U.S. Department
of Transportation

**Research and
Special Programs
Administration**

Strategies for Improving Bus Transit Service Reliability

Prepared for:
Office of University Research

SPA/DPB-50/81/27

Phase II: Final Report
Under Contract DOT-OS-80018
February 1982

This document is available to the
U.S. public through the National
Technical Information Service,
Springfield, Virginia 22161.

Technical Report Documentation Page

1. Report No. DOT/RSPA/DPB-50/81/27	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Strategies for Improving Bus Transit Service Reliability	5. Report Date February 1982
	6. Performing Organization Code
7. Author(s) Mark A. Turnquist	8. Performing Organization Report No.
9. Performing Organization Name and Address Northwestern University Evanston, Illinois 60201* *part of research was subcontracted to Cornell Univ.	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DOT-OS-80018
12. Sponsoring Agency Name and Address U. S. Department of Transportation Research and Special Programs Admin. Office of University Research Washington, D.C. 20590	13. Type of Report and Period Covered Final Report March 1979 - June 1980
	14. Sponsoring Agency Code DPB-50

15. Supplementary Notes

Technical Monitor: Mark D. Abkowitz
Trans. Systems Center

16. Abstract

This is the final report on a study of service reliability in bus transit networks. The objectives of the study were: 1) investigation of sources of reliability problems, 2) identification of potential service improvement strategies, 3) development of models to allow these strategies to be analyzed, and 4) evaluation of the the relative effectiveness of various strategies. This report is a sequel to report DOT/RSPA/DPB-50/79/5, which described the first phase of the project, concentrating mostly on sources and impacts of reliability problems.

Four major classes of strategies for improving bus transit service reliability are analyzed, including: 1) vehicle-holding strategies, 2) reducing the number of stops made by each bus, 3) signalization improvements, and 4) provision of exclusive right-of-way. The principle findings are: 1) strategies to improve service reliability can have very substantial impacts on overall service and quality, including improvements in average wait and in-vehicle time as well; and 2) the best strategy to use in the a particular situation depends upon several factors, but service frequency is the most important. For low frequency services (less than 10 buses per hour), schedule-based holding strategies or zone scheduling are likely to work best. For mid-frequency services (10-30 buses per hour) zone scheduling or signal preemption are likely to be most effective, although headway-based holding can also work well if an appropriate control point can be found. In high frequency situations (more than 30 buses per hour) an exclusive lane combined with signal preemption should be considered.

17. Key Words Computer simulation, service reliability, urban bus network, control strategy, passenger arrivals, vehicle travel times, transfers.	18. Distribution Statement Document is available to the U.S. Public through the National Technical Information Service, Springfield Virginia 22161

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

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

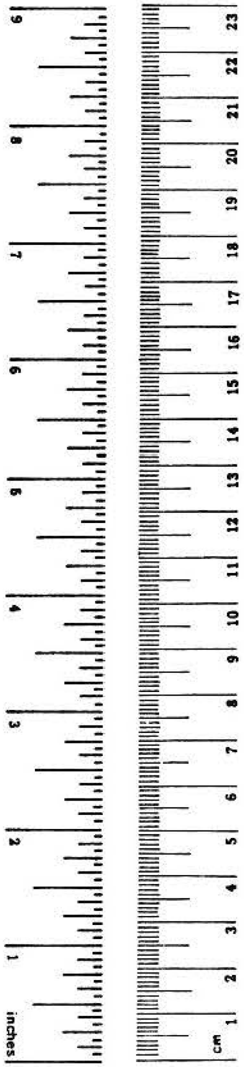
METRIC CONVERSION FACTORS

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

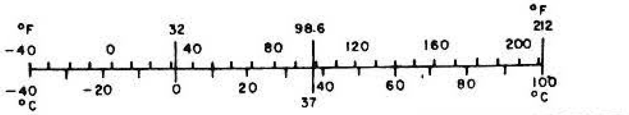
---

*1 in = 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



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	
l	liters	2.1	pints	
l	liters	1.06	quarts	
l	liters	0.26	gallons	
m^3	cubic meters	35	cubic feet	
m^3	cubic meters	1.3	cubic yards	
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



STRATEGIES FOR IMPROVING BUS TRANSIT SERVICE RELIABILITY

EXECUTIVE SUMMARY

The concept of "service reliability" has come into increasing prominence in recent years as an important characteristic of the quality of service provided by transportation systems. A basic definition of reliability, as the term is used here, is the variability of a system performance measure over time. The focus is on stochastic variation in performance, rather than on more traditional engineering concepts of probability of component or system failure. The level of service measure most clearly subject to variation is travel time, and this variability is often described in terms of non-adherence to schedule.

Service reliability is important to both the transit user and the transit operator. To the user, non-adherence to schedule results in increased wait time, makes transferring more difficult, and causes uncertain arrival time at the destination. The importance of some measure of reliability to trip-making behavior has been emphasized in several attitudinal studies.

In addition to its importance to transit users, unreliability in operations is a source of reduced productivity and increased costs for transit operators. This is due to the need to build substantial "slack time" into timetables in order to absorb deviations from the schedule. This leads to reduced utilization of both equipment and personnel.

In light of the current need for more cost-effective public transportation in urban areas, it is important to understand the sources of unreliability, and to investigate the potential of several alternative control strategies to improve both the quality of service provided and the productivity of the vehicles and personnel in the system.

This project has had four major objectives:

1) investigation of the sources of service reliability problems in bus transit networks;

2) identification of potential strategies for reducing the severity of the problem;

3) development of models to allow these various strategies to be analyzed and evaluated; and

4) general evaluation of the relative effectiveness of these strategies for improving service reliability.

In order to investigate the sources of service reliability problems, a substantial set of simulation experiments was performed. These experiments allowed for controlled variation of a number of important characteristics of bus networks, including network form (grid or radial), route density, frequency of service, degree of link travel time variability and the demand/capacity ratio. The results of these experiments can best be summarized in terms of vehicle bunching, variability of transfer times, and the relative impacts of service frequency and route density.

The experiments indicated how vehicle bunching is related to frequency of service, level of demand and the variability of link travel times. In particular, these results illustrate the importance of reducing link travel time variability in an effort to prevent bunches from forming. This represents an extension to the results of previous researchers, which placed primary emphasis on the demand/capacity ratio and boarding times.

The importance of transferring to overall trip reliability focuses attention on the trade-off between the length of the scheduled wait time at transfer points and the risk of missing the intended connecting bus. Where arrivals can be scheduled to coincide, as in radially structured networks, the application of controls to the operation of transit service has the potential to permit more closely scheduled arrivals on connecting services while maintaining a reasonable assurance that the intended connection will be successful.

Finally, it is clear from the experimental results that service reliability is much more sensitive to frequency of service than to route density. This implies that there are substantial reliability impacts of the trade-off between operating fewer routes at higher frequency or more routes at lower frequency, given a limited amount of vehicle resources. Traditionally, this trade-off has been evaluated using simplistic models of expected passenger wait time and the accessibility of transit service to users. However, the present work has shown that service reliability is also an important factor in this trade-off and should be included in the evaluation.

With these ideas in mind, a variety of potential strategies for improving service reliability have been considered. These strategies fall into four general classes: 1) vehicle holding; 2) stop reduction and zone scheduling; 3) changes in traffic signal operation; and 4) provision of exclusive right-of-way.

Within the category of holding strategies, two major subgroups of strategies are considered: schedule-based holding and headway-based holding. The schedule-based "checkpoint" strategy is very simple to implement and offers promise of significant benefits on long-headway routes where the schedule is sufficiently slack so as to make holding to schedule a reasonable procedure. The key elements of implementing such a policy are constructing a reasonable schedule as a goal and enforcing adherence to that schedule. This enforcement requires both proper incentives for drivers and a mechanism for accurate monitoring of their performance.

For routes operating with shorter headways, two near-optimal headway-based control strategies have been developed. One strategy holds a vehicle until its <u>pre</u>ceding headway is as close as possible to its <u>fol</u>lowing headway, allowing for an adjustment in consideration of the people delayed on the vehicle. Referred to as "Prefol", it requires a prediction of the arrival time of the following vehicle. A similar strategy that is dependent only on the known magnitude of the current headway, called the "Single Headway" strategy, is also proposed. The strategies are simple in form, require limited data about the route, and are near-optimal over a wide range of situations.

Models of the effectiveness of the strategies indicate that they are sensitive to three important characteristics of a control point: (1) the current level of unreliability, as measured by the headway coefficient of variation; (2) the relationship between successive headways, measured by the correlation coefficient; and (3) the proportion of passengers who must ride through the control point.

The Single Headway strategy performs less well than the Prefol strategy when vehicles arrive relatively independent of each other. As passenger loading delays increase and successive headways become more dependent on each other, the Single Headway strategy prediction capability improves and it approaches the performance of the Prefol strategy. In any event, however, the Single Headway strategy requires less information about the system, and could certainly be implemented without expensive AVM equipment.

Reducing the number of stops made by each individual vehicle is a second class of methods to improve service reliability. Two major ways of accomplishing this reduction have been examined.

Eliminating stops so as to increase average stop spacing can be useful if stop density is very high before reduction, and if traffic signal operation can also be changed to allow buses to take advantage of the potential for increased speed. It should be noted, however, that increasing stop spacing also has the effect of reducing the accessibility of the bus route to those who use it. Thus, there is a tradeoff of improvement in one (or more) dimension(s) of service quality and a degradation in another. The full implications of this can only be ascertained by including a demand analysis with the results of this work, in order to determine how travelers would react to this tradeoff. Such analysis remains for further study.

An alternative way of reducing the number of stops made by each vehicle, without increasing overall stop spacing, is by zone scheduling. A model has been developed to design optimal zone structures and allocate buses to zones, using dynamic programming. Application of this model to a route in Chicago illustrates the potential effectiveness of zone scheduling as a service improvement strategy. Substantial improvements in reliability, as well as in other measures, appear possible. Simulation of a particular zone strategy for two routes in Cincinnati has indicated the attractiveness of such a method in an alternative context as well, using a more detailed model.

The third class of potential strategies involves changes in traffic signal operation. Such changes can take the form of changes in signal timing or of allowing buses to preempt signals. The results of simulation experiments indicate the important interactions between stop spacing and signal timing. Both characteristics must be considered jointly in order to make changes in either effective. The limited experimentation done in this project indicates that this may be difficult to do, but that if the stop spacing and signal timing are appropriately "matched," the effects may be similar to those from signal preemption.

The signal preemption strategies tested resulted in 17-18% increases in both mean and standard deviation of vehicle speed, and 8-9% reductions in mean and standard deviation of waiting time. These are significant improvements, and indicate that signal preemption can be an effective means to improve reliability of service, as well as average travel time. These benefits to buses (and riders) able to preempt traffic signals must be balanced against additional delays which may be inflicted upon cross-traffic. The tests performed in this project have been based on signal settings designed to minimize total expected delay for both main and cross-direction traffic, and thus should provide a reasonable basis on which to judge the potential effectiveness of signal preemption, but more detailed assessment of the effects on cross-traffic would require a more detailed traffic simulation model.

The fourth class of strategies, provision of exclusive right-of-way, may be considered for situations in which there is a very high level of transit activity, say 30 or more buses per hour along a particular street. Empirical findings in several demonstrations in the U.S., Europe and Australia have indicated the effectiveness of reserved lanes in particular circumstances. Additional simulation experiments performed in this project have also indicated that the combination of a reserved lane for buses and signal preemption can be an effective means of improving both average bus speed and service reliability.

The selection of an appropriate strategy for a given situation appears to depend heavily on the frequency of service on the route(s) involved. For low frequency situations (less than 10 buses per hour) a checkpoint (schedule-based holding) strategy, or some form of zone scheduling will probably be most effective. In medium frequency cases (10 to 30 buses per hour), zone scheduling and signal preemption appear to be most effective, and in high frequency environments (more than 30 buses per hour) a reserved lane with signal preemption is likely to be effective. Headway-based holding can also be effective in medium and high frequency situations if an appropriate control point can be identified.

Zone scheduling works best when most passengers are destined for (or originate at) a single stop. Examples of such services are feeder routes to rapid transit stations, or radial routes terminating in the central business district (CBD) whose ridership tends to be mostly commuters.

Signal preemption is most effective when there are reasonably heavy bus flows in the main direction and little or no bus traffic on the crossing streets. To obtain the maximum benefits from preemption, it is necessary to adjust the signal cycle length and phase splits as well.

This project has provided a battery of analytic tools for the transit operator or planner to use in evaluating particular strategies for improving service reliability in a given situation. These tools include a method for easily evaluating whether headway-based holding will be effective in a given situation and identification of appropriate control points (if any), a dynamic programming procedure to design zone schedule structures, and a discrete-event network simulation model for more detailed testing of a variety of possible strategies. The general conclusions drawn in this report should provide guidance in the consideration of possible actions, but the most important product of the research is the tools made available to the operator or planner for detailed evaluation of specific changes in his/her own environment.

Preface

This report is a sequel to the earlier report by M.A. Turnquist and L.A. Bowman, entitled Control of Service Reliability in Transit Networks, report DOT/RSPA/DPB-50/79/5, dated March, 1979. Chapters 3 and 4 of this report summarize many of the major elements of that earlier document, but the interested reader may wish to refer to that report for additional details. There is also a Users' Manual for the computer simulation model described in Chapter 2 of this report. This manual, by W.C. Jordan and M.A. Turnquist, provides a companion document useful to persons who wish to utilize the model directly.

This research was supported by contract DOT-OS-80018 from the Office of University Research, U.S. Department of Transportation. The advice and assistance of the contract monitor, Mark Abkowitz, have been extremely helpful throughout the course of the project. I would also like to thank several present and former graduate students for their contributions to this work. In particular, credit is due Larry Bowman, whose work appears in Chapters 2, 3 and 4; Steven Blume, for work in Chapter 5; and William Jordan, for work in Chapters 2 and 6. Finally, I would like to thank Keith Armstrong of the General Motors Urban Transportation Laboratory in Cincinnati, and the people at Queen City Metro, for making available extensive data from their system.

Table of Contents

Table of Contents - Continued

STRATEGIES FOR IMPROVING BUS TRANSIT SERVICE RELIABILITY

CHAPTER 1

INTRODUCTION

The concept of "service reliability" has come into increasing prominence in recent years as an important characteristic of the quality of service provided by transportation systems. A basic definition of reliability, as the term is used here, is the variability of a system performance measure over time. The focus is on stochastic variation in performance, rather than on more traditional engineering concepts of probability of component or system failure. The level of service measure most clearly subject to variation is travel time, and this variability is often described in terms of non-adherence to schedule.

