORE SAMPLE

Description	Rank by Mean Importance	Mean Value
Arrive without accident	1	6.42
Safest vehicle	2	6.34
Avoid stopping for repairs	3	6.27
Protect from weather while waiting	4	6.01
Avoid changing vehicle	5	5.99
Vehicle unaffected by weather	6	5.95
One-way cost of 25¢ rather than 50¢	7	5.74
Arrive at intended time	8	5.67
Comfortable	9	5.65
Avoid walking more than a block	10	5.58
Clean vehicle	11	5.58
One-way cost of 25¢ rather than 35¢	12	5.56
One-way cost of 3¢ rather than 15¢	13	5.53
Avoid waiting more than 5 min.	14	5.40
Package and baggage space	15	5.33
Shortest distance	16	5.30
Cost	17	5.24
Fast as possible	18	5.23
Feel independent	19	5.18
Uncrowded vehicle	20	5.15
Avoid unfamiliar area	21	5.07
Travel when traffic is light	22	4.89
Shortest time	23	4.82
Friendly people	24	4.75
People you like	25	4.67
Take along family and friends	26	4.65
New Modern vehicle	27	4.62
Pride in vehicle	28	4.48
Ride with people who chat	29	4.15
Look at scenery	30	4.04
Avoid riding with strangers	31	4.04
Need not pay daily	32	3.79
Listen to radio	33	3.46

Source: Paine, et al (1)

Several studies have been conducted on user attitudes towards travel reliability on a particular transit mode. Attitudinal surveys were conducted among riders of the Shirley Highway express-bus-on-freeway in the Washington, DC area (Wachs, 4), and respondents were asked to rate attributes of the bus service which were of consequence to them. Multiple responses were permitted by the survey instrument. 90% of the users cited reliable schedules, while only 29% cited a five-minute saving in travel time. This result suggests that while elapsed travel time is important to bus travelers, reliability in arrival time may be of even greater importance.

A General Motors survey of potential users of a "demand responsive jitney" system in Warren, Michigan (Golob, 5), found "arriving when planned" to be the most essential attribute. In that survey, "arriving when planned" was found to be more important to potential users than "lower fare", "having a seat", "no transfer trip", "less wait time" and "short travel times". This is very significant in pointing out that service improvements which are focused on reliability rather than speed, convenience or reduced fares, may be more successful in boosting ridership. The results of a similar survey of dial-a-bus users and non-users in Columbia, Maryland, closely paralleled those of General Motors, with "arriving when planned" again being most important.

Park-and-ride users in Henrietta (Rochester, New York) (Keck and Cohen, 6) were asked to rank thirteen level of service variables; "reliability" was cited as the second most important system attribute ("convenience" was ranked first).

User attitudes on reliability of specific modes were also examined in Haddonfield, New Jersey (Winchester, 7). The Haddonfield data was separated into dial-a-ride user households, bus user households, and taxi user households to examine how users perceived the importance of reliability based on their experience with each mode. All user groups rated "reliability" as the most important attribute. There were only slight differences in the mean weighting of reliability for each mode, with taxi users placing the highest absolute weight on reliability, followed by bus users and then dial-a-ride users.

The results of the surveys of attitudinal response to reliability on specific modes point out the importance users place on travel reliability regardless of the mode concerned. This is a particularly significant finding when considering the vast difference in service characteristics of the different modes (e.g., dial-a-ride door-to-door service vs. express bus line-haul service).

There appear to be relationships between the demographic characteristics of travelers and their ranking of importance of reliability. The Chicago surveys of commuters (Stopher, 3) found that males ranked "arrive at the intended time" higher than females. Older riders placed a greater weight on this attribute than did younger riders. As the data is disaggregated by age group, the sample size is reduced so some doubt may be cast on the statistical significance of these results. However, a possible explanation for the above results may be that type of employment and level of position influence the importance of punctuality for the work trip and therefore the relative weighting of reliability with respect to other service attributes. These status characteristics vary with age and sex.

The GM survey (Golob, 5) also contrasted scalings for elderly, low income, and under 25 and single groups. "Arriving when planned" scored high for all groups. Only for the elderly were other attributes ranked more important ("having a seat", "no transfer trip" and "lower fares"). However, a higher weight on comfort and convenience by the elderly is not unexpected. Since the elderly might also be expected to make fewer or no work trips, the diminished importance of reliability may again be tied to the trip purpose, and perhaps to the fact that the elderly have fewer travel alternatives and therefore are less sensitive to variations in service.

In summary, attitudinal surveys show reliability to be among the most important service attributes for all travelers under a variety of travel conditions. For both work and non-work trips, reliability was considered more important than average travel time and cost, generally thought to be the predominant determinants of travel demand. Some differences are apparent between the work and non-work trip, and among modes and demographic backgrounds. Age and sex often correlate with work type and trip purpose, however, implying that many of these other differences may also be work-related.

One must recognize the limitations of the findings reported from the attitudinal studies. In many cases, the list of attributes did not include specific reliability attributes ("variability of travel time"), but rather reliability-related attributes ("arrive at the intended time"). In addition, there was no effort to validate the survey responses (e.g., observe the travel behavior of the respondents). One can interpret the findings as an indication that reliability is an important attribute in the travel decision process, and that direct reliability measures and inclusion of specific reliability variables in empirical demand models are needed to provide an accurate assessment of the importance of reliability in the travel decision-making process.

In light of the importance that users appear to place on reliability, it is worth examining current levels of satisfaction users experience regarding the reliability of particular modes. The Philadelphia survey (Paine, et al., 8) addressed this question and helps point to the significance of reliability as a factor influencing the attractiveness of transit service. The survey asked each respondent to express their satisfaction with both existing public transit and the automobile. Results for work and non-work trips appear in Tables 2.3 and 2.4 respectively. Among the 15 most important attributes, "arrive at the intended time" shows the fourth and sixth highest differences between satisfaction for auto and transit for work and non-work trips respectively.* These results support the view that transit is currently considerably less reliable than auto, and one could suggest that this performance deficiency is likely to be an important reason for the preference for auto trips when autos are available.

2.2 TRAVELER BEHAVIOR

Although attitudinal studies indicate the importance of reliability in the traveler decision-making process, insufficient research has been devoted to providing a conceptual framework for understanding the impact of changes in service attribute variability upon travel behavior. This section provides a theoretical discussion on travel behavior as it is affected by the reliability of transit service. It begins with a general discussion of the potential effect of reliability on the fundamental travel choices of frequency, destination, mode split, and route choice. Although it

---

*Part of the difference between ratings of auto and bus satisfaction may be due to cognitive dissonance on the part of auto users replying to transit attributes. For example, an auto user may not be aware of his transit alternative, but because he wants to defend his choice, he ranks transit reliability far below auto reliability. The reverse holds true for transit users, but because the sample is probably predominantly auto users, the effects of auto user cognitive dissonance is more pronounced in the survey results.

TABLE 2.3 PERCENTAGE OF SATISFIED RESPONSES TO EACH ITEM FOR AUTO VERSUS PUBLIC MODES WITH ITEMS RANGED BY MEAN IMPORTANCE: WORK TRIP

Description	Importance Rank	% Satisfied with Auto	% Satisfied with Public Transit	Difference
Arrive without accident	1	89	86	3
Arrive at intended time	2	94	63	31
Safest vehicle	3	94	91	3
Avoid Stopping for repairs	4	94	92	2
Shortest distance	5	97	64	33
Fast as possible	6	95	70	25
Avoid changing vehicle	7	98	69	29
Vehicle unaffected by weather	8	90	84	6
Protected from weather while waiting	9	96	54	42
Shortest time	10	92	65	24
Avoid waiting more than 5 minutes	11	98	55	43
One-way cost of 25¢ rather than 50¢	12	88	69	19
Comfortable	13	97	75	22
One-way cost of 25¢ rather than 35¢	14	88	67	21
Clean vehicle	15	92	73	19
Feel independent	16	89	58	31
Avoid walking more than a block	17	97	65	32
One-way cost of 3¢ rather than 15¢	18	84	67	17
Cost	19	91	62	29
Avoid unfamiliar area	20	93	78	15
Travel when traffic is light	21	73	64	9
Uncrowded vehicle	22	94	51	43
Package and baggage space	23	94	57	37
Pride in vehicle	24	88	60	28
New modern vehicle	25	86	68	18
Friendly people	26	91	67	24
People you like	27	93	66	27
Need not pay daily	28	82	61	21
Avoid riding with strangers	29	89	66	23
Listen to radio	30	84	44	40
Ride with people who chat	31	84	62	22
Look at scenery	32	74	72	2
Take along family and friends	33	87	57	30

Source: Paine, et al (8)

TABLE 2.4 PERCENTAGE OF SATISFIED RESPONSES TO EACH ITEM FOR AUTO VERSUS PUBLIC MODES WITH ITEMS RANGED BY MEAN IMPORTANCE: NON-WORK TRIP

Description	Importance Rank	% Satisfied with Auto	% Satisfied with Public Transit	Difference
Arrive without accident	1	86	77	9
Safest vehicle	2	89	81	8
Avoid stopping for repairs	3	89	81	8
Protected from weather while waiting	4	95	47	48
Avoid changing vehicle	5	96	68	28
Vehicle unaffected by weather	6	90	81	9
One-way cost of 25¢ rather than 50¢	7	86	65	21
Arrive at intended time	8	93	66	27
Comfortable	9	97	69	23
Avoid walking more than a block	10	95	57	38
Clean vehicle	11	91	63	28
One-way cost of 25¢ rather than 35¢	12	85	63	22
One-way cost of 3¢ rather than 15¢	13	82	60	22
Avoid waiting more than 5 minutes	14	93	45	48
Package and baggage space	15	95	52	43
Shortest distance	16	91	55	36
Cost	17	89	54	35
Fast as possible	18	90	64	26
Feel independent	19	84	51	33
Uncrowded vehicle	20	91	48	43
Avoid unfamiliar area	21	91	72	19
Travel when traffic is light	22	67	55	12
Shortest time	23	91	55	36
Friendly people	24	92	60	32
People you like	25	91	60	31
Take along family and friends	26	86	59	27
New modern vehicle	27	84	62	22
Pride in vehicle	28	89	53	36
Ride with people who chat	29	82	56	26
Look at scenery	30	78	76	2
Avoid ride with strangers	31	91	58	33
Need not pay daily	32	76	51	25
Listen to radio	33	79	32	47

Source: Paine, et.al. (8)

appears on theoretical grounds that service variability probably influences these travel choices, there is insufficient empirical evidence at the moment.

A fundamental travel choice which has been largely ignored by existing travel demand theory is the departure time decision. The departure time decision involves the timing of a trip departure from an origin in order to reach a destination by a specified time.*

An important finding of the study is that reliability is likely to be an important determinant of departure time behavior. There is convincing empirical evidence that travelers alter their departure time behavior in response to the degree of variability in service. This evidence of departure time behavior suggests a general theory of the travelers' response to improvements in transit service reliability which provides a framework for estimating the benefits of these improvements. Specifically, it is proposed that travelers will respond to service reliability improvements by altering their departure time behavior in accordance with their relative preference for travel speed, reliability, and schedule delay. This will result in travelers choosing to take some of the reliability benefits in the form of reduced mean travel or wait time and the remainder of the benefits in the form of improved reliability itself. To quantify the benefits of reliability, it is therefore necessary to measure the value travelers place on the variability of travel time components. It is important to understand this relationship, for it is possible that the net benefits from reliability improvements may be greater than the net benefits from a comparable investment in increased speed.

2.2.1 Reliability and Travel Choice

Travelers make travel choices based on the alternatives available and their preferences for the attributes of these alternatives. An individual traveler associates a disutility with a trip based on his/her perception of the service attributes of the transportation system. The traveler weighs the importance of these characteristics according to his/her travel demands and socio-economic

---

*The concept of departure time should not be confused with the choice of time of day of travel.

status.* The traveler, assuming he acts rationally (Manheim, 9), will select the travel alternative which minimizes his disutility.

In the interest of simplicity and for the purposes of this report, only the travel time related disutility of a given trip, U_T, will be considered. As evidenced in Section 2.1, it would seem that travelers generally do not base their behavior solely upon the mean values of travel time components, but also are influenced by the inherent and exogenously induced variability of these attributes of transportation services. The variability of travel time components is important to travelers in their decisions concerning a variety of travel choices, particularly mode choice and departure time.

The travel time components of the disutility of a given trip, U_T, will be expressed mathematically as:

$$U_T = U_{TM} + U_{TV} = \alpha T_M + \beta T_V \qquad (1)$$

where

U_{TM} = mean travel time disutility
U_{TV} = travel time variability disutility
T_M = mean travel time
T_V = travel time variability
α = disutility coefficient of mean travel time
β = disutility coefficient of travel time variability

α and β may vary among individuals and, for each individual, may vary by trip purpose, trip frequency, mode and time of day. Equation (1) does not differentiate between the different travel time components (e.g., walk time, wait time, in-vehicle time). The mean and variability of each component will have different disutility coefficients. T_M, T_V, α, and β are used in the following discussion to represent composite values of mean travel time, travel time variability, and the disutility coefficient of each, respectively.

In making the travel choice decision, the traveler may be thought of as having a set of indifference curves which represent levels of travel time variability and mean travel time combinations that result in the same amount of disutility. The traveler can be viewed as being indifferent

---

*The reader may wish to refer to Ben-Akiva (10, 11) or Manski (12) for a discussion of utility and choice theory as it relates to the travel decision-making process.

to all combinations of T_M and T_V whose total disutility, U_T, is the same (assuming that the trip purpose, frequency, origin, destination, cost, and other amenities of the trip associated with each combination of T_M and T_V are the same). An indifference curve, which plots travel speed ($1/T_M$) vs. travel time reliability ($1/T_V$) for a given value of U_T, reflects the traveler's tradeoff between T_M and T_V. Figure 2.1 shows two such curves, U_{T1} and U_{T2}. The traveler is indifferent between mean travel time and travel time variability along each of these curves. However, the traveler will always prefer U_{T1} to U_{T2} because it represents a lower disutility of travel.

If a traveler has two choices, A and B, for making a given trip, the traveler will choose B; the traveler is trading off T_M and T_V and views the greater travel time of B as being less onerous than the lower reliability of A. If the traveler has a third choice C, for making the trip, he will still choose B. The lower reliability of B is now viewed as being less onerous than the greater travel time of C.

This simple model suggests that whenever the traveler is faced with a travel choice decision for which such tradeoffs exist, the reliability of each alternative may influence that decision. Moreover, where the reliability of each alternative is very poor, the disutility associated with each alternative can be so high that the traveler may choose not to make the trip, make it less frequently, make it to a different destination, select a different mode, a different route or depart at a different time.

2.2.2 Reliability and the Departure Time Decision

Travelers often make departure time decisions for the trips they take. Frequently these decisions are made with the objective of reaching the trip destination at or by a specified time. This is particularly important for the work trip, where lateness can have severe penalties associated with it. In selecting a departure time, the traveler, whenever possible, attempts to minimize his/her travel time related disutility. The traveler trades off mean travel time components with travel time variability components in the same fashion as in making other travel choice decisions.

Each traveler, whether an auto or transit user, must consider total trip travel time and its variability in making the departure time decision. Total travel time variability will directly affect the time of arrival at the destination. Total travel time variability for the transit

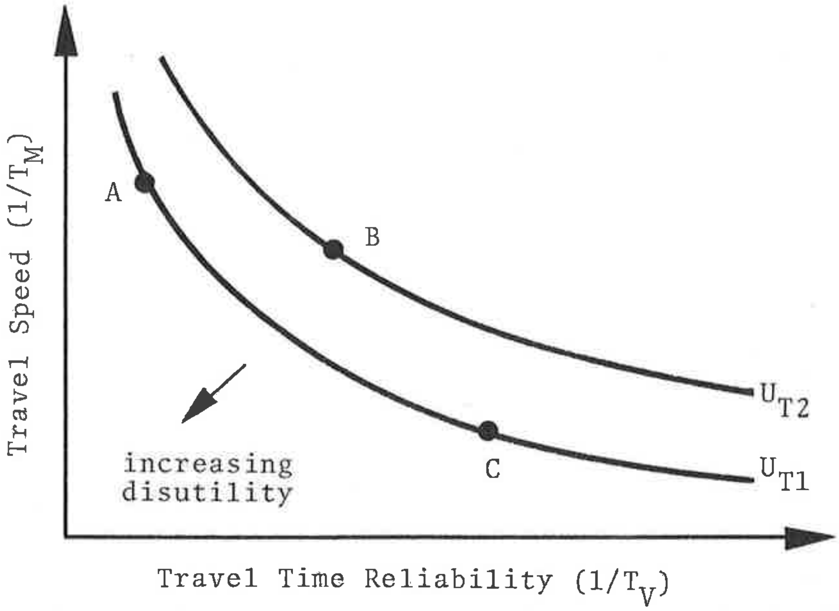


FIGURE 2.1 MEAN TRAVEL TIME - TRAVEL TIME VARIABILITY INDIFFERENCE CURVES

user is a function of both in-vehicle travel time variability and wait time variability. Thus, the transit user must also take into account the expected mean and variability of transit arrival time at his/her boarding point in an attempt to minimize some combination of the mean and variability of his/her wait time. The following discussions examine the role of the variability of these travel components in departure time decisions. The first discussion focuses on the role of the variability of total travel time, and the second examines the effect of the variability of transit vehicle arrivals at the boarding point on passenger wait times.

Travelers who place importance on arriving at their destination at a particular time adjust to the variability of total travel time by expending resources in the form of extra time allotted to the trip. Marfisi, et al., (20), explored this phenomenon in an analytical model. The traveler is assumed to be free to leave as early as desired in order to be sure to arrive "on time". The traveler chooses a premature departure time, P*, so as to minimize the disutility of the trip, which is made up of the disutility of the time he/she allows for the trip (the mean travel time plus the additional time required to account for the premature departure time), and the disutility associated with arriving late even though the traveler departed P minutes early. The disutility associated with arriving early is assumed to be less than the disutility of arriving late (i.e., the traveler is assumed to be risk averse). The resulting expression:

$$P^* = \frac{cv(t)}{2} \qquad c>0^* \qquad (2)$$

indicates that the traveler's premature departure time, P*, depends only on the variance of total trip time, v(t), and degree of risk aversion, c. The more risk averse the individual, the larger the value of c. For a given individual, the value of c will vary from trip to trip depending on the penalty the individual perceives for arriving late at the particular destination. Equation 2 does not assume that the traveler consciously computes the variance of trip time. It does assume that the traveler is sufficiently familiar with his/her trip to enable him/her to set a departure time based on past experience. This type of behavior is probably most characteristic of the journey to work and for travel to appointments and special events, but

---

*A complete derivation of the equation is presented in Appendix A.

may still occur for other trip types where the traveler may have a high disutility for late arrival.

In developing Equation 2, Marfisi assumes that travel time is normally distributed with mean, M, and variance, v(t). It is more likely that the travel time distribution is highly skewed in the direction of travel times higher than the mean. When this is the case, a measure of skewness (such as $E(h^3)$, the third moment of the headway distribution, or the percent of travel times greater than N minutes later than the mean) is needed in addition to v(t) to reflect the variability of travel time. Intuitively, just as $P^* = cv(t)/2$, it follows that when the travel time distribution is skewed, $P^* = f(c,v(t),S)$, where S is a measure of skewness.

Marfisi's supposition that the traveler's premature departure time, P*, depends only on the variability of his total trip time may, in some instances, require some modification. For fixed route transit services, passengers are attempting to minimize some combination of mean and variability of both in-vehicle travel times and wait times. Thus, travelers may arrive at the bus stop earlier than the time indicated by P* in an attempt to minimize mean wait time and its variability. The discussion which follows will focus in detail on the extent to which fixed route transit travelers adjust their departure time behavior to take into account the variability of arrival time of transit vehicles at passenger boarding points.

Variability of transit vehicle arrival time, as well as average headways, affects mean passenger wait time. As the variability in the headway distribution increases (while average headway remains unchanged), the expected wait time of passengers will increase. It can be derived that the expected wait time, E(w), of passengers who arrive randomly (i.e., as a Poisson process) at their boarding point is:

$$E(w) = \frac{H}{2} + \frac{V(h)}{2H} \qquad (3)$$

where H is the mean headway and V(h) is the variance of the headway distribution.* Friedman (14) demonstrated that for randomly arriving passengers the variance of wait time is given by:

$$V(w) = \frac{E(h^3)}{3H} - (E(w))^2 \qquad (4)$$

---

*See Appendix B for a derivation of this equation.

The variance of wait time is related to the third moment of the headway distribution, which is a measure of skewness. A measure of skewness is an important element of the variance in wait time since it reflects the proportion of extremely late arrivals, caused by vehicle breakdowns and major traffic delays.

Equations (3) and (4) are based on the assumption that bus arrivals are totally unpredictable and that regardless of when a traveler arrives at the bus stop, he/she can expect to wait an average of $H/2 + V(h)/2H$ minutes for the bus and can expect his/her wait time to vary from day to day by an amount reflected in $V(w)$. From these findings, one can conclude that expected passenger wait time is a function of: 1) the mean headway, 2) the headway variation, and 3) the skewness of the headway distribution.

However, bus arrivals are sometimes predictable, and, as a result, passenger arrivals at their boarding points are not always random. There is convincing evidence that passenger arrivals at bus stops are influenced by the reliability of transit service. Joliffe and Hutchinson (15) examined the arrivals of passengers at suburban bus stops in London in the summer of 1974 to determine whether most passengers were actually arriving randomly. They selected a single bus stop on each of ten routes and collected data on both passenger arrivals and bus arrivals over a period of one hour for eight days. The data was tabulated as shown in Table 2.5. At three of the stops (1 - 3), observations were made during the A.M. peak period, and during the off-peak hours at the others. The eight days on which the observations at any one stop were made were not consecutive, being spread over four weeks, but they covered each day of the week (Monday - Friday).

Joliffe and Hutchinson found that travelers tend to arrive randomly at the bus stop under the following conditions:* (1) when bus arrivals at a given stop are totally unpredictable, or (2) when bus arrivals are, to some extent, predictable, but travelers are not aware of the daily performance of the bus route and do not avail themselves of a posted schedule (where one exists). The first point deserves further elaboration. When headways are very frequent (i.e., six minutes or less (Holroyd and Scraggs, 16)) schedule adherence tends to deteriorate rapidly along a bus route. As a result, bus arrivals at all but the first few stops along the route are so random as to be unpredictable. The resulting "bunching" effect of

---

*Data from Seddon and Day (17) was also used.

TABLE 2.5 MEASURES DERIVED FROM OBSERVATIONS OF PASSENGER AND BUS ARRIVALS AT TEN SUBURBAN BUS STOPS IN LONDON

Stop	Mean Scheduled Interval between Buses	Per-centage of Buses Cancelled	Mean Observed Headway	Standard Deviation of Observed Headways	Standard Deviation of Bus Departure Times (averaged)	Average Lateness[a]	Expected Waiting Time for Random Passenger Arrivals	Minimum Expected Waiting Time	Potential Gain from Knowing the Service, $=\omega_{rand}$ $-\omega_{min}$	Mean Observed Waiting Time	Standard Deviation of Observed Waiting Time
	SI	c	μ	σ	Ω	L	ω_{rand}	ω_{min}	g	W	s
1	6.4	17	8.0	6.3	3.2	-1	6.5	4.6	1.9	5.2	4.6
2	11.2	7	14.2	8.4	3.4	1	9.4	4.1	5.3	6.2	5.3
3	23.0	0	23.9	2.2	2.4	2	12.9	4.3	8.5	5.8	5.4
4	15.5	3	16.2	4.9	2.6	-1	8.7	4.9	3.8	7.3	6.0
5	15.0	3	15.4	4.1	1.4	2	8.3	3.1	5.2	6.2	4.3
6	7.1	9	8.0	5.7	2.6	-1	6.1	4.5	1.6	5.3	4.8
7	7.7	4	8.7	4.9	2.1	-2	5.5	3.2	2.3	4.2	3.9
8	11.6	17	14.3	7.1	2.1	(1)	8.6	4.8	3.7	7.4	6.3
9	20.7	33	30.7	10.0	3.1	2	17.3	9.7	7.7	12.1	9.8
10	20.3	8	23.5	10.7	9.1	4	14.0	7.1	6.9	13.1	11.2

[a] Positive means late, negative means early. The timetable for the service at stop 8 is not published.

vehicles results in more pronounced service deterioration further along a route. This effect is discussed in more detail in Chapter 5. As headways become larger, schedule adherence may deteriorate as rapidly, but the magnitude of the variation in bus arrival patterns is generally not large enough relative to the size of the headways to completely randomize the bus arrivals. Schedule variation and some of its causes will also be discussed further in Chapter 5.

Joliffe and Hutchinson also found that where bus arrivals are predictable, travelers may be able to minimize their wait time at the bus stop by arriving there at a particular time instead of randomly. The term "predictable" is used to imply that buses arrive at approximately the same time each day, but may not be adhering to schedule. They observed in several cases the coincidence of high passenger arrivals with resulting waiting times lower than the expected waiting time, $\underline{w}_{rand}$, had the passengers arrived randomly. They calculated that for each bus stop, there was a particular arrival time, w_{min}, which would minimize the expected waiting time for the bus.*

Joliffe and Hutchinson's work presents an interesting outlook on traveler behavior relative to schedules. Where the user is aware of a constant schedule adherence problem, he/she adjusts his/her arrival pattern to receive more reliable service. For example, a bus which is consistently five minutes behind schedule is a predictable service which the traveler can adjust to by arriving at his/her bus stop five minutes later. In a sense, users do not view schedule adherence as a reliability problem, as long as the service is predictable on a daily basis. As will be seen in Chapter 3, the operator has a contrasting view of reliability.

The material presented in Sections 2.2.1 and 2.2.2 suggest deficiencies in the structure of current travel behavior models used in travel forecasting. Current travel behavior models generally do not include reliability measures, and treat level of service in terms of mean values. An estimated constant, the mean of omitted terms, is included which explains part of the variance in service attribute levels and the effect of other service amenities (comfort, convenience, etc). Part of the variance is also likely showing up in the coefficient for mean travel time. Failure to explicitly treat the variability of specific measures in travel demand model formulation is undoubtedly

---

*Further Joliffe and Hutchinson analyses of non-random and random arrivals appear in Appendix E.

introducing bias into the parameters of the mean time components (Starkie, 18), leading to forecast error.

Quantitative models of travel behavior are needed which can be used to quantify the impacts of service attribute variability upon travel choices. Not only are these models needed for improved patronage forecasts, but they are also required to evaluate the impacts of system improvements and to provide information on essential tradeoffs in service design and operating policies.

To date, few attempts have been made to model transit reliability in travel demand models. Cambridge Systematics (Lerman, et al., 19) attempted to incorporate both wait time and ride time reliability as a variable in a logit formulation for a dial-a-ride mode choice model. They had little success with this approach, citing three reasons for their difficulties:

1) There was a great deal of difficulty obtaining the proper data to accurately estimate the reliability variable. Repeated observations on an origin-destination basis were difficult to collect.

2) In the data collected, there was a high correlation between the mean and variance in ride times.

3) The amount of variation in wait time was often too low to allow for proper modelling of the reliability variable.

It is important to note that the Cambridge Systematics study was conducted using zonal averages for travel time reliability. The use of aggregate measures of supply clearly obscures the impact of reliability on demand. Cambridge Systematics also did not account for the disutility associated with late arrival at the destination.

Gardner and Reid (20) have attempted to compute disaggregate reliability data for individual trips. A network is defined using nodes and links. The level of unreliability experienced on a link is dependent on the level of congestion on the link. Unreliability experienced at nodes is a function of the volume of passengers wishing to board, transfer, or disembark at the node. An individual's trip from origin to destination can be traced through the network to determine the total trip time variability that will be experienced.

Gardner and Reid made several assumptions in their research. In particular, their method relies on the ability of transit schedulers to identify links and nodes which are known to have service reliability problems. In addition, the program itself places restrictions on the exact specification of the traveler's route, and computes link and node reliability at discrete rather than continuous periods of time. Furthermore, generalized time distributions (e.g., same distribution for all arterials with similar characteristics) are used for all the nodes and links on the network, thereby losing some of the information on node and link-specific effects on reliability. Nevertheless, this approach represents an improvement in disaggregating supply in order to measure the impact of reliability on demand.

Prashker (21) recently calibrated a multinomial logit mode choice model using objectively measured variables for mean travel time components and travel costs, and an attitudinal travel time reliability variable. Prashker used two specifications, one omitting the reliability variable and one including the reliability variable. Prashker found that inclusion of a perceived reliability variable was statistically significant and improved the predictive power of the model.

It is difficult to interpret the work done by Prashker. His reliability term is not related to the flexibility of an individual's arrival time at work, so he is not able to capture the loss (disutility) which a user perceives when faced with uncertain arrival times at work caused by unreliability. The other problem with Prashker's work is that the use of an attitudinal variable is not metric. Thus two travelers who each experience the same variability in travel time may rate their reliability differently (e.g., one may rank his/her level of reliability as a 3 on a scale from 1 to 5, while the other may rank the same level of reliability as a 5). Thus, the explanatory power of the variable is difficult to determine.

Spear (2) also estimated a mode choice model using a generalized reliability variable based on both the attitudinal importance users placed on reliability and their satisfaction with the service attribute. He too found a better goodness of fit with model specifications that included the generalized reliability variable than with model specifications that omitted the reliability variable. Again, for reasons cited above, it is difficult to interpret his research findings, although it does indicate that further research in modeling of attitudinal reliability variables is appropriate.

Very little travel demand model research has been directed at understanding the departure time decision. Cosslett (22) estimated a multinominal logit travel demand model for the departure time decision. Using data collected for the Urban Travel Demand Forecasting Project sample, Cosslett examined the individual's tradeoff between the mean and variability of travel times during the peak period and the probability of arriving early, on-time, or late for work. Earliness and lateness presumably have different disutilities associated with them. Cosslett used the following utility form:

$$U(T) = \alpha y(T) - \beta E - \gamma L(T)$$

where:

T = the time of arrival at work (discrete time period)
$y(T)$ = the in-vehicle time corresponding to arrival time T
E = "early" arrival time at work
$L(T)$ = the probability that the subject will arrive late for work, if he plans to arrive at time T.

α, β, γ are the coefficients of the respective variables.

Cosslett had trouble with the independence of error terms of the alternatives (because departure time alternatives are correlated) which is a necessary assumption when using logit. He couldn't resolve this problem, but performed a logit analysis anyhow, interpreting his results as an empirical fit rather than as a utility-maximizing choice model. Cosslett found that there is a small, but measurable, effect due to highway congestion, with travelers tending to schedule departure times away from the peak whenever possible.

There are several drawbacks to Cosslett's approach. First, as mentioned above, by using discrete intervals of time as alternatives, and because departure time periods are clearly correlated, the independence of irrelevant alternatives assumption, necessary when using logit, was violated. Secondly, the effect of reliability on the departure time decision is modeled by Cosslett assuming a normal distribution of travel times about the mean. One would expect the distribution of travel times to be skewed significantly to the right, with the degree of skewness being very important in the traveler's probability of extreme lateness. Thirdly, Cosslett's approach assumes that travelers perceive a constant amount of arrival time loss for every additional minute of late (or early) arrival at work. One would expect the perceived loss to increase proportionately as a traveler arrives further away from the official work start time. Finally, Cosslett's analysis examines departure time decisions for auto and carpool alternatives. Thus, transit, perhaps the most difficult mode to model (since both wait and ride times vary significantly), was not included in the analysis.

Further research is necessary to develop accurate techniques for modeling transit reliability and its effect upon travel demand. This is discussed in more detail in Chapter 7.

2.2.3 A Theoretical View of Departure Time Decisions for Fixed Route Transit

The research conducted to date suggests that travelers for whom it is important to arrive at their destination at a particular time adjust to the variability of total travel time by setting their departure time to allow extra time, P*, for their trip. It has also been demonstrated that on fixed route transit where headways are large but bus arrivals are fairly predictable, travelers may arrive at their origin bus stop earlier than the time indicated by P* in order to minimize wait time. Based on these findings, one can postulate a theory of departure time decison-making as it relates to fixed route service reliability. The discussion which follows is developed by examining travelers' bus stop arrival patterns for all combinations of the following factors:

(1) Whether or not it is important for travelers to be at their destination by a certain time.
(2) Whether or not travelers are familiar with the system.
(3) Bus arrival patterns (headway lengths and predictability of bus arrivals to travelers).

Table 2.6 summarizes the important relationships developed from the following theory.*

---

*In the scenario that is developed, it is assumed that mean in-vehicle travel time, in-vehicle travel time variability, and seat availability do not vary over the range of arrival times at bus stops that travelers might consider on a particular trip. Generally, the narrowness of the range ensures that these service attributes will not vary. When these attributes do vary, travelers will take into account (Continued)

TABLE 2.6 SUMMARY OF RELATIONSHIPS BETWEEN THE VARIATION OF BUS ARRIVAL TIMES AT BUS STOPS AND TRAVELER ARRIVAL TIMES

Is it important for travelers to be at their destinations by a certain time?	Are travelers familiar with the system?	Bus arrival patterns	Traveler bus stop arrival patterns
Yes	Yes	$h \leq 6$ min.; or $6 < h \leq 12$, where travelers view bus arrivals as unpredictable.	Arrival times based solely on T_1.
Yes	Yes	$6 < h \leq 20$, where travelers view bus arrivals as predictable.	Arrival times between T_w and T_1, based on traveler preferences.
Yes	Yes	$h > 20$	Arrival times based solely on T_w
Yes	No	All headways	Arrival times based solely on perceived value of T_1.
No	Yes	$h \leq 6$; or $6 < h \leq 12$, where travelers view bus arrivals as unpredictable	Arrival times generally based on considerations not related to bus level of service attributes.
No	Yes	$6 < h \leq 20$, where travelers view bus arrivals as predictable	Arrival times generally based on considerations not related to bus level of service attributes.
No	Yes	$h > 20$	Arrival times based solely on T_w
No	No	All headways	Arrival times generally based on considerations not related to bus level of service attributes.

T_1 = Latest premature arrival time at bus stop which ensures that the risk of arriving late at the destination is not too great.

T_w = Arrival time at bus stop which minimizes the traveler's combined disutility for wait time and wait time variability.

Travelers for whom it is important to arrive at their destination by a particular time and who are familiar with the characteristics of the bus service must select a premature departure time which ensures that their arrival time distribution at their destination is acceptable. Let T_1 be the latest possible passenger arrival time at the origin bus stop which ensures this. These travelers' departure time decisions may also be influenced by the variability of transit vehicle arrival times at their bus stop. Let T_W be the arrival time at the bus stop that will minimize travelers' wait time.

Where travelers do not perceive any savings to be gained in wait time by arriving at the bus stop at a particular time (e.g., where bus arrivals are totally unpredictable), T_W does not enter into their bus stop arrival time decisions. Travelers will select T_1 as their arrival time. Bus arrivals will appear to be unpredictable to travelers when the arrivals are quite random. As discussed earlier, bus arrivals are inherently random for headways under 6 minutes. For headways as high as 12 minutes, bus arrivals may also be random under certain circumstances (e.g., a congested, central city route).

Where travelers view bus arrivals as predictable, they may now be faced with some complicated tradeoffs in deciding between arriving at the bus stop at T_1 or T_W. Ideally, if T_1 and T_W coincided, there would be no tradeoffs to make. But typically, T_1 is later than T_W (the manner in which T_1 is defined ensures that T_1 can never be earlier than T_W). Arrival at T_W will ensure a low wait time, but may result in an early arrival at their destination. Arrival at T_1 will ensure an acceptable risk of arrival at the destination at the preferred time but will result in longer wait times and/or wait time variability. Arrival at any time in between T_W and T_1 involves similar tradeoffs. Each traveler will behave differently, depending on personal preferences. Where transfers to a second vehicle are involved, the tradeoffs can become extremely complicated (though, usually the arrival of one of the buses will be unpredictable, simplifying the tradeoffs significantly).

(Continued)
their variation as well as the variation of bus arrival times at the bus stop in making the decision as to what time to arrive at the stop. When these attributes do vary over the feasible range of arrival times at the bus stop, traveler behavior becomes quite complex, involving tradeoffs between several service attributes.

As headways become larger, and more is to be gained from arriving at the bus stop at T_W (i.e., wait time will be significantly greater at other than the optimal arrival time), it is less likely that any travelers will attempt to arrive at the bus stop later than T_W. For headways larger than 20 minutes, T_W may be the only feasible time to arrive at the bus stop. At the extreme, the single scheduled bus trip associated with T_W will cause travelers to always arrive at their destination much earlier than they would like. The premature arrival may have considerable disutility to the travelers, depending upon the use to which they can put this extra time. A similar phenomenon may occur on the return trip, where the scheduled bus associated with T_W will cause travelers always to arrive at their destination much later than they would prefer.

Travelers for whom it is important to arrive at their destinations at a particular time but who are not familiar with their bus service are faced with an interesting problem. They must set their departure time according to their uncertain perceptions of the bus service.

Travelers who do not have to arrive at their destination by a particular time and are familiar with the characteristics of the bus service will, where bus arrivals are predictable, attempt to minimize their disutility by arriving at the bus stop at T_W. Where bus arrivals are perceived as totally unpredictable to the travelers, or they lack familiarity with the bus service, the travelers who do not have to arrive at their destinations by a particular time will base their arrival at the bus stop on considerations not related to bus service attributes. For example, such travelers may leave for the bus stop as soon as they are finished with their previous activity.

2.2.4 A Theoretical View of Departure Time Decisions for Dial-A-Ride

A similar scenario can be developed for users of demand responsive transit (DRT). Because of the complete lack of empirical data on this subject, the discussion will be qualitative in nature. As with fixed-route transit users, DRT users can be divided into those for whom it is important to be at their destination by a certain time and those for whom it is not a concern.

In most DRT systems, three types of service requests are available to users: (1) requests for immediate service, (2) requests for service at some later time (called advance requests), and (3) requests for particular service on a daily basis (in some systems, vehicles are reserved for subscription work trips to and from a particular destination or destinations). For service requests (2) and (3), users will generally request a time at which they desire to be picked up, but they may request a desired arrival time at their destination. For all three service request types, users are generally given a "promised pickup time" to be picked up by a DRT vehicle at their origin, which may either be a discrete time, or a time range, referred to as a "window".

Travelers for whom it is important to be at their destination by a particular time request an advance pickup time which accounts for the variability of the total trip time. Travelers requesting advance service are generally given the promised pickup time they request.

Those who aren't sure in advance what time they want to be picked up can either (a) request immediate service and risk facing high and variable wait times, or (b) estimate what time they should be picked up and make an advance request, but at the risk of finding out later that they would have preferred another pickup time. Those requesting immediate service may be given promised pickup times several minutes or even hours in the future. Where this occurs frequently, only those travelers who do not find this treatment objectionable or who are not familiar with the service would make immediate requests. Some travelers may know in advance at what time they want to be picked up and will make an advance request for service. Others may decide to estimate the approximate time they want to be picked up and call in advance for service. Such estimates become very difficult for many trips, especially return trips. Moreover, once travelers have made the request for advance service, they have to be ready to board the DRT at a particular time. The only alternative is to take a chance with making an immediate request for service.

In making immediate and advance requests, travelers must take into account not only a highly variable wait time but also a highly variable in-vehicle travel time in determining their premature departure time. The large in-vehicle travel time variability is produced by the variability of the DRT vehicle's route between a traveler's origin and destination on different days since it depends on the number of other demands for service.

Subscription riders are subjected to similar reliability considerations as fixed route travelers. The major difference is that subscription riders must be ready to board the scheduled DRT vehicle, since there are no later vehicles to board if the traveler misses his/her assigned vehicle.

Travelers using subscription service are generally given the same promised time for each day. Where an entire vehicle is reserved for subscription service and the same people are picked up on the same route each day, reliability is much higher than on other forms of DRT. Because of this patterned service, travelers are generally locked into set pickup and dropoff times and can have these altered only if they are able to get scheduled onto a different vehicle being used for subscription service. Even so, with the discrete service choices available, travelers using this type of service will often have to accept service that gets them to their destination too early in the morning and/or picks them up too late in the evening, unless the vehicle is serving their specific shift at their place of work.

All DRT users, once they are given a "promised arrival time" for their vehicle, must be ready to board their vehicle by the <u>earliest</u> promised arrival time, to ensure that the vehicle will not leave without them. Users may even have to be ready earlier if the DRT system is sufficiently unpredictable that vehicles arrive before the promised pickup time and do not wait until the promised pickup time if the passenger is not yet ready to board. Mean wait time, which is defined as the mean of the difference between actual pickup time and promised pickup time, will tend to increase as the vehicle arrival time variability increases, since the DRT users must still be ready to board at the earliest promised pickup time. DRT users could conceivably "outguess" the system by not being ready to board until later than the earliest promised pickup time when that time is found to be unrealistically early. However, it is difficult to postulate a specific behavioral pattern with such a high level of variability in vehicle arrivals. Because of the door-to-door nature of DRT, wait time and its variability tend to be considerably less onerous when the individual requesting service is both not subjected to the weather while waiting and can put his/her waiting time to good use (e.g., an individual waiting at home).

2.3 BEHAVIORAL CHANGES AND POTENTIAL TRAVELER BENEFITS FROM IMPROVED RELIABILITY

The preceding discussion points to a theoretical structure for understanding the behavioral responses to reliability improvements and quantifying the resulting benefits to travelers. Travelers are likely to experience changes in travel characteristics from a reliability improvement. It is suggested that they may alter their travel behavior in response to these changes. The travel behavior changes will depend on the traveler's preferences, alternatives, and the magnitude of the reliability improvement offered.

Reliability improvements can effect changes in a variety of travel choices, including departure time behavior. Section 2.3.1 discusses possible changes in travel choice stemming from a transit reliability improvement; Section 2.3.2 treats specifically changes in departure time behavior when an improvement in reliability is offered. There may also be instances where the traveler does not make any behavioral changes in response to a reliability improvement, yet he/she is able to accrue some benefits anyhow. These situations are also discussed in the following sections.

2.3.1 Reliability Improvements and Travel Choice

Reliability improvements may enhance the mobility of transit dependent travelers and may potentially induce auto users to switch to transit. Reliability improvements effect a significant decrease in the disutility of trips for which these improvements are realized. Specifically, the attractiveness and viability of transit trips will be increased as a result of decreases in the following sources of trip disutility: total travel time variability, wait time variability, the total time travelers allocate for their trip, and mean wait time.

Substantial benefits may be realized by transit dependent persons who had previously curtailed their tripmaking because of the high disutility associated with making certain transit trips. The reduced disutility of the transit trips may enable these travelers to make trips more frequently than they did before, at more convenient times, to more desirable locations, and on different modes, or may even enable them to make entirely new trips, thus increasing their mobility.

It is suggested that travelers may change their mode choice in response to a reliability improvement, provided the improvement is significant and the traveler places a high value on travel time reliability. Figure 2.2 illustrates this possibility. Before the reliability improvement was implemented, the traveler selected mode 2 because both the speed and reliability of mode 2 (point A) were superior to those of mode 1 (point B). However, given

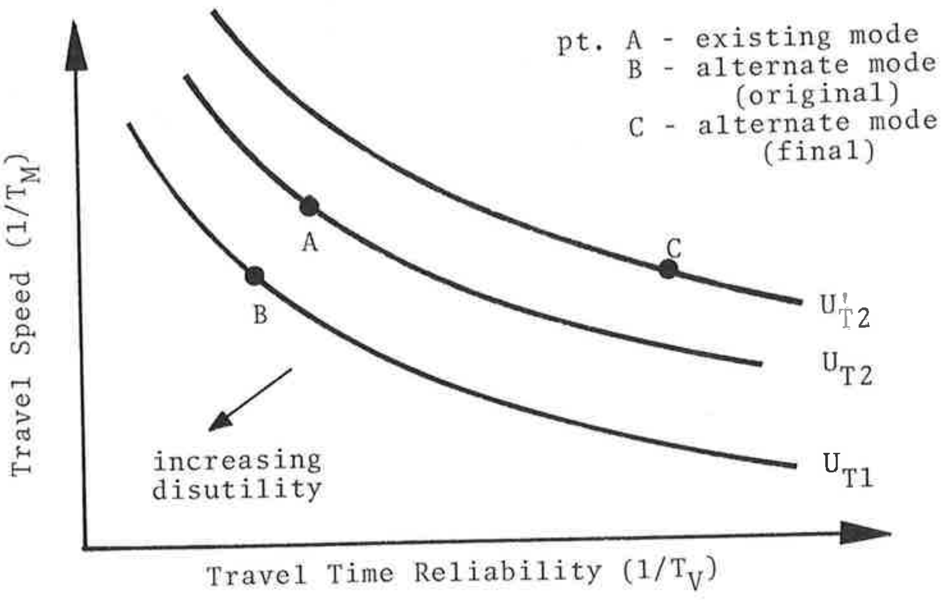


FIGURE 2.2 CHANGES IN TRAVEL CHOICE IN RESPONSE TO A RELIABILITY IMPROVEMENT

a reliability improvement on mode 1 (without any change in travel speed), the traveler will tradeoff the greater speed associated with mode 2 for the now better reliability of mode 1, and will choose mode 1 because the combined disutility is lower (point C).

2.3.2 Reliability Improvements and Departure Time

Improvements in reliability, such as increased probability of on-time arrival, or in some travel time components, such as reduced mean wait times, may accrue to transit travelers without any travel behavior changes on their part. Other improvements in mean wait time and improvements in the form of later trip departure times will result from changes in travel behavior in response to improvements in service reliability.

In response to a reliability improvement (assuming no change in mean travel time) travelers are likely to tradeoff increased probability of on-time arrival with later departure times. At one extreme, travelers will not alter their departure time, while increasing their probability of on-time arrival. At the other extreme, travelers may depart at a later time while maintaining their former probability of on-time arrival. Most travelers are likely to respond somewhere within this range depending on their attitudes towards late arrival and leisure time.

2.3.2.1 Increasing the Probability of On-Time Arrival

This benefit is derived from improved transit reliability for travelers who elect not to change their departure times. This benefit is illustrated in Figure 2.3a. Before the improvement, the traveler departing at T_D experiences an arrival time distribution corresponding to distribution 1, with a mean travel time of t and a mean arrival time of T_A. With the reliability improvement, the traveler, without changing his/her departure time T_D, experiences arrival time distribution 2. Although the traveler has the same departure time and mean travel time as before, the probability of arriving late has been reduced significantly. (The area under each distribution after T^*, the time of late arrival, represents the probability of arriving late).

In many cases, reduced wait time and in-vehicle travel time variability benefits automatically accrue to travelers without any action on their part. For example, where wait time variability and in-vehicle travel time variability have

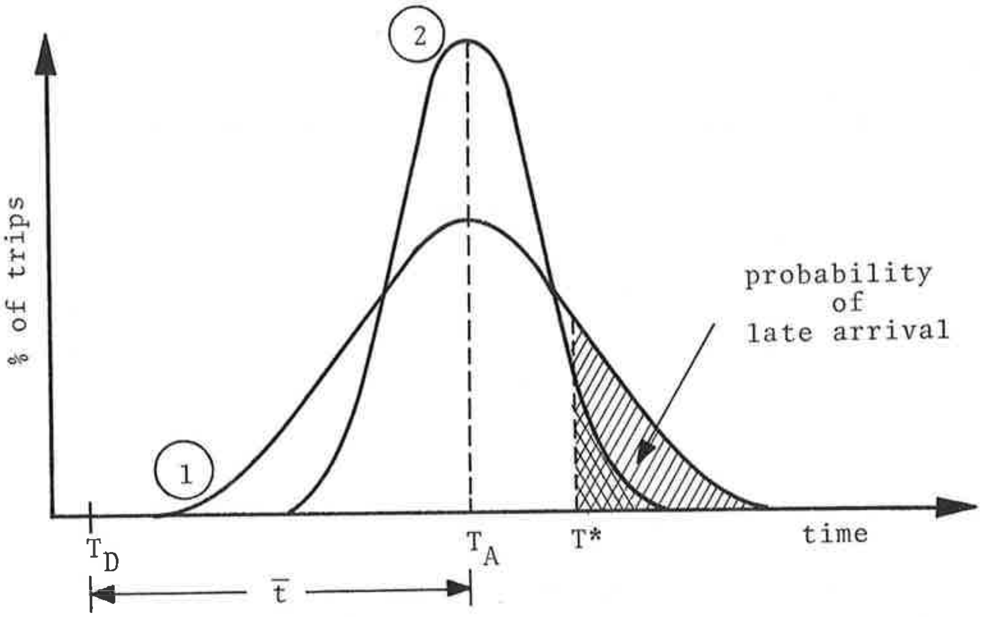


FIGURE 2.3a

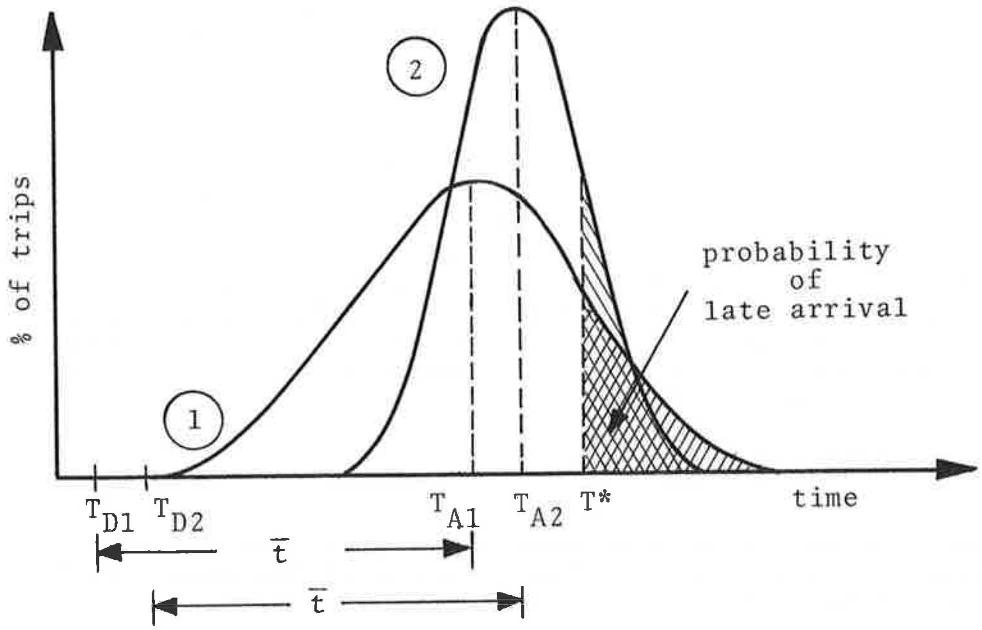


FIGURE 2.3b

FIGURE 2.3 CHANGES IN DEPARTURE TIME BEHAVIOR IN RESPONSE TO A RELIABILITY IMPROVEMENT been reduced system-wide on DRT (without increasing the mean values of these components), all DRT users, will, over the long run, find the DRT system more attractive because of the reduced uncertainty associated with taking a trip.

2.3.2.2 Enabling Transit Travelers to Leave Later on Their Trip

Some travelers may choose to take the reliability improvement by departing later on their trip while maintaining their former probability of on-time arrival. Recall that travelers for whom it is important to arrive at their destination at a particular time adjust to the variability of arrival time at a destination by choosing a premature departure time, P*, which ensures that the risk of arriving late at their destination is not too great. For each individual and trip, P* is directly proportional to the individual's degree of risk aversion and to the total travel time variability of the trip. Reducing this variability will enable each person to decrease P* and thus depart later on their trip. This case is illustrated in Figure 2.3b. Again, distribution 1 represents the traveler's arrival time distribution before a reliability improvement is instituted. With the reliability improvement, the traveler can depart later (at T_{D2}) and maintain the same probability of on-time arrival, even though mean travel time remains the same (see distribution 2).

The effect that reduced variability of transit vehicle arrivals at passenger boarding points has on reducing passenger mean wait time will also enable many transit users to leave later on their trip. All transit travelers who arrive randomly at their bus stop, either because bus arrivals are unpredictable or because the travelers are not familiar with the system, will automatically experience reduced mean wait times when the variability of bus arrival times at their boarding points decreases. Recall that the expected wait time E(w) of passengers who arrive randomly at their boarding point is:

$$E(w) = \frac{H}{2} + \frac{V(h)}{2H}$$

where H is the mean headway and V(h) is the variance of the headway distribution. Thus, E(w) will decrease when the variation in the headway distribution decreases.

Reductions in mean wait time may also be realized where bus arrivals are predictable and travelers who are familiar with bus arrival patterns set their arrival time at the bus stop to minimize their combined disutility for wait time and wait time variability. When the predictability of bus arrivals is increased, these travelers will, by resetting their arrival time, be able to decrease either their wait time, their wait time variability, or both, depending upon the particular bus arrival patterns and individual preference.

2.3.3 Quantifying Reliability Improvements

Benefits derived from transportation improvements have traditionally been quantified using measures of travel costs, travel time savings, and value of travel time. Reliability improvements will, in addition to reducing travel costs and times, introduce a new set of traveler benefits which must be quantified. The reduction in variability of travel times will have a different value than the reduction in mean travel time. In many cases the improved predictability from more reliable service will decrease the number of late and missed appointments, reduce schedule delay (e.g., early arrival at work), lessen the occurrence of excessively long waits for transit service, etc. Thus, new measures are needed to quantify the benefits derived from improved reliability, an issue which is addressed in more detail in Chapter 4.

2.4 CONCLUDING REMARKS

Chapter 2 has identified the importance of reliability in the travel decision-making process, and has established a theoretical basis for understanding the relationship between reliability and travel choice. In making both mode choice and trip departure time decisions, travelers appear to make tradeoffs between alternatives with shorter travel times and those with improved reliability (among other factors) based on their relative preferences. For work trip departure time decisions (as well as other decisions), travelers are hypothesized to account for the probability of being late (or early) in addition to the mean and variability of travel times. To date, there has been a paucity of empirical work directed at understanding the effects of reliability on mode choice and departure time decisions. Further research is needed in this area; several proposed efforts are described in Chapter 7.

REFERENCES

1. Paine, F.T., A.N. Nash, S.J. Hille and G. A. Brunner. Consumer Conceived Attributes of Transportation. College Park, MD, Dept. of Business Administration, Univ. of Maryland, 1967.

2. Spear, Bruce D. "Generalized Attribute Variable for Models of Mode Choice Behavior." Transportation Research Record 592. 1976, pp. 6-11.

3. Stopher, Peter R., B.D. Spear and P.O. Sucher. "Toward The Development of Measures of Convenience for Travel Modes." Transportation Research Record 527. (1974), pp. 16-32.

4. Wachs, Martin. "Consumer Attitudes Toward Transit Service: An Interpretive Review." AIP Journal. (January 1976), pp. 86-104.

5. Golob, Thomas F., et al. An Analysis of Consumer Preferences for a Public Transportation System. (GMR-1037), General Motors Research Laboratory, Warren MI, 1970.

6. Keck, Carol A. and G.S. Cohen. "Characteristics and Attitudes of Dial-a-Bus and Park-and-Ride Users in New York State." Transportation Research Record 590. (1976), pp. 21-25.

7. Winchester, S.A., "Haddonfield Dial-a-Ride Demonstration, Third Household Survey." MITRE Corp., McLean, VA, 1976.

8. Paine, F.T., A.N. Nash, S.J. Hille and G.A. Brunner. "Consumer Attitudes Toward Auto vs. Public Transport Alternatives." Journal of Applied Psychology. Vol. 53, No. 6. (1976), pp. 472-480.

9. Manheim, Marvin L. Fundamentals of Transportation Systems Analysis. MIT, Preliminary Edition, Third Revision, August 1975.

10. Ben-Akiva, Moshe and S. Lerman. Forecasting Models in Transportation Planning, paper prepared for presentation at the Conference on Population Forecasting for Small Areas, May 1975.

11. Ben-Akiva, Moshe. "Review of the Structures of Current Approaches to Travel Demand Modelling." (PhD. Thesis, MIT, Ch. 5, June 1973).

12. Manski, Charles F. The Stochastic Utility Model Of Choice. October 1973.

13. Marfisi, et al. Economic Criteria for Federal Aviation Agency Facility and Equipment Expenditures. WP-210-U4-108, U.S. DOT, July 1975.*

14. Friedman, R.A. "The Effects of Transit Reliability on Passenger Waiting Times." (M.S. Thesis, The Transportation Center, Northwestern Univ. December 1975).

15. Jolliffe, J.K. and T.P. Hutchinson. "A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop." Transportation Science. 94. November 1975, pp. 248-282.

16. Holroyd, E.M. and D.A. Scraggs. "Waiting Times for Buses in Central London." Traffic Engineering and Control. 8(3)., (July 1966), pp. 158-160.

17. Seddon, P.A. and M.P. Day. "Bus Passenger Waiting Times in Greater Manchester." Traffic Engineering and Control. 15 (9). (1974), pp. 442-445.

18. Starkie, D.N.M. "Modal Split and the Value of Time, A Note on Idle Time." Journal of Transport Economics and Policy. Vol. 5, No. 2, (1971).

19. Lerman, Steven, et al. A Method for Estimating the Patronage on Demand-Responsive Transit Systems. Cambridge Systematics, DOT-TJT-77, 1977.

20. Gardner, Wendall and F. Reid. Transit Reliability Data for Disaggregated Traveler Demand Models. Working Paper No. 7603, Univ. of California. 1977.

21. Prashker, Joseph N. "Development of a Reliability of Travel Modes Variable for Mode-Choice Models." (PhD. Dissertation, Northwestern Univ., Dept. of Civil Engineering, Evanston, IL, June 1977).

22. Cosslett, Stephen. "The Trip Timing Decision for Travel to Work by Automobile." Demand Model Estimation and Validation. The Urban Travel Demand Forecasting Project, Phase I. Final Report, Vol. V, Institute for Transportation Studies, Univ. of California, June 1977.

*Material on file: U.S.Department of Transportation,Transportation Systems Center, Cambridge MA.

3. RELIABILITY AND OPERATOR BEHAVIOR (FIXED ROUTE BUS SYSTEMS)

Bus operators as well as travelers are affected by the inherent and exogenously induced variability of travel time components. The effects of bus service variability must be accounted for in scheduling drivers and vehicles on a fixed route system. Operators must allocate additional drivers and vehicles to maintain scheduled headways. Otherwise, the service provided may be considerably less predictable and the operator may risk losing riders and revenue. Whatever they do, bus operators are faced with a situation where reliability has a direct impact on the net cost of bus operations.

This chapter, organized in a manner similar to Chapter 2, will explore the effects of service reliability on fixed route bus operations and the operator decision-making process. Section 3.1 provides an examination of the bus operator's view of reliability based on discussions with bus operators of fixed route systems in ten metropolitan areas. Section 3.2 presents a more theoretical discussion of operator behavior. The material in this section explores how operators might deal with the tradeoffs between average travel time and the variability of travel time that affect travel behavior. Section 3.3 provides a discussion of the potential operator benefits from improved bus reliability.

The material presented in this chapter focuses entirely on fixed route bus systems. However, the theoretical discussion of operator behavior and potential operator benefits of improved reliability appear applicable to all forms of fixed route transit, and perhaps, for demand responsive transit systems as well.

3.1 OPERATOR VIEW OF RELIABILITY

In order to understand how fixed route bus operators in the United States address the problem of unreliable service, discussions with bus properties in ten U.S. cities were conducted in February, 1977. The bus systems in all but two of these cities serve areas of over one million persons. Table 3.1 describes many of the important characteristics of each system: fleet size, number of routes, route types, route miles, headways, in-service speed, yearly ridership, and population in service area.

Most of the information for each bus property was provided by their Scheduling and Operations departments. Discussions concerned the bus operator's perception of

TABLE 3.1 CHARACTERISTICS OF THE BUS SYSTEMS INTERVIEWED BY TELEPHONE

	IN-SERVICE VEHICLE FLEET SIZE SIZE		NUMBER OF ROUTES		ROUTE MILES ANNUALLY	TYPES OF ROUTES	AVERAGE HEADWAYS (minutes)		IN-SERVICE SPEED	YEARLY RIDERSHIP	POPULATION IN SERVICE AREA
CITY	BASE PERIOD	PEAK PERIOD	BASE PERIOD	PEAK PERIOD	(millions)		BASE PERIOD	PEAK PERIOD	(miles per hour)	(millions)	(millions)
1	260	745-827	Less than 20	120	25.7	9% suburban local 9% express 82% urban local	10-60	3-20	12	100.0	2.0
2	270	722-732	142	142	25.0	40 express 102 local	10-60	3-30	13.5	65.0	2.0
3	504-1,376	2,066-2,849	175	195	18.5	12.5% express 12.5% urban 75% suburban-urban	5-15	2-10	14.1	127.8	7.0
4	215	461-489	73	89	22.7	16 commuter 45 city local 28 county local	15-60	3-30	11	42.8	1.25
5	375	776	150	171	37.4	28% express 72% local	10-15	5	12.5	109.6	1.6
6	30-129	313-375	39	45	10.4	11% express 78% local 11% local crosstown	30	9	N.A.	30.6	0.75
7	300	900	85	120	28.0	N.A.	30-60	2-5	N.A.	63.0	2.0
8	70	150	10	10	3.95	3 express 4 urban 3 suburban-urban	20	10	11.8	11.0	0.35
9	130-224	620-671	Less than 126	126	26.8	6 crosstown 3 express remainder local	20-30	3-9	14	77.0	1.1
10	310	815	178	195	24.8	13 express 117 urban 65 suburban	7-120	2-60	12	143.0	2.8

N.A. = not available

reliability and its importance to his operation. The discussions focused on three key issues operators would face when dealing with unreliable bus service:

(1) The treatment of travel time variability in bus system design.

(2) The methods used to identify reliability problems.

(3) The strategies used to prevent and correct reliability problems.

Reliability from the bus operator's perspective is viewed mainly as adherence to the schedule. In fact, many of the operators consider service reliability and schedule adherence to be identical. As will be seen in this chapter, the operator's view of service reliability differs somewhat from the traveler's view of service reliability. Schedule adherence is considered to be very important in maintaining efficient bus system operations. Only one operator downplayed the importance of reliability relative to other facets of system operation. While all of the operators view reliability mainly as an operational consideration, most do believe that reliability is important as part of providing good service to passengers.

3.1.1 The Treatment of Travel Time Variability in Bus System Design

All of the operators build some amount of recovery time into their schedules (recovery time is the time allowed in the schedule between the scheduled arrival time at the destination and the scheduled departure time of the return trip). Recovery time enables buses to leave on schedule for their next run in spite of the variability of arrival times of the buses at the end of the previous run. Recovery time is also required by labor union contract. Only four of the operators stated that the aforementioned reliability consideration was a primary reason for scheduling recovery time. Two of the operators stated that recovery time is largely the result of chance, being the amount of time needed to maintain clock face schedules (e.g., if one-way running time is 27 minutes, a recovery time of three minutes is allowed to make clock face scheduling easier). Four other operators stated that recovery time is set to the minimum values required in labor union contracts. Many operators noted that minimum values required by labor agreements are often the governing factors. Recovery time is generally kept to the minimum possible because this is thought to be an efficient operating strategy.

None of the operators build slack time into the scheduled running time to compensate for the variability of bus running time. Running time is purposely kept tight to operate the system close to its limit. All required slack time is built into recovery time. This also helps prevent early departures which are more onerous to passengers.

Another operator choice which affects reliability is the availability of spare drivers and vehicles. Spares are needed to maintain system reliability in case of vehicle breakdowns or other contingencies requiring more vehicles than scheduled. Each of the operators allocates at least eight percent of its vehicle fleet for backup duty. Each of the operators has spare drivers available, but operator policy is not always clearly discernible. In many cases, the spare drivers are obtained from "extra boards", which are driver assignments designed to cover some or all of the following: other drivers' sick leave and vacation, short runs (i.e., lines too short to be considered for "open bid" by drivers), vehicle breakdowns and other contingencies. The number of spare drivers available to be deployed in addressing reliability problems varies from day to day depending on the number of drivers required to cover other contingencies. Some operators reserve a percentage of spare drivers for addressing reliability problems. In one operation, drivers in the garage (whose runs have terminated) are informally drafted into service when the need for a spare vehicle arises. This technique may be restricted on some systems due to labor rules.