Service reliability is important to both the transit user and the transit operator. To the user, non-adherence to schedule results in increased wait time, makes transferring more difficult, and causes uncertain arrival time at the destination. The importance of some measure of reliability to trip-making behavior has been emphasized in several attitudinal studies. For example, Paine et al. (1966) found that potential users ranked "arriving when planned" as the single most important service characteristic of a transit system. This finding has been substantiated in further studies by Golob et al. (1972) and Wallin and Wright (1974).

In addition to its importance to transit users, unreliability in operations is a source of reduced productivity and increased costs for transit operators. This is due to the need to build substantial "slack time" into timetables in order to absorb deviations from the schedule. This leads to reduced utilization of both equipment and personnel. The recent report by Abkowitz et al. (1978) provides an excellent summary of the major issues in transit service reliability from the perspectives of both the user and the operator.

In light of the current need for more cost-effective public transportation in urban areas, it is important to understand the sources of unreliability, and to investigate the potential of several alternative control strategies to improve both the quality of service provided and the productivity of the equipment and personnel in the system.

This project has had four major objectives:

1) investigation of the sources of service reliability problems in bus transit networks;
2) identification of potential strategies for improving reliability of service;
3) development of models to allow these strategies to be analyzed and evaluated; and
4) general evaluation of the relative effectiveness of these strategies.

The primary tool used for investigation of the sources of service reliability problems is a network simulation model. A basic prototype model was developed under a previous grant from the National Science Foundation. During the current project, this model has been greatly extended and revised, to make it more flexible, more comprehensive, and easier to use. This model has allowed investigation of the fundamental relationships between system characteristics (i.e. route length, route density, frequency of service, network connectivity, etc.) and the level of service reliability provided. Understanding of these relationships is vital to the formulation of effective control strategies. The structure of this model is described in Chapter 2, and the process of model validation is discussed in Chapter 3. Experiments with the model to investigate sources of reliability problems are described in Chapter 4. Chapters 3 and 4 summarize work described in greater detail in an earlier report from this project (Turnquist and Bowman, 1979), with the addition of some new material in Chapter 3. The interested reader is referred to the earlier report for more detail on these topics.

Chapters 5 through 8 provide detailed analyses of four major classes of strategies for improving service reliability. Chapter 5 discusses vehicle holding strategies. Chapter 6 treats methods for reducing the number of stops made by each vehicle, including increasing stop spacing and zone scheduling. Signal preemption and timing strategies are discussed in Chapter 7, and Chapter 8 describes strategies involving exclusive right-of-way for buses.

Conclusions from the research and practical implications of the results are presented in Chapter 9.

CHAPTER 2
THE SIMULATION MODEL

Construction of a prototype simulation model was accomplished as part of a previous project (Turnquist, 1978). In the course of the current project, several major refinements and extensions have been made to the model. First, the component of the model dealing with the times of passenger arrivals at bus stops has been improved greatly. Second, the specification of link travel time distributions for buses has been improved, including an option for both "macroscopic" and "microscopic" simulation of vehicle movements. This is described in greater detail in Section 2.3. Third, additional logic to reflect the enactment of various vehicle control policies has been added. Finally, the model has been validated using data from Evanston, Illinois, and Cincinnati, Ohio.

Section 2.1 describes the basic logical structure of the model. Sections 2.2 through 2.6 discuss the major elements of the model individually. A more detailed explanation of the actual mechanics of the simulation program is contained in the Users' Manual (Jordan and Turnquist, 1980).

2.1 Overview of Model Structure

The computer simulation model of bus transit network operations is written in SIMSCRIPT II.5 (CACI, 1976). The model is very flexible, so as to allow representation of a wide variety of systems. It can be used to analyze single routes (including those which branch) or entire networks of routes.

The bus system is modeled by defining classes of entities which correspond to the major physical elements of the system. For example, bus stops are defined as an entity class. Each entity has attributes associated with it which describe the current state of that entity. For example, the bus stop entity mentioned previously would have attributes indicating the time that the last bus departed, the number of passengers currently waiting, the number of routes serving the stop, etc.

Entities may belong to sets. A set is a group of entities which is "owned" by another entity. For instance, a bus stop entity may have (own) a set of passenger entities associated with it (i.e., passengers that are waiting at the bus stop). Entities can move into and out of a set, and may belong to several sets at once.

The physical elements of the bus system are modeled by this entity-attribute-set structure. The dimension of time is introduced through events. Events are incidents that are of importance to the operation of the system being modeled - a bus arrives at a stop, a bus departs from a stop, etc. Certain entities, such as buses and passengers, move through the system through the occurrence of series of events.

Bus routes are modeled as an ordered sequence of links, connecting pairs of bus stops. Buses move over these links, interacting with the general traffic stream, and are delayed at stops for boarding and alighting passengers. Passengers originate in the system at the time of their arrival at a bus stop, where they are placed in a queue to wait for an appropriate bus. When such a bus arrives, they board and ride to either their destination or a transfer stop. If they must transfer, they are placed in another bus stop queue, and the process repeats until they arrive at their destination.

Five major elements characterize the simulation model and distinguish it from other bus transit simulations:

1) flexibility in network representation;
2) user options with respect to simulating bus movements over links;
3) thoroughness in modeling passenger arrivals at bus stops, and resulting wait time;
4) ability to track passengers through the system and collect detailed origin-destination trip times and other statistics; and
5) ability to model the effects of various strategies for controlling schedule adherence in transit systems.

These elements are discussed individually in the following subsections.

2.2 Network Representation

The flexible manner in which bus routes are specified in the model allows a wide range of systems to be specified easily. Three of the basic entity types in the model are bus stops, links, and routes. Routes are defined simply as an ordered sequence of links connecting pairs of bus stops. Routes in the model are one-way, so that various types of special cases can be handled as easily as possible. For example, peak-period express services that operate in one direction only, routes that operate on different (perhaps one-way) streets in the opposite direction, or routes with one-way loops can all be represented very easily. A normal two-directional route is simply represented by a pair of one-way routes connected through vehicle assignments.

Vehicle runs are input in a timetable which specifies times of dispatches for each run on each route, and the next route to which each vehicle is assigned on completion of a given route. This type of specification allows many variations of through-routing or interlining, short turns, and branches of routes to be represented quite easily.

Throughout the model, an attempt has been made to minimize assumptions required about the structure of the network being modeled. This has been done so as to make the model as easily applicable as possible in a wide variety of situations. As a result, the transit operator or planner as a potential user of the model should be able to represent most of the important idiosyncracies of the network being analyzed without needing to make modifications to the model itself.

2.3 Bus Movement Over Links

The model contains two basic options for simulating bus in-motion times over links in the route network. In the basic, or macroscopic, model bus in-motion times on links are assumed to be random variables following a shifted gamma distribution. The amount of the shift corresponds to the minimum, or free-flow, time in which a bus could traverse a link. This value is assumed to be equal to the link length divided by the speed limit. The gamma distribution corresponds to the delay experienced by the bus. The distribution has two parameters. The scale parameter, z, reflects the duration of an average delay. The shape parameter, k, indicates the average number of interferences the bus will encounter per mile.

While z is relatively constant over different street types, obviously the k parameter can vary significantly for a bus traveling on, say, an urban street versus a bus traveling on an expressway. Hence the model allows for the input of the k and z parameters and the speed limit for several different types of links. Additional detail on this issue can be found in Turnquist and Bowman (1979).

In the macroscopic model then, if link j is of type i, the expected value and variance of in-motion time are as follows:

$$E(\gamma_j) = L_j(k_i z_i + 1/S_i) \tag{2-1}$$

$$Var(\gamma_j) = L_j\ k_i(z_i)^2 \tag{2-2}$$

where: γ_j = in-motion time of bus on link j

L_j = length of link j

S_i = speed limit on link type i

k_i, z_i = gamma distribution parameters for link type i .

At the option of the user, an alternative method of simulating link travel times can be used. This is referred to as the microscopic option. For a bus traveling on a route segment which is being simulated microscopically two important sources of bus delay are reflected explicitly in the model: signalized intersections and traffic congestion.

Traffic influences bus travel time through a relationship between traffic density and bus speed. This relationship is a modified version of the function suggested by Radelat (1973), and is illustrated in Figure 2-1. This function is used to compute the expected value of bus cruising speed. The actual bus cruising speed is assumed to be a random variable normally distributed around the expected value with a standard deviation of 4 miles per hour (Radelat, 1973, pg. 79). The traffic density is determined by dividing the traffic volume for any given block by the length and number of lanes of the block.

Just as traffic volumes affect bus speeds, the presence of buses affects the movement of traffic. This is reflected in the model through representation of activities at intersections and bus stops. A bus stopping at an unprotected bus stop will often block one lane of traffic. This is reflected in the model in the simulation of traffic movement. Additional detail on the model representation of activities at protected and unprotected bus stops, and at intersections, is contained in the Simulation Model User's Manual (Jordan and Turnquist, 1980).

Explicit representation of traffic congestion and intersections is important both because it provides a description of major sources of variation in bus travel times, and also because it provides the ability to test several promising strategies for reducing this variation, including changes in signal progression, signal preemption, etc. Such tests are discussed more fully in Chapters 5 through 8.

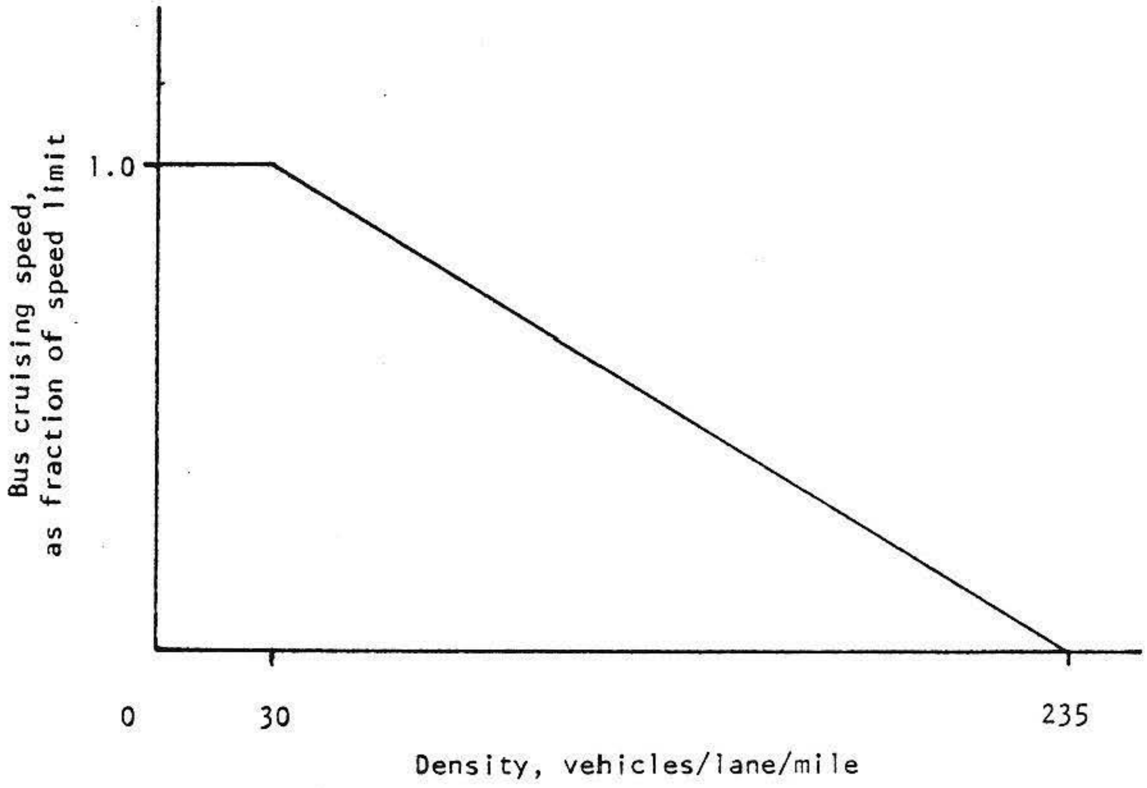


Figure 2-1. Expected bus speed as a function of traffic density.

2.4 Passenger Arrivals at Bus Stops

Passenger arrivals at a bus stop are allowed to follow one of two different processes in the simulation model. The simplest is the traditional one, assuming arrivals follow a Poisson process and are uncoordinated with the bus schedule for the stop. This random arrival pattern is an accurate representation of reality only when bus headways are short or service is very unreliable (Jolliffe and Hutchinson, 1975). It is assumed that the average arrival rate, the parameter of the Poisson process, does not vary over the simulation period.

Alternatively, passenger arrivals may be coordinated with the bus schedule. This passenger-choice model has been described previously by Turnquist and Bowman (1979). Considering a schedule headway interval of length H, a passenger arriving in this interval is assumed to have a probability density function of arriving at time t, $0 \leq t \leq H$, given by:

$$f(t) = \frac{e^{u(t)}}{\int_0^H e^{u(t)}} \qquad (2\text{-}3)$$

where: $u(t)$ = the passenger's utility for arriving at time t.

This utility function is assumed to be related to the passenger's expected wait time when arriving at time t. Based on empirical tests described by Turnquist and Bowman (1979), a useful functional form is:

$$u(t) = [E(W_t)]^c \qquad (2\text{-}4)$$

where: $E(W_t)$ = expected wait time for a passenger arriving at time t

c = constant coefficient.

When aggregated over a population of users, this probability density function becomes a time-varying arrival rate of passengers at the bus stop. This rate function is used as the basis for simulating passenger arrivals as a non-stationary Poisson process.

The importance of this passenger-choice model is that it provides a better estimate of the effects of service frequency and reliability on passenger wait time than does the traditional random arrival model. This is illustrated graphically in Figures 2-2 and 2-3.