Table 3.2 summarizes the treatment of travel time variability for each bus system.

3.1.2 Identification of Reliability Problems

The operators generally identify and address three categories of reliability problems: (1) in-service vehicle breakdowns and accidents, (2) vehicles not adhering to schedule on a given day because of an abnormal circumstance, and (3) vehicles not adhering to schedule on a "recurrent" basis.

About half of the operators define non-adherence to schedule in a precise fashion (e.g., when the vehicle is a set number of minutes late). All would consider any level of schedule non-adherence to be a problem when it results in poor system operation or considerable passenger and driver complaints. While all of the operators monitor passenger overcrowding at the same time they monitor schedule adherence, the operators do not typically view overcrowding

TABLE 3.2 TREATMENT OF TRAVEL TIME VARIABILITY

City	Recovery Time/ Layover	Spare Vehicles	Spare Drivers*
1	Specific to each route as needed	8%	10-15%
2	3-5 min. minimum depending on day of week (based on union rules)	9%	15%
3	5 min. minimum (based on union rules)	10%	38%
4	5 min. minimum (based on union rules)	10%	no rule on spare drivers
5	5-20 minutes	20%	some, but not specified
6	3 min. minimum (based on union rules)	20	floaters=25%
7	5 minutes	10%	2 per garage
8	10 min. average	10%	varies by time of day
9	Specific to each route to as needed	9%	informal drafting of anyone available
10	Function of headways	15%	12%

*This may include "extra board" assignments.

as a reliability-related problem. Overcrowding is viewed as the result of insufficient capacity, the solution to which is the addition of more vehicles.

The operators employ a number of systematic methods to identify reliability problems: road supervisor reports, driver reports, and periodic monitoring. These methods are summarized in Tables 3.3 and 3.4. In addition, passenger and driver complaints, and noticeable operational problems (e.g., buses always returning late to the garage) will alert the operator to possible reliability problems. The comprehensiveness and frequency of the operator's efforts vary considerably among operators (even taking the differences in system size into account).

In all systems, a number of supervisors in radio-equipped cars generally devote full-time effort to monitoring en route performance. Supervisors are on duty during all hours of bus operation. The number of supervisors varies considerably among bus systems. Generally, a supervisor is assigned to a particular quadrant of the city or to the downtown area. The supervisor cruises the assigned area, checking whether the buses are running according to schedule and identifying vehicle breakdowns, accidents, unusual delays and overcrowded conditions. The road supervisor's role is mainly to oversee the day-to-day operation of the bus system and to assist in maintaining its smooth operation.

Much of the information the road supervisor collects could also be obtained through radio contact between bus drivers and dispatchers. All of the bus properties have equipped, or are in the process of equipping, their entire bus fleet with two-way radios. While not all of the operators have so stated, it would appear that all bus drivers would immediately notify their dispatcher via radio in the event of an accident or breakdown. A few of the operators implied that their bus drivers might notify the dispatcher if they were running very late. However, only two operators reported that bus drivers are required to notify the dispatcher if they are more than a certain number of minutes late.

Another technique for identifying reliability problems is through periodic monitoring. Periodic monitoring may consist of point checks and/or riding checks. All but one of the bus systems use point checks. Point checks generally involve stationing checkers at points along a bus route, usually maximum passenger loading points, to record bus arrival times and determine whether the buses along the route are adhering to schedule. In three systems, passenger

TABLE 3.3 DYNAMIC MONITORING OF RELIABILITY PROBLEMS

City	Type of Dynamic Monitoring	Contact with Dispatcher	Action
1	road supervisors	none	dynamic rescheduling of buses through shorter recovery times or shorter bus routes
2	road supervisors	via radio	rescheduling for breakdowns, but not for periodic tie-ups
3	road supervisors	via radio	rescheduling for breakdowns; written reports for periodic delays
4	road supervisors	via radio	action dependent on situation
5	road supervisors	via radio	action dependent on situation
6	road supervisors	via radio	immediate action for a breakdown; more analysis done for periodic problems
7	road supervisors	via radio	immediate action for breakdowns; more analysis for periodic problems
8	stationary and road supervisors	via radio	action dependent on situation
9	road supervisors	via radio	rerouting, adding extra buses, scheduling modifications
10	road supervisors	via radio	adding extra buses

TABLE 3.4 PERIODIC MONITORING OF RELIABILITY PROBLEMS

City	Type of Periodic Monitoring	Frequency of Monitoring	Action
1	Point checks	three day check for each route	Riding checks done for identified problem routes, resulting in more buses or schedule modifications
2	Point and riding checks	all routes quarterly	Schedule changes on problem routes
3	Riding checks	all routes three times annually	More checks for problems, resulting in addition of more buses or schedule modifications
4	Point and riding checks	all routes at least twice annually	Evaluation of problem routes, resulting in schedule changes
5	Point and riding checks	all routes quarterly	More checks done on problem routes, resulting in more buses or schedule changes
6	Point checks	all routes monthly major routes weekly	More checks done on problem routes; recommendations to Scheduling department
7	Point checks	several days on each route	More analysis of routes with consistent schedule deviation
8	Point checks	all routes once a year	No systematic process, schedule changes sometimes recommended
9	Point and riding checks	all routes twice annually	Schedule changes for chronic problems
10	Point and riding checks	all routes quarterly	Quarterly review. Further analysis for problem routes. Eventual schedule and route changes.

counts and bus arrival times are recorded at the same time. Only one bus system relies entirely on riding checks, which involve placing checkers on buses to record the bus running time profile and passenger load characteristics. Five other operators regularly use both riding checks and point checks.

The amount of periodic monitoring varies considerably among operators. Some operators monitor regularly twice a year, and others three or four times a year; one operator monitors specific routes on a weekly basis. Monitoring is done by one operator at unexpected dates because the element of surprise is considered to be important. Some operators monitor or attempt to monitor every scheduled bus in their system, and others select only the major routes or some representative sample of buses. Both off-peak period routes and peak period routes are generally selected. During a particular monitoring period, each bus is monitored only once, although two operators construct a time profile from multiple-day checks of a route.

Identification of reliability problems (i.e., buses not adhering to schedule and severe overcrowding) with periodic monitoring requires a systematic review of the data collected, which is, in fact, done by all but one of the operators. The review is performed by the Scheduling department either directly after the monitoring effort has been completed or continually as the data comes in. A few operators have programs which tabulate the data. The quality of the review is a function of both the comprehensiveness of the monitoring effort and the quality of the data collected. Adherence checks at a number of points along a route coupled with passenger counts provide better data for analysis than an adherence check at a single point without passenger counts. Even better data for analysis is provided by the operator who obtains data for a particular route over several days. Qualitative data obtained by one operator makes a systematic review difficult. None of the operators have detailed criteria for using the data obtained from periodic monitoring to identify potential problems other than a generalized rule that if a bus is more than a set number of minutes behind schedule, there is a potential problem.

The operators are also made aware of potential reliability problems through passenger complaints of highly variable wait and/or travel times on a particular route or bus. Some operators view passenger complaints as the single most important source for identifying potential reliability problems. Frequent complaints from bus drivers who regularly run late also alert the operator to potential problems.

For many of the potential reliability problems spotted by periodic monitoring and some of those spotted by other means, neither the cause of the problem nor its recurrent nature is known. Both pieces of information are needed for the operator to decide whether to correct a potential problem and how to correct it. None of the ten operators will seek a permanent solution to a reliability problem unless the problem is recurrent. Each operator appears to have his own definition of a recurrent problem. For example, some operators have mentioned that they don't consider delays produced by "occasional" traffic tie-ups to be recurrent problems.

To determine the cause and recurrent nature of a potential problem, several operators have indicated that they continue to monitor problem areas that they have spotted. One operator performs three days of riding checks on problem buses identified through the periodic monitoring process, noting bus arrivals and passenger loads at selected points. One operator performs a single riding check on problem buses spotted during periodic monitoring, taking another check if the problem still crops up. This same operator will take three days of riding checks on problem buses spotted by other means. Another operator performs point checks for two or three days (passenger counts included) on problem buses spotted by road supervisors, drivers, or passengers. With this operator, the fact that periodic monitoring of a particular route is itself taken over a three to four day period obviates the need for further checks after data obtained from periodic monitoring has been reviewed. Two other operators perform a single riding check on problem buses spotted during periodic monitoring. One operator indicated that checkers were sent out when the cause of an identified problem was unknown. Another operator stated that a review of the data obtained from periodic monitoring may lead to personal monitoring by the Director of Plans and Schedules or his staff.

The operators who did no further monitoring seem to have placed less emphasis on the review of periodic monitoring for spotting potential problems and more emphasis on road supervisors and passenger complaints. One operator places less emphasis on spotting reliability problems in general and feels that reliability is not as important relative to other facets of his operation.

3.1.3 Correction of Reliability Problems

Efforts to correct reliability problems also vary considerably among the operators. All operators stated that when a bus breakdown or accident has been identified, either by the bus driver or road supervisor, the road supervisor would determine whether corrective action is needed. A spare vehicle would, where feasible, be dispatched to the location of the incident to finish out the route. Several operators noted that they are currently upgrading their fleet with new buses, hopefully leading to fewer vehicle breakdowns.

Only five operators stated that they might respond immediately to situations where a bus or buses are running far behind schedule, or there is a severe overcrowding condition, as observed by a road supervisor or bus driver. Generally, this would involve sending out an extra bus on the entire length of the route or only for the most heavily utilized portion of that route (this latter bus is called a "header"). These five operators said they might reschedule future runs or even detour buses when a bus was sufficiently behind schedule. The other five bus operators take no immediate action when a road supervisor or bus driver reports that a bus or buses are running far behind schedule, or there is severe overcrowding. However, if the operators found these problems recurring, they would take action.

All the bus operators attempt to remedy recurrent reliability problems. Where the recurrent problem is caused by a traffic bottleneck, rerouting the bus around the bottleneck may solve the problem. Some operators will assist the bus drivers in speeding up their runs when the drivers are regularly running behind schedule. In the more general case, the operators indicated that when buses were consistently running behind schedule, they would adjust their schedules to reflect actual running time, or, in other words, would set scheduled arrival times to minimize the deviation of actual arrival times from the scheduled times. None of the operators mentioned that recovery times would be increased to counteract the effects of late arrivals at the termination of runs. When severe chronic overcrowding conditions are found to exist, an additional bus or buses may be scheduled on the route in question. Recurrent overcrowding is treated as a problem caused solely by a lack of capacity and not by the unreliability of particular buses.

3.1.4 Summary of Operator Policy

It is apparent that at least nine of the ten bus operators consider the maintenance of stable route operating characteristics to be one of their major concerns, and they often take specific actions to achieve this end.

Reliability from the operator's perspective is treated primarily as an operational consideration (i.e., adherence to schedule). It is within this framework that operators attempt to provide reliable service to their patrons.

When schedule adherence problems have been found to be chronic, and neither rerouting nor bus driver diligence can improve the situation, the operator typically adjusts the schedule to prevent the propagation of schedule adherence problems from run to run. For example, where buses are <u>consistently</u> arriving at their destination X minutes late, the operator may increase the scheduled arrival by X minutes. However, this may have little effect on bus users, whose view of the reliability of the buses will not change if they are aware that the buses consistently arrive X minutes late at their destination. Operators tend to take little action when buses arrive at their destination on time on the average, but with a great degree of day-to-day variability. Unfortunately, bus users view this situation as being unreliable because of the high degree of variability in day-to-day arrivals.

When chronic overcrowding conditions exist, the operator may schedule additional buses on a route if they are available. If the overcrowding is caused by unreliable service, the addition of buses may do little to improve the situation (see Chapter 5 for a better understanding of this point).

In day-to-day situations where bus users are experiencing extreme delays and overcrowding, only half of the bus operators contacted indicated that they would take immediate corrective action. The general solutions suggested (sending out an extra bus over either a portion of or the entire length of the problematic route, rescheduling future runs, and detouring buses) are aimed mostly at restoring schedule adherence on the system and relieving overcrowding. While generally such strategies will improve reliability experienced by users, in some instances these modifications could actually worsen users' reliability. For instance, rescheduling future runs may increase the waiting times of those individuals who set their arrival time at the bus stop according to an expected bus arrival time distribution.

As is evident from the previous discussion, a variety of operator decisions influence the reliability of bus service. Although many of these decisions implicitly involve tradeoffs between travel time and travel time components (such as wait times and ride times), and component variabilities experienced by bus users, the

T_V that can be provided with the given allocation.* Any point in the feasible region not lying on line AB represents suboptimal utilization of vehicle and driver deployment. For example, if the operator is providing service at interior point E, it is suboptimal, because the operator could provide service at point F, where $T_{VF} \leq T_{VE}$ and $T_{MF} \leq T_{ME}$.

To maximize the level of service produced, the operator should operate along AB. The precise combination of travel times and reliability that the operator should produce to maximize the utility of bus travelers depends upon the travelers' preferences for these attributes of bus service. If we assume that all travelers have the same travel demands and preferences, then $\overline{U}_T$, the sum of the disutilities of all trips by all bus travelers on the specific bus route can be expressed mathematically as:

$$\overline{U}_T = N(\alpha T_M + \beta T_V)$$

given that the travel time disutility of a given trip by a given traveler, U_T, is:

$$U_T = \alpha T_M + \beta T_V \text{ *}$$

where:

N = is the total number of trips taken on the bus route
α = disutility coefficient of mean travel time
β = disutility coefficient of travel time variability.

Indifference curves, similar to those in Figure 2.1 of Chapter 2, can be plotted for different values of $\overline{U}_T$. As shown in Figure 3.2, $\overline{U}_{T1} > \overline{U}_{T2}$ (we assume that these user indifference curves are concave). Superimposing the feasible operating region and user indifference curves (Figure 3.3), one can visualize how the operator should set his policy to maximize user satisfaction. In Figure 3.3, the operator can use his vehicles to provide a level of service at point C, point G, or at point H. Since the operator is indifferent to operating at a particular point on line AB, it would be in his best interests to provide service at point H. Operating at that point would maximize user satisfaction, and hence ridership.

---

*Physically, as shown in Figure 3.1., line AB represents a frontier of the <u>maximal</u> combination of $1/T_M$ and $1/T_V$ that can be provided with the given allocation. This is equivalent to the minimal combination of T_M and T_V that can be provided.

---

*See Section 2.1 for a discussion of this equation.

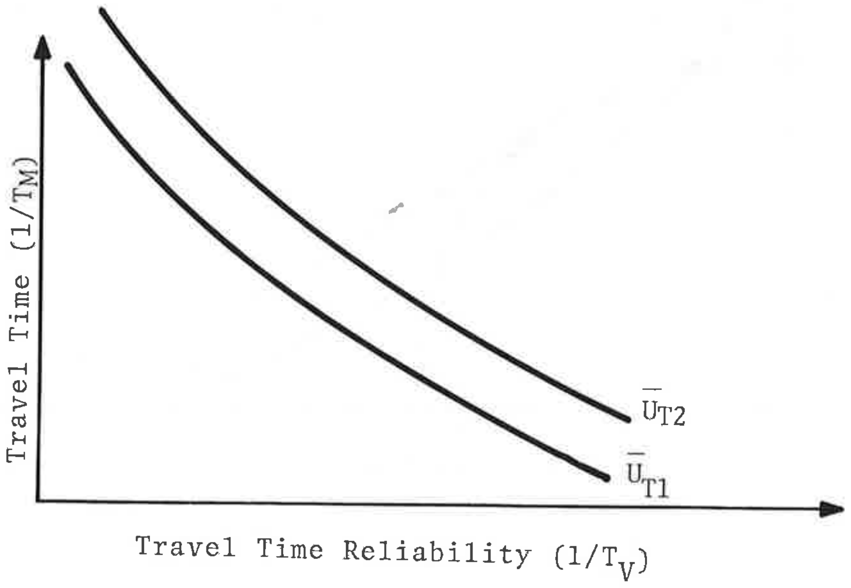


FIGURE 3.2 MEAN TRAVEL TIME - TRAVEL TIME VARIABILITY INDIFFERENCE CURVES FOR A HOMOGENEOUS GROUP OF TRAVELERS

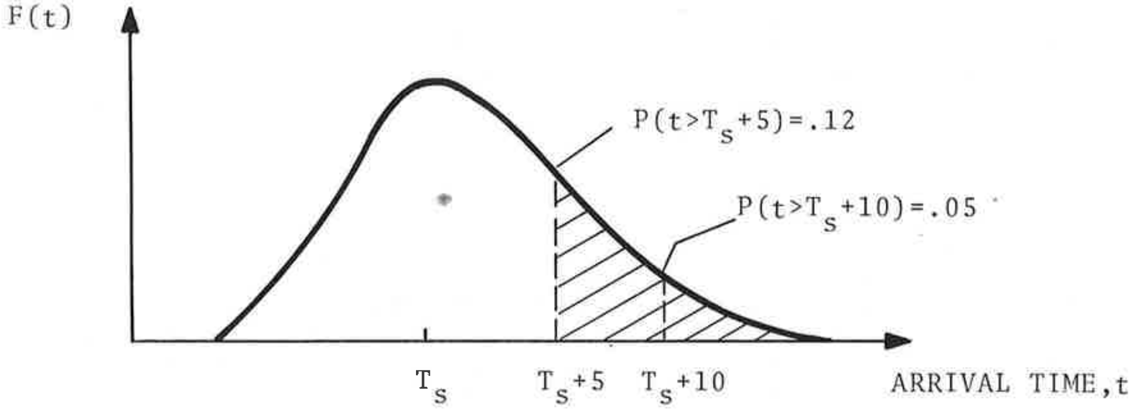


FIGURE 3.4 BUS ARRIVAL TIME DISTRIBUTION

performance.* Only seven buses would be needed and one spare vehicle and driver would be available.

(3) Operate at 12 minute headways and 10 minutes of recovery time to maintain 95 percent on-time performance for the start of the next trip.* Seven buses would be needed and one spare would be available.

From the bus users' viewpoint, the three options involve tradeoffs between average travel time and travel time variability. Some of the more obvious of these tradeoffs that will affect bus user's disutility are those between the potential lower reliability resulting from vehicle breakdowns with Option 1, the potential lower reliability resulting from insufficient recovery time with Option 2, and the longer wait times with Option 3. If the operator has no preference between the three options, implying he is operating along the frontier of his feasible region, he should select the one which will provide bus users with the lowest trip disutility.

This simple example illustrates one of several kinds of tradeoffs the operator makes implicitly in setting his operating policy. Each option will result in a different type of service being provided from the bus users' perspective. To bus users, the combinations of travel time components and travel time component variability provided by certain options may have lower disutility than other options.

In the real world, several of the assumptions made in this theoretical discussion do not hold. Average travel time and travel time variability experienced by individuals would vary along a route. In addition, traveler disutility coefficients (α and β) would vary among individuals and for different trip purposes. The assumption that the bus operator is indifferent to operating at any point along his service frontier curve generally does not hold. The operator may prefer particular options for one or more of the following reasons: they provide more service, adhere more to schedule, generate higher ridership, or produce a lower deficit than the other options. The operator may also

---

*In fact, the on-time performance in all these cases will be worse than indicated here, because whenever a later than planned arrival occurs, the bus arrival time distribution will be shifted to the right for the following run, thus increasing the probability of late arrivals in the future.

be constrained by public service commitments for minimum levels of service.

The operational environment in the real world is also quite different. Bus arrivals along a route are rarely independent, and there is generally a conditional relationship between successive arrivals (bunching effect). In addition, operations along a particular route are affected by interactions with other routes in the network. Finally, because T_M and T_V vary along a route, there will be a different operating region diagram for each point, and the diagrams will be linked together in a very subtle and complex way. Thus, it is not clear that the operator can realistically adopt a strategy of always moving to the frontier of one diagram, because it may not result in operation along the frontier of other diagrams on the route.

While the graphical portrayal of the theoretical discussion would be significantly more complex without the assumptions, it nevertheless demonstrates a valid conceptual approach. Existing operator behavior may have some effect in reducing the trip disutilities of individual bus users and the unreliability of service provided them. It is hypothesized that the operator can further reduce the disutility of bus trips by taking into account the behavior of travelers. In order to test this hypothesis, considerable research is needed on both traveler behavior and bus operations.

3.3 POTENTIAL OPERATOR BENEFITS FROM IMPROVED BUS RELIABILITY

While bus operators may not be currently addressing reliability problems in the most effective fashion, they nonetheless must devote a considerable amount of their available resources to counteract its effects. Unreliable service can significantly impact the operator's costs and may significantly affect system ridership and revenue. Improved bus reliability can impact the bus operator's profitability by affecting both his costs and revenues; the extent to which each of these is affected depends largely on operator behavior.

An improvement in bus reliability provides the operator with the opportunity to increase the availability of drivers and vehicles in some of the following ways. For fixed route service, reduced vehicle travel time variation can enable the operator to reduce the amount of recovery time normally built into schedules. Recall in the theoretical discussion that these layover periods are included in schedules to account for the probabilistic nature of route travel time and are set to provide reasonable assurance that the bus will arrive in time for a succeeding run. Fewer vehicle breakdowns and reduced headway variation (or improved schedule adherence) can enable the operator to reduce the number of spare vehicles and drivers kept on hand. Recall that spare vehicles and drivers are included in the operation in case of vehicle breakdowns and accidents, and may be used to counteract the snowballing effects of resulting temporary instabilities in the bus system's operation. Fewer day-to-day reliability problems can enable the operator to reduce the number of road supervisors needed.

Given the opportunity to increase the availability of drivers and vehicles provided by an improvement in bus reliability, and assuming the operator does not incur any costs to implement the reliability improvement, the operator has three courses of action available to him:*

(1) Do nothing at all.

(2) Take advantage of the increased availability of drivers and vehicles by reducing the driver and vehicle fleet to provide the existing service (or one with comparable level of service) to the existing service area.

(3) Take advantage of the increased availability of drivers and vehicles by increasing frequency and/or coverage of service with the existing driver and vehicle fleet.

The operator who does nothing at all (or is unable to do anything because of labor constraints) maintains the existing dispatching strategy and realizes improvements in schedule adherence. Some of the benefits of improved schedule adherence are, of necessity, benefits to travelers as, for example, increased probability of on-time arrival, reduced mean wait time, and later trip departure times. Improved schedule adherence may enhance the mobility of bus dependent travelers and may induce auto users to switch to transit. These benefits, which are described in Section 2.3, may in turn produce benefits for the operator in terms of increased ridership, and thus, increased revenues. In addition, drivers may realize greater actual layover times.

If the operator selects the second option, it could reduce both his capital costs and operating costs by

---

*In reality, the operator does not have three distinct options available, but rather a much larger set which will encompass varying degrees of each option.

reducing the required fleet size and labor force. Because the operator is providing the same, or similar level of service as before, ridership and revenue should remain unchanged.

If the operator chooses the third option, the improved frequency and/or coverage of service may result in better service for existing bus users and may attract new ridership, thus increasing the operator's revenue. At the same time, operating costs should change little because no additional resources are being expended.

By taking advantage of the increased availability of drivers and vehicles with either the second or third options, the operator is converting improved reliability benefits into totally different benefits (i.e., cost savings with option 2 and improvements in other level of service components with option 3). This may result in some travelers experiencing less reliability than before the reliability improvement was made. For this reason, operator policy decisions are extremely complex in their dimensions and consequences. In designing strategies for improving bus reliability (as may be possible, for example, with Automatic Vehicle Monitoring), it may be particularly important to ensure that an appropriate balance is struck between easily measured benefits to the operator and less quantifiable benefits to travelers.

3.4 CONCLUDING REMARKS

This chapter has examined the impact of reliability on transit operations. The transit operator is likely to incur increasing operating costs and decreasing revenues when faced with a reliability problem. While transit operators recognize the need to control reliability problems, efforts to date have lacked a systematic approach.

In developing an approach to control reliability problems, it is extremely important for the operator to recognize the interrelationship between transportation supply and demand. The operator, to improve efficiency, should be fully cognizant of the demand for travel, so that better service is offered at the lowest possible operating cost. This, unfortunately, is not always the case in bus operations today. All too often, the operator will make decisions which do not take into account the demand for travel, and neither the operator nor the user fully benefits from this practice.

4. MEASURES OF RELIABILITY

The importance of transit reliability to both travelers and operators indicates the need to identify and develop meaningful measures of transit reliability. Developing a set of workable measures would aid transportation planners and the transit industry in:

(1) identifying and undertstanding reliability problems,

(2) identifying and measuring actual improvements in reliability,

(3) relating such improvements to particular strategies, and

(4) modifying these strategies, methods, and designs to obtain greater reliability improvements.

This chapter addresses the selection of appropriate measures of reliability. Section 4.1 presents a discussion of current approaches and measures which have been used in an applied environment. This includes previous, ongoing, and planned Service and Methods Demonstration projects. Section 4.2 addresses several improved approaches and measures that have been developed in theory but have received little attention in an applied context. Measures proposed and discussed in this section were selected by criteria which include explicitness of definition, controlability, expense and accurate observability, and independence. Section 4.3 suggests reliability measures which are thought to be most appropriate for analyzing the effects of transit reliability.

4.1 CURRENT APPROACHES AND MEASURES

Recent Service and Methods evaluations, particularly those involving priority or exclusive bus lanes or freeway entrance ramps, have included measurements of service reliability. In some cases, due to the chosen measures or exogenous influences, few conclusions could be reached about the effects on reliability of the demonstrated methods or strategies. These measures are described along with appropriate comments in Table 4.1.

<u>I-35W Urban Corridor Demonstration</u>: In the evaluation of the I-35W Urban Corridor (Minneapolis) Demonstration of express buses on a metered freeway, transit dependability was defined as providing service to transit patrons that

TABLE 4.1 RELIABILITY MEASURES USED IN PREVIOUS AND ON-GOING SERVICE AND METHODS EVALUATIONS

Demonstration	Measures	Results
I-35 W Urban Corridor	-absolute deviation between actual and scheduled bus departure and arrival times -proportion of buses "on-time"	The impact of metering and express buses will affect both the mean and variability of travel time.A shift in shorter mean arrival time will show improved service for buses classified initially as "late", but will provide poorer service if several buses are now "early", which initially were "on-time". Evaluation of bus reliability based on artificial schedules can distort results. Thus, no conclusions on improved reliability were made.
Shirley Highway Exclusive Busway	-percent over x minutes late. -"on-time" performance -seat availability	Reliability improved for express bus segment. No change noted in downtown travel. Attempts to measure changes in seat availability were unsuccessful due to growth in ridership.
Golden Gate	-reduced lateness from scheduled arrival times -standard deviation of arrival times -change in number of relay runs	Again, there was difficulty in distinguishing improvements in mean travel times and improvements in the variability of travel times. Some improvement was noted from the standard deviation measure. A decrease in relay runs was noted.
San Bernardino	-coefficient of variation of travel time	This was an effective measure because the improvement in standard deviation could be measured relative to changes in the mean. Some improvement was noted.
Sacramento	-change in travel time -change in standard deviation of travel time	Significant improvements occurred. A more than 50% reduction in standard deviation was noted, while a 23% reduction in mean travel time was recorded.

TABLE 4.1 RELIABILITY MEASURES USED IN PREVIOUS AND ON-GOING SERVICE AND METHODS EVALUATIONS (CONTINUED)

Demonstration	Measures	Results
Los Angeles	-mean and standard deviation of "on-time" performance -coefficient of variation of "on-time" performance	No significant difference in mean on-time performance for all inbound, a.m.-trips and for outbound, p.m. L.A. downtown trips. Outbound west L.A. trips were significantly earlier as a result of the Diamond Lane.
Miami	-schedule adherence	Improved schedule adherence with coordinated signal operation and with reversable exclusive bus lane. Signal preemption produced a lower degree of schedule adherence when compared to pre-demonstration service.
Rochester	-mean and standard deviation of system response times, pickup deviations, ride times, and transfer times	Pickup time deviations of over five minutes with coefficient of variation over one for immediate requests, advance requests, and transfers. System response times and ride times were less variable relative to the mean.
	-% of pickups more than x minutes late	15% of pickups were found to be 20 or more minutes early or late (assuming normal distribution about mean).

closely corresponds to that described in published timetables. The evaluation focused on the following objectives (Bather, et al., 1):

> To measure the absolute deviation between actual and scheduled bus departure and arrival times.
>
> To measure the proportion of buses that operate within acceptable deviations from scheduled bus times. A quantifiable measure, "on-time performance", was selected and defined as the probability (percentage of buses) arriving between exactly on schedule and 5 minutes after scheduled arrival time. (In accordance with standards of service for metropolitan-wide transit service, it is preferable that buses be late rather than early so that passengers will not miss their bus and have to wait for the next bus).

Defining measures in terms of deviations from schedules caused problems for evaluation. Since metering or other strategies aimed at improving reliability may also affect mean travel time, the distribution of arrival time may be shifted as well as modified in shape. A shift resulting in shorter mean arrival time will cause the "on-time" performance measure to indicate that service was improved if there were previously a large proportion of late buses (and that service has deteriorated if many on-time buses are now classified as early), even if no change in variance occurs. Thus evaluation relative to artificial and irrelevant schedules may distort results. In fact, no correlation between ramp metering and improved reliability could be made.

<u>Shirley Highway Exclusive Busway</u>: In the evaluation of the Shirley Highway exclusive busway, adherence to schedules was measured for freeway buses before and after implementation of the busway. To determine the effects of the busway on schedule reliability, schedule adherence was examined for buses arriving at their first stop upon leaving the busway. A drastic reduction in "percent over 7 minutes late" was noted (from 67% to 3-4%). It was also noted that heavy patronage increases on busway routes and longer suburban collection segments increased delays during the collection (non-busway) portions of the routes, preventing an even greater improvement in schedule adherence. The use of a different measure such as the variability of travel time between the last stop before the busway and the first stop after the busway might have provided a better estimate of the effects of the busway on reliability (McQueen, et al., 2).

In order to measure the effects of the downtown priority lanes, on-time performance was measured as the travel time between the first and last downtown stops. No significant improvements were evident. This may be due to the effects of subway construction which disrupted downtown traffic. An attempt to measure changes in seat availability was unsuccessful primarily due to growth in busway bus ridership and resulting overcrowding.

Golden Gate Corridor: In this locally-sponsored priority bus lane project, reduced lateness with respect to existing (old) bus schedules was used as the measure for evaluating effects of reliability. However, this reduced lateness could be more a result of reduced mean travel time than reduced variation in travel times. Standard deviation of arrival times was also used to measure reliability. Some improvement was evident, although data was limited (Bigelow-Crain, 3).

Evidence of improved reliability was noted in the fact that operators reported a decided decrease in the number of "relay-runs" required since the advent of the with-flow bus lanes (relay runs serve late morning inbound trips when early morning buses encounter round trip delays and additional buses are needed).

San Bernadino Busway: This locally sponsored evaluation utilized variability ratio (coefficient of variation) of travel time as the reliability measure.* A comparison of busway, downtown and regular freeway bus runs showed variability ratios for busway buses to be consistently smaller, over different time periods of the day. Downtown buses had variability ratios as high as .24 while those for busway routes were in the 0.05-0.07 range. The second year report notes that the decreased variability in travel time produced by the busway permits schedules to be tightened, thus improving operator efficiency (Bigelow-Crain, 4)

Sacramento: A locally sponsored project in Sacramento, California, examined the effects of bus pre-emption of traffic signals along an arterial on the reliability of express bus service. Reliability was measured in terms of changes in the mean and standard deviation of travel time. Express buses equipped with pre-emption devices were able to

---

*The coefficient of variation is expressed mathematically as travel time standard deviation divided by the travel time mean.

reduce travel time along the 3.8 mile stretch of arterial by 23% and reduce the standard deviation of travel time by over one half (Elias, 5).

A number of demonstrations are currently undergoing evaluation. Among these are priority-lane schemes in Los Angeles (Santa Monica Freeway) and Miami, and integrated transit in Rochester.

Los Angeles: On the Santa Monica Freeway, on-time performance was monitored by checkers for a.m. inbound and p.m. outbound trips (SYSTAN, 6). The percent early and late buses were recorded, as well as the mean and standard deviation of the difference between actual and scheduled arrival times. Comparisons for before and during the Diamond Lane were conducted using the coefficient of variation. The hypothesis that no difference in reliability occurs was tested*; no significant difference in mean on-time performance was noted for a.m. inbound trips or the p.m. outbound trip from downtown Los Angeles. However, the p.m. outbound trip from West L.A. was found to be significantly earlier during Diamond Lane operation.