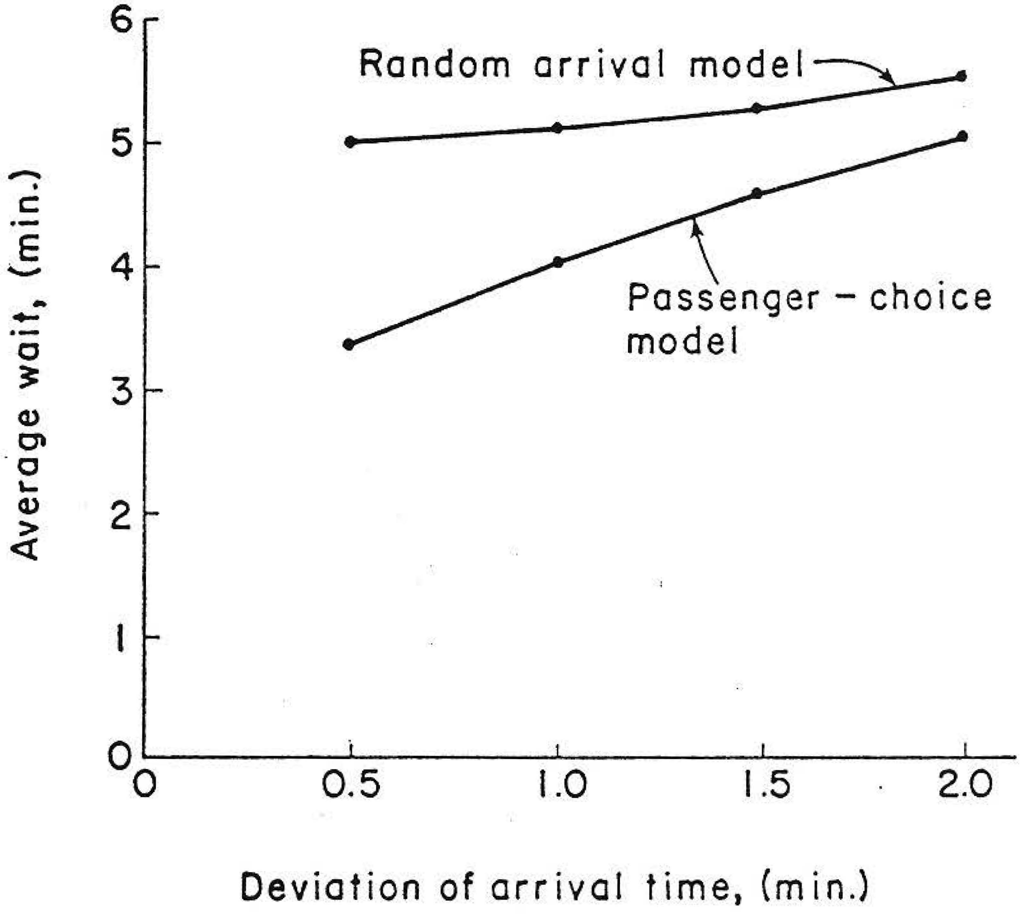


Figure 2-2. Average wait as a function of standard deviation from schedule of bus arrival time, for a service on 10 minute headways.

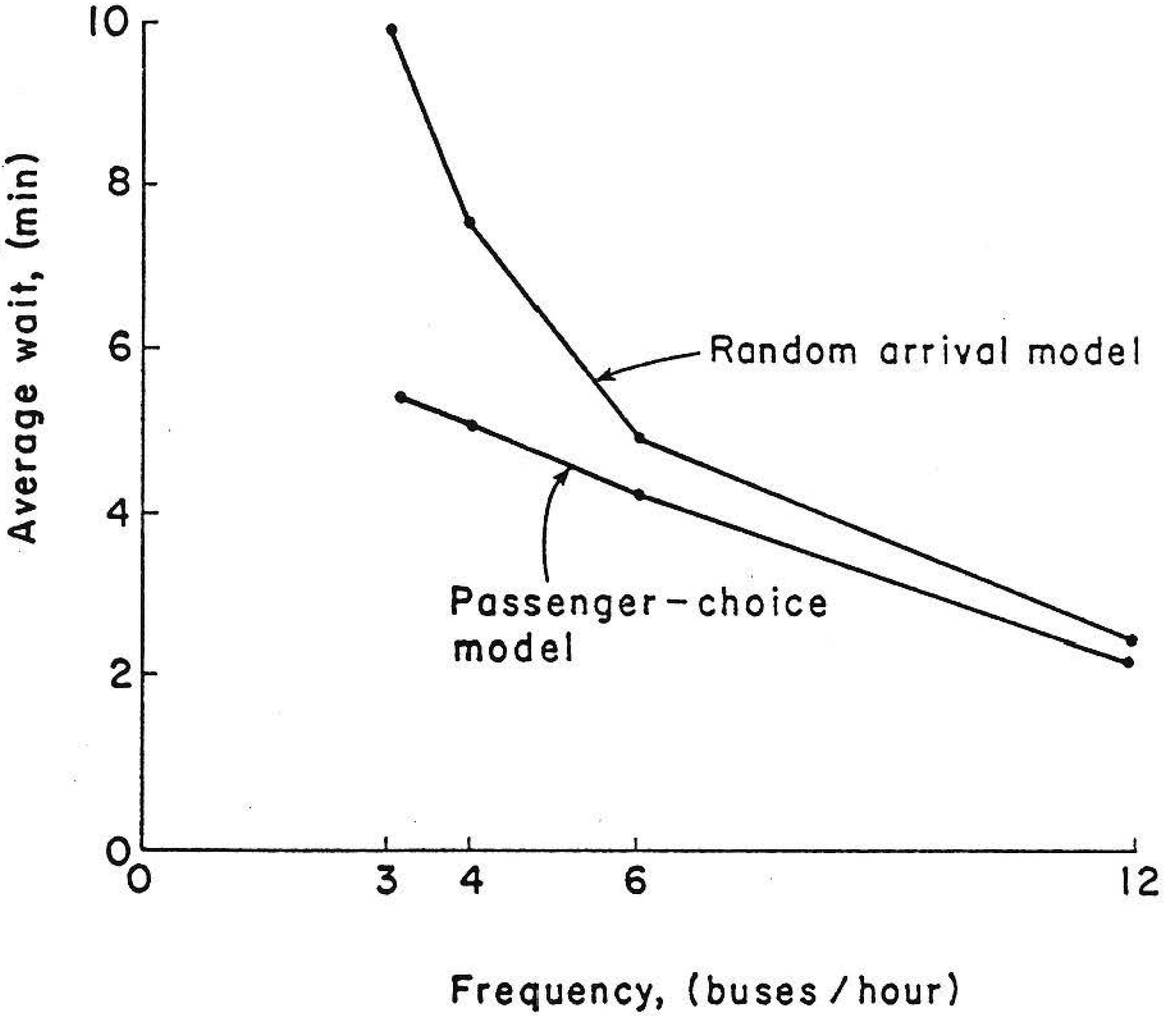


Figure 2-3. Average wait as a function of bus service frequency, for constant standard deviation from schedule of 1.0 minute.

Note that the passenger-choice arrival model indicates a much greater sensitivity to schedule deviation, and a much lower sensitivity to frequency, than does the random arrival model. This difference is extremely important, for it indicates that the benefits of service reliability improvements may be much greater than previously predicted. On the other hand, the wait time reductions from simply adding vehicles to improve frequency of service without attempting to control schedule adherence, are likely to be much less than predicted by a simple random arrival model. The implications for appropriate service improvement strategies by transit operator could be very significant.

2.5 Passenger Origin-Destination Trip Statistics

The ability to trace passenger movements from origin to destination, and to develop statistics related to this entire trip, is an important feature of the simulation model. Many previous models have either been single-route oriented, or network models which deal with "unlinked" trips. In either case, the ability of the model to test implications of various service changes on transfer trips is extremely limited.

Kulash (1971), for example, made use of a network simulation model in the analysis of system capacities, but due to his different needs he permitted the steps in a passenger's trip through the network to be "disconnected." The advantage of breaking a traveler's trip into sections is the ability to represent a multi-step trip as several independent no-transfer trips. Zonal trip generation rates are adjusted to account for additional trips created by passenger transfers. Then, instead of deciding whether each individual passenger is to be terminated or transferred, a percentage of the passengers remaining on the vehicle are simply terminated at each stop. In this research, however, analysis of the effects of service reliability on transfers requires statistics on transfer times and overall origin-destination (O-D) travel times. Each passenger, therefore, is followed through the system until arrival at the final destination.

The model produces several summary statistics on transfers and O-D level-of-service automatically. One of the most useful outputs is a histogram of effective O-D trip speed through the system. An example of this output is shown in Figure 2-4.

```
.....STATISTICS ON THE EFFECTIVE SPEED OF PASSENGERS FROM
     ORIGIN BUS STOP TO DESTINATION BUS STOP.....

          MEAN =  5.585 MPH
          STD. DEV. =  1.885 MPH
          MINIMUM =   .328 MPH

          HISTOGRAM OF EFFECTIVE SPEED
     .....INTERVAL....  .... OCCURENCES.....

           0 TO  1 MPH           12   *
           1 TO  2 MPH          117   *************
           2 TO  3 MPH          174   *******************
           3 TO  4 MPH          413   **********************************************
           4 TO  5 MPH          372   *****************************************
           5 TO  6 MPH          447   **************************************************
           6 TO  7 MPH          754   ************************************************************************************
           7 TO  8 MPH          558   **************************************************************
           8 TO  9 MPH          146   ****************
           9 TO 10 MPH           46   *****
          10 TO 11 MPH            6   *
          11 TO 12 MPH            0
          12 TO 13 MPH            0
          13 TO 14 MPH            0
          14 TO 15 MPH            1
          15 TO 16 MPH            0
          16 TO 17 MPH            0
          17 TO 18 MPH            0
          18 TO 19 MPH            0
          19 TO 20 MPH            0
          20 TO 21 MPH            0
          21 TO 22 MPH            0
          22 TO 23 MPH            0
          23 TO 24 MPH            0
          24 TO 25 MPH            0
          25 TO 26 MPH            0
          26 TO 27 MPH            0
          27 TO 28 MPH            0
          28 TO 29 MPH            0
          29 TO 30 MPH            0
          30 TO 31 MPH            0
          31 TO 32 MPH            0
          32 TO 33 MPH            0
          33 TO 34 MPH            0
          34 TO 35 MPH            0
          35 TO 36 MPH            0
          36 TO 37 MPH            0
          37 TO 38 MPH            0
          38 TO 39 MPH            0
          39 TO 40 MPH            0
```

Figure 2-4. Example of histogram of effective passenger speed, output from simulation model.

The advantage of using speed as a measure, rather than time, is that aggregating trip time statistics for O-D pairs of different trip length tends to confuse rather than clarify the meaning of the output. However, speed measures are not subject to this problem, and thus provide useful summary statistics. By looking for strategies which tend to improve both the mean and variance of the speed distribution, the model user can identify promising service improvement policies.

In addition to the summary O-D trip statistics produced automatically, the model also writes a separate, detailed passenger movement file which can be analyzed off-line. This file contains a record for each passenger that used the system during the simulation run, and includes information on origin stop, destination stop, transfer stops (if any), and arrival and departure times for each stop in the trip. This data can provide the basis for detailed analysis of wait times, transfer times, in-vehicle times, number of transfers, etc. It also allows such analyses to be segmented by trip length, number of transfers, O-D pair, etc. Because of the size of this data file, it is not printed with the simulation output, but is written to a scratch unit which can be saved on disk or tape.

2.6 Simulation of Control Strategies

Several possible strategies are available for the control of buses that have potential for improving reliability. Some of them, such as signal preemption, attempt to reduce the number and size of delays a bus must face. Others, such as vehicle holding strategies, attempt to minimize the consequences of those delays. Many of the strategies have other benefits in addition to reliability improvement, principally reduced travel time. Several are among the Transportation Systems Management (TSM) options that are being encouraged by the U.S. Department of Transportation.

Table 2-1 indicates a number of potential strategies, including both those which are "planning-oriented" and those which would be active in "real-time." In general, the distinction is that planning strategies involve changes in operations of a persistent nature. For example, changes in route structure, such as zone scheduling, have substantial long-term effects on the character of operations, and the decision to make such a change is the result of the service planning process. On the other hand, real-time control measures, such as the holding strategies, are designed to act quickly to remedy specific problems. These actions have immediate effects, but seldom exert any influence on the general nature of operations over a longer time period.

Table 2-1. Potential Strategies for Improvement of Reliability

1. Headway-based holding strategies.
2. Schedule-based holding strategies.
3. Zone scheduling of routes.
4. Reductions in the number of stops per mile of route.
5. Changes in signal timing to make progressive signalization oriented to speed of buses.
6. Reserved lanes for transit vehicles.
7. Signal preemption for buses.
8. Rescheduling for lower frequency of service and holding extra buses for fill-in duty.
9. Reduce policy volume/capacity ratio at peak load point.
10. More precisely timed dispatching at terminals.

Each of the strategies in Table 2-1, or combinations thereof, can be tested quite readily with the simulation model. Thus, the model developed in this project provides a more flexible environment for testing of a variety of service-improvement strategies than has been available previously. Chapters 5 through 8 of this report describe detailed testing of many of these strategies using both analytic models and the simulation model.

2.7 Summary

The simulation model developed in this project is the result of an effort to construct a flexible, understandable and relatively easy-to-use model of bus transit network operations. The preceding sections have described the major elements of the model, concentrating on those which distinguish it from earlier simulations of bus systems. These elements are: 1) the flexibility of network representation; 2) options available to the user for simulating vehicle movements over links; 3) the passenger-choice arrival model; 4) the ability to trace passengers through the network; and 5) the ability to test a wide variety of potential control strategies. Additional detail on all the elements of model structure, and specific instructions for use of the model, are contained in the User's Manual (Jordan and Turnquist, 1980) which complements this report.

An important aspect of model-building in general, and simulation modeling in particular, is the validation of the models against real observations. This is especially crucial in simulation modeling because the models tend to be large, complex, and incorporate many component models. Chapter 3 discusses the tests performed on the simulation model developed in this project in order to evaluate the validity of the model structure.

CHAPTER 3

SIMULATION MODEL VALIDATION

Validation of the simulation model has been based on tests involving two existing systems - Evanston, Illinois, and Cincinnati, Ohio. In each case, comparison of observed and simulated performance of the system provides the basis for assessing model validity.

Initial experiments were conducted with the Evanston route network. This network is illustrated in Figure 3-1.

Seven simulated morning rush hour periods were compared to seven days of observations collected in Evanston in 1974. Data were available for a total of 21 stops in the system. Analysis of the performance of the model addressed questions in three categories:

1. Schedule Adherence: Do vehicle schedule deviations resemble those observed in the real system?
2. Transfer Volumes: Do zonal generation-attraction rates result in the desired loadings? Do origin-destination flows correspond to those observed? Are passenger route-choice decisions modeled correctly?
3. Transfer Delay: How do the transfer times in the model correspond to those observed in the field?

The tests have been described in detail previously (Turnquist and Bowman, 1979) and will not be repeated here.

In general, the simulated results conformed quite well to the observed data from the Evanston system. The comparisons illustrated that it is easier to predict means than to predict deviations from the mean, but even when dealing with measures of deviation the simulation results were reasonably accurate.

The validation exercises indicated one aspect in which the model could be improved significantly. This involved the degree of schedule adherence on departure from the dispatching point at a route terminal. This change was incorporated into the model before further tests were run.

The second phase of model validation has involved testing the extended model with both macroscopic and microscopic representations of bus movement over links. This second phase has been based on data collected in Cincinnati, Ohio. The test network involves three routes in the "Reading Road" corridor,

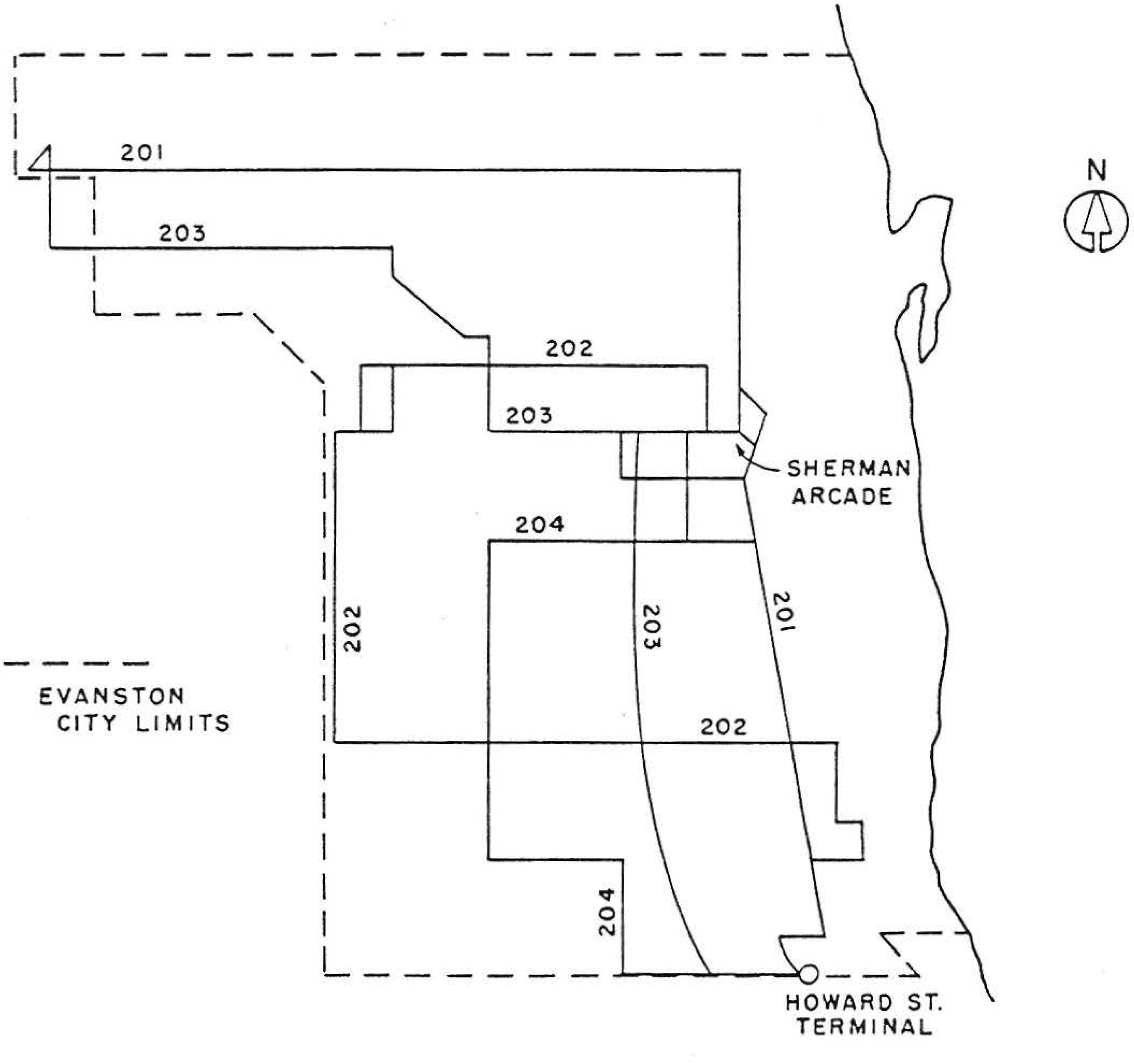


Figure 3-1. Evanston Bus Routes.

along which the General Motors Urban Transportation Laboratory has installed automatic vehicle monitoring (AVM) equipment. This has provided an excellent data base against which to conduct further validation tests. The test network is illustrated in Figure 3-2.

The set of measures used in the second phase of testing is somewhat different from those used in the first tests in Evanston. This is a result of the heavy emphasis on transfers in the first phase, and the fact that transfers are uncommon in the Cincinnati sub-network being examined. In addition, some data were available in Cincinnati that were not available in Evanston, so measures relying on these data could be used. The four principal measures used in the Cincinnati tests are:

1) mean deviation of buses from scheduled arrival time at each stop;
2) standard deviation of the deviations from schedule by stop;
3) average running time in each direction along each route; and
4) average bus load leaving each stop.

The first two of these measures reflect the degree to which the model represents the distribution of bus arrival times at each stop. These distributions are of great importance in assessing the impacts of unreliable service on passenger wait times. The degree to which the model reflects the first two moments of these distributions is an important test of its ability to indicate the effects of control strategies. The third measure is a test of the link travel time distributions and stop dwell time relationships embedded in the mode

The last measure is somewhat different from the first three, in that it constitutes more a verification of internal consistency in the model, than of valid model representation of reality. Because origin-destination (O-D) trip data were not available in Cincinnati, the O-D trip table input to the model was synthesized. Thus, the bus load measure provides a check on whether or not this model input was nearly correct, rather than being a true validation measure.

Initial experiments were run on the model using the macroscopic link travel time option, as had been the case in the initial tests in Evanston.

Figure 3-3 shows the mean deviations from schedule at each stop, plotted against the observed values. In each case, the means are taken over bus arrivals in a 3-hour period from 6:00-9:00 A.M. In general, the model predictions are quite good. A line to represent "predicted-observed" is shown for comparison. The correlation coefficient between the model values and the observed values is .73.

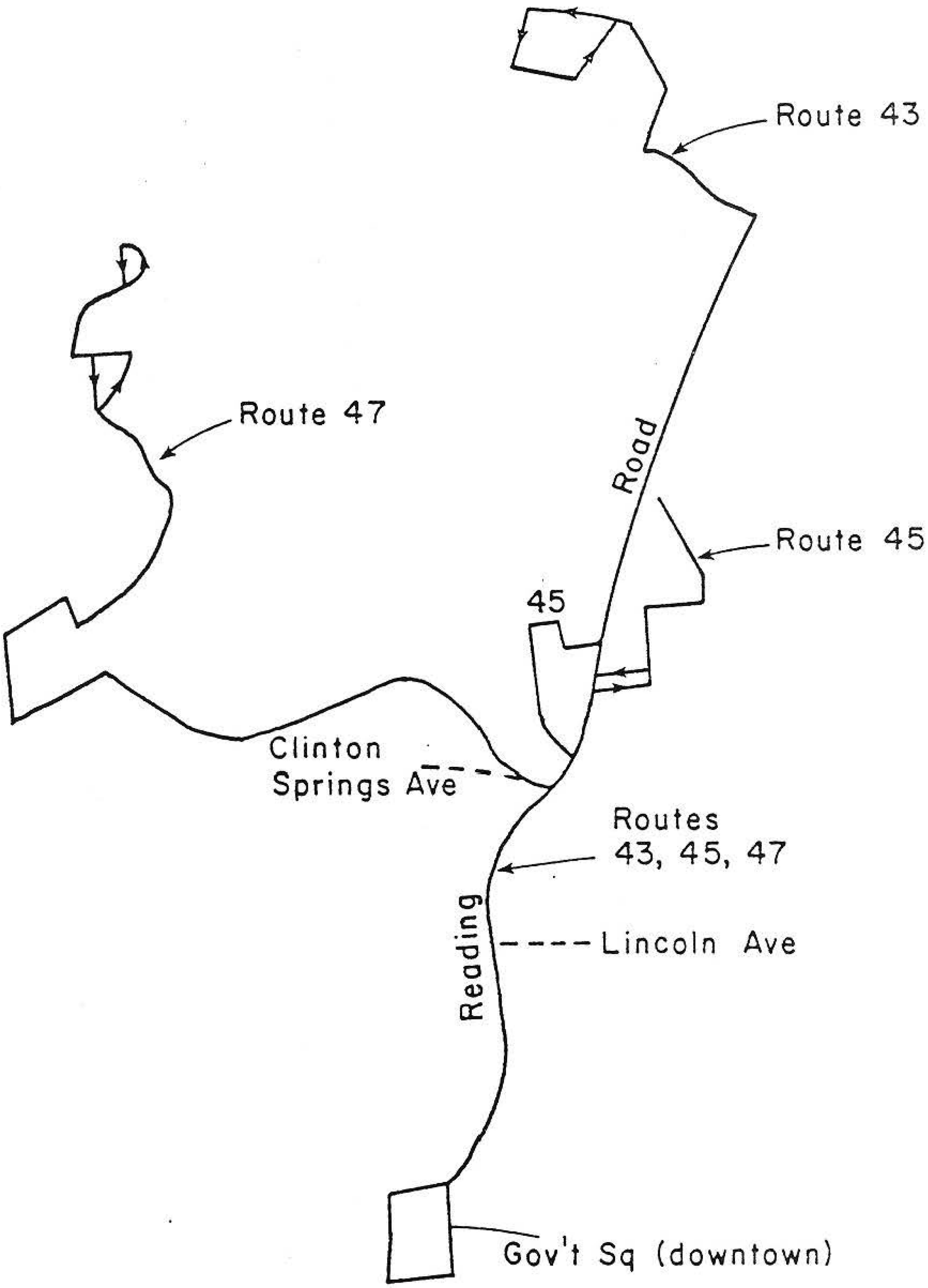


Figure 3-2. Route network from Cincinnati.

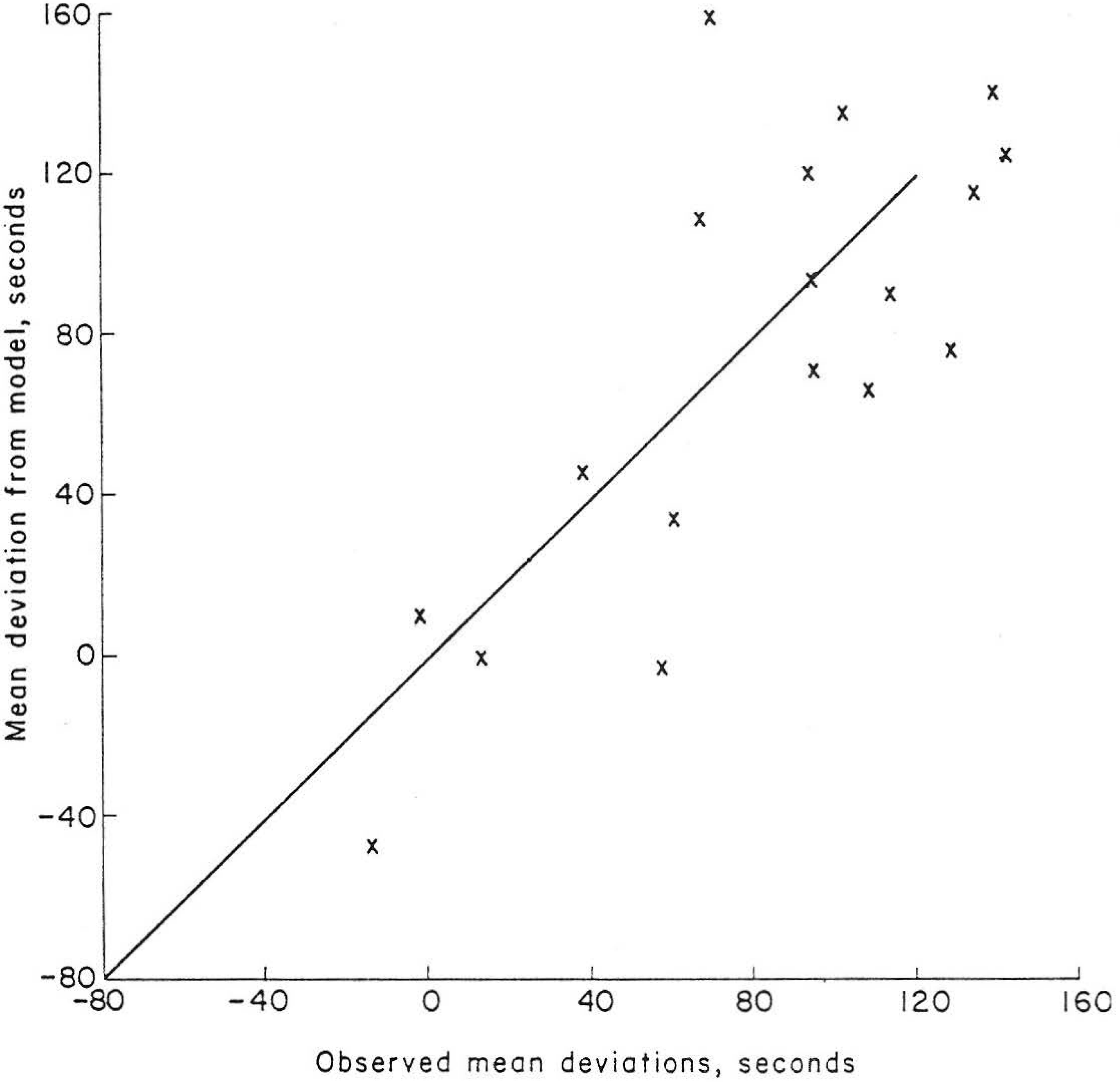


Figure 3-3. Mean bus deviation from schedule by stop - macroscopic model predictions vs. observations.

Figure 3-4 shows the relationship between predicted and observed values of the standard deviation of the distribution of departures from schedule at each stop. The correlation between predicted and observed values is .48. The model is clearly reflecting some of the major differences among stops, but is not representing all of these distributions as accurately as might be desired. This will be discussed further, in the context of the microscopic model.