Miami: Bus priority treatments, coordinated signals, and bus/carpool lanes on the freeway were important components of this locally sponsored demonstration project on I-95 and NW 7th Avenue. For each staged operational change schedule adherence was measured (Wattleworth, et al., 7).

The comparative analysis of schedule adherence was conducted by specifying the distribution of bus arrival times according to five categories: (1) early (more than 5 minutes), (2) early (less than 5 minutes), (3) on time (less than 2 minutes late), (4) late (between 2 and 7 minutes), and (5) late (greater than 7 minutes). The evaluators found that schedule adherence was improved with coordinated signal operation and with the reversible exclusive bus lane. Signal premption by the buses tended, on the other hand, to produce a lower degree of schedule adherence.

Rochester: Evaluation of integrated transit in Rochester, New York presents some different problems. This is due to the fact that dial-a-ride (DAR) operation is a more complex service and therefore involves more reliability factors than conventional bus service. As discussed in Chapter 2, the dynamic routing inherent in DAR will produce

---

*Hypothesis tests using statistical techniques were employed at the =.05 level.

considerable variability in both wait time and in-vehicle travel time for the same trip on different days. Although dial-a-ride has no published schedules, calltakers generally establish promised times of pickup or delivery. Not only will these promised times vary from day to day but so will the accuracy of these predictions. It might also be difficult to distinguish between unreliability caused by variation in demand and unreliability due to a poorly managed system (e.g., poor dispatching) (Wilson and Colvin, 8).

Many aspects of reliability were measured in the evaluation of Rochester (Greece) DAR service (SYSTAN, 9):

> mean and standard deviation of system response time (time between request and pick-up) for immediate requests.
>
> mean and standard deviation of pick-up deviation (time between promised and actual pick-up) times for immediate requests and advance requests.
>
> mean and standard deviation of ride times for immediate requests and advance requests.
>
> mean and standard deviation of fixed route - DAR transfer times.

During the evaluation period, the actual pick-ups for immediate service occurred an average of five minutes after the promised pick-up times, and the typical advance request pick-up occurred seven minutes late. Furthermore, the actual pick-up times varied widely around the promised times, making it difficult for users to plan their travel schedules (15% of pickups were 20 or more minutes early or late*). In the transfer study, 14% of the reported outbound transfers to DAR resulted in waiting times of at least 30 minutes. System response times and ride times had coefficients of variation around 0.6; the transfer time coefficient of variation averaged around 0.8; and the pick-up time coefficients of variation were 1.5 to 2.4 for advance and immediate requests, respectively.

Reliability on dial-a-ride might also be measured by comparing actual travel times for similar trips as they vary from day to day. Comparison of promised and actual times may show service to be good, while in reality, promised times may vary considerably from day to day for a given

---

*Assuming a normal distribution about the mean.

trip, leaving the user with a poor perception of reliability. Because of the sensitivity of DRT reliability to demand, perhaps reliability should be represented as a function of demand density and fleet availability.

In conclusion, some evaluations have already noted the potentially substantial effects of transit service improvements on reliability. Yet some of these evaluations utilized measures that were inappropriate and troublesome. The wide variety of measures of lateness and deviations from published schedules which were often not adjusted properly to reflect the impacts of the project and exogenous influences have led to inconsistencies which make comparison of reliability levels between projects virtually impossible. The measures used in San Bernadino and Sacramento have attempted to address this problem. Furthermore, few measures have attempted to capture the projects' effects on transit reliability as it is experienced by travelers. Most measures have focused on bus travel time and its relationship to operator oriented measures of scheduled travel and arrival times. The mean and variability of passenger travel and wait times have only been considered in the evaluation of Rochester. Finally, the theoretical discussions of the previous two chapters illustrate that measures of reliability should be longitudinal with respect to time (i.e., taken across several days) to address the variability of transit service which influences travel and operator behavior. This consideration has been totally lacking in current and previous demonstration evaluations.

4.2 IMPROVED APPROACH AND MEASURES

This section defines a framework for selecting an improved set of measures which may prove to be useful for reliability studies. A set of relevant service attributes is selected and alternative means to measure their variability are examined. For this preliminary study of measures, discussion will focus on fixed route bus with the intention of extending the results to other modes in subsequent discussions.

4.2.1 Relevant Service Attributes

The set of attributes must capture variability effects on traveler behavior as well as variability impacts on transit operators. Because the effects on travelers and operators are not the same, a separate set of attributes is needed for each. A short list of service attributes whose variability should be important to transit travelers include:

- total travel time
- wait time (including transfer time)
- in-vehicle time
- seat availability.

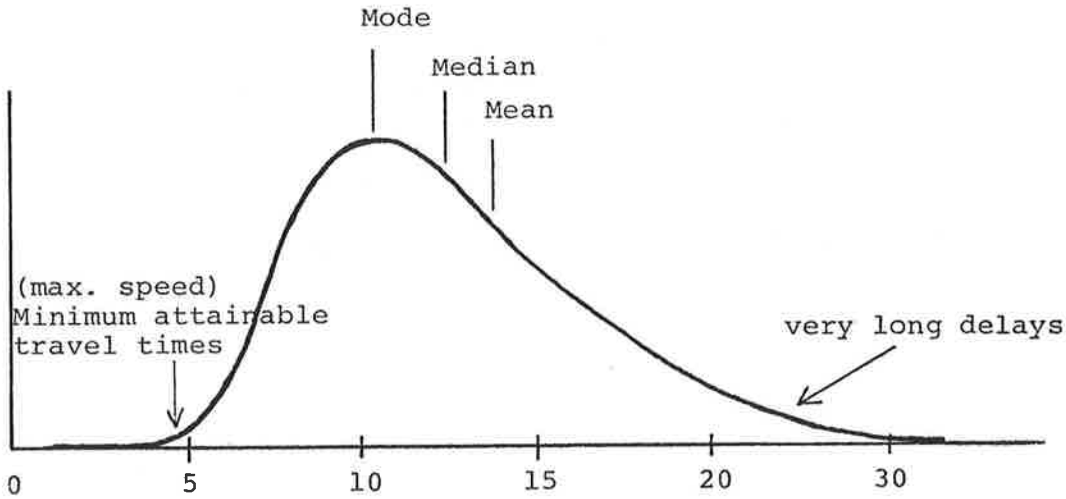
A short list of service attributes whose variability should be important to transit operators include:

- adherence to schedule at origin, destination and intermediate points
- headways
- seat availability.

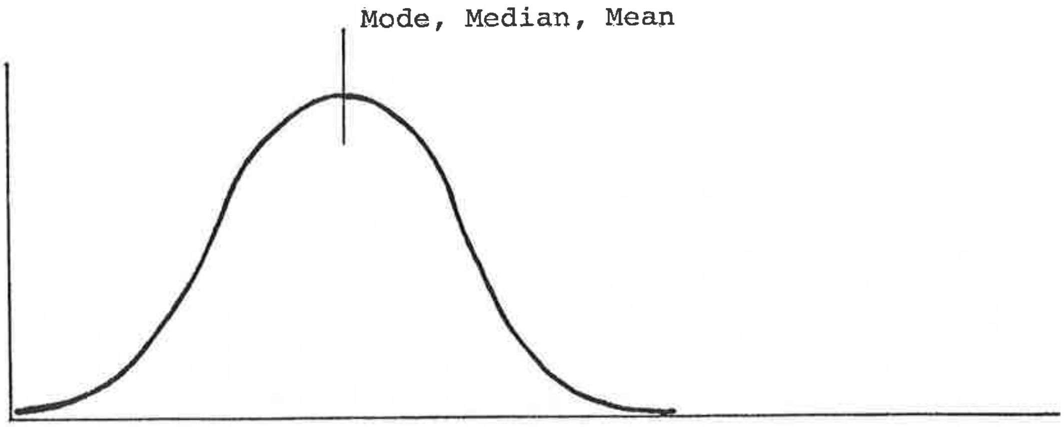
4.2.2 Variability Measures

The suggested service attributes tend to have skewed distributions. Travel times and arrival times are limited on one side by speed constraints while very large delays may be present on the other. Headways and wait times may approach large values but cannot fall below zero. For infrequent service, however, headway distributions may be approximated by non-skewed distributions (see Figure 4.1).

In measuring the variability (unreliability) of bus arrivals, the question arises whether measures should focus on 1) deviation from some fixed time (the schedule) or 2) compactness of the arrival distribution. Although riders who make infrequent use of the service may be totally dependent on printed schedules (if any), regular users may know the system from first-hand experience. In any case, where compactness of the distribution is evident yet deviations from the scheduled time are very large, it is clear that the scheduling and not the inherent unreliability in the system is at fault. Measures should be selected that clearly filter the effects of faulty schedules. Awareness of faulty schedules is important to the operator in that faulty schedules represent a system deficiency; however, because they are not a result of variability in bus operations, they are not an indication of unreliability to the traveler. In particular, in evaluations of strategies to improve reliability, where mean trip times may change and schedules need to be revised, deviation from printed schedules is not an effective measurement of the resulting reliability improvements. Consequently it is important to select measures that reflect compactness of distributions rather than deviations from the schedule in order to measure the variability of service attributes.



a. Skewed Distribution



b. Non-Skewed Distribution

FIGURE 4.1 SKEWED AND NON-SKEWED DISTRIBUTIONS OF TRAVEL TIME

The next problem, then, is to select appropriate measures of compactness. While several measures (both rigorous and intuitive) of the compactness of a distribution exist, the "acid test" is the relation between the performance measure and the service perceived by the user and operator. For example, a bus that is five minutes ahead of schedule (or mean arrival time) may create a much greater delay for the passenger than a bus five minutes late, yet most measures of deviation from central tendency (standard deviation, variance) would weigh both cases equally.

The differences between various measures are typically accentuated by the presence of skew in the observed distribution. Many measures of compactness will not realistically portray the impact of skew for each performance measure. Consequently different measures may be necessary for different performance characteristics.

Another criterion is the selection of a reliability measure which can be used in comparative analyses at a particular site or between sites. Both the user and operator will experience changes in the mean and variability of a service attribute when a service improvement is implemented. Thus, to evaluate the "true" impacts of a service change on reliability, it may be necessary to normalize the reliability measure by including the mean value of the service attribute.

In addition to a measure reflecting compactness of the distribution, it would be useful to have a separate measurement of the likelihood of extremely long delays. For users and operators such measures have significance in identifying the likelihood of a real system failure. In such instances, a user will suffer a serious consequence as a result of delay or have to rely upon an alternate mode. Likewise the operator might find it necessary to add a reserve vehicle to the system or suffer the consequences of system failure. Again, this measure should be normalized by the mean.

Alternative measures are discussed below (Martland, 10), noting the advantages and disadvantages of each. The measures are discussed in terms relating to the travel time distribution for a given bus trip. Similar measures, however, may be developed for other reliability-related attributes.

1. Standard deviation - Transit travel time distributions may be skewed to the right due to minimum travel times and possible long delays. Standard deviation reflects both compactness and

long delays in a single measure, and in doing so, masks the individual effects of each.

2. Standard deviation excluding extreme values - Eliminating the longest travel times before computing the standard deviation vastly increases its usefulness as a performance measure. For skewed distributions the normal interpretation of the standard deviation does not hold; for such distributions, the measure may still not give a reliable estimate of compactness.

3. Variance - Although it has the same problems as the standard deviation, it has the advantage that in cases where trip components are independent, variances of individual trip components may be added to obtain total trip time variance.

4. Coefficient of Variation - An attractive measure because it normalizes the effect of standard deviation relative to the mean and makes comparative analysis between projects possible.

5. Shortest interval in which X% of all arrivals occur - This measure is free of distortion caused by extreme values and when used with the mean, gives a good indication of the compactness of the distribution. Selection of X should be large enough to measure the effects of all but the extreme values of bus arrivals.

6. Maximum % of arrivals in an 'N' minute period - This measure is equivalent to the previous one. It is more appropriate in cases where data is discontinuous. Selection of N, however, can introduce misleading results if the distribution is quite compact relative to the magnitude of N.

7. Time by which X% of all arrivals occur - This measure gives an idea of very long travel times when used in conjunction with the mean. Using this measure in conjunction with the mean and measures of compactness (e.g., coefficient of variation, the maximum % of arrivals in an N minute period, or the shortest interval in which X% of all arrivals occur) yields a set of measures which appropriately describes the distribution. Again, X should be set at a high level.

8. % of trips taking 'N' minutes longer than median or mean (% N minutes late) - This is a useful

measure of very long travel times. Using this measure in conjunction with the mean and measures of compactness yields a set of measures which appropriately describes the distribution. The value of N selected should be large enough to capture solely extreme value effects.

9. <u>% of arrivals within 'N' minutes of scheduled time</u> - Scheduled time may not accurately reflect user perceptions of vehicle arrivals and may not filter out the effects of faulty schedules.

10. <u>% Late</u> - This measure has the same problems as the preceeding measure.

4.3 RECOMMENDED MEASURES

On the basis of the comparison above, and drawing from ideas developed by Martland (10), a number of measures are tentatively suggested as possible alternatives to represent the variability of many of the service attributes listed in Section 4.2.1. Reliability measures important to travelers will also be of significance to transit operators and vice-versa. Thus, the selection of reliability measures should be closely related to user and operator impact analyses.

The following measures are recommended to quantify reliability effects for distributions of travel times (total travel time, in-vehicle travel time, wait time) for travelers:

1. mean time
2. coefficient of variation of time (for distributions which are significantly skewed use the standard deviation excluding extreme values, defined as those exceeding 2.32 standard deviations greater than the mean*)
3. % of observations taking at least 2.32 standard deviations longer than the mean.

Measure #2 gives an indication of the variability of the distribution about its mean, and measure #3 indicates the skewness of the distribution. Values exceeding 2.32 standard deviations greater than the mean can be considered

---

*Thus, standard deviation should be calculated twice, first on the entire distribution to identify the values exceeding 2.32 standard deviations greater than the mean, and then on the restricted distribution.

extreme values since, if the distribution were perfectly normal, only 1% of the observations would be included in this category. For skewed distributions, more than 1% of the observations would fall in this category.

Alternative measures discussed in Section 4.2.2 could be used in place of measures #2 and #3 above. Although the mean is part of the expression for the coefficient of variation, it is important to retain the mean as a separate measure as well. Otherwise, cases where the coefficient of variation increased, only because the percent reduction in the standard deviation was less than the percent reduction in the mean, might be misinterpreted as an increase in unreliability.

The parameters in the recommended measures should be selected so as to provide meaningful yardsticks for evaluating service. Selection of extreme values to exclude, and of the value of parameter N, will vary depending on the compactness and the degree of skewness in the distribution, and what are considered acceptable levels to the user.

Another traveler-related attribute, seat availability at a particular stop, can be measured simply in terms of the percentage of trips for which travelers are forced to stand.

For the operator oriented measure of schedule adherence at origin, destination, and intermediate points, the following parameters are recommended:

1. average difference between mean and scheduled arrival time
2. standard deviation (about mean arrival time; <u>not</u> about printed schedule times)
3. % of arrivals 2.32 standard deviations later than mean arrival time.

The above measures should include the effects of cancelled bus runs, vehicle breakdowns, and added buses; the average number and percentage of buses disabled and runs cancelled and added should be tabulated. These measures would identify the impact of non-scheduled vehicle changes on schedule adherence.

For distributions which are significantly skewed, the same parameters as used for the traveler oriented measures are appropriate. Where service is infrequent, early departure may be intolerable. In such cases early departures should be measured in addition to very late departures.

Parameters may also be developed for other operator oriented measures. Measures of the headway distribution could be:

1. mean headway (h)
2. coefficient of variation of h

   or

   % of headways in interval of kh*
3. % of headways ≥ nh*.

Seat availability at a particular stop can be measured as the percentage of buses where the load factor exceeds one.

Both traveler and operator oriented measures should be taken across several days to uncover the day-to-day variability of service attributes. Moreover, both disaggregate demand and supply data are needed to assess the impact of reliability on travel. Thus, it is recommended that traveler data be collected and segmented by time-of-day, trip purpose, bus route, and the traveler's origin and destination bus stops. Operator data should be collected and segmented by time of day, route, and bus stop.

It seems logical that a series of mathematical relationships can be developed to relate measures of traveler and operator reliability. We know that wait times are a function of the headway distribution, and that schedule adherence can be related to user travel time. Also, by comparing printed scheduled arrival times to the mean arrival time, the effects of faulty schedules can be filtered out, and a more direct relationship between user travel time and "true" schedule adherence can be established. Some of the relationships developed by Joliffe and Hutchinson (11) that appear in Chapter 2 fall in this category.

Unfortunately, disaggregate demand modeling has only begun to address reliability as a component of travel mode and departure time choice, and the supply side implications

---

*k and n are constants determined by the analyst. The value of k may vary; the value of n should be greater than one.

of reliability for the transit operator have not been addressed. The measures selected here will hopefully contain enough information to be useful when progress is made in these related areas.

4.4 CONCLUDING REMARKS

A review of reliability measures used in previous evaluations of transportation service changes indicated several weaknesses in the current approaches to measuring transit reliability. In light of the limitations of many of these reliability measures, criteria were developed for selecting a set of summary statistics which are useful for comparison in before/after evaluations at a site or for comparison between sites. Based on these criteria, a set of reliability measures were defined and recommended for use in evaluation. The use of these measures in future studies is discussed in more detail in Chapter 7.

REFERENCES

1. Bather-Ringross-Walsfeld, Inc., Metropolitan Transit Commission, and Minnesota Highway Dept. I-35W Urban Corridor Demonstration Project, Final Report, August 1975.
2. McQueen, James, et al. Evaluation of the Shirley Highway Express-Bus-On-Freeway Demonstration Project-Final Report. U.S. DOT, Report No.: DOT/UMTA 7, August 1975.
3. Bigelow-Crain Associates. The Golden Gate Corridor Bus Priority System. Transportation Systems Center, May 1975.
4. Bigelow-Crain Associates. Second Year Report: San Bernardino Freeway Express Busway Evaluation, prepared for Southern California Association of Governments, September 1975.
5. Elias, Wilbur G. "The Greenback Experiment-Signal Preemption For Express Buses." Sacramento County, CA, April 1976.
6. SYSTAN, Inc. The Santa Monica Freeway Diamond Lanes: An Evaluation. Vol. II. UMTA Service and Methods Demonstration Program, UMTA-MA-06-0049-77-13, December 1977.
7. Wattleworth, Joseph A., et al. I-95/NW 7th Avenue Bus/Carpool Systems Demonstration Project. Report I. U.S. DOT, UMTA-FL-06-0006, Transportation Research Center, Univ. of Florida, September 1977.
8. Wilson, Nigel, H.M. and N.T. Colvin. "Computer Control of the Rochester Dial-a-Ride System". MIT Center for Transportation Studies Report 77-22, July 1977.
9. SYSTAN, Inc. The Rochester, New York Integrated Transit Demonstration: Suburb of Greece. Draft Final Report, UMTA Service and Methods Demonstration Program, UMTA-NY-06-0048-78,Vol.1,2,3 (Forthcoming).
10. Martland, Carl D. Rail Trip Time Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data. Vol. 2, MIT R72-37, 1972.

11. Jolliffe, J.K. and T. P. Hutchinson. "A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop." *Transportation Science*. 94. November 1975, pp. 248-282.

5. CAUSES OF UNRELIABLE SERVICE

This chapter addresses the basic causes of unreliability in transit services. The variability of transit level of service attributes can be viewed as the results of environmental and inherent factors. Environmental factors include all factors related to the setting in which the system concept is applied and all exogenous random influences. These include traffic controls, variation in transit demand and vehicular traffic, and vehicle availability. Inherent factors include those initial causes or magnifying properties that result from the basic transit service concept. Since inherent factors can only be discussed in the context of a particular service concept, the chapter is organized accordingly. Section 5.1 addresses causes of unreliability common to all transit operations. Sections 5.2 and 5.3 discuss causes of unreliability specific to fixed route bus and dial-a-ride respectively. While fixed route bus and dial-a-ride do not represent the entire range of transit services nor necessarily its extremes, the two cases should identify some basic issues relevant to most transit service concepts of interest. In each section, the sources of service attribute variance and evidence of the dominance of particular sources will be discussed.

5.1 CAUSES COMMON TO ALL TRANSIT OPERATIONS

A number of factors contributing to the variability of performance characteristics are common to all transit operations. Foremost among these factors, which are basically environmental in nature, are general traffic conditions, controlled intersections, variation in demand, and availability of drivers and vehicles.

As part of the general traffic stream, transit vehicles are subject to travel time variations resulting from interference among all motor vehicles. Interference can result from accidents, breakdowns, turning movements, lane speed changes, merging, stopping, parking, illegally parked cars, cross street traffic, weather, and driver characteristics. In a given physical setting or environment, many of these effects may be viewed as directly related to traffic volume.

Travel time and traffic volume may be plotted in a volume delay curve, shown in Figure 5.1. It should be noted that the relationship between <u>mean</u> travel time and traffic density is well established in traffic research; that between variability of travel time and traffic density is

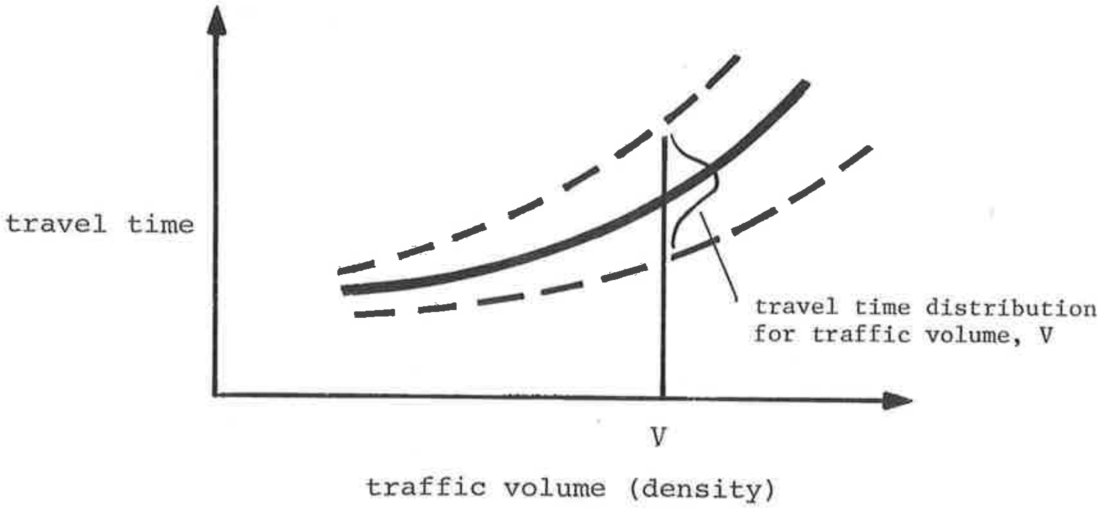


FIGURE 5.1 TRAVEL TIME AND TRAVEL TIME VARIABILITY AS A FUNCTION OF TRAFFIC VOLUME

not. Unreliability, as defined, is a function of the variability of travel time, rather than the mean. Obviously, variation in traffic volume will cause variation in travel time and, therefore, unreliability. A priori, one would expect the greater degree of interference between vehicles at higher volumes to cause a greater variance in travel times than less interference at lower volumes. Therefore, it is hypothesized that the variance of travel time also increases with traffic density.

Another major cause of variation in travel time along an arterial is the interrupted flow produced by the presence of intersections. Welding (1) found that buses spent 19% of their journey time stopped at traffic signals. A signal introduces variability by delaying some vehicles which happen to arrive at the intersection during the red phase. Increasing the number of unlinked signals increases the overall variance of travel time.*

Hourly, daily, and seasonal variations in transit user demand can also cause significant variation in level of service. For some of this demand variation, patterns may allow the operator to schedule service accordingly (e.g., peak vs. off-peak, and seasonal demand shifts). However, some variation will remain as random fluctuations over time causing service variability. These variations may affect seat availability, wait time, and ride time. The extent of these effects is dependent on the transit operation itself.

Service quality will be affected by two other characteristics of the "environment" in which the service concept is applied: vehicle and driver availability. As internal components of the system, unlike traffic, controlled intersections, or demand, these factors are at least partially under the control of the operator. The extent of these availability problems are affected by vehicle design (selection) and maintenance policies; employee policies, benefits and morale; and the use of reserve vehicles and drivers. The reader should refer to Chapter 3 for a discussion of operator policy and the tradeoffs of the travel time and travel time variability that result from different operator policies.

Finally, some variation in bus travel times is attributable simply to differences in driver behavior. Some

---

*If delay at each signal $t_1 = N(\mu_i, \sigma_i)$, and assuming that each signal is independent, then the variance of the total t for n signals = $\sum_{i=1}^{n} \sigma_1^2$.

of these differences may be reduced through better management; others may be unavoidable.

5.2 FIXED ROUTE BUS SERVICE

In addition to the environmental factors affecting transit described in the previous section, fixed route service experiences an inherently unstable headway distribution. The instability of headways in fixed route bus systems tends to become more pronounced further downstream. Data collected on Boston's Harvard-Dudley line show the following difference between bus arrivals at the first stop (Harvard) and an intermediate stop (Auditorium) (Kulash, 2):

	Mean Headway	Variance
Harvard	4.3 Min.	0.8 Min.^2
Auditorium	5.0 Min.	26.1 Min.^2

Exogenous factors, such as delays between stops due to traffic conditions, or delays at stops due to unusual loading conditions, trigger an initial deviation from schedule or scheduled headways. The inherent instability lies in the fact that any delay in arrivals results in an increased dwell time at that stop due to the increased passenger load that has accumulated over the previous long headway. Thus late buses get later and early buses (because of reduced demand) get earlier, eventually resulting in bunching and imbalanced loading (Newell and Potts, 3).*

While headway distribution at the dispatch point may be characterized by a deterministic or, more realistically, a low variance uniform distribution, the type of distribution which results downstream tends to approach an exponential distribution (Kulash, 2).

An important first step in determining improvement strategies is identifying those causes which contribute the greatest portion of variability to the system. A number of studies have already examined some of these causes.

In their evaluation of bus control strategies through simulation analysis of a case study route in Bristol, U.K., Bly and Jackson (5) found that variation in run times

---

*See Appendix D (Cohn, 4) for a mathematical derivation of the effect of the propogation of an initial deviation on the headway at a particular bus stop on a route.

between consecutive buses was more significant in contributing to system variability than variation in boarding times at each bus stop.

In a London Transport study of a central London bus route (Shields, 6), the following causes of irregularity were observed:

1. cancellations of scheduled journeys due to staff and bus shortages (20% cancelled).
2. journey times on the road varying considerably, partly attributable to traffic congestion.
3. buses leaving garages and terminals late.
4. difficulties in supervision.

It was not possible to rank the above by their degree of importance. Investigation of the percentage of journey time spent in delays at bus stops, signals, etc. (see Table 5.1), showed that time of day variations and average delays themselves were primarily due to signals. It should be noted that the relatively lower importance of bus stop dwell delays may be due to the fact that two-man operated buses are used on the route studied. In addition, there is evidence that considerable lateness is due to unreasonably short scheduled journey times during congested periods. While adjustments to the schedule may reduce lateness, there is still considerable variation in any time period as shown in Figure 5.2 (variation also exists between days). Finally run times themselves vary considerably, suggesting that even if all delays were removed, variability will remain due to individual driver differences or, possibly, traffic effects.

Research at Northwestern University (Sterman, 7) attempted to demonstrate correlation between unreliability and the number of signalized intersections, link length, and passenger boarding and alighting volumes. Under different conditions, such as peak or base time periods, or geographical location in the metropolitan area, different factors were found to be most important. It was difficult to single out any one factor as the most important in affecting reliability. Considerable variation was not "explained" by the regression models developed. It should be noted that traffic congestion was not accounted for explicitly in the regression models.

In a simulation study of a single bus route, Kulash (2) examined the sensitivity of mean passenger wait time (a significant effect of unreliability) to various factors,

TABLE 5.1 COMPONENTS OF JOURNEY TIME (Shields,6)

SOURCE OF DELAY	% OF JOURNEY TIME IN PERIOD					
	Before 0730	0730 -0900	0900 -1200	1200 -1630	1630 -1800	After 1800
BUS STOPS	9.5	11.8	10.6	11.6	11.8	11.1
TRAFFIC SIGNALS	13.6	20.4	24.0	25.7	27.1	24.7
OTHER CAUSES	0.4	1.5	3.0	1.9	1.3	1.0
TOTAL DELAYS (%)	23.5	33.7	37.6	39.2	40.4	36.8
JOURNEY TIMES (mins)						
N/B Schedule	63	71	72	72	73	63
N/B Actual	52.7	72.9	74.4	79.0	79.0	65.4
S/B Schedule	62	67	69	69	75	62
S/B Actual	53.6	63.4	67.6	71.2	74.3	70.6

Figures given are averages over the 3 days.

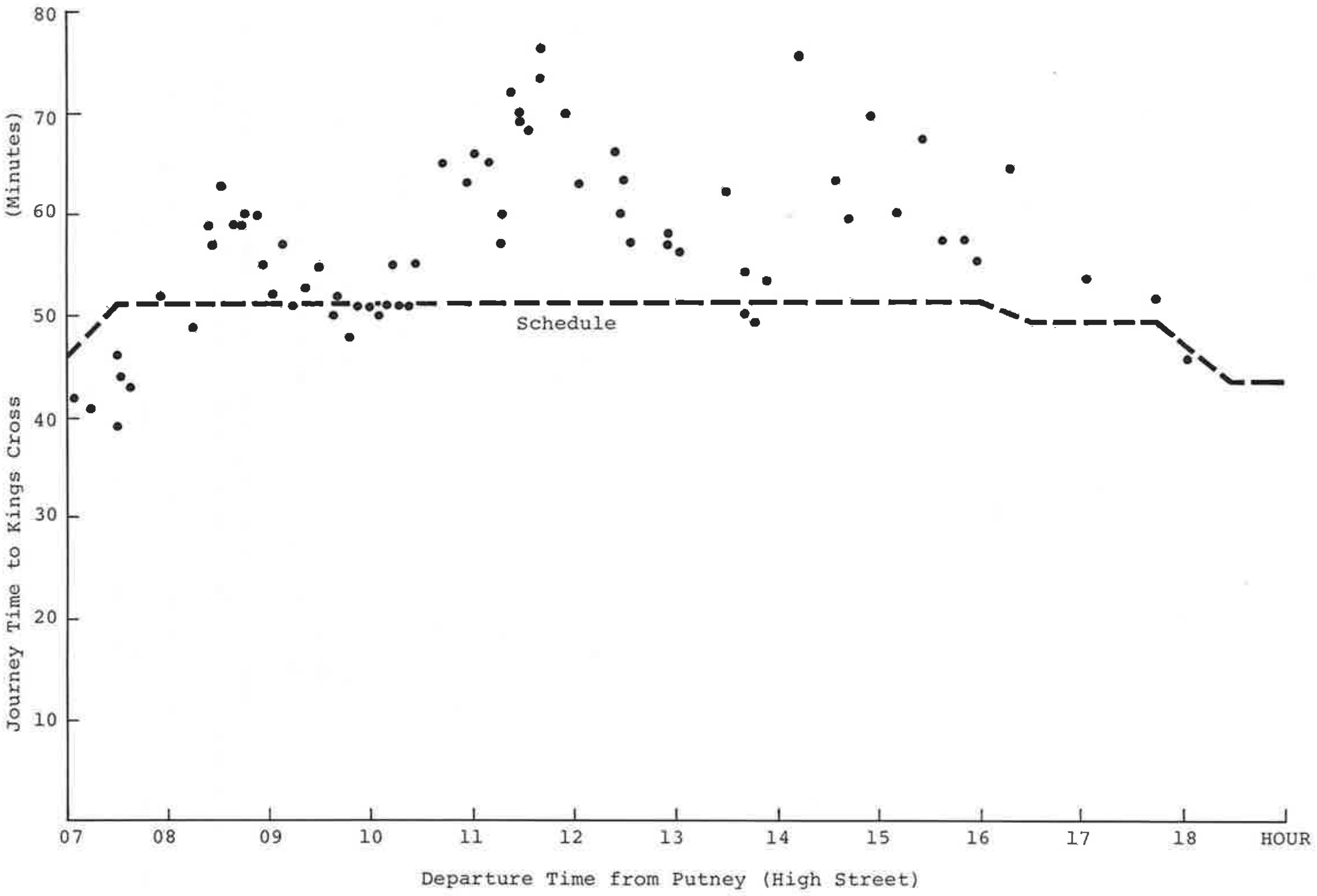


FIGURE 5.2 BUS JOURNEY TIMES (NORTHBOUND) - WEDNESDAY (Shields,6)

including demand, load and unload times, travel time variation and headway variation. He found that average wait time was closely tied to headway irregularity while other factors were less important. Increasing demand had little effect until the 80% saturation level was reached at some point on the route and large numbers of passengers had to wait for a second bus due to overcrowding. A zero variance in interstop travel time reduced wait time by only 10%. Instantaneous loading also resulted in only a 10% reduction in wait time. Kulash concluded that wait time can only be seriously reduced by dispatching vehicles regularly and designing routes and holding points which serve to reinforce that regularity.

Although considerable work has been done in the fixed route area, the results are somewhat inconclusive. However, there appears to be basic agreement that variation in journey times is a major cause of unreliability with variation in boarding volumes being of less importance. Further research is necessary to determine the relative importance of traffic, signals, boarding time, driver differences, etc., to aid in the design of strategies to improve reliability.

5.3 DIAL-A-RIDE

The dynamic, demand responsive nature of dial-a-ride makes it an inherently unreliable service. Table 5.2 summarizes operating data for seven days in the Rochester dial-a-ride system. The variability both within a given time period, and across different hours of the day is particularly striking. Standard deviations of wait times are typically about half the mean, indicating the wide range of service quality a regular user is likely to receive. For example, between 8 and 9 a.m., assuming a normal distribution of service times, a user can be 95.4% confident that his/her service will fall in the following ranges:

Wait time range	5.3 - 47.1 mins.
Ride time range	5.9 - 24.6 mins.
Pick-up lateness range	-4.9 - 11.6 mins.

These ranges clearly demonstrate the importance of devising strategies for improving the reliability of dial-a-ride.

Dial-a-ride operations introduce several unique sources of variance that can cause unreliability. Among these are dynamic routing, the sensitivity of dial-a-ride to demand, passenger cancellations, and passenger no-shows.

TABLE 5.2 SERVICE QUALITY BY TIME: IMMEDIATE RESULTS

	WAIT TIME			RIDE TIME			LATENESS		
PERIOD	AVERAGE	MAXIMUM	ST.DEV.	AVERAGE	MAXIMUM	ST.DEV.	AVERAGE	MAXIMUM	ST.DEV.
700 - 800	29.857(7)	41.000	9.69	12.714(7)	26.000	8.10	8.167(6)	16.000	6.67
800 - 900	26.161(56)	47.000	10.45	14.982(55)	47.000	9.08	3.316(38)	30.000	8.25
900 -1000	21.055(91)	69.000	10.78	13.022(90)	40.000	8.34	0.747(79)	20.000	6.82
1000 -1100	22.553(94)	56.000	11.12	11.559(93)	37.000	7.40	2.964(84)	36.000	8.27
1100 -1200	23.095(74)	73.000	19.44	14.554(74)	40.000	9.26	6.017(60)	48.000	13.11
1200 -1300	21.075(67)	57.000	11.96	15.075(67)	44.000	9.39	2.130(54)	37.000	9.33
1300 -1400	25.125(40)	46.000	10.57	13.825(40)	44.000	9.67	3.767(30)	26.000	8.93
1400 -1500	28.404(44)	61.000	10.92	22.256(43)	44.000	10.86	7.071(42)	36.000	9.52
1500 -1600	22.083(36)	61.000	12.87	15.583(36)	39.000	8.55	3.485(33)	36.000	10.71
1600 -1700	24.025(40)	64.000	13.16	17.025(40)	55.000	10.22	4.139(36)	29.000	8.70
1700 -1800	21.091(22)	40.000	10.28	16.333(21)	43.000	11.22	2.000(20)	15.000	7.57
1800 -1900	25.684(19)	45.000	10.97	17.579(19)	34.000	10.15	4.167(18)	20.000	9.84
1900 -2000	14.447(9)	27.000	8.68	9.111(9)	16.000	3.02	0.714(7)	14.000	9.63
2000 -2100	16.888(9)	31.000	8.71	16.778(9)	35.000	9.37	1.875(8)	9.000	4.73
2100 -2200	10.500(2)	12.000	2.12	15.000(2)	16.000	1.41	-2.000(2)	0.000	1.41

Note: Sample size in ( ) is total for 7 days.

Vehicle operations in dial-a-ride are determined by a formal or informal set of rules which are applied to assign riders to vehicle tours. For a specific request, the actual wait time and ride time will depend on the current positions and activities of all vehicles, the set of competing users currently active on the system, and the set of new users who will request service in the immediate future. Clearly all these elements are highly stochastic, as are the resulting wait and ride times. Thus level of service will vary for similar trips, over the day, and between days even if the overall level of demand and the number of vehicles are constant.

While variation in demand has already been identified as an environmental factor affecting level of service in all transit operations, for dial-a-ride this sensitivity merits special attention. In fixed route scheduled service, the addition of a passenger affects overall service only by the extra dwell time, which is virtually insignificant. In dial-a-ride, a new passenger will typically involve added vehicle-miles as well as a much longer dwell time. The net result is that dial-a-ride service quality is extremely sensitive to level of demand.

Because of the sensitivity of dial-a-ride service to demand level, the cancellation of a trip which has been scheduled can create significant service quality changes for other passengers. In general, a cancellation will result in early arrival at subsequent stops on the affected vehicle's tour. These arrivals may be before the advertised time for some passengers, thus increasing the probability that they will not be ready and hence become no-shows.

Passengers generally will be no-shows either when the vehicle arrives much earlier or much later than anticipated. This can have two effects downstream: (1) a longer than usual dwell time at the no-show stop can delay the vehicle at subsequent stops and, (2) the resulting cancellation of the associated delivery stop can result in the vehicle being early beyond this point in its tour. It is possible to imagine instabilities being created because of these no-show properties, particularly where some large initial perturbation is introduced such as a vehicle breakdown or major traffic delay. Figure 5.3 defines a range of the relationship between no-shows and system wait time based on observed data from Haddonfield. The relationship between system lateness and cancellations and no-shows is presented conceptually in Figure 5.4.

For dial-a-ride, very little empirical work has been performed on causes of unreliable service, although it is

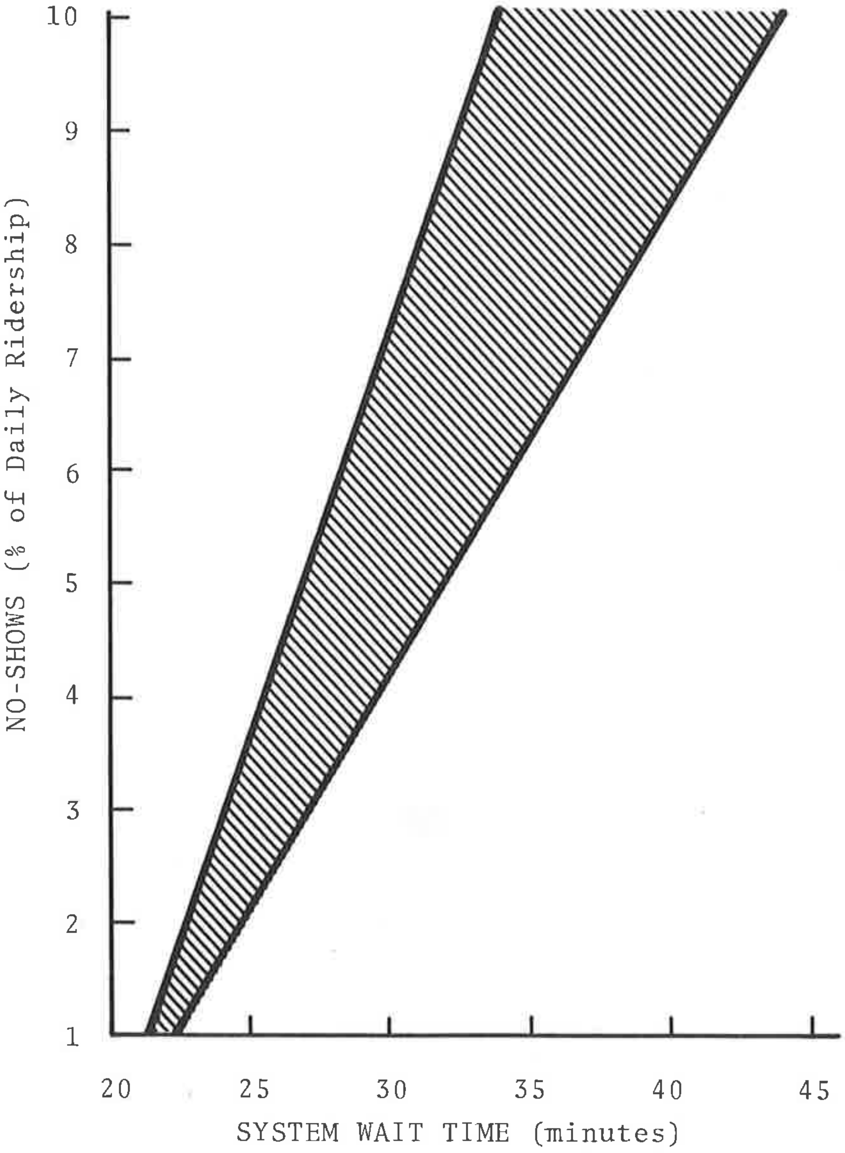


FIGURE 5.3 HADDONFIELD WAIT TIME VS. NO-SHOWS (Haddonfield,8)

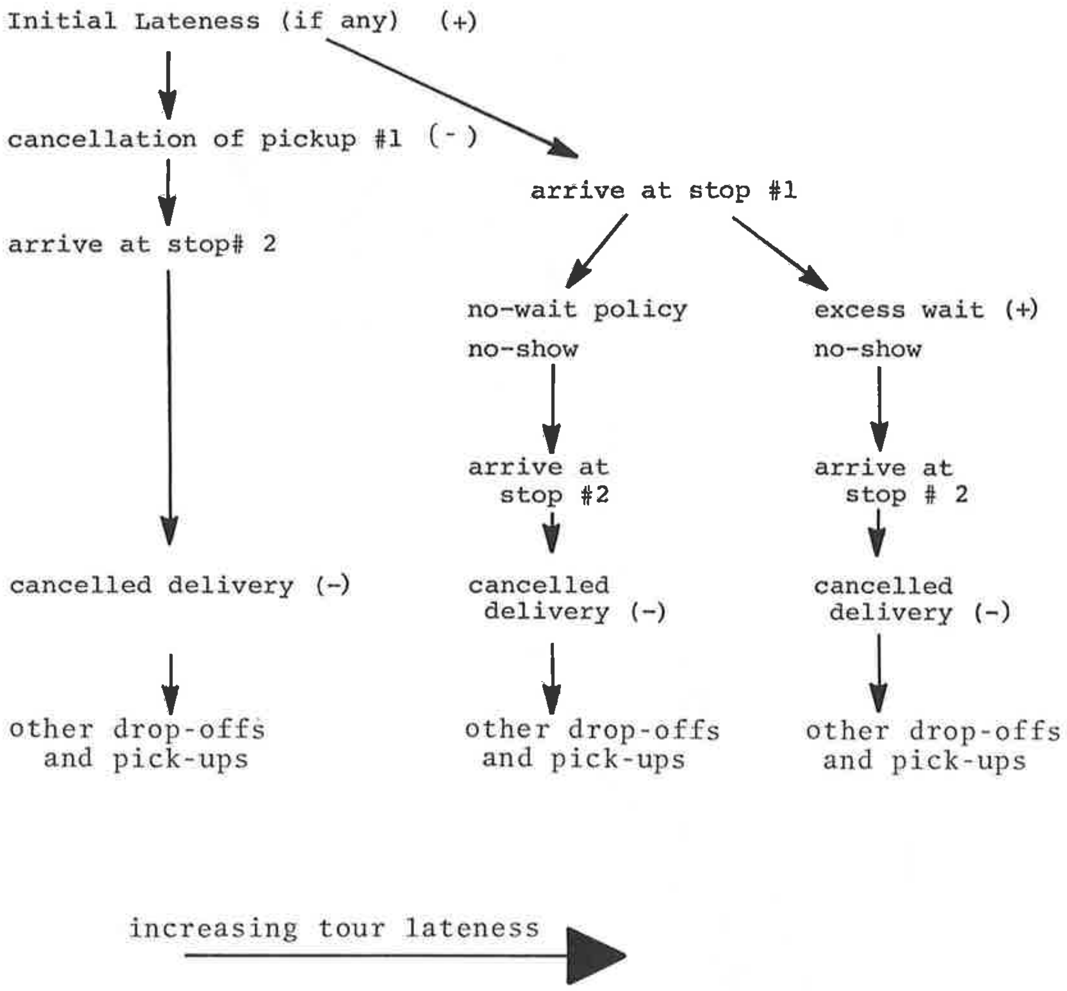


FIGURE 5.4 EFFECTS OF CANCELLATIONS AND NO-SHOWS

hypothesized that inherent characteristics of the dynamic routing, demand-responsive concept are the primary causes of unreliability. Further research is necessary to determine effects of no-shows, cancellations, demand variation, and dispatching procedures and algorithms on service reliability.

In addition to fixed route bus and dial-a-ride, other service concepts have specific reliability problems. For example, deviating bus systems have the variability of demand responsive services while trying to meet fixed schedules. No studies have yet addressed the reliability of such innovative service concepts.

5.4 CONCLUDING REMARKS

This chapter has identified potential causes of unreliability in transit operations, focusing in particular on fixed-route and dial-a-ride services. Although several studies have been conducted on the relative importance of each cause, results have been inconclusive to date. Nevertheless, for the causes of unreliability that have been identified, several strategies have been proposed to eliminate or negate the effects of these causes. These strategies are the subject of the following chapter.

REFERENCES

1. Welding, P.I. "The Instability of a Close-Interval Service." Operations Research Quarterly. (1957), 8(3), pp. 133-142.

2. Kulash, D. "Rerouting and Scheduling in Public Transit Systems." (MIT PhD. Thesis, 1971).

3. Newell, G.F. and R.B. Potts. Maintaining a Bus Schedule. Preceedings of the Second Conference, Austrlian Road Research Board, II, (1964), pp. 388-393.

4. Cohn, J., et al. "Qualitative Analysis of Bus Location and Control." Motorola, Inc., paper presented at ORSA Spring 1976 Meeting, Philadelphia.

5. Bly, P.H. and R.L. Jackson. Evaluation of Bus Control Strategies by Simulation. TRRL Laboratory Report 637, Transport and Road Research Laboratory, Dept. of Environment, Crowthorne (Great Britain), 1974.

6. Shields, L.R. Reliability Study of a Central London Bus Route. London Transport, Great Britian, 1974.

7. Sterman, Brian P. "Factors Affecting Reliability of Urban Bus Services." (M.S. Thesis, Northwestern University, IL, 1974).

8. "Haddonfield Dial-A-Ride Project." UMTA-NJ-06-0002-74-7, N.J. DOT U.S. DOT, 1974.

6. STRATEGIES TO IMPROVE RELIABILITY

This chapter discusses strategies for improving service reliability problems resulting from the sources identified in Chapter 5. There are three basic methods or techniques for improving reliability: priority, control, and operational. Priority methods involve special treatment of transit vehicles apart from other motor vehicles; control strategies concern dynamic modification of bus operations from the original schedule; operational methods deal with route, schedule, and resource allocation changes. Potential strategies are grouped by method in Table 6.1.

A second categorization of strategies is related to where in the process of unreliability development they are applied. Two categories are considered: preventive (type 1), and corrective/restorative (type 2). Preventive strategies are directed at reducing the probability of initial deviations; corrective/restorative strategies attempt to correct deviations and restore operations to the initial schedule.

This chapter examines improvement strategies and considers their usefulness based on empirical studies. An attempt is made to examine the effects of these strategies both on the variability and the mean values of service attributes.

The material in this chapter is organized in a manner similar to Chapter 5. Section 6.1 addresses strategies directed at improving reliability problems common to transit operations in general. Again, much of this discussion is based on experiences with fixed route bus systems. Section 6.2 and 6.3 discuss improvement strategies specific to fixed route bus and dial-a-ride service respectively.

6.1 STRATEGIES COMMON TO ALL TRANSIT OPERATIONS

While many strategies may be feasible for a particular service type (e.g., exclusive lanes for fixed route bus), there are a few strategies that are applicable to transit operations in general.

Two strategies can be used to combat delays resulting from vehicle breakdowns. Reserve vehicles can be set aside to deal with emergency situations. Another approach is through day-to-day rescheduling, where the available fleet can be utilized to stabilize headways. Driver illness or absenteeism can be handled by keeping either reserve drivers on hand or through day-to-day rescheduling.

TABLE 6.1 STRATEGIES GROUPED BY METHOD

1. Priority
   - exclusive lane
   - signal pre-emption
   - merge/freeway access

2. Control
   - holding
   - passing
   - turnback
   - skip stop
   - speed modification

3. Operational
   - reserve fleet
   - reserve drivers
   - modify frequency of service
   - express service
   - schedules
   - fare collection (exact fare, passes, fare free)
   - improved access (into the vehicle)
   - dwell time policies
   - dynamic rerouting
   - computer dispatching
   - eligibility for deviations/doorstep service
   - restructure service

Modification of schedules, routes, layover time, vehicle allocation and promised times of pick-up and drop-off can be made to counteract the effects that <u>predictable</u> variations in demand and traffic congestion will have on the variability of service attributes. For example, variations in seat availability on fixed route buses due to predictable variations in demand can be alleviated by scheduling additional buses during the higher demand periods.

6.2 FIXED ROUTE BUS SERVICE

Figure 6.1 illustrates the development of unreliable service along a route, and the application of different types of strategies at different points in the development process. Strategies specifically oriented to fixed route service are identified by type in Table 6.2. Each strategy is described below; those for which studies have been performed are analyzed in more detail in Section 6.2.2.

6.2.1 Summary of Strategies

The instability of fixed route bus systems is partly due to the variation in passenger loads. Since the rate of passenger arrivals at bus stops is not constant but varies over time (even over short intervals), the boarding and alighting volumes and corresponding bus stop dwell times will vary. This may result in deviations from scheduled arrival times downstream. Such variation in headway will cause a change in bus stop passenger boarding volume and consequently bus stop dwell time. This results in further deterioration of the service.

While it is impossible to control random variation in passenger arrivals, a few strategies--express service, fare collection techniques, improved access into the vehicle--may reduce the effects of variable boarding and alighting volumes. Express service will reduce the number of stops and increase the length of line haul segments. As a result, the probabilities of slight deviations in travel time and of increased wait times at downstream bus stops are reduced. Fare collection strategies will reduce the per passenger boarding time and decrease the overall effects of boarding volumes. Wider, more accessible vehicle entrances can also reduce per passenger vehicle boarding time. This may be especially significant for elderly passengers, who often constitute a significant percentage of off-peak riders and who have great difficulty boarding typical transit vehicles.

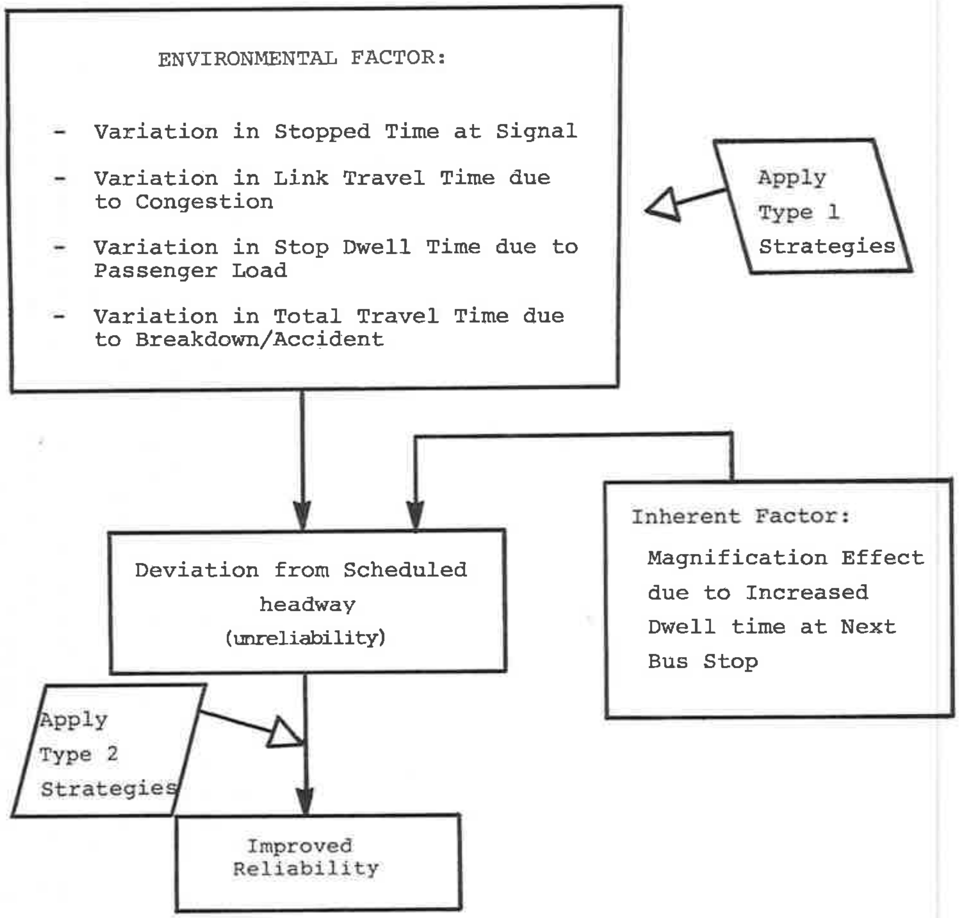


FIGURE 6.1 APPLICATION OF IMPROVEMENT STRATEGIES (FIXED ROUTE)

TABLE 6.2 CLASSIFICATION OF STRATEGIES (FIXED-ROUTE)

Strategies	Type 1: Reduce Probability of (initial) Deviations	Type 2: Corrective/ Restorative
a) Exclusive Lanes/ Merge Access	X	
b) Signal Priority	X	X
c) Fare Collection	X	
d) Improved Access	X	
e) Express Service	X	
f) Speed Modification	X	X
g) Holding Strategies		X
h) Turnback		X
i) Skip Stop		X
j) Reserve Vehicles	X	X
k) Multiple Route Vehicle Pool		X
l) Pass First Bus		X
m) Imbalanced Slack in Bus Schedules	X	X

Another important cause of unreliable service on fixed routes is variation in traffic volumes and traffic congestion. Link travel times are increased as a result of increased traffic on roadways. In addition, link travel time variance is likely to be greater at higher traffic volumes and where random interference occurs. Interference includes effects of speed changes, lane changing, stopping, parking, parked cars, turning, merging, accidents, and breakdowns. The resulting deviations from scheduled bus arrival times downstream will produce variations in passenger loads, leading to further deterioration of the service.

Providing exclusive lanes for buses may reduce both the variance and mean of link travel times due to traffic effects. Unless priority for buses is provided for entering the exclusive lane and for passing through intersections, however, a potentially large part of travel time variance will remain.

Another strategy, providing priority to buses at signals (pre-emption), should reduce the mean and variance of stopped time at signals. Stopped time at signals is likely to vary due to the random arrival of buses with respect to signal phase, and variation in the queue length of traffic waiting at the light. With signal pre-emption, although some variance will remain due to traffic conditions, some of the effects of traffic on bus travel times and headway variation might be alleviated.

There are a number of strategies, mainly corrective, that attempt to restore the system to regular service once deviations and disruptions have occurred on a bus route (i.e. headways have become irregular and buses are bunching). They include modifying the speed of buses, holding buses at control points, adding reserve vehicles, dispatching vehicles from a multiple route vehicle pool, turning back buses before the end of a route, skipping bus stops, passing delayed buses and utilizing unevenly distributed slack time.

Instructing drivers to change speeds according to deviations from the desired headway or schedule corrects for any initial deviation. This is possible on exclusive lanes or uncongested streets where fixed schedules are observed or where vehicle monitoring is in operation.

Control point holding strategies are designed to stabilize the variation in headways. By holding vehicles in mid-route at specified control points, they reduce the magnification of deviations further downstream which might result in more serious service deterioration. The trade-off between increased in-vehicle time for passengers on board and wait time for passengers downstream must be analyzed further to determine optimal holding strategies and location of holding points.

Adding a reserve vehicle is one of several corrective strategies employed once significant deviations have developed. A trade-off exists between maintaining a regular lower frequency service (with reserve vehicles) and providing a more variable, but higher frequency service (utilizing all available vehicles).

A multiple route vehicle pool requires orienting all radial routes to a central point and exchanging vehicles between routes to allow shorter layover times. Early vehicles from one route may be used as reserve vehicles to cover another route. Union work rules regarding assigned runs may render this strategy infeasible.

Turnback strategies, turning a vehicle around before it reaches the terminus, may correct headway deviations upstream (fill in a gap) at some inconvenience to passengers who must deboard in order for the vehicle to turn back. The long-run effects of this strategy could be disruptive, since the future schedule of bus operations has been altered.

Skip stop strategies involve passing those stops where no one onboard wishes to alight. This allows gaps to be reduced, at some inconvenience to waiting passengers who are passed by. Such a strategy might be used when buses are bunched, although operationally this may be a difficult strategy to implement.

Passing strategies, changing the order of two buses which are bunched, are also oriented toward correcting bunching of buses. They restore balance to the system by allowing the inherent instability to work in a correcting direction. Passengers onboard the first bus may be slowed down as the first bus assumes the schedule of the second bus.

The selective use of providing priority to buses at signals, previously discussed in terms of its preventive capability, may also be useful as a corrective measure for buses behind schedule or where gaps in service appear.

It is also possible to reduce the overall effect of the reliability problem by allowing for scheduling adjustments between the arrival and departure of a bus at the terminal. One strategy is to schedule imbalanced slack in schedules, where the more unreliable routes (particularly during the peak period) have added slack time built into the schedule. Thus, a late bus arriving at the terminal can depart on the next run without propagating the delay leftover from the previous run.

Another set of strategies directed at improving fixed route service reliability is major routing and scheduling changes. Because reliability tends to deteriorate as a bus proceeds further away from the dispatching point, the implementation of shorter routes might improve reliability. However, the effects of extra transfers might counteract the potential benefits of more reliable service. An alternate strategy may be to examine the tradeoffs of fewer routes (lower route density) and higher frequency service as opposed to more routes (higher route density) and lower frequency service. Turnquist and Bowman are currently analyzing these tradeoffs at Northwestern University's Transportation Center (1).

6.2.2 Analysis of Strategies

In the more detailed analyses that follow, the impacts of bus service reliability improvement strategies have generally been measured only in terms of changes in the mean values of travel time components. Studies have largely ignored the possible changes in the variability of travel time components. Yet, as we have seen in Chapter 2, the variability of these components is hypothesized to play a key role in travelers' mode choice and departure time decisions. Furthermore, it cannot be assumed that if a strategy results in decreases in the mean values of travel time components, the variability of these components will also decrease. In fact, some improvement strategies may produce tradeoffs between the mean values and variability of travel time components. The results of the analysis presented here cannot be viewed as complete until a more detailed examination is made of the effects of these strategies on the variability of travel time components.

6.2.2.1 Priority Strategies

Exclusive lanes have been a major element in the Service and Methods Demonstration Program as described in Chapter 4 of this report. That chapter also described a project involving bus pre-emption of traffic signals. Bus priority strategies have also been demonstrated in the U.K. and Ireland with particular attention given to impacts on reliability (Chapman, 2). Some results are presented in

Table 6.3. In Southampton, linked signals and other bus priority strategies did not lead to significant changes; in Derby, standard deviation was reduced while coefficient of variation actually increased (indicating the care which must be taken in using coefficient of variation as a reliability measure). The primary effect of such strategies was a reduction in mean travel time. The strategies do, however, exhibit promise in improving reliability.

Signal preemption is another priority strategy in which buses may overcome the delays resulting from signalized intersections. This strategy is also generally directed at mean travel time. By preventing random delays at signals, the preemption strategy should reduce travel time unreliability as well.

Signal preemption may also be used as a control device specifically oriented toward improving reliability. Preemption may be selectively used to speed up a bus behind schedule. With information on other bus locations, headways may be regulated in this manner. In his simulation studies, Koffman (3) found such preemption strategies to decrease both wait time and its variability, and line haul travel time and its variability. However, he cautioned that such benefits might be cancelled out by a potential worsening in traffic congestion.

In addition to signal preemption schemes which extend the normal green phase, priority can be given to transit vehicles by separate signal phases and exclusive lanes in the vicinity of a signal (as for right turning lanes). This type of priority scheme is common in Europe and is beginning to come into use in the United States. Moreover, buses can be given a form of priority by retiming traffic signals in a manner so as to facilitate bus movement. The results of a series of simulations performed at the Transportation Systems Center (Muzyka, et al., 4) suggest that such retiming of traffic signals can reduce bus delays at intersections almost as much as signal preemption strategies.

6.2.2.2 Control Strategies

Utilizing a simulation model of a London bus route, Jackson and Stone (5) studied the effectiveness of three methods of control: roadside inspectors, central control with radio-telephone communications, and an automatic bus location system (AVM). Control actions included trip curtailments, trip projections, adjustments to departure times from terminals and the authorization of additional overtime payments to get late crews back on time for their

TABLE 6.3 PRIORITY STRATEGIES in U.K./IRELAND (Chapman, 2)

Strategy	Location	Date	Effect
Priority for buses at a signal	Derby	1972	(through intersection) Reduced mean travel time: 50 sec → 31 sec Reduce std. dev. of travel time: 26 sec → 22 sec. Reduced range: 9-160 sec → 9-135 sec.
Exclusive Lanes (concurrent flow)	Dublin	1971	Reduced mean route travel time: 10.6 min → 8.4 min. Reduced range: 5-22 min → 5-11 min. Reduced % headways > 5 min. 12% → 5% Average wait time also decreased as a result of improved regularity and extra buses.
Linked signals with bus priority	Southampton	1972	Insignificant changes in mean wait time

next run. Comparing mean wait times for the alternative control methods with the no control alternative revealed the following: roadside control reduced mean wait by 11%, radio-telephone central control reduced mean wait by another 3%, and automatic bus location reduced mean wait by an additional 2%. It is possible that the results would have been different if the strategies were applied in a different order.

While there may be some additional manpower savings achieved through implementation of central control, the savings were considered to be dependent on the organizational structure, and they were not estimated. Revenue increases and overtime payments to drivers were included in a benefit-cost analysis, in addition to value of time estimates, which led to the conclusion that the cost effectiveness of automatic bus location systems has not been proven.

In Dublin, a system of radio-telephone control was implemented in which a central controller called each bus in turn and asked the driver for his location. This information was then plotted on a preprinted control sheet containing the scheduled locations of each bus for a point in time. When buses suffered delays or were found to be bunching, a number of strategies were available to the controller. These included:

(1) transferring passengers and turning buses,
(2) injecting standby buses,
(3) running out-of-service buses to particular points on the route,
(4) switching crews,
(5) holding buses at particular points,
(6) using standby crews if available, and
(7) altering bus routes on a temporary basis.

The radio-telephone control system reduced mean wait time from 3.7 to 2.8 minutes, and also reduced passenger complaints.

Bly and Jackson (6), in their simulation study, investigated a number of strategies on a seven mile crosstown bus route with 10 minute headways. These strategies, which depend upon a bus location monitoring system to give the position of each bus in service, included:

(1) replacing a late bus at the city center by a reserve bus (at approximately the midpoint in the bus route),

(2) bypassing stops to reserve capacity on a bus for later stops,

(3) skipping stops to close gaps in service,

(4) turning back buses to split bunches,

(5) injecting reserve buses into regular service at the city center into large gaps, and

(6) holding back early buses at stops.