Figure 3-5 illustrates the results for mean bus running time in each direction on each of the three routes. Once again, the results are generally good, with a correlation of .60. There is one observation (route 43 outbound) for which the model predicts running time approximately 20% too long. However, this does not seem to indicate a serious problem with the model. The other predictions are quite close, and the one outlier represents the off-peak direction on this route, so it is likely to result from a local peculiarity on that route rather than from a substantial model shortcoming.

The results on mean bus load are shown in Figure 3-6. The correlation coefficient between preducted and observed values is .99, indicating that the model is predicting link flows in the system very accurately, and that the O-D trip table is approximately correct.

On the basis of these tests, the macroscopic model has been judged to be acceptable. Its major weakness appears to be in representing the variability of deviations from schedule of bus arrivals at stops. It is precisely in this area that we would expect the microscopic model to provide an improvement. By reflecting the interactions of buses and the general vehicle stream more completely, and by explicit representation of signalized intersections, etc., we would expect the variability in bus travel times along a route segment to be reflected more accurately. This should translate into better representation of the variability of vehicle arrival times at stops.

Figure 3-7 illustrates that this is indeed the case, showing predicted standard deviations by stop from the microscopic model plotted against the observed values. The correlation coefficient for these values is .78, indicating a substantial improvement over the macroscopic model. It should also be noted in this regard that the microscopic representation was applied to only a section of the network, from Clinton Springs Ave. to Government Square, a section used by all three routes. The remainder of the network had the same macroscopic representation as in earlier runs. Thus, by applying the micro-

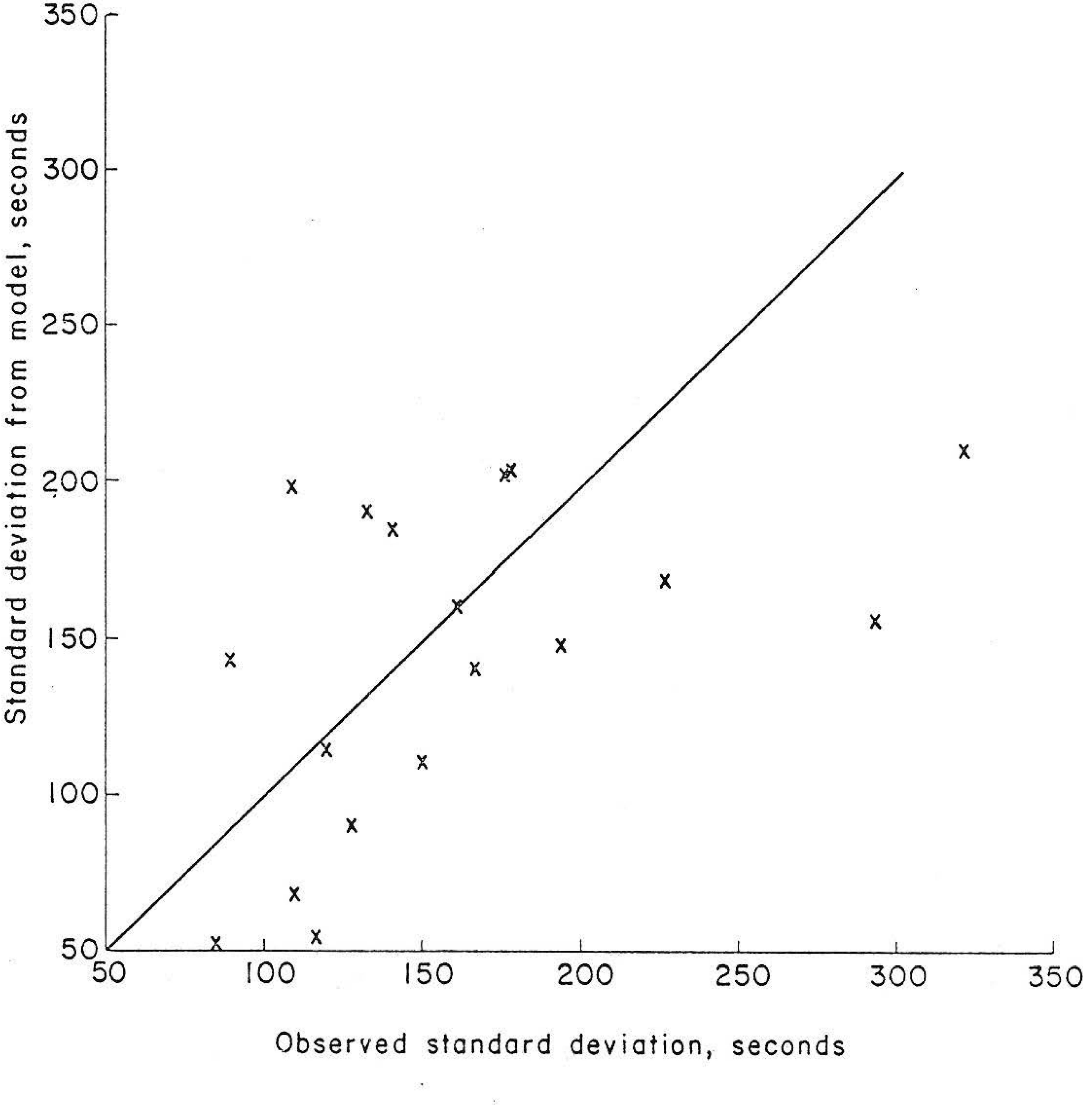


Figure 3-4. Standard deviations of the distributions of schedule deviation, by stop - macroscopic model vs. observed values.

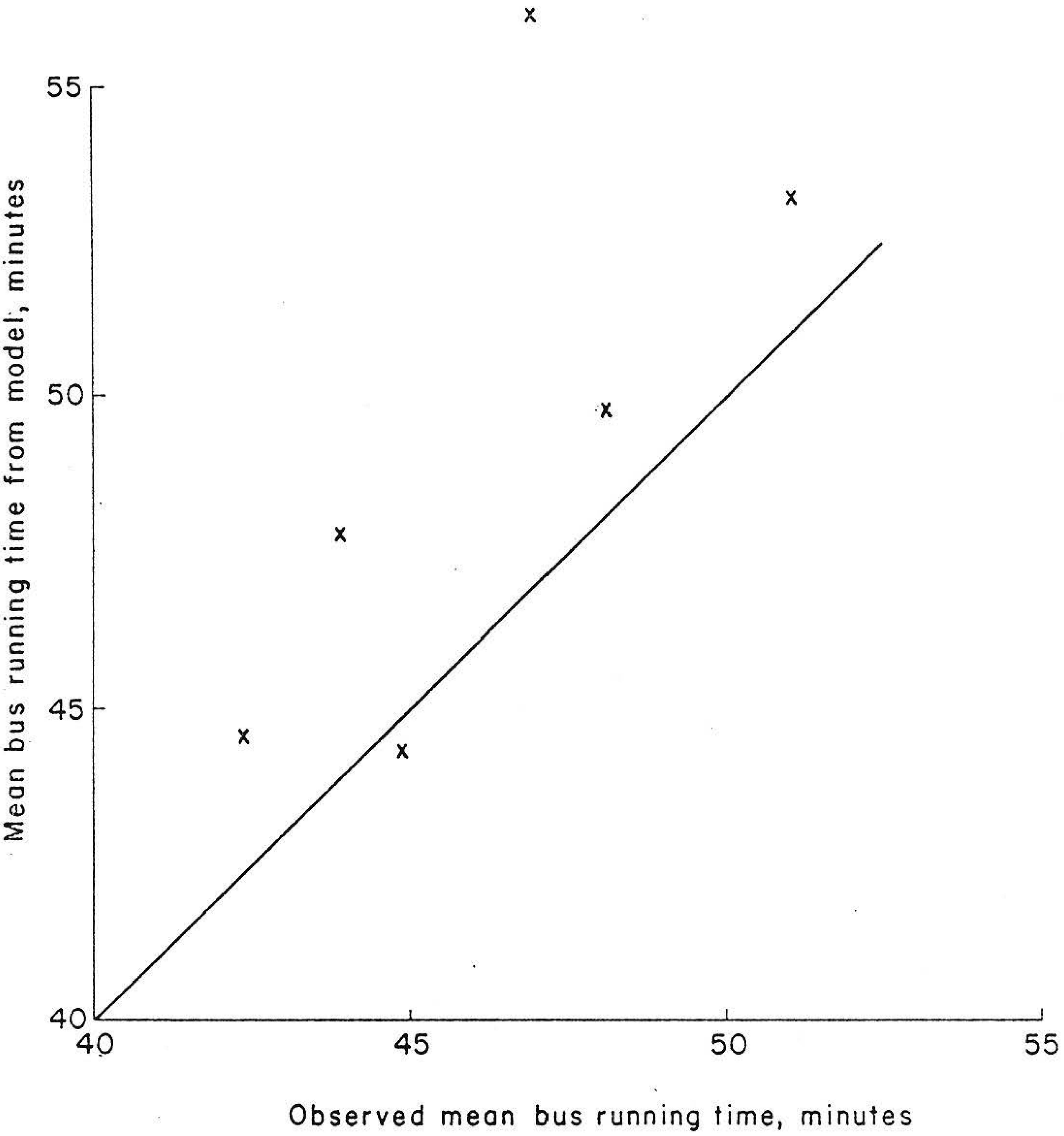


Figure 3-5. Mean bus running time, by route - macroscopic model vs. observed values.

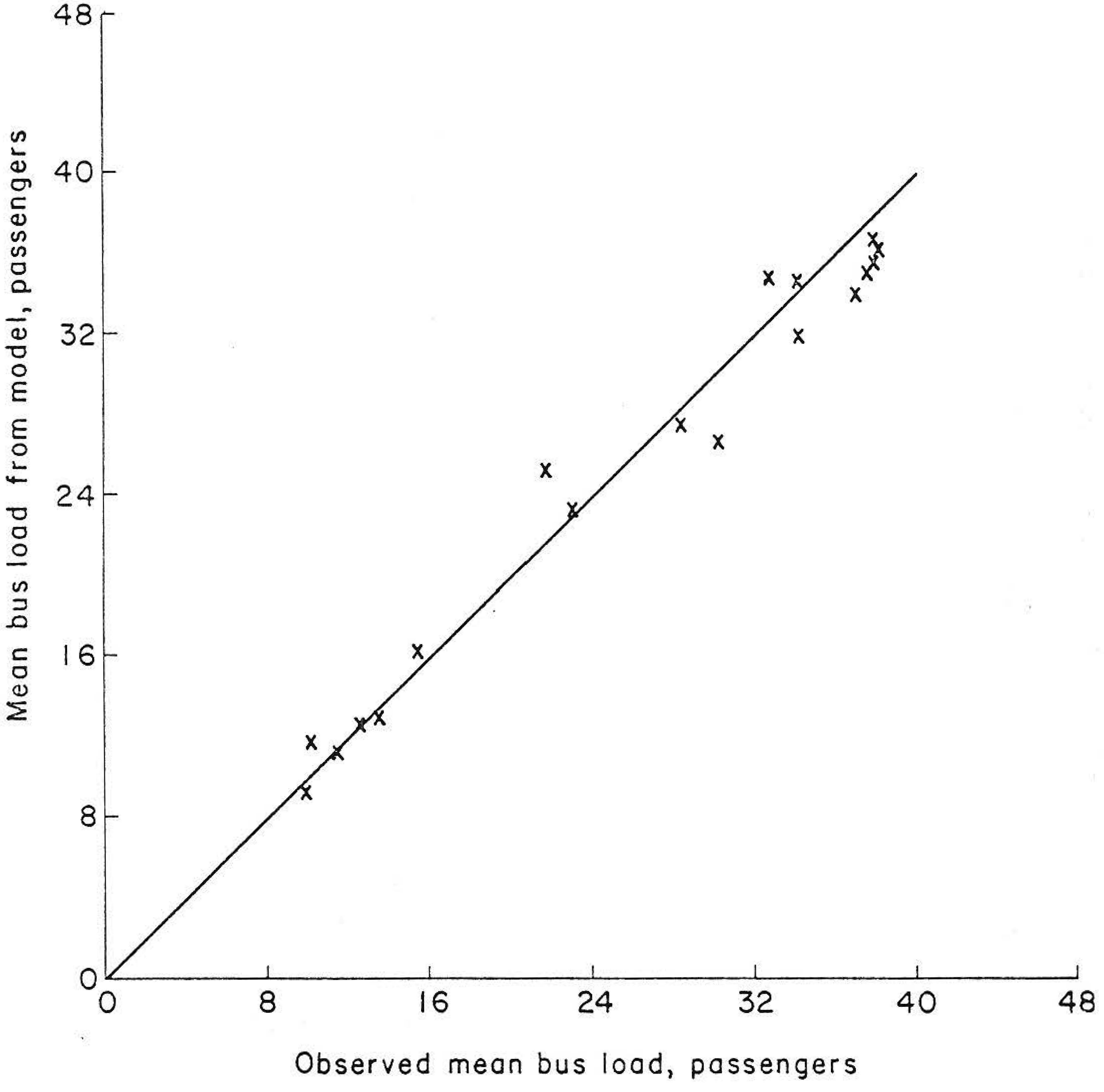


Figure 3-6. Mean bus loads by stop - macroscopic model vs. observed values.

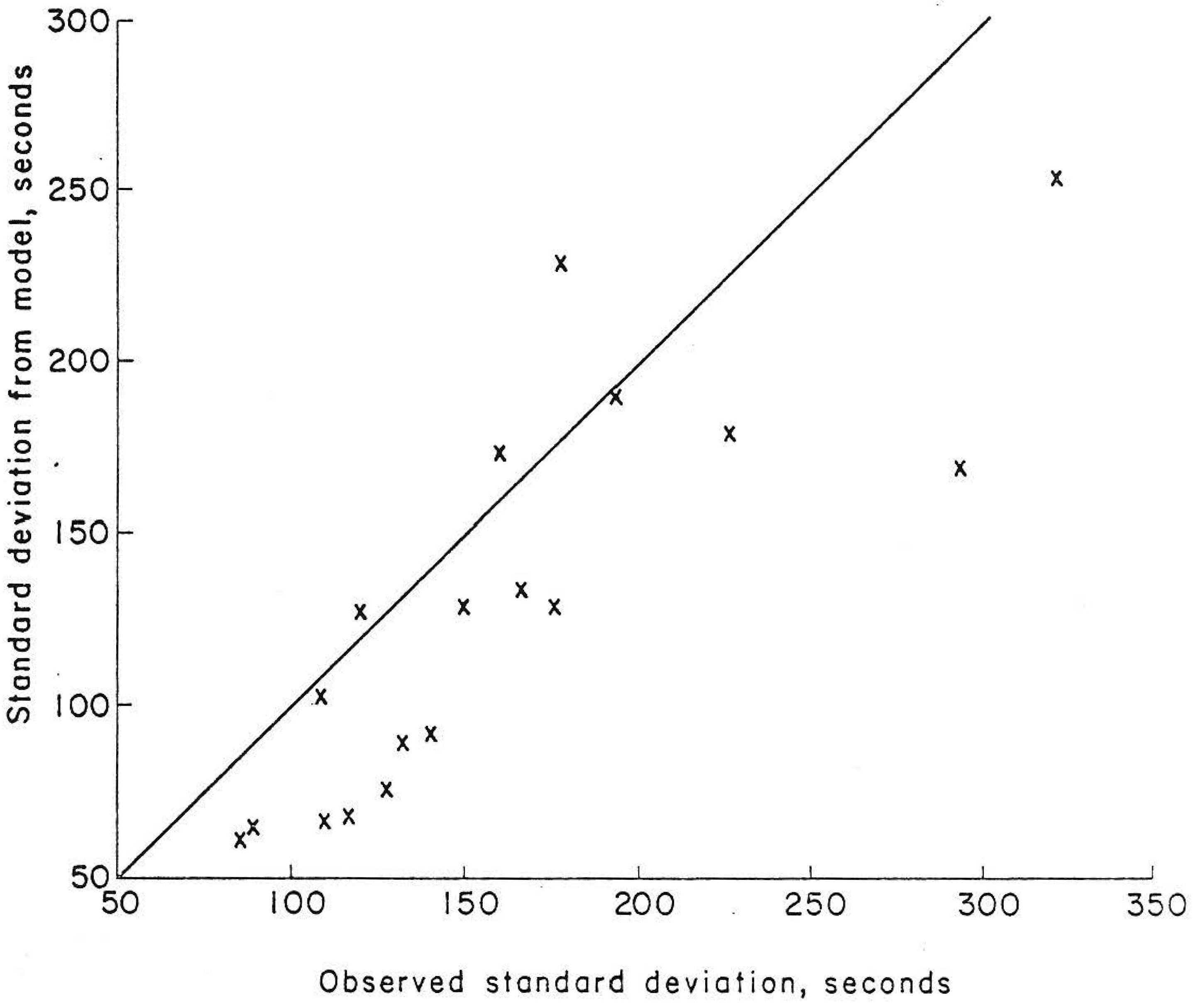


Figure 3-7. Standard deviations of distributions of bus deviations from schedule - microscopic model vs. observed values.

scopic representation to only a small key portion of the network, the overall results were improved substantially. It seems quite clear that the capabilities provided by the microscopic model are an important and valuable addition to the overall simulation.

This test has also illustrated the ability of the model user to combine macroscopic and microscopic representations on different parts of a single network. By selecting key parts of the total network for more detailed representation, the user can achieve much improved results with only a modest increase in data requirements and computer running time.

The results presented here are the culmination of a sequence of tests, model revisions followed by more tests, etc. Based on these measures, the resulting model appears to be a generally accurate representation of observed system performance, and may be accepted as a basis for further experiments.

CHAPTER 4

SOURCES OF UNRELIABILITY AND IMPLICATIONS FOR CONTROL STRATEGIES

Several previous studies (Newell and Potts, 1964; Vuchic, 1969a; Barnett, 1974) have focused on reliability at single stops and along linear routes. In addition, there are a number of fundamental questions regarding the interactions of schedule reliability and system characteristics at the network level, which have received virtually no attention. One of the objectives of this research project has been to focus on the ways in which network characteristics affect schedule reliability, and hence the level of service experienced by the users.

A set of experiments have been conducted to examine two relationships which seem to be of primary importance:

1) the effect of factors contributing to the tendency for vehicles to bunch together as they travel; and

2) the effect of network configuration, as exemplified by grid versus radial networks,

The first relationship to be considered has previously been addressed by Vuchic (1969a) using a deterministic model to explore the propagation of schedule disturbances along a transit line. This model attempts to explain the pairing of successive vehicles, or "bunching," in terms of the arrival and boarding rates of passengers at stops. The conclusion reached is that the most effective means of controlling these schedule disturbances is to reduce boarding times. The present work extends that research by including the effect of "batch" passenger arrivals from connecting routes, and more importantly, the variability in link travel times.

Grid and radial networks represent fundamentally different patterns of service. They will result in different trip routings, different lengths of trips on the network, and different transfer characteristics. Thus, it is vital to contrast the levels of service reliability offered by the two types of network structure.

In order to reach conclusions on the two major relationships indicated above, a set of experiments was designed involving five factors: 1) frequency of service (buses/hour); 2) coefficient of variation of link travel time; 3) demand/capacity ratio (total passenger-miles per hour divided by available "space" miles per hour -- both seated and standing -- on all vehicles); 4) route density (miles of two-way route per square mile); and 5) network orientation (grid or radial). Frequency of service was assumed to be the same for all routes, and the coefficient of variation in link travel the same for all links in the network.

The experimental design and details of the experimental results have been discussed at length by Turnquist and Bowman (1979), and will not be repeated here. However, a summary of the major findings of the experiments is contained in the following section.

4.1 Vehicle Bunching

The effects of the factors which increase or decrease the tendency of vehicles to bunch when traveling along a bus route have a dominant impact on service reliability. Previous research on this problem suggests that higher frequency services should be more susceptible to the tendency for vehicles to group together in their travel along routes (Osuna and Newell, 1972).

Uneven spacing of vehicles is aggravated by the differences in service delay resulting from more passengers arriving during longer service intervals, and relatively few passengers arriving during the shorter intervals. Thus, there is a tendency for long intervals to get longer and short intervals shorter. This suggests that the level of user demand also plays an important role in affecting the bunching of vehicles. The relative impact of passenger service delay will be the greatest when the spacing between vehicles is small.

Thus, the effects of demand level and service frequency tend to reinforce each other. Frequent service is generally associated with a high level of demand, and both of these factors tend to cause degradation of service reliability, as the delays for passenger boarding and alighting become a larger and larger portion of headway between vehicles.

The experiments performed in this research have also indicated an additional important factor - the coefficient of variation in link travel times. This effect is somewhat more subtle, but important. When link travel time deviations are small, the primary source of variability in bus travel times is boarding/alighting delays. Under these conditions, the interactions of demand level and service frequency, as described above, are the primary source of schedule disruptions. However, when travel times over links are also highly variable, this presents an additional source of disruption to the schedule, as well as exacerbating the effects of boarding/alighting delays.

Furthermore, the formation of vehicle bunches has an interesting counter-effect on service disruptions when link travel times are highly variable. While bunching results in service disruption when link travel time deviations are small, it can also serve to limit the extent to which the service can deteriorate. As the deviations in link travel time increase, higher frequency services will experience bunching before low frequency services. As vehicles form clusters and "leap-frog" down the route, the collective platoons of buses will be less easily influenced by individual deviations in travel times than the buses were singly. The reason bunching under high link deviations helps to constrain the magnitude of the deviations from schedule is because clustering will occur sooner on high frequency routes. Vehicles on low frequency routes have to get further off schedule before clustering begins to limit the effects of the standard deviation of link travel times.

Thus, the simulation experiments indicate that a third factor, link travel time variability, also plays an important role with respect to vehicle bunching, through its interactions with the frequency of service and level of demand. This is in addition to the obvious direct effect that travel time variability has on vehicle arrival time deviations at stops.

The formation of vehicle bunches leads to a deterioration of the distribution of headways between vehicles, as illustrated in Figure 4-1. We would generally expect the headway distribution to be symmetric and approximately normal. However, when vehicles have bunched, the distribution becomes multi-modal, corresponding to relatively large probabilities of very short headways (within bunches) and very long headways (between bunches). It is clear that this causes a substantial increase in the variance of the headway distribution, even though the mean value may remain unchanged.

This increase in headway variance has a marked impact on passenger wait time, and the tendency of bunches to form on higher frequency routes means that increases in headway variance as service frequency is increased counteract at least some of the anticipated benefits of reduced wait time for the lower average headways. Thus, simply changing service frequency is likely to have less impact on passenger wait time than would be predicted by a simple model.

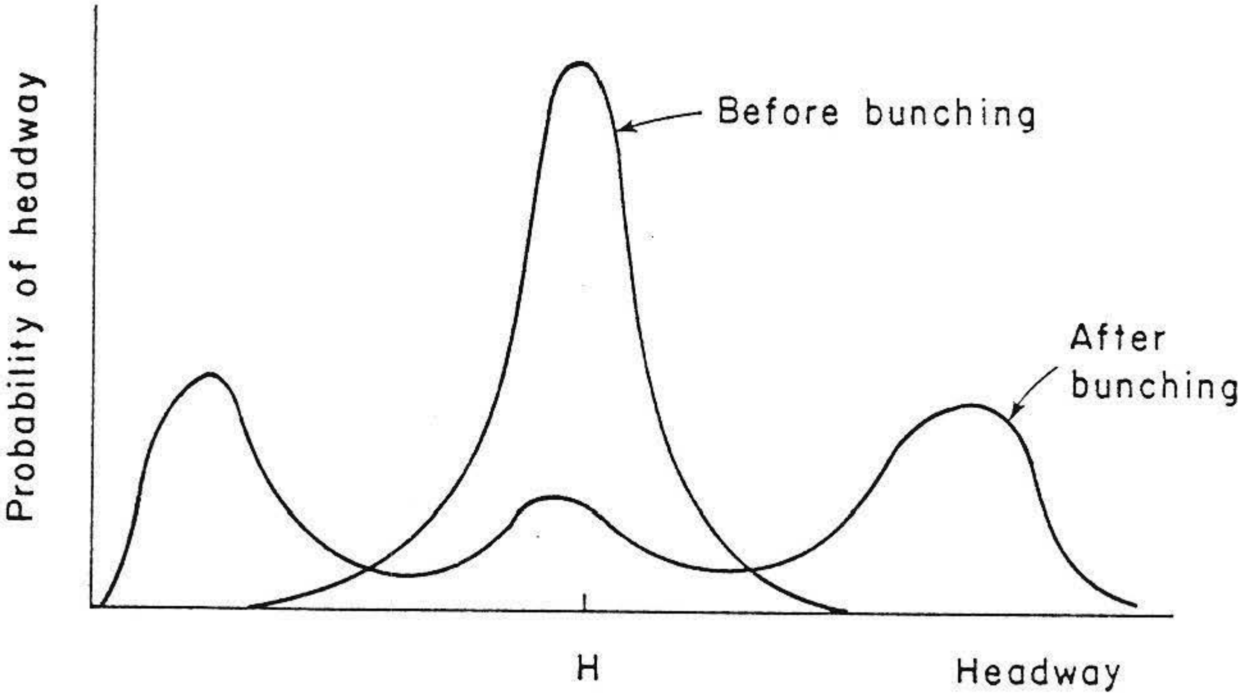


Figure 4-1. Effects of vehicle bunching on the distribution of headways.

On the other hand, policies to improve service reliability directly are likely to be much more beneficial than would be predicted by simple passenger wait time models. This is the point discussed in Section 2.4, and illustrated by Figure 2-2 (reproduced here for convenience as Figure 4-2). Therefore, transit operators seeking to allocate scarce resources most effectively to improve service would be well advised to consider efforts to improve service reliability rather than simply increasing frequency of service.

4.2 Effect of Route Configuration on Passenger Transfers

Network form and route density can be expected to have little effect on vehicle travel time reliability. However, they do contribute to uncertainty of travel through their influence on transfers. Uncertainty in the arrival times at the transfer point of both the bus from which the passenger is transferring and the bus he/she is boarding, suggests that transferring is an important source of overall trip unreliability. Thus, it is important to understand the effects of network structure on both the number of transfers which must be made, and the degree of reliability associated with each transfer.

The effects of network form and route density on the expected number of transfers can be described most easily by considering the average number of transfers as the product of two terms: the probability of transferring, and the expected number of transfers given that a transfer must be made. This is shown in equation 4-1:

$$E(X) = P(X \geq 1) \cdot E(X|X \geq 1) \tag{4-1}$$

where: X = number of transfers per passenger trip.

As route density increases, the probability of transferring also increases, since it is more likely that a traveler will utilize more than one route in the shortest path from origin to destination. This results from the fact that there are more routes serving the same total service area. However, as route density increases, it is also more likely that a single transfer will connect any given origin and destination, and thus the conditional expected number of transfers decreases. Hence, the unconditional expected number of transfers is the product of one term which is increasing with increasing route

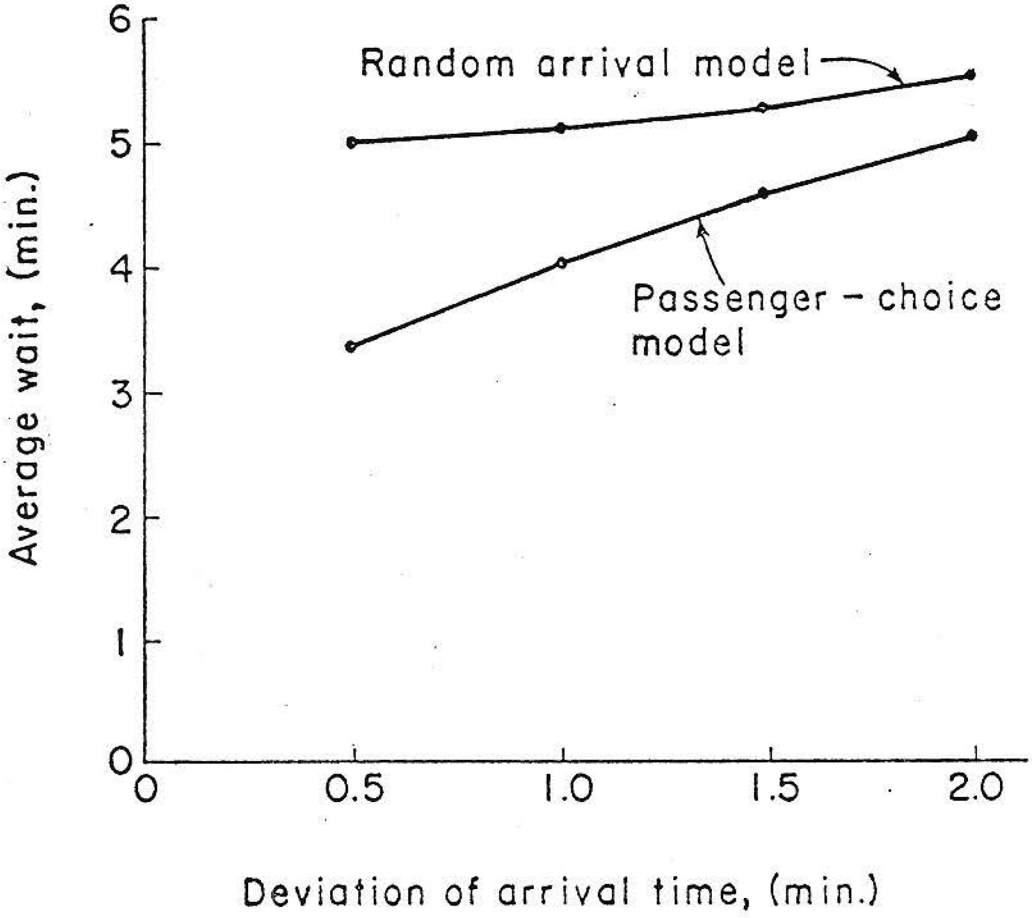


Figure 4-2. Average wait as a function of bus deviations from schedule for a 10 minute average headway.

density, and a second which is decreasing. As a result of this balancing, the overall effect of route density on the expected number of transfers is relatively small. In our experiments on grid networks, expected transfers varied between .82 and .87 as route density changed. The range for radial networks was between .74 and .87.