Few of the strategies proved very effective. The use of reserve buses in regular service was found to improve service more than the use of reserve buses to substitute for late buses. At the city center, there was an inconvenience to through passengers when a late bus was replaced by a reserve bus and the late bus was cancelled. While this problem could be overcome by substituting reserve buses at a terminus instead of the city center, the reserve buses could be assigned only to individual routes. Economic use of a reserve pool requires that the pool be available to service several routes. This would be possible at a location where many routes intersect, such as the city center. Using reserve buses to fill in large gaps should be used only for unpredictable gaps. Predictable gaps in service would be handled more efficiently through rescheduling. In addition, a reserve bus strategy is most beneficial where headways are small so that removal of a bus from regular service has little adverse effect on wait time.

Bypassing stops to reserve capacity for later stops was not found to be effective in achieving goals and had harmful effects on wait time at bypassed stops. Bypassing stops to close gaps had similar harmful effects, outweighing any possible gain. (Koffman, in his simulation work, also found this to be the case.) Turnback strategies, while solving the immediate problem, can have potentially undesirable long term effects due to the resulting location and availability of the fleet.

Finally, holding strategies were found to be very useful in improving service regularity, at the expense, however, of increased travel time. Therefore holding should be restricted to a few stops along the route. If implemented at all stops, the increase in travel time outweighs the benefits of reduced wait time (assuming that travel time and wait time are valued equally).

The study also included a run in which buses waited a specified layover of 2 minutes at termini without reference to schedule adherence. This policy reduced the layover time for many buses. However, the resulting deterioration in schedule adherence produced a 40% increase in mean wait time despite a 22% reduction in mean headway.

In summary, the study was skeptical on the cost-effectiveness of automatic bus location systems. However, examination of a longer route with shorter headways might yield different results. The study also noted the importance of accurate and efficient schedules. Much variation in bus travel time is predictable and can be built into schedules by varying scheduled travel times and layover times to reflect these predictable variations. Thus layover times and scheduling are important strategies for counteracting the effects of the inherent instability of fixed route bus systems.

As described earlier in this section, holding strategies at a few control points (bus stops) proved to be particularly useful in controlling headway variation. A procedure has been developed to construct an approximate-optimal dispatching strategy for the control point (Barnett, 7).

The procedure examines the delay of individual vehicles at a control stop B. An objective function is specified consisting of delay incurred by through passengers at stop B and expected wait time for passengers who board at B, along with appropriate weights which also consider passengers downstream. The optimal holding strategy is one that minimizes the value of this objective function. The vehicle arrival lateness distribution is approximated by two points and holding strategies are examined for a sequence of vehicle arrivals at stop B. Barnett's procedure deals only with mean travel and wait times; it does not include consideration of changes in the variance of travel and wait times.

Barnett's procedure was applied in a model of the northbound Red Line in Boston at the busiest stop on the line, using real data from this stop. The optimal holding strategy resulted in a reduction in average wait time at the stop from 2.8 minutes to 2.53 minutes (close to the ideal wait time of 2.5 minutes associated with the five minute scheduled headways) and an average delay of 25 seconds for passengers passing through the stop. A more detailed description and application of Barnett's procedure appears in Appendix C as well as a discussion on extensions to his approach.

On-line control strategies were tested in Toulouse, France, with a simulation model and a heuristic algorithm (minimization of the sum of squares of headway variables, subject to constraints) to determine "near-optimal" dynamic schedules for bus departures for specific termini and control point stops (Giraud and Henry, 8). Simulation results for control at termini only and at termini plus two stops were compared with the results for an operating procedure used in the Toulouse bus system (referred to as the reference procedure). The reference procedure consisted of adherence by the crew to official schedules if possible; otherwise delayed buses left the terminus at once. Improvements in mean wait time and standard deviation of wait time resulted from the use of controls (see Figure 6.2). Passenger travel time was slightly increased by control at intermediate stops; more than one intermediate stop per direction was not recommended. On-line control strategies were then experimented with in the Toulouse bus system; however, not enough data was collected to draw definite conclusions.

6.2.2.3 Operational Policy Strategies

A highly unreliable service can result when many buses are missing from scheduled service because either drivers or buses are not available to make scheduled runs. This problem is particularly severe in cases where buses run infrequently. To alleviate this problem, the operator may provide a reserve pool of drivers and buses to replace buses that are cancelled. However, the availability of fewer drivers and vehicles for regularly scheduled service will result in a new service scheduled at lower frequency. An alternate solution is to reschedule on a day-to-day basis, utilizing all available vehicles and providing equal headways in a schedule which would change daily. A TRRL study (Bly and Jackson, 9) utilizing analytic and simulation models investigated these strategies to determine potential effects on passenger wait time. The model results indicated that both the strategy of rescheduling buses and the strategy of reserve bus pools provided similar reductions in passenger waiting time (when compared to original schedule, 10% of buses missing) for service headways below about 12 minutes. The percentage reductions obtained were roughly equal to the mean percentage of buses missing from the original scheduled service. At longer headways, using reserve bus pools provided still larger waiting time reductions (21% for a 30 minute headway, when 10% of the buses were missing, on average) while rescheduling buses caused waiting times to increase (see Figures 6.3, 6.4).

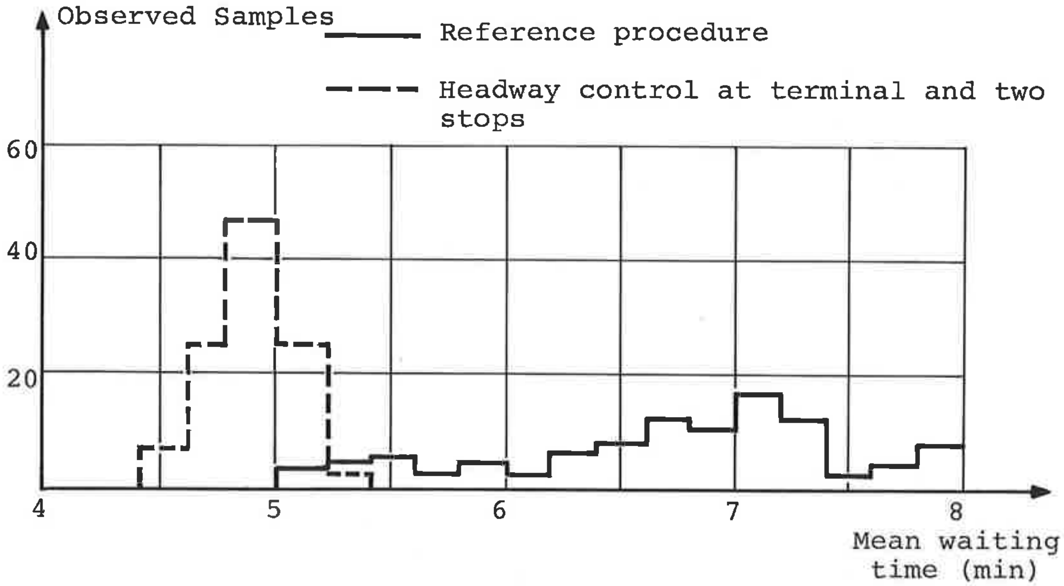


FIGURE 6.2 HISTOGRAM OF THE MEAN WAITING TIME
(ON 100 EVENING PERIODS 4-10 P.M.)
(Giraud and Henry, 8)

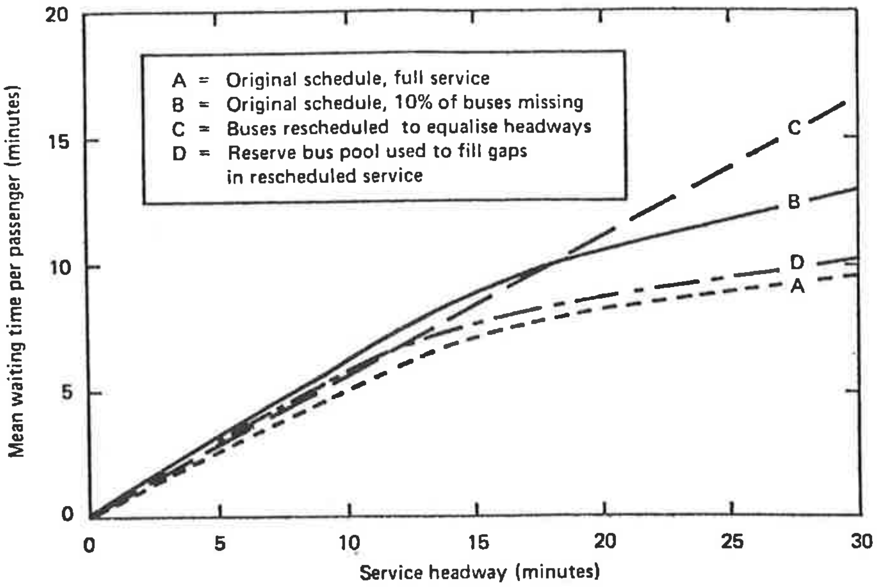


FIGURE 6.3 COMPARISON OF EFFECT OF RESCHEDULING STRATEGIES ON PASSENGER WAITING TIME, WHEN THE PROBABILITY DISTRIBUTION OF MISSING BUSES HAS A MEAN OF 10%, WITH STANDARD DEVIATION OF 2.5% (Bly and Jackson, 9)

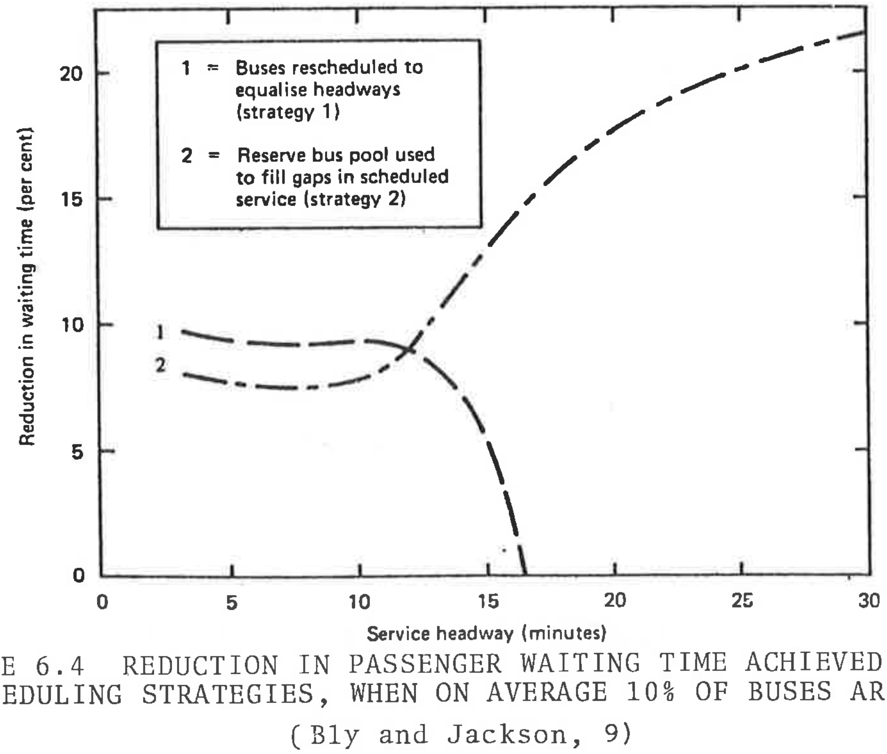


FIGURE 6.4 REDUCTION IN PASSENGER WAITING TIME ACHIEVED BY USE OF RESCHEDULING STRATEGIES, WHEN ON AVERAGE 10% OF BUSES ARE MISSING

(Bly and Jackson, 9)

Due to the difficulty of constant rescheduling, the reserve bus pool strategy may be preferable at all headways.

Where reserve pool strategies are used, the analysis showed that the number of buses in fixed schedule service should equal the average number of available buses for short headway services, and slightly less than that for longer headway services. Model predictions have not been verified in practice and there would be value in experimenting with these strategies in existing operations which have availability problems.

6.3 DIAL-A-RIDE

Since dial-a-ride has no specified route or headways, it does not experience the same error magnification process as does fixed route service. However, dial-a-ride has its own inherent unreliability due to dynamic routing and scheduling in response to requests for "immediate" service. Figure 6.5 illustrates the development of unreliability in dial-a-ride service and the application of preventive and corrective strategies. Note the addition of a feedback loop from unreliability to the likelihood of cancellations and no-shows. On dial-a-ride, if a pick-up is sufficiently late, the rider may refuse service (cancellation or no-show) and thus reduce further lateness. Cancellations and no-shows may have various effects on other users depending on initial lateness, dwell time policies, and the scheduled time of other users' pickups and dropoffs with respect to the cancelled trip.

A number of strategies may counteract some of the effects of demand variation, dynamic routing and scheduling, and cancellations and no-shows. Table 6.4 classifies these strategies by type. Each strategy is summarized below, and strategies for which studies have been performed are later analyzed in more detail.

6.3.1 Summary of Strategies

Dial-a-ride is very sensitive to demand level. Variation in demand from day-to-day may inconvenience regular users. One strategy may be to guarantee service to those who make advance requests (especially for work and school trips which require a high degree of reliability) and accept other requests only when conditions permit. As demand increases relative to supply, a decision must be made whether to let service degrade for everyone or to establish service norms for certain classes of users and to refuse

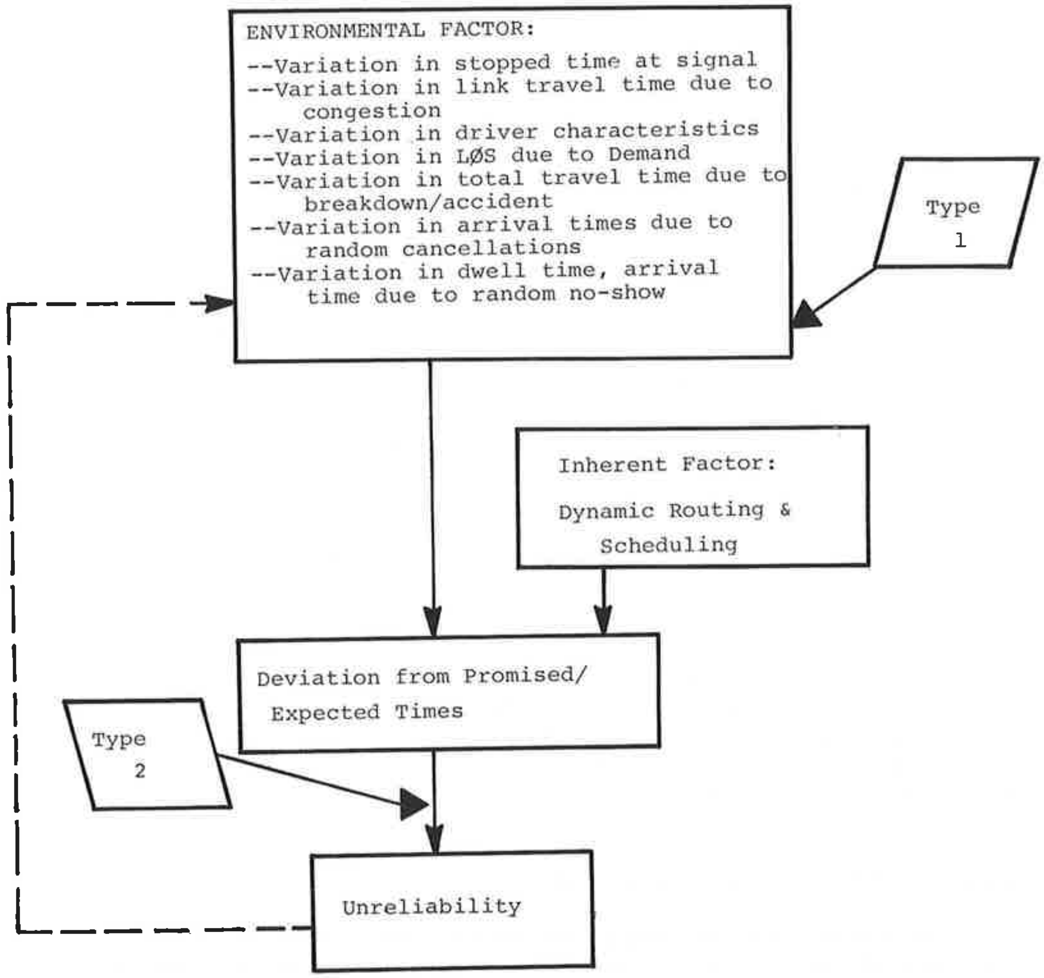


FIGURE 6.5 APPLICATION OF IMPROVEMENT STRATEGIES (DIAL-A-RIDE)

TABLE 6.4 CLASSIFICATION OF STRATEGIES (DIAL-A-RIDE)

Strategies	Preventive Type 1	Corrective/ Restorative Type 2
a) Dwell Time Policies	X	
b) Cancellation: Restructure Routes/Schedules		X
c) Guarantee Advance Requests Only	X	
d) Computerized Dispatching for DRT	X	X
e) Transition from DRT to Scheduled Service	X	
f) Signal Priority	X	
g) Speed Modification	X	X
h) Reserve Vehicles and Drivers	X	X

service to others (or offer unacceptably bad service to them), rather than allow these norms to be broken.

Another inherent problem with dial-a-ride is that the dynamic nature of demand responsive dispatching introduces variation, making the pick-up time estimates unreliable. Computerized dispatching of dynamically-routed transit vehicles might reduce variance of waiting times through better assignment and estimation of arrival times.

A cancellation of a request for service of a dial-a-ride user will cause early arrivals at future pick-up points. The simplest operating policy in the event of a cancellation is to arrive early at the next stop and wait until the promised pick-up time if necessary. This will likely annoy passengers waiting onboard. An alternate policy might be to attempt to reroute the vehicle to utilize the extra time productively (e.g., dropping off people on board before proceeding to the next pickup).

The occurrence of no-shows introduces additional variance to the system. Long unexpected dwell times will affect travel times for passengers on-board and wait times for future pickups. To minimize these dwell times, maximum dwell time policies should be set.

It should be noted that scheduled services are typically more reliable than DRT and it is unlikely that the innate unreliability of DRT can be fully overcome. As a result, transition to subscription or route-deviation schemes may be desirable where trip densities permit and when the market served is particularly sensitive to reliability (i.e., work trips).

Other strategies for improving dial-a-ride service are similar to those for fixed route bus service. Signal priorities and speed modifications could be used to reduce lateness in promised pick-ups and drop-offs. Reserve vehicles and drivers are a good strategy for combatting potential breakdowns or when demands are too high for the scheduled fleet to handle them effectively.

6.3.2 Analysis of Strategies

In this section, the following strategies for improving reliability in DRT systems will be examined in detail: dial-a-ride dispatching algorithms, alternative service options, dwell time policies and cancellation strategies.

6.3.2.1 Dial-A-Ride Dispatching Algorithms

Computer dispatching may have significant impacts on level of service due to the speed and efficiency of processing and storing information. In addition, the reliability of service (invariability of level of service) may be improved by more complex dispatching policies that would not be possible with manual procedures. For example, M.I.T. research on dial-a-ride algorithms (Wilson, et al., 10) found that a quadratic objective function produces assignments giving a more uniform level of service. This quadratic form (now being used in Rochester) provides higher mean service times but with lower standard deviations than the linear form used in Haddonfield.

As shown in Figure 6.6, for a quadratic utility function, the disutility of attribute $x = (a+b)/2$ is less than the average of the individual utilities of $x=a$ and $x=b$. This preference may be characterized as being "risk-averse". In the case of transit travel time, wait time, etc., this type of objective function may favor those distributions with smaller variances despite higher mean values. For many-to-many service, the provision of relatively uniform wait times or overall travel speeds is a feasible method for improving the actual and perceived reliability of estimated service times for individual trip makers.

Additionally, computer dispatching may improve reliability by reducing the variability of the service characteristics weighed most heavily by travelers for different types of demand. For instance, delivery time may be most signficant to advance callers; response time may be most important for immediate service users. Transfer passengers have special constraints, since missing a connection may have serious penalties. Thus deviation in arrival time at the transfer point may be intolerable on the late side and inconvenient if very early. The inconvenience of early arrivals may be alleviated by the scheduling of fixed transfer stops where the dial-a-ride vehicles wait for the connecting bus (causing delays to other on-board passengers). Advanced computer algorithms have been developed for services which utilize these fixed stops. Finally, the assurance of more reliable service to advance request passengers will encourage users to call in advance and help reduce the inherent variability of immediate requests for dial-a-ride. This would facilitate providing guaranteed service for regular customers who would be largely protected from demand fluctuations.

Preliminary evidence from the Rochester demonstration project tends to support the hypothesis that computer

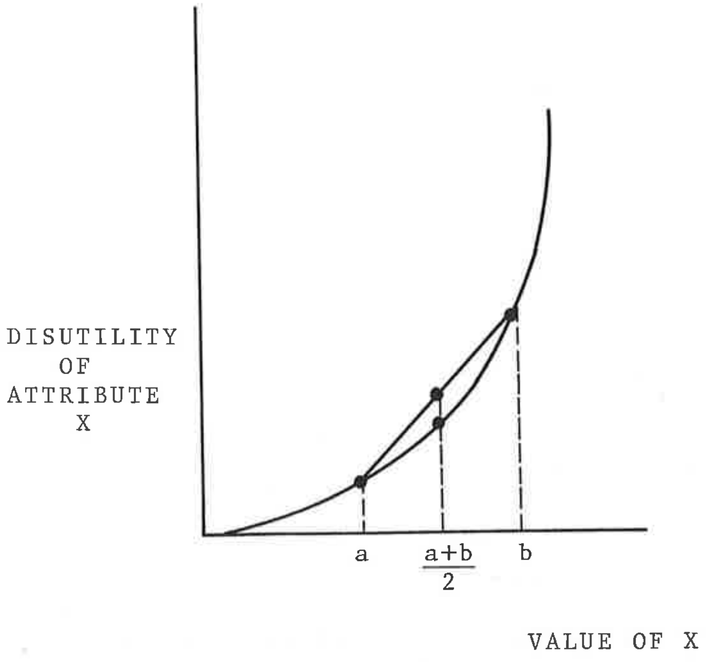


FIGURE 6.6 QUADRATIC DISUTILITY FUNCTION

dispatching may be able to provide generally improved service reliability over manual dispatching (Wilson and Colvin, 11). In the second dial-a-ride system implemented as part of the project in Irondequoit, NY, level of service data were gathered for a sample of passengers carried under both manual and computer dispatching. These results are summarized in Table 6.5 which also indicates whether the difference was statistically significant at the 95% level of confidence.

Focusing only on those differences which were statistically significant, computer dispatching appears to provide better service for immediate requests in terms of lower mean wait times, and for advanced requests in terms of lower pickup time deviations. On the other hand, manual dispatching appears to provide lower ride times for advanced requests. Four of five standard deviations whose difference were statistically different were lower with computer dispatching (standard deviations for individual wait times and pickup deviations for immediate requests, and for individual and mean daily pickup deviation for advanced requests). However, these standard deviations measure aggregate variability of level of service components between individuals and not the variability of level of service components for a given individual or individuals taking the same trip different times. Thus the significance of these measures is somewhat suspect, though the very much lower standard deviation of wait time and pickup deviation for immediate requests under computer dispatching suggests that computer dispatching is superior in this regard. Analysis of disaggregate level of service data would have been preferred. Finally, no statement about the overall superiority or inferiority of computer control can be made in the absence of reliable information on customer preferences. Given the flexibility in computer control, the mix of service characteristics could relatively easily be controlled to reflect real passenger preferences, once they were known.

6.3.2.2 Alternative Service Options

Dial-a-ride service may also provide an opportunity to improve the user's satisfaction with system service levels by placing more importance on factors of greater consequence to users. Generally, dial-a-ride is offered as advanced request or immediate request service; advanced request is most widely used for work trips. Surveys in Rochester (Menhard, 12) found advanced request users weigh maximum travel time and bus arrival prediction time accuracy (reliability of arrival and travel time) more heavily than

TABLE 6.5 COMPARISON OF MANUAL AND COMPUTER DISPATCHING IN IRONDEQUOIT,NEW YORK[1]

Passenger Type	Measure of Service[2]	Manual	Computer	Better	%Improvement	Statistically Significant Difference at 95% Level
Immediate	$M(W)$	29.69	17.44	C	41%	*
Requests	$\sigma(W)$	20.42	10.23	C	50%	*
	$\sigma(W_D)$	5.93	4.12	C	31%	
	$M(P)$	5.51	6.25	M	(13%)	
	$\sigma(P)$	16.24	8.53	C	47%	*
	$\sigma(P_D)$	3.94	2.49	C	37%	
	$M(R)$	10.41	10.72	M	(3%)	
	$\sigma(R)$	5.49	7.19	M	(31%)	*
	$\sigma(R_D)$	1.35	1.28	C	5%	
Sample Size		228	303			
Advanced	$M(P)$	7.12	4.24	C	40%	*
Requests	$\sigma(P)$	15.29	10.87	C	29%	*
	$\sigma(P_D)$	5.99	2.32	C	61%	*
	$M(R)$	10.60	13.05	M	(23%)	*
	$\sigma(R)$	7.11	7.27	M	(2%)	
	$\sigma(R_D)$	1.61	1.83	M	(14%)	
Sample Size		396	473			

[1] Source Wilson and Colvin (11)

[2] W, W_D are the individual and mean daily wait times.
P, P_D are the individual and mean daily pickup time deviations.
R, R_D are the individual and mean daily ride times.
$M(x), \sigma(x)$ are the mean and standard deviation of the observation x.

mean travel time. Immediate request users were found to weigh mean travel time and the mean and variance of wait times more heavily than bus arrival prediction accuracy (variance of arrival time). User preferences were found to vary considerably. Thus, a multi-class service would seem to be advantageous, where each user could pick the type of service that suits him/her best, in a similar way to mode choice decisions. Computer assignment may provide management with greater control over the service provided each passenger and thus may enable the operator to offer multi-class service dial-a-ride. Users may not only trade-off different service attributes, such as arrival accuracy and travel time, but may tradeoff higher fares for priority service with lower fares for low priority service.

6.3.2.3 Dwell Time Policies and Cancellation Strategies

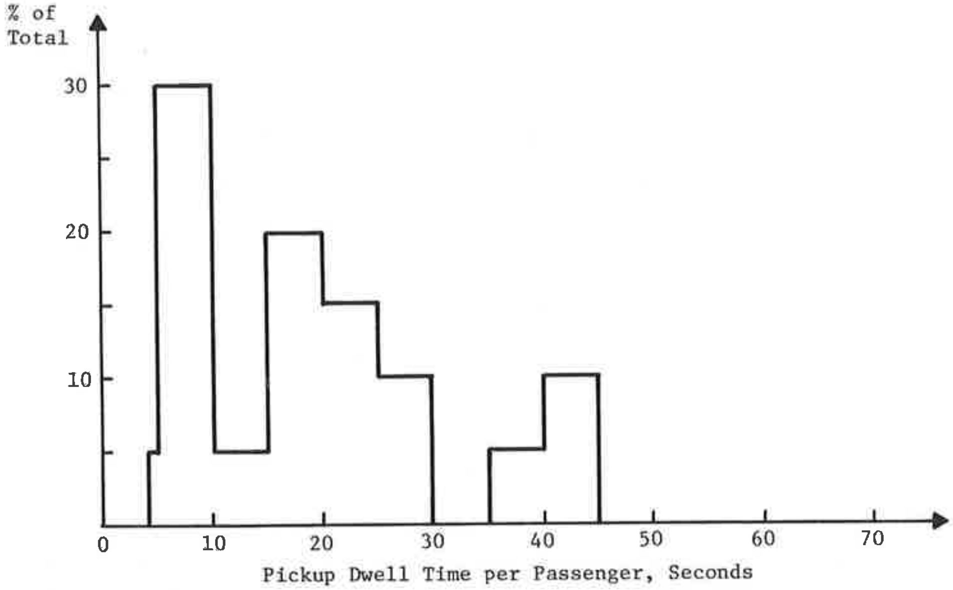
In an effort to reduce the effects of no-shows and improve vehicle productivity, strategies to minimize dwell time for pick-ups may be employed. In Rochester, New York, wait policies are liberal. The driver waits for the customer to board when the bus arrives at the pick-up location. The driver contacts the control center for instructions when no one has appeared after some period of time (usually 2-5 minutes). Usually the control center will attempt to call the user, and only then will the vehicle depart.

In Gaithersburg, Maryland, a tighter policy is employed. The driver sounds the horn on arrival, waits 30 seconds, and then calls the control center for instructions if no one appears. In addition, at other than single unit dwellings, users are requested to wait outside. Of course, if vehicles are early, such policies cannot be implemented. If buses tend to arrive slightly later than the advertised time, dwell times may be reduced at the expense of user wait time.

Gaithersburg and Rochester pick-up dwell times are exhibited in Figures 6.7 and 6.8 respectively. It is clear that the Gaithersburg policy has reduced very long dwell times significantly. Adoption of a Gaithersburg-type policy is recommended in a MITRE report (Chung, 13) to improve vehicle productivity. These policies also seem useful in efforts to improve reliability.

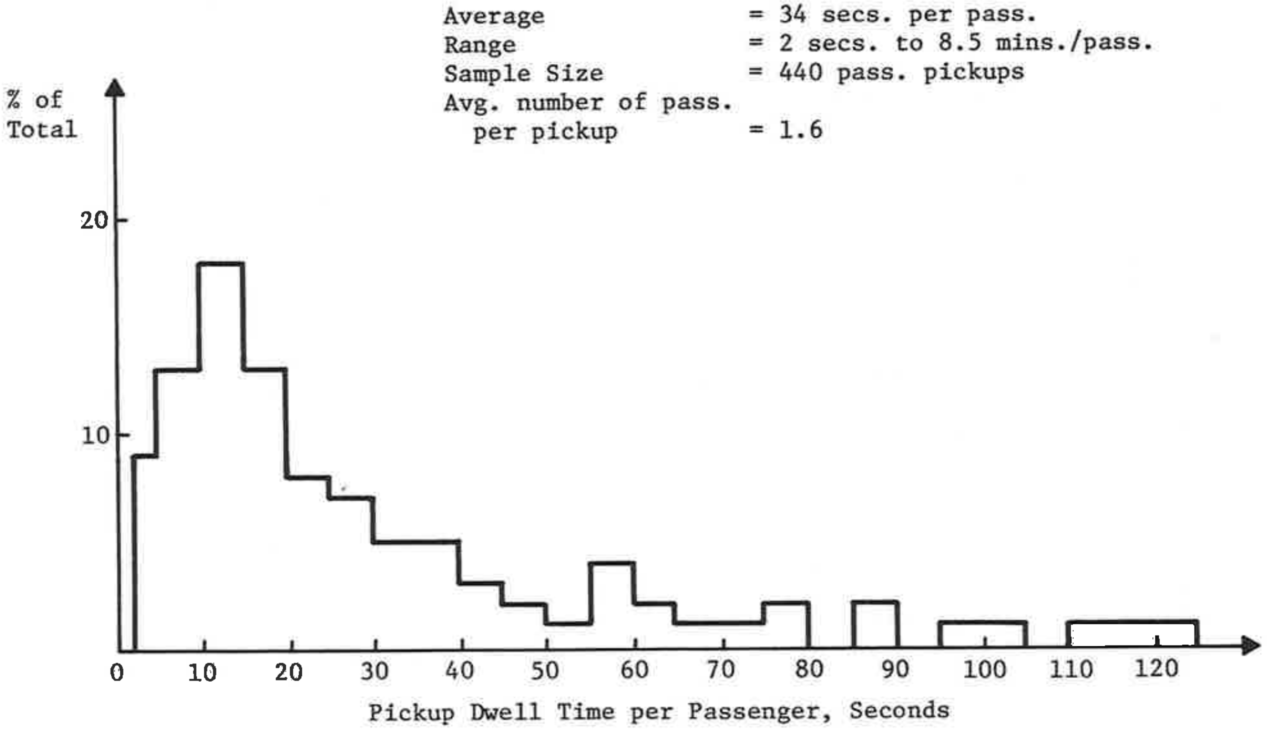
In the event of a customer cancellation, with manual control systems, the stops will typically be removed from the affected vehicle tour. As discussed earlier, this may

Average = 20 seconds/pass.
Range = 5 to 45 seconds/pass.
Group Pickup Factor = 1.7 pass./pickup



Note: From a sample of 20 passenger pickups. Data collected between 9:00 a.m. and 4:00 p.m. on four days: 1/1/76 (Thurs.), sunny; 1/21/76 (Wed.), cloudy with occasional snow; 2/10/76 (Tu.), sunny; and 2/11/76 (Wed.), variably cloudy. The prescribed policy is for the driver to wait a maximum of 30 seconds.

FIGURE 6.7 GAITHERSBURG ZONAL DIAL-A-RIDE PICKUP DWELL TIMES AT SINGLE UNIT HOUSES (Chung, 13)



Average = 34 secs. per pass.
Range = 2 secs. to 8.5 mins./pass.
Sample Size = 440 pass. pickups
Avg. number of pass. per pickup = 1.6

Note: Data collected in July 1975.
Percentages less than one percent are not shown.

FIGURE 6.8 ROCHESTER GENERAL DIAL-A-RIDE PICKUP DWELL TIMES (Chung, 13) lead to no-shows further down the tour. In the Rochester computer dispatch system, a passenger cancellation will automatically result in the tour of that vehicle being restructured if any improvements can be made.

6.4 CONCLUDING REMARKS

This chapter has examined several proposed strategies to improve the reliability of transit service. This included a description and analysis of strategies which are applicable to transit operations in general, as well as those specifically oriented to fixed route and dial-a-ride services.

The UMTA Service and Methods Demonstration Program has investigated a number of priority strategies for fixed route services; European studies have been more extensive. Generally, these studies have indicated that some strategies appear to be successful in combating unreliability. However, limitations in the implementation of several strategies and the lack of proper measures to evaluate reliability improvements have hampered research efforts to date. Studies of dial-a-ride reliability improvement strategies have been very limited.

Further studies of reliability improvement strategies are clearly needed before any definitive conclusions can be reached. Chapter 7 discusses suggestions for further work in this area.

REFERENCES

1. Turnquist, Mark A. and L.A. Bowman. "The Effects of Network Structure on Reliability of Transit Service." Northwestern Univ., presented at International Symposium on Travel Supply Models. Montreal, 1977.

2. Chapman, R.A. "Bus Reliability-Definition and Measurements." (Working Paper No. 18). Transport Operations and Research Group, Univ. of Newcastle Upon Tyne (Great Britain), 1976.

3. Koffman, David. "A Simulation Study of Alternative Real-Time Bus Headway Control Strategies." Presented at the 57th Annual Meeting of the Transportation Research Board, Washington, DC, January 1978.

4. Muzyka, A., D. Schmieder, J. Fantasia, and G. Feeney. *Traffic Signal Strategies for Bus Priority - A Simulation Study*. Transportation Systems Center, Urban and Regional Research Series, (Forthcoming).

5. Jackson, R.L. and D. Stone. *Experiments in Bus Service Control Using an Interactive Model of a Typical Urban Bus Route*. Univ. of Newcastle Upon Tyne (Great Britain), 1976.

6. Bly, P.H. and R.L. Jackson. *Evaluation of Bus Control Strategies by Simulation*. TRRL Laboratory Report 637, Transport and Road Research Laboratory, Dept. of Environment, Crowthorne (Great Britain), 1974.

7. Barnett, Arnold. "On Controlling Randomness in Transit Operations." *Transportation Science*. Vol. 8, No. 2, May 1974.

8. Giraud, Alain and J.J. Henry. *On-Line Control of a Bus Route - Studies and Experiments in Toulouse*. Third International Symposium on Control in Transportation Systems, Columbus, OH, 1976.

9. Bly, P.H. and R.L. Jackson. "Scheduling for Depleted Services." Univ. of Newcastle Upon Tyne (Great Britain), 1975.

10. Wilson, Nigel, R. Weissberg and J. Hauser. "Advanced Dial-a-Ride Algorithms Research Project: Final Report." UMTA-MA-11-0024, MIT Report R76-20, March 1976.

11. Wilson, Nigel H.M. and N. J. Colvin. "Computer Control of the Rochester Dial-a-Ride System." MIT Center for Transportation Studies Report 77-22, July 1977.

12. Menhard, Henry R. "User Preferences Pertaining to Dial-a-Ride Service Options." (M.S. Thesis, MIT, 1976).

13. Chung, Cady C. "Dial-a-Ride Vehicle Time Analysis and Vehicle Productivity." MITRE, VA, 1976.

7. PROPOSED DEMONSTRATIONS AND FUTURE STUDIES

The earlier chapters of this report have provided a review of the state-of-the-art of knowledge regarding transit service reliability and have developed a conceptual approach for understanding the relationship between transit reliability and user and operator behavior. The material presented in this report has provided supportive evidence that transit reliability is an important issue to transit users, operators and planners, and that transit reliability improvement strategies can impart significant benefits to both transit users and operators.

However, it is also clear that our present knowledge of transit reliability is still quite limited and needs to be extended. This chapter proposes a set of activities to supplement our knowledge about reliability and to hasten its application to the SMD Program. Two types of projects are described in this chapter: 1) analytical studies to examine, in more detail, the importance of reliability and its effects on operator and traveler behavior, and 2) implementation and evaluation of SMD projects explicitly directed at improving transit service reliability. It is likely that several of the analytical studies described in this chapter can be conducted using data collected from the proposed SMD projects.

7.1 RESEARCH PROJECTS AND RELATED STUDIES

7.1.1 Causal Reliability Relationships

Understanding the determinants of unreliable service in a particular operating environment is critical in selecting and designing the most effective reliability improvement strategies. Development of causal empirical models of the determinants of transit service reliability is proposed.

Utilizing simulation models and the empirical data collected before and during the proposed demonstration projects, the following impacts on reliability ought to be explored and incorporated in the empirical relationships: (1) demand variation in time and in space, and in mode and in path, (2) traffic networks and volumes, (3) vehicle availability, (4) traffic management, and (5) weather. Careful attention should be devoted in the models to the interdependence between the reliability of different service attributes. From the relationships, it may be possible to develop planning guidelines for incorporating reliability improvement strategies into transit operations.

7.1.2 Reliability Measures

The reliability measures developed in Chapter 4 of this report should be empirically tested and, where necessary, modified as the demonstration projects are implemented and evaluated. Measures should be examined for appropriateness as tools for evaluating service attributes from both the traveler's and transit operator's perspective, and for feasibility and cost-effectiveness of required data collection. In addition to quantitative mesaures of service attribute distributions, measures of the benefits of reliability improvements should also be developed (e.g., reduction in the number of late or missed appointments).

7.1.3 Travel Behavior Models

Models of the impact of transit reliability on travel behavior are needed to understand and evaluate the effects of transit reliability improvement strategies on travelers and to fine tune and improve the strategies themselves. More generally, such models could assist transportation planners in designing transit systems. From a theoretical viewpoint, these models would represent an important advance in the fields of travel behavior and demand modeling. Empirical data collected during the demonstration projects and results obtained in deriving causal reliability relationships might be used to develop these models. The models should be capable of representing the role of reliability in influencing travel frequency, destination choice, mode choice, route choice, and departure time decisions.

In developing new travel behavior models, consideration must be given to identification of variables that can accurately measure transit reliability and its impact. Mean travel time is the only service attribute vaguely related to reliability in current travel behavior models. Based on the material discussed in Chapter 2, the variability of travel time and its effect on arrival time should be incorporated into future travel behavior models; measures proposed in Chapter 4 might serve well as reliability variables in travel behavior models.

The incorporation of reliability measures into travel behavior models also requires a more disaggregate view of transit supply. Information on service attribute variability which has been used in model estimation has been based on zonal averages; for example, all individuals making trips from zone 1 to zone 2 are assigned a travel time variance equivalent to the travel time variance associated with travel from the zone 1 centroid to the zone 2 centroid. Much information on the variability of individual travel is lost when zonal averages are used, often introducing bias into model estimates. Furthermore, individual travelers may value reliability differently for different trip purposes. Research should be directed at developing a practical strategy for incorporating disaggregate supply information (for individual trips and across individual trip purposes) in travel behavior models.

It is hoped that an important result obtained from the models will be the determination of the traveler's value of reliability relative to his/her value of mean travel time. This relationship would explicitly define the basic traveler tradeoff between travel time and its variability and would give planners a better understanding of potentially effective service improvement strategies. The models could be validated using the "after" data from the SMD projects.

7.1.4 Operator Behavior Relationships

The reliability of service attributes plays a significant role in operator behavior as well as in travel behavior. From the operator's viewpoint, unreliability results in decreased productivity for both vehicles and drivers, and thus increased costs of operation. Understanding the effects of unreliability on transit operators and other producers of urban transportation services may enable these producers to allocate their resources more efficiently. Thus, it is proposed that further research be conducted on (1) the impact of unreliability on transit providers, and (2) provider responses to changes in reliability levels, with the aim of developing improved operator policies. Empirical data collected before and during the proposed demonstrations would be the major source of data for this study.

7.1.5 Simulations

In some or all of the demonstration projects, it is proposed that simulation models be used during both the development and evaluation phases of the project. The simulation models could be used to assist in the final design of the innovative project features and the details of their deployment. During the demonstration period, the models could be used to fine tune the innovative features (i.e., when something unaccounted for in the original demonstration design is affecting the outcome adversely). The models could assist in understanding the outcome of the project by being used for parametric studies of the influence of several project variables. As discussed earlier, simulation could also be used as a tool in developing empirical models of the determinants of transit reliability.

In selecting route simulation models that would be useful in this capacity, reference should be made to a TSC staff study by Waksman and Schmieder (1). The report examines existing bus route simulation models, focuses on their strengths and weaknesses, and recommends the set of simulation tools which are capable of carrying out an effective analysis of bus service improvement strategies. For simulating the operations of a demand responsive transit system within an urban area, the Massachusetts Institute of Technology's Dial-A-Ride Simulation Model appears to be useful.

7.2 PROPOSED DEMONSTRATION PROJECTS

Improvement in transit service reliability is a major objective of the SMD Program and is also intimately related to the SMD objectives of improved travel time and increased vehicle productivity. Strategies directed at controlling transit unreliability can result in significant time savings for both users and operators, and improvements in reliability are likely to produce an increase in revenues and a decrease in operating costs, desirable objectives for any transit agency to pursue.

A number of reliability improvement strategies have been proposed which might be of interest to the SMD Program. These strategies include holding vehicles as they proceed along a route, employing signal pre-emption on congested arteries, and using other routing and scheduling techniques to regulate service. The following discussion proposes some potential demonstrations directed at designing and measuring the effectiveness of specific strategies in improving transit reliability.

Evaluation of these demonstrations would require the collection and synthesis of before-demonstration data and during-demonstration data for comparative analysis. At a minimum, the impacts of reliability improvements on both the user and the operator would be assessed. User impacts would include an analysis of mode share, mean wait and ride times, and the variability in wait and ride times. Operator impacts would include an examination of cost savings, headway distribution, and a comparison of scheduled and actual arrival times.

7.2.1 Control Point Holding Strategies

From previous research it appears that holding strategies may be the most promising for improving transit reliability. There are many variables involved in such strategies: number and location of control points, criteria for application, length of delay, stages of application, and the quantity of information required for these decisions. It has been found that the number of control points should be kept small; otherwise increased travel time resulting from holding buses will likely outweigh the reduced wait time benefits. Control points should be located where significant passenger turnover occurs so that total passenger delay is minimized. Where few passengers continue through such a point, splitting bus routes into two separate routes may be a better strategy. Distributing corrective action over a number of consecutive buses may result in more flexible policies and less inconvenience to passengers. Different policies may be optimal at various points along a route, where different passenger boarding patterns prevail. The effectiveness of control strategies will clearly depend on the quantity and quality of data available for decision-making.

The demonstration would entail selection of a fixed route bus system which is significantly affected by service unreliability. Two or three of the more unreliable routes would be examined, and location of control points selected. Buses would be equipped with radio telephone equipment and all drivers would maintain communication with a bus dispatch center. Each bus would report to the dispatcher when reaching a control point, relaying information on time of arrival, volume on the bus and the number of people waiting to board at the control point. The dispatcher, aware of the status of the buses prior to and succeeding the bus located at the control point, would make a decision of whether to hold the bus and the duration of time to hold the bus. The effects of this strategy would be measured in terms of benefits to the users and operators and the tradeoffs between benefits to people waiting for a vehicle and those delayed by a holding strategy.

Control point holding strategies are currently in effect in Dublin, Ireland and Toulouse, France. In both cases, the transit operating agencies have reported decreases in the mean and variance of passenger wait times. Neither agency has conducted an analysis of passenger in-vehicle times, partronage, and operator costs.

Depending upon the success of employing a holding strategy, the demonstration could be extended to include other control strategies which would involve radio-telephone communication and a dispatching process. This would include coordination of the following strategies:

1. Transferring passengers and turning particular buses.
2. Injecting standby buses.
3. Running buses out of service until reaching particular points on the route.
4. Altering bus routes on a temporary basis.

Considerable work has yet to be done on holding strategies. The simple holding strategy based on a two-point lateness model (Barnett, 2) might be refined. Strategies could then be tested with alternative models, different numbers of control points, and different sources of information.

It might be most efficient to coordinate such a study with the SCRTD-AVM demonstration in Los Angeles. When AVM equipment is installed, holding strategies will be applied and effects evaluated. A comprehensive reliability demonstration might include the testing of numerous strategies with various types of monitoring: manual, radio and automatic. However, in using the Los Angeles site, care must be taken to separate the effects of the AVM system from the effects of the control strategies, a potentially difficult task.

7.2.2 <u>Priority Schemes</u>

Priority schemes, particularly exclusive right-of-way for buses, have been significant elements in the Service and Methods Demonstration Program. These schemes have been predominantly directed at reducing mean transit travel times to encourage modal shifts to transit. Exclusive lanes or freeways have been successful in attracting riders by effecting such travel time reductions and by doing so in view of auto users. Kulash found the reliability effects of such strategies to be minor, based on simulation results, and concluded that such improvement schemes should be justified on the basis of reduction of mean travel time rather than variance. Since numerous demonstrations of exclusive lanes have already been undertaken, it is recommended that in future demonstrations of this nature, efforts be taken to insure that reliability effects are measured to assess the validity of Kulash's statement.

Signal preemption is another priority strategy which is generally directed at mean travel time. By preventing major delays at signals, the preemption strategy should reduce travel time unreliability as well. Signal preemption may also be used as a control device specifically oriented towards improving reliability. Preemption may be selectively used to speed up a bus behind schedule. The effects of these alternate uses of the signal preemption capability should be studied and measured in a demonstration. Moreover, buses can be given a form of priority by retiming traffic signals in a manner so as to facilitate bus movement. The low cost involved makes retiming of traffic signals particularly attractive.

The following demonstrations are suggested:

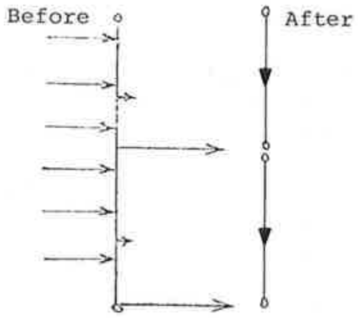
1. Signal pre-emption only (selective and general use)
2. Exclusive lanes only
3. Exclusive lanes with signal preemption (selective and general use)
4. Exclusive lanes with separate signal phases
5. Exclusive lanes with speed modification or other control strategies
6. Retiming traffic signals to facilitate bus movement.

7.2.3 Restructuring Routes

As discussed earlier in this report, service attribute variability increases along a route and that splitting long routes into shorter routes, or into express and local routes may be valuable strategies to improve reliability. If, in fact, reliability can be improved by redesigning routes at little or no additional cost, such strategies may be a first step in improving reliability. Expensive right of way, priority and automated control schemes should only be considered after low cost strategies, such as route restructuring, have been attempted.

While the aim of a route redesign demonstration would be to determine the effects of redesign on reliability, an important element to examine will be the effects of redesign on other aspects of service. For example, the inconvenience of transfers and the predominance of particular travel patterns must be considered in route redesigns. Where major turnover occurs, routes may be split into two entirely separate routes (see Case I in Figure 7.1). The Harvard-Dudley line of the MBTA in Boston, for example, has a significant turnover in the Auditorium area where Dudley-Back Bay transfer passengers alight and Back Bay-Harvard

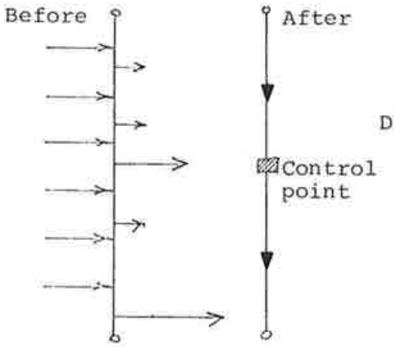
Case I - Split Route



Benefit: Improved Reliability of Arrival Time, Headway Regularity (Wait Time)

Disbenefit: Transfers required for all through trips

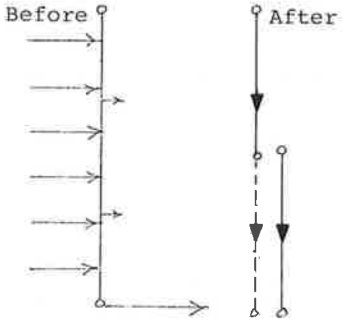
Case II - Holding Point



Benefit: Improved Reliability of Arrival Time, Headway Regularity (Wait Time)

Disbenefit: Through passengers are subjected to delays at the control point

Case III - Express/Local Service



Benefit: Improved Reliability of Arrival Time, Headway Regularity (Wait Time)

Disbenefit: Reduced frequency, Transfer required for intra-route trips

Note: Arrows to left and right of the line under the "Before" heading indicate the initial boarding and alighting distributions respectively. Each distribution is appropriate for the particular reliability improvement strategy suggested.

FIGURE 7.1 STRATEGIES FOR RESTRUCTURING ROUTES passengers board. Boarding counts would give an accurate assessment of the inconvenience caused by splitting the route at Auditorium versus the improved service which would result from shorter more reliable routes. Where turnover is less significant, holding strategies may be utilized to provide through service at some smaller inconvenience (delay) due to the possibility of layover (Case II). Finally where long routes experience constantly increasing loads with little major turnover, the use of express and local services may be the ideal strategy (Case III). In both the Case I and Case III redesigns, consideration should be given to minimizing transfer times.

7.2.4 Rescheduling

Simple planning by the operator with regard to schedules may also lead to improved reliability. Increasing recovery time between runs and increasing the number of spare drivers and vehicles available for various contingencies will, from the transit user's perspective, involve tradeoffs between reliability and other service components. The net effect of rescheduling may be that traveler benefits derived from the improved reliability outweigh the disadvantages associated with a degradation of other service components. Such rescheduling may have little or no effect on the costs incurred by the operator.

A rescheduling demonstration would determine the actual effects on transit users and operators of making some simple changes on specific routes and systems. The key aspects of such a demonstration would be identifying and measuring all of the level of service tradeoffs experienced by transit users, and the cost implications for the operator.

7.2.5 Traffic Engineering Strategies

In addition to strategies which give buses priority over other vehicles, there exist other methods for improving bus flow and reliability. It would appear that any traffic improvement which removes causes of delay will have a beneficial effect on transit travel time variance as well. Thus, general improvements of roadway facilities should improve transit reliability.

Bus movements should be explicitly considered in the preparation of plans for general traffic improvements. It is not enough to locate bus stops so as not to interfere with other traffic and to remove pedestrians from interfering with the vehicular flow. Each traffic element, automobile, bus, and pedestrian must be specifically considered to design an "optimal" system. This consideration should, however, occur within the context of general improvements in the Transportation Systems Management (TSM) Element.

Specifically, improvements which may prove valuable in increasing reliability of transit operation include:

Improving physical condition of curb lanes -- to permit its use and thus speed up boarding in addition to reducing interference with other traffic.

Providing bus stops of adequate length -- to permit buses to load/unload immediately upon arrival and ease transfers between buses.

Providing adequate turning radii at corners -- to ensure smooth and quick turns into appropriate lanes.

Reorganizing bus stops -- to ease transferring and to reduce delay at stops.

Enforcing parking restrictions -- to clear bus lanes and stops.

Channelization where bus turning movements are heavy (even if few other vehicles make the turn) -- to reduce delays to buses at intersections.

Improving the location/phasing/control of pedestrian movements -- to minimize interference with bus flow.

Coordinating bus stop location and signal timing -- to reduce delays at signals (e.g., alternating near-side and far-side stops where two linked signals occur).

Improving bus stop location -- far side stops have been found to be preferable due to the following:

1. They reduce conflicts between right-turning vehicles and stopped buses.
2. They provide additional intersection capacity by making the curb lane available for traffic.
3. They eliminate sight-distance deficiencies on approaches to intersections.
4. They encourage pedestrian crossings at the rear of the bus.

5. They require shorter maneuvering distances for the buses to enter and leave moving traffic (where there is curb-side parking).

6. At signalized intersections, buses can find gaps to re-enter the traffic stream (where there is curb-side parking).

There are, however, some disadvantages to far-side stops and in some cases, near-side or midblock stops may be more appropriate.

In conclusion, there are a number of traffic engineering strategies that may improve bus flow and reliability. The effectiveness of each strategy has yet to be determined. Study of a sample corridor may assist in examining the impact of TSM strategies.

7.2.6 Reconfiguration of Demand Responsive Transit Service

Demand responsive transit (DRT) service has been identified in this report as having certain characteristics which can produce unreliable service. These include totally flexible and thus variable routing, immediate service with promised times, and the fact that an additional trip adds significantly to vehicle travel time (due to doorstep pick-up and drop-off). A possible strategy for reducing unreliability of demand responsive transit, short of conversion to fixed route, is the development of new hybrid service concepts which offer many of the advantages of dial-a-ride but which overcome some of its reliability problems.

Such modifications to the original many-to-many dial-a-ride service concept include:

1. Scheduled fixed stops with deviation service.
2. Many-to-few service.
3. Integrated service (zonal dial-a-ride & fixed route).
4. Restricted doorstep service to special needs groups/specified time periods.

Existing dial-a-ride services experiencing similar reliability problems might be selected to adopt these modified policies to test their effectiveness in improving reliability and to determine if perceived level of service is acceptable to current and potential users. In some cases these modifications may be made in a staged process to determine effects of each step in a natural progression of service changes. As a control, these changes could be implemented in a sector of a larger dial-a-ride service area.

7.2.7 DRT Multiple Service Options

Algorithms to assign passengers to vehicle tours based on specific preferences of service levels should be developed and tested. A demonstration could assess the ability of the computer dispatch system to improve total passenger satisfaction with service and possibly expand the market to those who previously found the service too unreliable in one respect or another. Interesting conclusions may be drawn from the resulting sub-"modal split" among available DRT options. The relative weighting of the variance and mean of service components for different user groups may become evident. Included in this demonstration could be the implementation of pricing schemes for the different service options provided. The goal of different prices for different services could be the maximization of either overall level of service, ridership on specific services, net revenue, or some combination of these. The demonstration could provide insight into the more effective pricing policies.

7.2.8 Restricted DRT Service

This demonstration would test the acceptability and efficiency of a service which strives for reliable service levels for regular users and puts other users on a stand-by basis at a lower level of service. It is related to the previous demonstration in distinguishing between user groups. This service would evolve towards subscription service with marginal or separate service for other users.

7.2.9 DRT Dwell Time/Rerouting Policies

This demonstration would experiment with alternate policies to deal with cancellations and no-shows on an existing dial-a-ride system. The objective is to develop policies which will minimize the effects of such perturbations on the other users. Such policies may include sounding the horn at arrival and leaving if no response is made within 30 seconds, or calling in to the dispatcher to request instructions if no one shows up within some specified time. If a no-show or cancellation occurs, policies may include proceeding to the next stop, reordering to drop-off some passengers first so as not to arrive early at a pick-up, or adding another pick-up to the tour list. The effects on service reliability would be analyzed for the alternative policies including any significant change in the occurrence of no-shows and cancellations which can be related to the change in policy.

7.3 CONCLUDING REMARKS

The findings of this study suggest that transit reliability has a significant impact on both the traveler and operator, and that transit service reliability improvement strategies can have significant beneficial impacts on both traveler behavior and transit operator efficiency.

There appear to be a variety of ways in which reliability improvements can be obtained. Some of these strategies have exhibited promise in the limited number of empirical and analytic studies which have been conducted.

However, major gaps in our knowledge of transit service reliability problems, improvements, and impacts still remain. The authors hope that the ideas and suggestions contained in this report will prove useful in enhancing our understanding of the effects of service reliability on various elements of travel, and will promote the design and implementation of more effective policies directed at improving transit travel and operator efficiency.

REFERENCES

1. Waksman, Robert and D. Schmieder. "Bus Routes Simulation Models for Studying Service Improvement Strategies." Transportation Systems Center, SS-24-U.3-137, November 1977. *

2. Barnett, Arnold. "On Controlling Randomness in Transit Operations." <u>Transportation Science</u>. Vol. 8, No. 2, May 1974.

---

*Material on file: U.S.Department of Transportation, Transportation Systems Center, Cambridge,MA.

APPENDIX A

DERIVATION OF RELATIONSHIPS BETWEEN TRAVEL TIME AND TRAVEL TIME VARIABILITY*

Total actual trip time (T) experienced by the transit user is composed of the sum of mean trip time (M) plus random delays (D). The random delay facing the user can be expressed as the following probability distribution:

$$g(D) = \frac{1}{\sqrt{2\pi}\, V} \exp\left(-\frac{D^2}{2V^2}\right) \qquad -\infty < D^2 < \infty .$$

That is, D is approximately normal with mean zero and variance V^2. The traveler is assumed to be free to leave as early as he likes in order to be sure to arrive "on time." His premature departure time is denoted by P. His actual trip time is T = M+D. The total time he allows for his trip is A = P+M. Therefore, on average, he will arrive E(A-T) = E(P+M-M-D) = E(P-D) = E(P) = P hours ahead of time.** In any given instance, however, he may be very late if D is large positive or very early if D is large negative, for then P-D will be correspondingly large negative or positive respectively.

The traveler's tastes are assumed to be of the following form. The first component of disutility originates from the fact that his total time allowed for the trip (A) will be wasted time. Since this wasted time becomes harder to fit into his schedule as this block of time increases in length, this component of utility can be expressed as:

$$-wM^{\alpha}\ (M+P) \qquad \alpha \geq 0,\ w > 0$$

where w is his value of time, M^{α} is a factor which allows the loss per hour to increase as the average total trip time increases, and (M+P) is the total time allowed for the trip. Graphically, this component of disutility encompasses an entire family of functions as seen in Figure A.1. Note that the disutility per hour always increases.

The second component of disutility reflects the fact that since actual trip times have an element of randomness,

---

*Summarized from Marfisi, et al., (1).

**E( ) is the expected value of the term in parentheses.

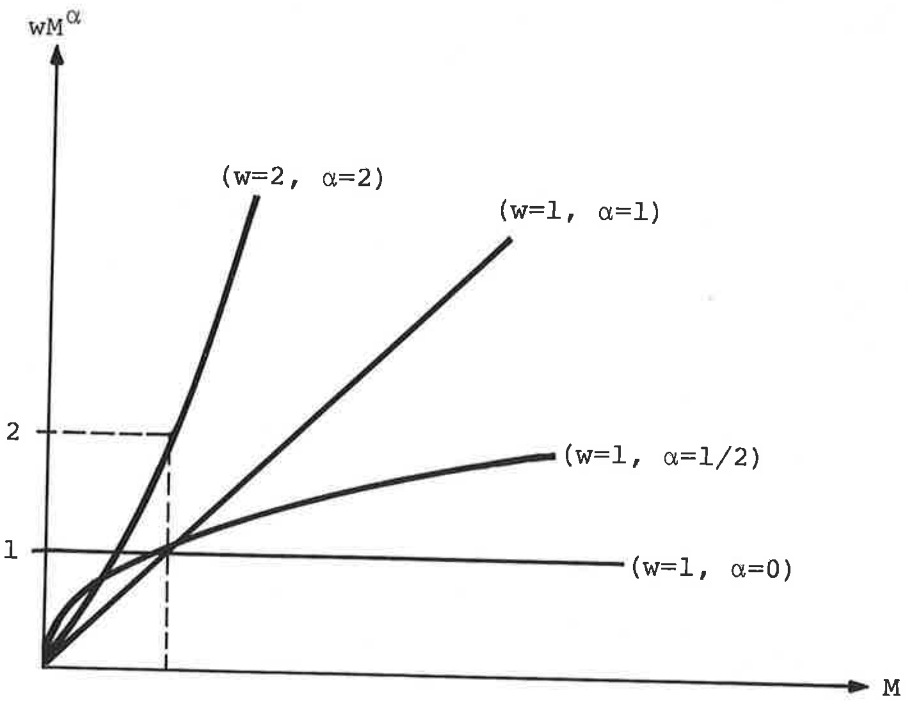


FIGURE A.1 GRAPH OF POSSIBLE DISUTILITY FUNCTIONS FOR DIFFERENT VALUES OF w AND α g(D), the traveler could easily arrive either earlier or later than he expected. If the latter occurs, there may be unfortunate consequences. If the former occurs, the traveler will have some free time which, to be sure, has some value, but is not likely to be as usefully spent as if he could count on having that extra time beforehand. The value the traveler places on a particular random delay can be represented by the following von-Neumann-Morgenstein utility function:

$$b\left\{1 - e^{c(d-P)}\right\} \qquad b,\ c > 0$$

This function is constructed so that if the actual delay (D) is precisely equal to premature departure time (P) there is no loss or gain from this component. The function embraces a wide family of functions depending upon the choice of b and c as is depicted in Figure A.2.

This function can be made as nearly linear as desired merely by allowing c to approach 0 while setting (-bc) at the desired constant slope -k. The rationale for this functional form is that longer delays in excess of premature departure times are more than proportionally costly to shorter ones. Similarly, longer "prearrival" times are less proportionally valuable to shorter prearrival times.

The second component of disutility can now be expressed as:

$$\int_{-\infty}^{\infty} b \cdot \left\{1 - e^{c(D-P)}\right\} \cdot g(D)dD$$

where each value of D is multiplied by the probability of its occurrence.

The traveler's total disutility, B, can now be expressed as:

$$B = -wM^{\alpha}(M+P) + \int_{-\infty}^{\infty} b \cdot \left\{1 - e^{c(D-P)}\right\} \cdot g(D)dD + K$$

$$= -wM^{\alpha}(M+P) + b - b\exp\left(-cP + \frac{c^2V^2}{2}\right) + K$$

where K is constant scaled so that B=0 when M and V are such that the trip is just worth making.