The major effect with which we must be concerned is then the variability in time associated with each transfer, and not the total number of transfers encountered. The standard deviation of transfer times is most directly influenced by frequency of service, since this factor largely determines the length of wait associated with a missed connection. However, if this deviation is normalized by dividing by mean headway, the effects of network form and route density become more clear.

In the experiments run, radial networks required 5.3% fewer transfers than grid networks on average, but the normalized uncertainty in the length of a transfer delay was 12.7% higher. The combined effect was an estimated 6.7% greater uncertainty in the amount of delay to users from transferring in radial networks. The greater concentration of transfers at the center node of radial networks has a more disruptive effect on reliability than does the more dispersed assignment of transfers which occurs in grid networks.

4.3 Conclusions

Understanding of the network relationships which influence the reliability of service provides insight into potential approaches for service improvement. Reliability of service, as affected by vehicle bunching, may be improved either by preventing bunches from forming, or by breaking them up after they form. The experiments performed in this research have indicated how vehicle bunching is related to frequency of service, level of demand and the variability of link travel times. In particular, these results illustrate the importance of reducing link travel time variability in an effort to prevent bunches from forming. This represents an extension to the results of Vuchic (1969a), which placed primary emphasis on the demand/capacity ratio and boarding times.

It is clear from the experimental results that service reliability is much more sensitive to frequency of service than to route density. This implies that there are substantial reliability impacts of the trade-off between operating fewer routes at higher frequency or more routes at lower frequency, given a limited amount of vehicle resources. Traditionally, this trade-off has been evaluated using simplistic models of expected passenger wait time and the accessibility of transit service to users. However, the present work has shown that service reliability is also an important factor in this trade-off and should be included in the evaluation.

This research has several practical implications for transit operators attempting to improve the level of service provided to passengers. First, the presence of large variability in link travel times can reduce substantially the benefits resulting from increasing frequency of service, due to the tendency of vehicles to bunch together along the route. In such cases, it is well worthwhile to investigate techniques for reducing this travel time variability.

The influence of transfers on level of service points out the need to pay special attention to the on-time arrival of vehicles at major transfer stations. This is especially true for radially-oriented network structures. As a rule, providing excess slack time in the route schedule is to be avoided, due to its effect on slowing down travel time and vehicle productivity. However, where a large number of passenger transfers can be aided by creating enough slack time to assure successful connections, allowing a short delay may be highly beneficial.

4.4 Potential Strategies for Improving Service Reliability

As discussed above, the major sources of reliability problems in transit service are vehicle bunching and poor connections at transfer points. In a broad sense, then, the major objectives of control strategies are to keep bunches from forming (or to break them up after they have formed) and to ensure that scheduled arrival times at transfer points are met. At a more detailed level, deviations from schedule, which lead to bunching and poor transfer connections, can be traced to excessive variability in either link travel times between stops, or dwell times at stops. Therefore, potential control strategies should be focused on reducing one or both of these sources of variability.

This investigation has concentrated on four general classes of strategies:

1) vehicle holding;
2) reductions in the number of stops served by each vehicle;
3) modifications to traffic signal settings and operation; and
4) provision of exclusive right-of-way for transit vehicles.

Such a classification provides a useful framework for discussion of many individual strategies, and a comparison of their relative effectiveness in particular situations.

The following chapters provide detailed discussions of each of these classes of control strategies, including both theoretical derivations and empirical test results.

CHAPTER 5

VEHICLE HOLDING STRATEGIES

Vehicle holding strategies attempt to prevent bunches from forming, and serve to break up bunches that may already have formed. When enacted at major transfer points, such strategies can also be useful in ensuring that scheduled connections are made.

Two important sub-classes of strategies can be distinguished. One type is oriented toward holding vehicles to a particular schedule, and the second is focused on maintaining constant headways between successive vehicles. The schedule-based approach will be discussed first, followed by a discussion of headway-based methods.

5.1 Schedule-Based Holding

A schedule-based holding strategy is nothing more than creating "checkpoints" or "time points" along a bus route, and insisting that no vehicle leave a time point before its scheduled departure time. This is probably the simplest form of schedule control possible, and is practiced (at least in theory) by many transit operators. Theory and practice often differ, however, because of lack of enforcement.

The keys to successful implementation of a schedule-based checkpoint strategy are: 1) to have a schedule to which vehicles have a reasonable chance to adhere, and 2) to enforce the rule of no early departures from the checkpoint. It is important that the mean arrival time of buses at the checkpoint be approximately the scheduled time. If the schedule is unrealistic, so that vehicles are consistently late, this strategy will have little or no effect since the control actions directly affect only those vehicles which are ahead of schedule. On the other hand, it is inadvisable to have a schedule so slack that almost all vehicles are early, since delaying all these vehicles to meet the schedule of the slowest vehicles imposes penalties on a large number of passengers and reduces overall vehicle speed and productivity.

The issue of enforcement is also important. Successful implementation of the policy requires that drivers have incentives to adhere to schedule, and that they are able to monitor their own degree of adherence. Appropriate incentives can often be provided by imposing penalties (monetary or otherwise) for early arrival at the end of a route. However, if such sanctions are to be imposed, it is also important to provide the means for drivers to monitor their own performance accurately. It is likely that if drivers must rely on their own watches for timekeeping, substantial variation from schedule will still exist even if they are trying to maintain the schedule as closely as possible. This problem could be remedied through installation of accurate clocks on the buses. Electronic digital clocks are extremely accurate, and relatively inexpensive. This would provide a standard for timekeeping, and a basis for judging whether or not penalties should be imposed on a given driver for departing from schedule.

A schedule-based holding strategy can be particularly useful on suburban routes, or in other instances where headways are quite large. When service is relatively infrequent, passengers tend to learn the schedule and coordinate their arrival at the bus stop with the scheduled arrival time of the bus, so as to minimize wait time. This has been described in detail in Chapter 2. In such cases, adherence to schedule by the buses is very important in provision of quality service to the passengers. The discussion in Chapter 2 also indicated the magnitude of potential benefits from increasing adherence to schedule, as might be accomplished by a checkpoint strategy.

5.2 Headway-Based Holding Strategies

In situations where service is quite frequent, we might expect headway-based holding strategies to be effective. If service is frequent enough so that passengers may be assumed to arrive randomly in time at a given bus stop without regard to the schedule of service, the average waiting time, E(W), is [Welding, 1957]:

$$E(W) = \frac{E(H)}{2} + \frac{V(H)}{2E(H)} \qquad (5\text{-}1)$$

where : $E(H)$ = expected headway between successive vehicles

$V(H)$ = variance of headways.

It is clear from (5-1) that making the headways more regular (i.e., reducing the variance) will tend to reduce average waiting time. This is the motivation for headway-based control strategies.

We will consider the case of control in "steady-state" time periods in which the headways scheduled are constant in response to a constant demand for the service through the period. The term "steady-state" is used to describe periods over which the underlying conditions affecting bus operations and passenger arrivals do not change, at least to an approximation. We also assume that the appropriate number of vehicles has been scheduled for the period that they all must be dispatched within that period to provide sufficient capacity.

In general, the objective of control is to minimize a weighted sum of headway variability and expected delay due to the holding strategy. Consider specifically the weighted sum of passenger waiting time and in-vehicle delay, $E(\psi)$, when control has been instituted:

$$E(\psi) = (1 - b)\, E(W') + bE(D) \tag{5-2}$$

where: $E(W')$ = expected wait time after control

$E(D)$ = expected delay to passengers who must ride through the holding point

b = a weighting constant, $0 \leq b \leq 1$.

The expected wait time after control can be written as follows:

$$E(W') = \frac{E(H')}{2} + \frac{V(H')}{2E(H')} \tag{5-3}$$

where H' is a random variable describing the headways after control. Since we are assuming that all vehicles must be dispatched during the time period under analysis, we can assert that:

$$E(H') \equiv E(H). \tag{5-4}$$

That is, the average headway will not be affected by the control actions. However, if the control is effective, we should find that:

$$V(H') < V(H) \tag{5-5}$$

which would indicate a reduction in expected passenger wait time.

One natural interpretation for the parameter, b, is that it is the proportion of passengers delayed, out of all those affected by the control strategy. However, the formulation is really quite general, since b can be set to reflect any set of priorities on reduced variability vs. increased delay. It could take into account knowledge of the actual likely improvement in variance some distance after the control point, different values of bus stop waiting time and in-vehicle delays, or operating priorities independent of passenger time considerations. Throughout this analysis, b will be assumed to be the proportion of passengers who must travel through the holding point. Also, it will be assumed that each stop downstream of the control point experiences the same reduction in variance. With these assumptions, we can write the net improvement in delay of the average passenger as follows:

$$\begin{aligned} E(\Delta\psi) &= (1 - b)\ [E(W) - E(W')] - b\ E(D) \\ &= \frac{(1 - b)}{2E(H)}\ [V(H) - V(H')] - b\ E(D)\ . \qquad (5\text{-}6) \end{aligned}$$

We wish to maximize $E(\Delta\psi)$ by determining appropriate holds to be applied to vehicle i, $i = 1,2,\ldots,n$. Let us begin by finding the net passenger minutes saved, ξ_i, achieved by holding the i^{th} vehicle X_i minutes. We define the difference in arrival times at the control point between the i^{th} and the $i\text{-}1^{th}$ bus to be the i^{th} headway, H_i.

The people who benefit from the holding of vehicle i are those who now can board bus i and otherwise would have had to wait for vehicle i + 1. The people who arrive at their stops just after the time bus i would have come had it not been held each save a wait $H_{i+1} - X_i$ because they can now board the i^{th} bus instead of bus i + 1. (They save an additional amount of time if the $i + 1^{st}$ bus is held, but this factor is considered at the $i + 1^{st}$ holding decision.)

Let Q be the total arrival rate (passengers/minute) of all passengers on the route who travel beyond, or board the vehicle beyond, the control point. (The control has no effect on those people who disembark before the control point.) Then, on average, there are $(1 - b)\ QX_i$ people who benefit, so the expected passenger-minutes saved is $(1 - b)\ QX_i\ (H_{i+1} - X_i)$.

The people who are penalized by the holding of bus i include those who were already waiting for the i^{th} bus and now must wait another X_i minutes for it. They are the customers who arrived since vehicle i - 1 passed; on average, there are $(1 - b)\ Q(H_i - X_{i-1})$ of these people. Those already on the bus are also delayed an amount X_i by the hold. The expected number of passengers on the bus is bQH_i.

The net reduction in passenger minutes of delay from holding decision i is then:

$$\begin{aligned} \xi_i &= (1 - b)\ QX_i(H_{i+1} - X_i) - (1 - b)\ Q(H_i - X_{i-1})\ X_i - bQH_iX_i \\ &= (1 - b)\ QX_i(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1} - X_i). \qquad (5\text{-}7) \end{aligned}$$

Note that:

$$\begin{aligned} E(\Delta\psi) &= \frac{\text{expected net change in total delay during period}}{\text{expected number of passengers during period}} \\ &= \frac{nE[\xi_i]}{QnE(H)} \\ &= \frac{1}{E(H)\cdot Q}\ E[\xi_i] \quad , \qquad (5\text{-}8) \end{aligned}$$

so that maximizing $E[\xi_i]$ is equivalent to maximizing $E(\Delta\psi)$, the overall improvement in average passenger delay.

5.2.1 Optimal and Near-Optimal Strategies

Theoretically, the decision on how much to hold the i^{th} vehicle should depend not only on the immediate benefits achievable, but also on the effect the hold will have on possible future benefits. (Note that equation (5-7) for ξ_i contains X_{i-1}.) If all headways were known, the exact solution would require a quadratic programming algorithm which would simultaneously determine all the expected future holds and X_i.

To see what this solution would look like, consider the hypothetical case in which all the headways are known and it is desired to find the set of holds that maximizes the total wait reduction. This problem may be written as follows:

$$\max \sum_{i=1}^{n} \xi_i = (1 - b)\, QX_1 (H_2 - H_1 - \frac{b}{1-b} H_1 + X_0 - X_1)$$
$$+ (1 - b)\, QX_2 (H_3 - H_2 - \frac{b}{1-b} H_2 + X_1 - X_2)$$
$$\vdots$$
$$+ (1 - b)\, QX_{n-1} (H_n - H_{n-1} - \frac{b}{1-b} H_{n-1} + X_{n-1} - X_n) ,$$

subject to

$X_i \geq 0,\ i = 1,2,\ldots,n-1$
$X_0 = 0,$
$X_n = 0.$

The problem is similar to the dispatching problem discussed by Bisbee et al. (1968) and Newell (1971) for the minimum wait solution to dispatching a given number of vehicles in response to a known demand pattern. The solution is certainly feasible, but not particularly interesting, since in a practical situation, the strategy must be implemented through time, and decisions on X_i must be made before H_{i+1}, H_{i+2}, ..., H_n are known.