Assuming that the traveler desires to minimize his disutility, B, by an appropriate selection of a premature time P, $\frac{\partial B}{\partial P}$ is set equal to zero, and the solution for P is computed.

$$\frac{\partial B}{\partial P} = -wM^{\alpha} + bc\exp\left(-cp + \frac{c^2V^2}{2}\right) = 0 \ .$$

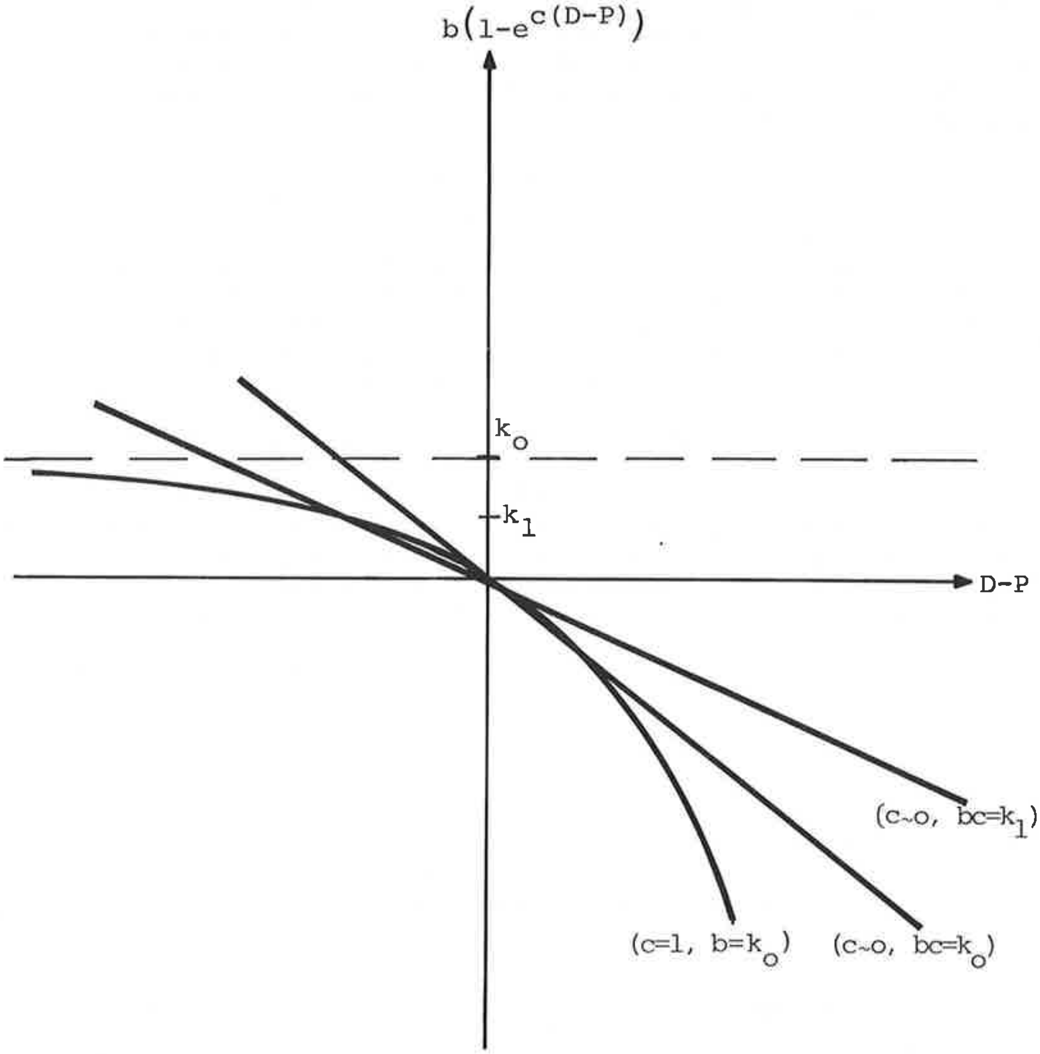


FIGURE A.2 GRAPH OF POSSIBLE DISUTILITY FUNCTIONS FOR DIFFERENT VALUES OF b AND c

Denoting P* as the optimal value of P,

$$P^* = -\frac{1}{c} \ln \left(\frac{wM^{\alpha}}{bc} \right) + \frac{cV^2}{2}$$

If the variance in actual delay approached zero, it would be expected that the traveler would schedule neither a positive nor a negative premature depature time. That is, as $V^2 \longrightarrow 0$, $P^* \longrightarrow 0$. This will occur if $b = \frac{wM^{\alpha}}{c}$.

This allows eliminating b from the set of parameters. The equation for P* can now be rewritten as:

$$P^* = \frac{cV^2}{2} \qquad c > 0$$

This equation indicates that the passenger will always schedule a positive premature departure time in the only case considered here, namely, that of positive "risk aversion" to delays.

APPENDIX B

DERIVATION OF TRAVELERS' MEAN WAIT TIME

Let $g_r(h)$ = headway distribution on route r

$\bar{w}(h)$ = average wait time for headway h

$n(h)$ = number of persons arriving in headway interval h

for interval h:

$$\bar{w}(h) = h/2 \qquad n(h) = kh$$

$$\text{Total \# of users} = \int_0^\infty n(h) \cdot g_r(h) \cdot dh$$

$$\text{Average wait time} = \bar{w} = \frac{\int_0^\infty n(h) \cdot \bar{w}(h) \cdot g_r(h) \cdot dh}{\int_0^\infty n(h) \cdot g_r(h) \cdot dh}$$

$$= \frac{k/2 \int_0^\infty h^2 \cdot g_r(h) \cdot dh}{k \int_0^\infty h \cdot g_r(h) \cdot dh} = \frac{E(h^2)}{2E(h)}$$

Since $Var(h) = E(h^2) - h^2$

$$E(w) = \frac{Var(h) + \bar{h}^2}{2\bar{h}} = \frac{\bar{h}}{2}\left(1 + \frac{Var(h)}{\bar{h}^2}\right. .$$

Source: (Kulash, 2)

APPENDIX C

BARNETTS'S HOLDING POINT MODEL

A procedure has been developed to construct an approximate-optimal dispatching strategy for the control point (Barnett, 3). A description, application, and extensions of this procedure are discussed below.

The interarrival interval (or headway) between two consecutive vehicles at a point B, with respective lateness ℓ_1 and ℓ_2 and scheduled headway a, is:

$$h = a + \ell_2 - \ell_1$$

While ℓ_1 and ℓ_2 both have the same unconditional density function, it is unlikely that they are independent; presumably from the available data one can obtain the conditional distribution of ℓ_2 given by ℓ_1.

If passenger arrivals are assumed Poisson distributed, the average waiting time is:

$$\bar{w} = \frac{E(h^2)}{2E(h)} = \frac{a^2 + E(\ell_1^2 + \ell_2^2 - 2\ell_1\ell_2)}{2a}$$

The only form of control being considered is the delay of individual vehicles at the control point B. One can set the goal of a holding strategy to minimize an objective function, J, of the following form:

$$J = \alpha\ E_b(d) + (1 - \alpha)E_b(w)$$

where: $E_b(d)$ is the expected delay at B to passengers who boarded earlier and for whom B is an intermediate stop

$E_b(w)$ is the expected wait time for passengers who board at B

α is a weighting constant $(0 \leq \alpha \leq 1)$

Since wait time varies linearly with headway variance, and since headway variation will build up proportionally along the route, reducing headway variation is the objective at point B most beneficial to passengers further down the route. Their expected wait can be shown to be proportional to that time at point B, and therefore, the expected delay and wait times at point B appropriately weighted are sufficient input to the objective function.

To simplify the problem, a two point model is developed for the lateness distribution, using parameters c, d, and p. c and d are the points and p the probability weight assigned to c $(c \le d)$.

Selection of c, d, and p are constrained by the following relationships:

$$u = cp + d(1-p)$$

$$v = c^2 p + d^2(1-p)$$

$$\begin{aligned} w &= c^2(pp_{cc}) + 2cd(pp_{cd}) + d^2(1-p)p_{dd} \\ &= c^2 p + d^2(1-p) - pp_{cd}(d-c)^2, \end{aligned}$$

where: p = probability of delay c
p_{cc} = probability of delay c, given delay c on previous vehicle
u = mean lateness
v = second moment of lateness
w = expected product of two consecutive latenesses in the actual vehicle arrival pattern at point B.

We define $L = d-c$. The model is formulated as follows:

Possible Arrival Delay Sequences	Headway	Probability
c,c	a	p_{cc}
c,d	$a+L$	p_{cd}
d,c	$a-L$	$(1-p)\cdot p_{dc} = pp_{cd}$
d,d	a	$(1-p)\cdot p_{dd}$

Average wait at B satisfies

$$E_b(w) = \frac{E(h^2)}{2E(h)} = \frac{a^2 + E(\ell_1^2 + \ell_2^2 - 2\ell_1\ell_2)}{2a}$$

$$= \frac{a^2 + (2v-2w)}{2a} = \frac{a^2 + 2pp_{cd}(d-c)^2}{2a}$$

$$= \frac{a^2 + 2L^2 pp_{cd}}{2a}$$

in the model.

Thus variations around mean headway 'a' have increased expected wait time by $(L^2 pp_{cd})/a$ over its ideal value of $a/2$.

A simple strategy to reduce average wait time at B would be to regularize intervals by holding the second vehicle in a (d,c) sequence by some amount x (c_1 refers to a vehicle of lateness c immediately after one of lateness d).

Arrival Sequence	Associated Departure Headway	Prob. of Occurrence
(d,d)	a	$(1-p)\cdot p_{dd}$
(d,c_1)	$a-L+x$	$(1-p)\cdot p_{dc}$
(c_1,c)	$a-x$	$(1-p)\cdot p_{dc}\cdot p_{cc}$
(c,c)	a	pp_{cc}^2
(c_1,d)	$a+L-x$	$(1-p)\cdot p_{dc}\cdot p_{cd}$
(c,d)	$a+1$	$pp_{cc}\cdot p_{cd}$

The resulting value of the objective function is:

$$J(x) = \alpha x(1-p)p_{dc} + (1-\alpha)\,\frac{E(h^2)}{2E(h)}$$

J(x) is a minimum when:

$$x = \max\left\{0,\ \left(\frac{L(1+p_{cd})}{2} - \frac{\alpha}{1-\alpha}\left(\frac{a}{2}\right)\right)\right\}$$

(Note: for $L < a$, $0<x<a$)

The simple strategy described above leads to an overall reduction in average wait time. The overall average headway is unchanged and thus the system does not require additional vehicles or layover time. Some vehicles, however, are delayed and at some points in the sequence headway interval variation is increased.

This procedure was applied to a model of the northbound Red Line in Boston at Washington Street, the busiest stop on the line. Actual route operating data was used as input to the model. The scheduled headway, a, was 5 minutes for the analysis period. Lateness compared to scheduled time had a mean, u, of 0.7 min. and a second moment, v, of 1.7 min.^2. The expected product of two consecutive latenesses, w, was 0.3 square minutes.

An approximate two-point model was created. From the data, a graph of p_{cd} as a function of p was obtained and utilizing the values and equations for u, v, and w, the values of c, d, and p were obtained as well as p_{cd}, p_{cc}, p_{dc}, and p_{dd}:

$c = -0.4$
$d = 1.9$
$p = 0.48$
$p_{cd} = 0.58$
$p_{cc} = 0.42$
$p_{dc} = 0.54$
$p_{dd} = 0.46$

Average wait at Washington Street was:

$$- a/2 + (L^2 pp_{cd})/a = 2.8 \text{ min.}$$

Choosing $\alpha = 0.1$ to reflect the small number of passengers continuing through Washington Street, a one stage holding strategy of 1.54 min. can be determined. This results in a reduction of average wait to 2.62 min. and an average delay of 25 seconds for passengers passing through Washington Street.

After a few iterations of the model, multi-stage strategies yield a 10% reduction in average wait, to 2.53 min., very close to the ideal wait time value, 2.5 min. This is a drop of 90% in excess wait time.

Average holding time at Washington Street is under one minute, only a small fraction of typical buffer periods at terminals. An important side effect is the reduction of the average number of standees leaving Washington Street northbound by about 70%.

The simple control strategy that has been described above can be extended to include multi-stage holding strategies in which more than one vehicle is held to reduce the adverse effects. A more complex procedure has been developed to construct such multi-stage strategies and is easily programmable. The procedure for finding an approximate best dispatch strategy from the central stop is as follows:

1. The arrival delay distribution for vehicles is estimated by a two-point distribution that preserves those characteristics of the arrival pattern most relevant to the objective function.

2. An algorithm is used to obtain the best holding policy under the two-point model of vehicle arrival.

3. An elastic holding scheme is devised to implement as closely as possible in the actual situation the policy found best in the simplified model.

The procedure described above could be extended to include:

1. a more precise approximation of the vehicle lateness distribution (more than two points)

2. the use of more detailed information about future arrivals at the control stop, and

3. more control stops.

APPENDIX D

DERIVATION OF THE HEADWAY AT A PARTICULAR BUS STOP ON A ROUTE

Propagation of an initial deviation with number of stops along a route may be expressed as follows (Cohn, 4):

$$e_i = e_{i-1} + e_{i-1}(P \bullet b)$$

$$e_n = e_o(1 + P \bullet b)$$

where:
e_n = error in arrival at stop n
e_o = initial departure error
P = arrival rate of passengers at stops
b = boarding time per passenger.

This model can be extended to include:

1. probabilistic nature of passenger arrivals (let the number of passengers waiting to board be a function of the headway and a random variable P, the instantaneous rate of arrival);

2. probabilistic nature of travel times on all segments of the route.

Thus the expression for the error may be expanded to:

$$e_i = (e_{i-1} + t_i) + (e_{i-1} + t_i)(P \bullet b)$$

where:

t_i = a random variable which is a function of distance from i-1 to i. This represents a deviation from an expected travel time from i-1 to i.

P = a random variable representing passenger arrival rate at a stop.

Furthermore, an expression for headway may be developed:

Let: $H_i^{k-1,k}$ = headway at stop i between buses k and k-1
s = scheduled headway

Then: $H_i^{k-1,k} = s + e_i^k - e_i^{k-1}$

$$e_i^k = (e_{i-1}^k + t_i^k - e_i^{k-1})(1 + P \bullet b)$$

$$H_i^{k-1,k} = s + (1 + P \bullet b)(e_{i-1}^k + t_i^k - e_i^{k-1} - e_{i-1}^{k-1} - t_i^{k-1} + e_i^{k-2}).$$

APPENDIX E - FURTHER ANALYSIS OF JOLIFFE AND HUTCHINSON RESEARCH

A major finding of Joliffe and Hutchinson was that where bus arrivals are predictable, travelers may be able to minimize their wait time at the bus stop by arriving there at a particular time instead of randomly. They observed in several cases the coincidence of high passenger arrivals with resulting waiting times lower than the expected waiting time, w_{rand}, had the passengers arrived randomly. They calculated that, for each bus stop, there was a particular arrival time which would minimize the expected waiting time for the bus. This waiting time was called w_{min}.

On all of the bus routes observed, it was possible for a traveler to reduce his/her mean wait time by some amount. $g = w_{rand} - w_{min}$ was defined as the potential gain in reduced waiting time that could be obtained by a passenger with knowledge of the bus service who arrives at the optimal time.* Joliffe and Hutchinson observed that the actual waiting time, W, fell in between w_{rand} and w_{min}. They hypothesized that only a certain percentage of travelers were taking advantage of the time savings and the rest were arriving randomly. They developed the following scenario: a small percentage of the passengers, q, ran to catch the bus when it was in view, and had no waiting time at all; "p" percent of the remaining passengers, or $(1-q)p$ percent of all of the passengers, arrived at the optimal time and waited w_{min} for the bus; the remaining $(1-q)(1-p)$ percent arrived at random and waited w_{rand}. p was plotted against q for both peak and off-peak routes (see Figure E.1). Given the observed points and given that p (1) should always be in the range 0 to 1, (2) should be zero when g is zero, and (3) should be continuously increasing, a suitable function was found to be $p = 1 - e^{-\lambda g}$. The curves in Figure E.1 have this form, where λ is 0.131 for the peak and 0.015 for the off-peak.

It seems logical that when g is large, p is large, for two reasons: (1) the fact that a potential gain is available will be more easily perceptible, and (2) the incentive for arriving at the optimal time rather than randomly is larger. It is also expected that p is larger in the peak than in the

---

*Optimal time refers to the particular recurrent point in time at which a passenger would arrive at the bus stop so as to minimize his wait time.

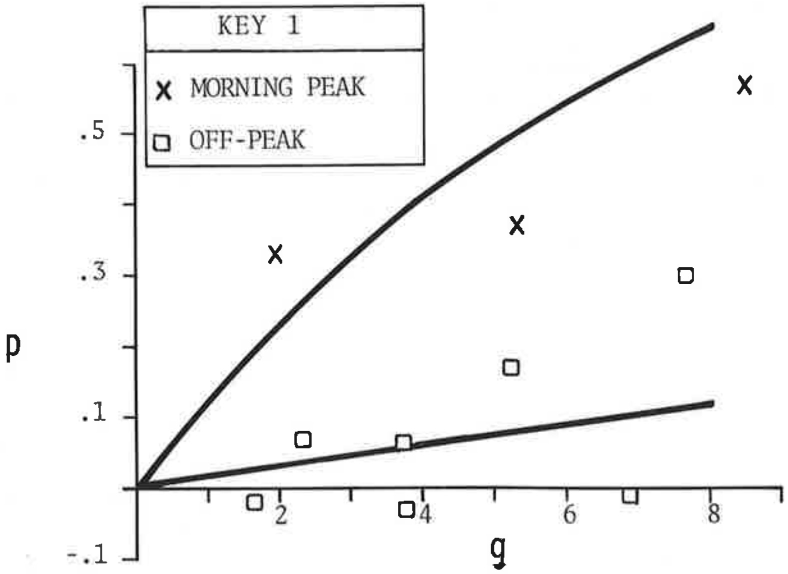


FIGURE E.1 PERCENT OF PASSENGERS ARRIVING AT THE OPTIMAL TIME AS A FUNCTION OF THE POTENTIAL GAIN IN REDUCED WAITING TIME off-peak, because of the greater proportion of regular travelers in the peak hours.

Joliffe and Hutchinson also found that w_{min} is positively correlated with Ω, the standard deviation of bus departure times <u>excluding</u> the effect of cancellations (see Figure E.2). If Ω were zero, then w_{min} should be zero also, since passengers could arrive at the same time as the bus. But this is only true if there are no cancellations. If some buses were cancelled, w_{min} would be positive even if Ω were zero, implying a positive intercept in Figure E.2. An exceptionally high proportion (33 percent) of buses at stop 9 were cancelled, giving rise to the outlying point at (3.1, 9.7). The authors of this report feel that this analysis would be strengthened if Ω', the standard deviation of bus departure times <u>including</u> cancellations, were used instead of Ω. With $\Omega' \geq \Omega$, all points would be shifted to the right, with the outlying point shifted the most (as indicated by the arrow in Figure E.2), and w_{min} would tend towards zero when Ω' was zero.

Joliffe and Hutchinson assumed that many travelers set their arrival time so as to minimize their expected wait time. However, at that time, the variability of wait time may not be minimized. Joliffe and Hutchinson did not explore this possibility. It might be that travelers will choose the time which minimizes their combined disutility for wait time and wait time variability.

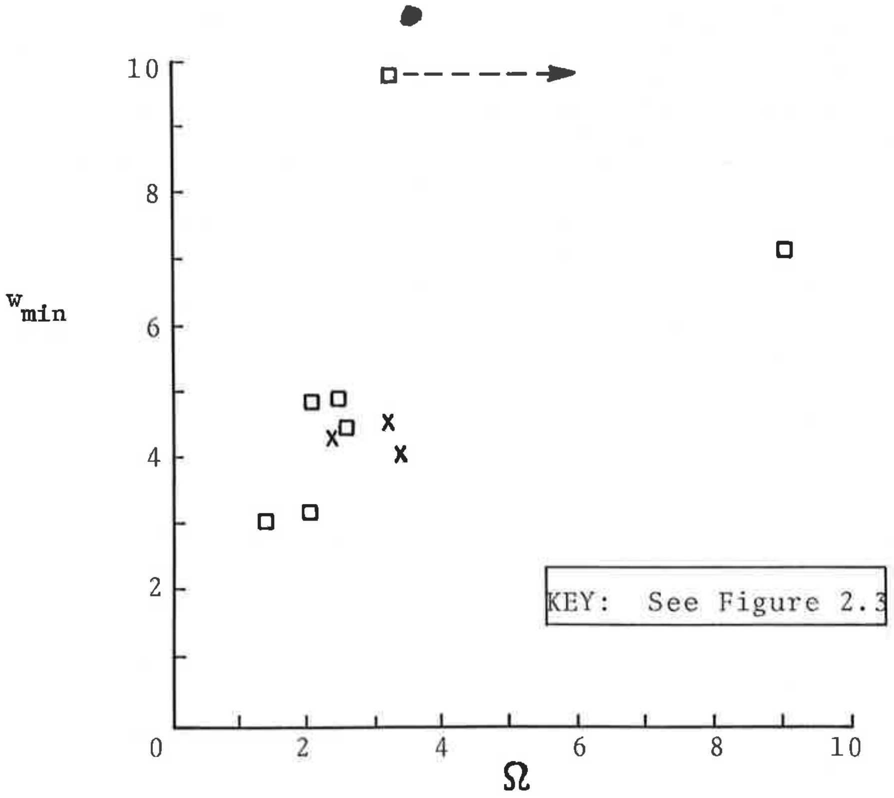


FIGURE E.2 RELATION BETWEEN STANDARD DEVIATION OF BUS DEPARTURE TIMES AND MINIMUM WAITING TIME

REFERENCES

1. Marfisi, et al. Economic Criteria for Federal Aviation Agency Facility and Equipment Expenditures. WP-210-U4-108, U.S. DOT, July 1975.*

2. Kulash, D. "Routing and Scheduling in Public Transit Systems." (MIT PhD. Thesis, 1971).

3. Barnett, Arnold. "On Controlling Randomness in Transit Operations." Transportation Science. Vol. 8, No. 2, May 1974.

4. Cohn, J., et al. "Qualitative Analysis of Bus Location and Control." Motorola, Inc., paper presented at ORSA Spring 1976 meeting, Philadelphia.

---

*Material on file: U.S.Department of Transportation, Transportation Systems Center, Cambridge,MA.

BIBLIOGRAPHY

Barnett, Arnold. "On Controlling Randomness in Transit Operations." Transportation Science. Vol. 8, No. 2, May 1974.

Bather-Ringross-Wolsfeld, Inc., Metropolitan Transit Commission, and Minnesota Highway Dept. I-35W Urban Corridor Demonstration Project, Final Report, August 1975.

Ben-Akiva, Moshe. "Review of the Structures of Current Approaches to Travel Demand Modelling." (PhD. Thesis, MIT, Ch. 5, June 1973).

Ben-Akiva, Moshe and S. Lerman. Forecasting Models in Transportation Planning. Paper prepared for presentation at the Conference on Population Forecasting for Small Areas, May 1975.

Bigelow-Crain Associates. The Golden Gate Corridor Bus Priority System. Transportation Systems Center, May 1975.

Bigelow-Crain Associates. Second Year Report: San Bernardino Freeway Express Busway Evaluation. Prepared for Southern California Association of Governments, September 1975.

Bly, P.H. and R.L. Jackson. Evaluation of Bus Control Strategies by Simulation. TRRL Laboratory Report 637, Transport and Road Research Laboratory, Dept. of Environment, Crowthorne (Great Britain), 1974.

Bly, P.H. and R.L. Jackson. "Scheduling for Depleted Services." Univ. of Newcastle Upon Tyne (Great Britain) 1975.

Chapman, R.A. "Bus Reliability-Definitions and Measurement." (Working Paper No. 18). Transport Operations and Research Group, Univ. of Newcastle Upon Tyne (Great Britain), 1976.

Chung, Cady C. "Dial-a-Ride Vehicle Time Analysis and Vehicle Productivity." MITRE, VA, 1976.

Cohn, J., et al. "Qualitative Analysis of Bus Location and Control." Motorola, Inc., paper presented at ORSA Spring 1976 Meeting, Philadelphia.

Constantino, Don P., T.F. Golob and P.R. Stopher. "Consumer Preferences for Automated Public Transportation Systems." Transportation Research Record 526, (1974), pp. 81-93.

Cosslett, Stephen. "The Trip Timing Decision for Travel to Work by Automobile." Demand Model Estimation and Validation. The Urban Travel Demand Forecasting Project Phase I Final Report, Vol. V. Institute for Transportation Studies, Univ. of California, June 1977.

Elias, Wilbur G. "The Greenback Experiment-Signal Preemption for Express Buses." Sacramento County, CA, April 1976.

Friedman, R.A. "The Effects of Transit Reliability on Passenger Waiting Times." (M.S. Thesis, The Transportation Center, Northwestern Univ. December 1975).

Gardner, Wendall and F. Reid. Transit Reliability Data for Disaggregated Traveler Demand Models. Working Paper No. 7603, Univ. of California, 1977.

Giraud, Alain and J.J. Henry. On-line Control of A Bus Route-Studies and Experiments in Toulouse. Third International Symposium on Control in Transportation Systems, Columbus, OH, 1976.

Golob, Thomas F., et al. An Analysis of Consumer Preferences for a Public Transportation System. (GMR-1037), General Motors Research Laboratory, Warren, MI, 1970.

Golob, Thomas F. and R. Dobson. "Assessment of Preferences and Perceptions Toward Attributes of Transportation Alternatives." Transportation Research Board Special Report 149. General Motors Research Laboratory, Warren, MI, 1974, pp. 58-81.

Gustafson, Richard L., H.N. Curd and T.F. Golob, "User Preferences for a Demand-Responsive Transportation System: A Case Study Report." Highway Research Record 367. 1971, pp. 31-45.

Gustafson, Richard L. and F.P.D. Navin. "User Preferences for Dial-a-Bus." Highway Research Board Special Report 136: Demand-Responsive Transportation Systems. 1973, pp. 85-93.

"Haddonfield Dial-a-Ride Project." UMTA-NJ-06-0002-74-7, NJ DOT, US DOT, 1974.

Holroyd, E.M. and D.A. Scraggs. "Waiting Times for Buses in Central London." Traffic Engineering and Control. 8(3), (July 1966), pp. 158-160.

Jackson, R.L. and D. Stone. Experiments in Bus Service Control Using an Interactive Model of a Typical Urban Bus Route. Univ. of Newcastle Upon Tyne (Great Britain), 1976.

Johnson, M. "A Comparison on Several Methods of Collecting Travel Time Data in Analysis of Urban Travel Behavior." Technical Memorandum to the Metropolitan Transportation Commission, Berkeley, CA, February 1975.

Jolliffe, J.K. and T.P. Hutchinson. "A Behavioral Explanation of the Association between Bus and Passenger Arrivals at a Bus Stop." Transportation Science. 94. November 1975, pp. 248-282.

Keck, Carol A. and G.S. Cohen. "Characteristics and Attitudes of Dial-a-Bus and Park-and-Ride Users in New York State." Transportation Research Record 590. (1976), pp. 21-25.

Knight, Trevor E. "An Approach to the Evaluation of Changes in Travel Unreliability: A 'Safety-Margin' Hypothesis." Transportation. Vol. 3, No. 4, December 1974, pp. 393-407.

Koffman, David. "A Simulation Study of Alternative Real Time Bus Headway Control Strategies." Presented at the 57th Annual Meeting of the Transportation Research Board, Washington, DC, January 1978.

Kulash, D. "Routing and Scheduling in Public Transit Systems." (MIT PhD. Thesis, 1971).

Lerman, Steven, et al. A Method for Estimating the Patronage on Demand-Responsive Transit Systems. Cambridge Systematics, DOT-TJT-77, 1977.

Manheim, Marvin L. Fundamentals of Transportation Systems Analysis. MIT, Preliminary Edition, Third Revision, August 1975.

Manski, Charles F. The Stochastic Utility Model of Choice. October 1973.

Marfisi, et al. *Economic Criteria for Federal Aviation Agency Facility and Equipment Expenditures*. WP-210-U4-108, U.S. DOT, July 1975.*

MARTA. "The Effect of Fare Reduction on Transit Ridership in the Atlanta Region: Analysis of Non-Transit User Data." Technical Report #3, December 1974.

Martland, Carl D. *Rail Trip Time Reliability: Evaluation of Performance Measures and Analysis of Trip Time Data*. Vol. 2, MIT R72-37, 1972.

Menhard, Henry R. "User Preferences Pertaining to Dial-A-Ride Service Options." (MS Thesis, MIT, 1976).

McFadden, D. and F. Reid. "Aggregate Travel Demand Forecasting From Disaggregated Behavioral Models." *Transportation Research Record 534*. pp. 24-237, 1975.

McQueen, James, et al. *Evaluation of the Shirley Highway Express-Bus-On-Freeway Demonstration Project-Final Report*. US DOT, UMTA, 7, August 1975.

Muzyka, A., D. Schmieder, J. Fantasia, and G. Feeney. *Traffic Signal Strategies for Bus Priority-A Simulation Study*. Transportation Systems Center, Urban and Regional Research Series, (Forthcoming).

National Cooperative Highway Research Program Report 49. "National Survey of Transportation Attitudes and Behavior, Phase I, Summary Report." 1968.

National Cooperative Highway Research Program Report 82. "National Survey of Transportation Attitudes and Behavior, Phase II, Analysis Report." 1970.

Newell, G.F. and R.B. Potts. *Maintaining a Bus Schedule*. Proceedings of the Second Conference, Australian Road Research Board, II, (1964), pp. 388-393.

Nicolaidis, Gregory C. and R. Dobson. "Disaggregated Perceptions and Preferences in Transportation Planning." Transportation Research, Vol. 9, No. 5, October 1975, pp. 279-295.

Paine, F.T., A.N. Nash, S.T. Hille and G.A. Brunner. "Consumer Attitudes Toward Auto vs. Public Transport Alternatives." *Journal of Applied Psychology*. Vol. 53, No. 6 (1976), pp. 472-480.

*Material on file: U.S.Department of Transportation, Transportation Systems Center, Cambridge MA.

Paine, F.T., A.N. Nash, S.J. Hille and G.A. Brunner. Consumer Conceived Attributes of Transportation. College Park, MD, Dept. of Business Administration, Univ. of Maryland, 1967.

Prashker, Joseph N. "Development of a Reliability of Travel Modes Variable for Mode-Choice Models." (PhD Dissertation, Northwestern Univ., Dept. of Civil Engineering, Evanston, IL, June 1977).

Seddon, P.A. and M.P. Day. "Bus Passenger Waiting Times in Greater Manchester." Traffic Engineering and Control. 15(9), (1974), pp. 442-445.

Shields, L.R. Reliability-Study of a Central London Bus Route. London Transport, Great Britain, 1974.

Spear, Bruce D. "Generalized Attribute Variable for Models of Mode Choice Behavior." Transportation Research Record 592. (1976), pp. 6-11.

Starkie, D.N.M. "Modal Split and the Value of Time: A Note on Idle Time." Journal of Transport Economics and Policy. Vol. 5, No. 2, (1971).

Sterman, Brian P. "Factors Affecting Reliability of Urban Bus Services." (M.S. Thesis, Northwestern University, IL, 1974).

Stopher, Peter R., B.D. Spear and P.O. Sucher. "Toward the Development of Measures of Convenience for Travel Modes." Transportation Research Record 527. (1974), pp. 16-32.

SYSTAN, Inc. The Rochester, New York Integrated Transit Demonstration: Suburb of Greece. Draft Final Report, UMTA Service and Methods Demonstration Program, UMTA-NY-06-0048-78,Vol.1,2,3 (Forthcoming).

SYSTAN, Inc. The Santa Monica Freeway Diamond Lanes: An Evaluation. Vol. II. UMTA Service and Methods Demonstration Program, UMTA-MA-06-0049-77-13, December 1977.

Turnquist, Mark A. and L.A. Bowman. "The Effects of Network Structure on Reliability of Transit Service." Northwestern Univ., presented at International Symposium on Travel Supply Models. Montreal 1977.

Wachs, Martin. "Consumer Attitudes Toward Transit Service: An Interpretive Review." AIP Journal. (January 1976), pp. 86-104.

Waksman, Robert and D. Schmieder. "Bus Routes Simulation Models for Studying Service Improvement Strategies." Transportation Systems Center, SS-24-U.3-137, November 1977.*

Wallin, Rex J. and P.H. Wright. "Factors Which Influence Modal Choice." Traffic Quarterly. Vol. 28, No. 2, April 1974, pp. 271-290.

Wattleworth, Joseph A., et al. I-95/NW 7th Avenue Bus/Carpool Systems Demonstration Project, Report I, US DOT, UMTA-FL-06-0006, Transportation Research Center, Univ. of Florida, September 1977.

Welding, P.I. "The Instability of a Close Interval Service." Operations Research Quarterly (1957), 8(3), pp. 133-142.

Wilson, Nigel, H.M. and N.J. Colvin. "Computer Control of the Rochester Dial-a-Ride System." MIT Center for Transportation Studies Report 77-22, July 1977.

Wilson, Nigel, R. Weissberg and J. Hauser. "Advanced Dial-a-Ride Algorithms Research Project: Final Report." UMTA-MA-11-0024, MIT Report R76-20, March 1976.

Winchester, S.A. "Haddonfield Dial-a-Ride Demonstration, Third Household Survey." MITRE Corp., McLean, VA, 1976.

---

*Material on file: U.S.Department of Transportation, Transportation Systems Center, Cambridge,MA.

660 Copies

☆U. S. GOVERNMENT PRINTING OFFICE: 1979 — A-1093/299