An alternative formulation is to consider the problem as a dynamic programming problem, treating future unknown headways as random variables. Define $g_i(X_{i-1})$ as the maximum expected reduction in wait for vehicles i, i+1,...,n, as a function of the previous hold, X_{i-1}.

Then:

$$g_i(X_{i-1}) = \max_{X_i} \{\xi_i(X_{i-1}) + E[g_{i+1}(X_i)]\} \quad , \qquad (5\text{-}9)$$

where: $\xi_i(X_{i-1}) = (1 - b)\, QX_i (H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1} - X_i).$

Given the last hold X_{i-1}, we would like to find the X_i that maximizes the reduction in wait at this stage and the expected reduction at future stages.

Rigorous solution of this recursion, however, is difficult. To illustrate this, consider the case where H_{i+1} is known at the time of decision i. Then

$$g_i(X_{i-1}) = \max_{X_i} \{ \xi_i(X_{i-1}) + \int g_{i+1}(X_i)\ f(H_{i+2} | H_{i+1})\ dH_{i+2} \} , \qquad (5\text{-}10)$$

where $f(H_{i+2}|H_{i+1})$ is the probability distribution of H_{i+2}, given H_{i+1}. This is difficult to solve because the optimum X_i depends on the optimum X_{i+1}, which in turn depends on the yet unknown H_{i+2} and X_{i+2} and so on. Thus, the dynamic programming formulation does not lead directly to an implementable solution.

It does, however, indicate the nature of some appropriate solutions that are likely to be near-optimal. Suppose, for example, we assumed that:

$$\frac{d\int g_{i+1}(X_i)\ f(H_{i+2}|H_{i+1})\ dH_{i+2}}{dX_i} \simeq 0 \quad . \tag{5-11}$$

That is, we neglect the effect the hold X_i has on subsequent holding decisions and wait reductions. This "decouples" the problem by making the stages independent, and allows us to attack the maximization of each stage separately.

We find the X_i that maximizes the immediate benefits ξ_i:

$$\frac{d\xi_i}{dX_i} = (1 - b)\ Q(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1} - 2X_i) = 0 \tag{5-12}$$

$$\Longrightarrow \quad X_i = .5(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}). \tag{5-13}$$

If H_{i+1} is small enough, the computed quantity on the right will be negative. Since vehicles can only be delayed, $X_i \geq 0$. Because ξ_i is concave in X_i, application of the additional Kuhn-Tucker condition necessary to ensure $X_i \geq 0$ yields the result that the optimum hold when the quantity $H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}$ is negative, is to not hold at all. Thus, one proposed near-optimal control strategy is

$$X_i^p = \max[0,\ .5(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}^p)] \ . \tag{5-14}$$

A strategy similar to X_i^p has been referred to by Jackson (1977) and Turnquist and Bowman (1979) as the "Prefol" policy because it splits the difference between the <u>pre</u>vious and <u>fol</u>lowing headways, $(H_i - X_{i-1})$ and H_{i+1}. However, the policy evaluated previously did not incorporate the adjustment factor $\frac{-b}{1-b} H_i$, which reflects a consideration for the delay of the people on board the vehicle.

This policy requires a prediction of the arrival time of the following ($i+1^{st}$) vehicle. Automatic train control systems could provide train location in rapid transit applications. For bus systems, location may be determined by Automatic Vehicle Monitoring (AVM) technologies. A projection of its speed to the control point would also be required.

A less reliable, but much less expensive, prediction of the following headway would be its statistical expectation, $E(H_{i+1}|H_i)$, based on the current headway. This suggests an alternative control policy:

$$X_i^s = \max\{0,\ .5[E(H_{i+1}|H_i) - H_i - \frac{b}{1-b}H_i + X_{i-1}^s]\}\ , \qquad (5\text{-}15)$$

which will be referred to as the "Single Headway" policy. It is dependent only on the known current headway and previous hold.

These near-optimal strategies neglect the effect of the current hold on subsequent holding decisions. To determine the impact of this, consider, for example, a situation in which two successive vehicles are to be held. In this case, the holds should be such that all three headways affected should be equal. The Prefol policy would only make two of them equal. However, given the large variability in transit operations, the random errors in the predictions of future headways -- no matter how carefully made -- are likely to overcome any benefits of a more sophisticated calculation and/or more extensive monitoring equipment. The two policies proposed here are simple in form and near enough to optimality to provide a large portion of the benefits possible with headway control.

5.2.2 Characteristics of the Strategies

It is illustrative to consider a simple example of how each strategy would work. Let us first examine an example of the Prefol strategy. Suppose buses are scheduled to arrive every six minutes at a downtown stop in the afternoon peak. Because of reliability problems that have accumulated through the day, the buses arrive erratically, resulting in long passenger queues and some underloaded and some overloaded vehicles. Because few people are yet on the buses traveling outbound, some type of headway control may be effective; few people would be delayed. The enactment of a Prefol strategy would require monitoring equipment capable of locating the following bus. Because there are few people boarding before the control stop, the time of arrival of the following bus can be predicted with relative confidence. The prediction technique and the Prefol policy, with $b = 0$, gives:

$$X_i^P = \max[0, \ .5(H_{i+1} - H_i + X_{i-1}^P)].$$

This is equivalent to saying that the i^{th} vehicle is held until the previous headway $H_i - X_{i-1} + X_i$ is equal to the projected following interval $H_{i+1} - X_i$. The control might be carried out manually by a "starter" on the street, or displayed on a variable-message sign at the stop, directing the driver when to leave.

An example sequence of vehicle intervals is shown in Figure 5-1. A 7-minute headway is projected to follow a 3-minute headway. The vehicle is held until both the previous and following headways are equal to 5 minutes ($X_i^P = .5(7 - 3 + 0) = 2$). The next vehicle is not held because the decision criterion, $.5(3 - 7 + 2) = -1$, is negative.

The Single Headway policy might be enacted if information on the current position of the following vehicle is unavailable. The only required capability is a device or a person at the control point to recognize the presence of a vehicle from the controlled line, remember successive arrival times, and calculate the hold. In a bus system, this person or device may also indicate to the driver when he can proceed from the holding point.

Consider again the sequence of headways in Figure 5-1. If the following headway is predicted simply to be the average (6 min.) the Single Headway policy would hold the first vehicle:

$$.5[E(H_{i+1}|H_i) - H_i + X_{i-1}] = .5(6 - 3 + 0) = 1.5 \text{ min.}$$

The next vehicle would be held .25 minutes. Note that the Single Headway strategy does not reduce variability as much as the "more informed" Prefol strategy in this example, although the difference is relatively small.

The proposed Prefol policy has several appealing features. First of all, it is a <u>robust</u> strategy. If the prediction of H_{i+1} is accurate, and if the expected number of passengers arrive, the immediate reduction in delay, ξ_i, is always positive or zero. Strategies based on less information -- such as the Single Headway policy -- will actually increase the total delay (i.e. ξ_i will be negative) in a certain portion of the decisions; they are only optimal "on the average."

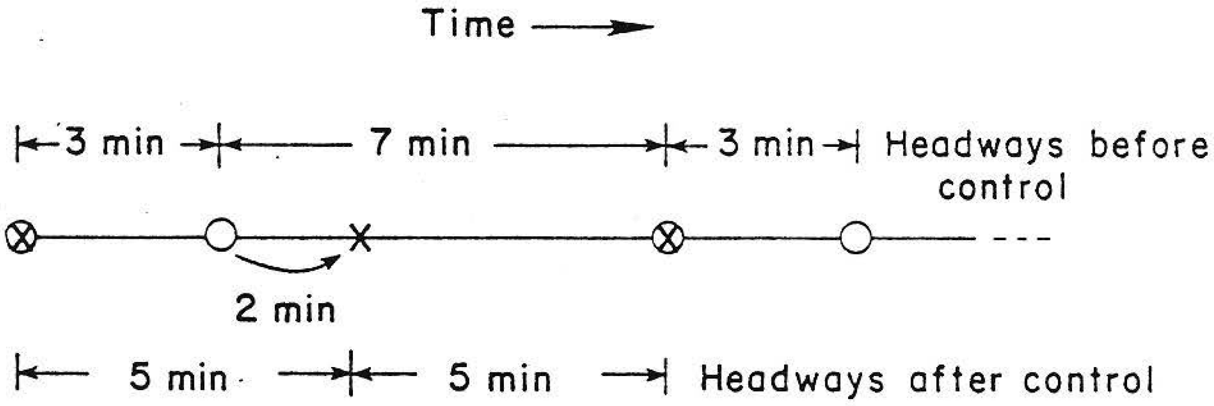


Figure 5-1(a). Example holding by Prefol strategy.

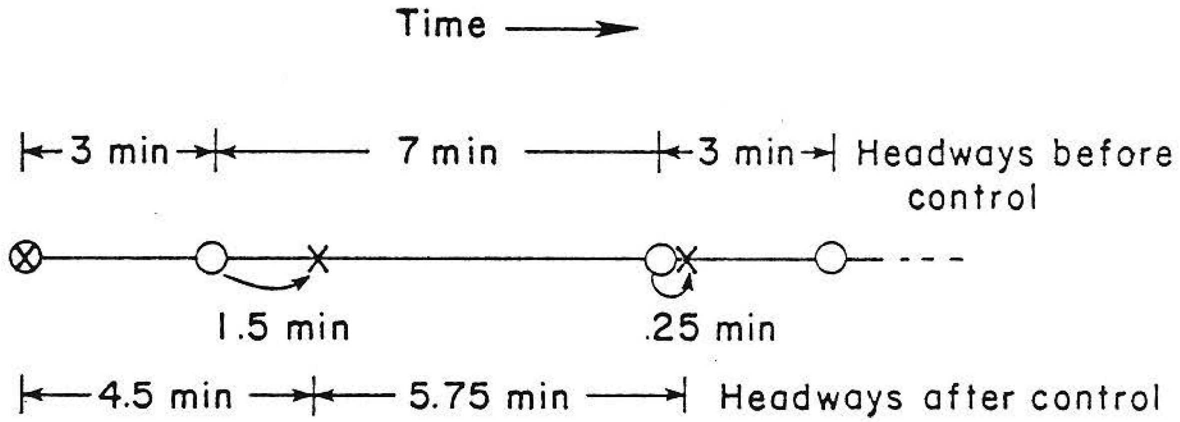


Figure 5-1(b). Example holding by Single Headway strategy.

The only parameter of the Prefol policy is b and the same decision rule is applied regardless of the average headway. This implies that the strategy may work reasonably well in changing conditions when statistics necessary for other strategies may be unreliable. In a non-steady-state period, when passenger arrivals and scheduled headways are changing, the Prefol strategy will continue to work well as long as the passenger arrival rate doesn't change too rapidly from one bus to its follower.

The success of the strategy will depend on the accuracy with which the arrival time of the following bus is predicted. This might be based, as in the above example, on the follower's position (e.g., as shown on a computer display) or, more crudely, on a measurement of the headway at an upstream monitoring point. An upper bound on the benefits of control is available by assuming that H_{i+1} is known exactly. The Single Headway strategy, since it substitutes an expected value for H_{i+1}, provides a lower bound on the success of Prefol by representing the case where the prediction accuracy is poor.

Note that a Prefol implementation would not increase the average headway, since no vehicle is held past the arrival of its follower. (Technically, the average headway is increased by a small amount if the last bus is held past the end of the period.)

Both the Prefol and the Single Headway policies are likely to be more effective than "threshold-based" holding strategies. The strategy "hold until the headway reaches a minimum threshold" has often been suggested, and modeled in the literature (Barnett, 1974; Jackson, 1974; Koffman, 1978; Turnquist and Bowman, 1979). Simulation work has indicated that this strategy tends to delay too many vehicles too long, increasing the average headway and sometimes actually lengthening passenger wait time. See, for example, Jackson (1974), Koffman (1978), and Turnquist and Bowman (1979).

By imposing two additional constraints on the Single Headway policy, we can derive a threshold-type policy. We assume first, that $E(H_{i+1}|H_i)$ is equal to the unconditional mean, $E(H)$, and second, that the term $\frac{b}{1-b} H_i$ is equal to its expected value, $\frac{b}{1-b} E(H)$. The Single Headway policy then becomes

$$X_i^s = \max[0,\ .5(\frac{1-2b}{1-b}\ E(H) + X_{i-1}^s)] \tag{5-16}$$

which is a threshold-type policy. Turnquist and Blume (1980) have shown this to be an optimum threshold, among holding strategies of this type, but it should be noted that such policies impose two additional constraints (assumptions) beyond those required for derivation of the Single Headway policy. Since the Single Headway policy is less constrained, it will yield at least as good a solution as any threshold-type policy, and in general will be superior. The Single Headway policy, in turn, provides a lower bound on the effectiveness of the Prefol policy.

5.2.3 Potential Effectiveness of Headway Controls

In the last section, two near-optimal headway control strategies were derived: the Prefol and Single Headway policies. In this section, we focus on the magnitude of the improvement in service quality that can be expected if the strategies are implemented on a given transit line. The benefits depend on the current unreliability of the route, measured by the coefficient of variation in headways; the degree of statistical correlation between successive headways; and the proportion of passengers on board at the control point.

5.2.3.1 Improvement in Average Passenger Delay With the Prefol Strategy

As defined earlier, the expected reduction in overall delay to all passengers from holding vehicle i X_i minutes is:

$$E(\Delta\psi) = \frac{1}{E(H)\cdot Q}\ E(\xi_i) \tag{5-8}$$

where: $$E(\xi_i) = (1 - b)\ QE[X_i(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1} - X_i)]\ . \tag{5-17}$$

Recall that ξ_i is the net savings in total passenger-minutes achieved by holding the i^{th} bus X_i minutes. It is the reduction in total wait, minus the increase in total delay, from holding vehicle i.

We want to find $E(\Delta\psi)$ for the Prefol policy of $X_i = X_i^p$, where X_i^p is given by:

$$X_i^p = \begin{cases} \max[0,\ .5(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}^p)], & i = 1,2,\ldots,n-1 \\ 0 & , \quad i = 0,n. \end{cases} \tag{5-14}$$

The assumption that we don't hold the last bus, $X_n^p = 0$, is a modeling convenience to ensure that $E(H_i^p)$ is exactly equal to $E(H_i)$. Let us denote $E(\Delta\psi)$ for the Prefol policy as $E(\Delta\psi^p)$.

In order not to obscure the major points, we give only the result here. The details of the derivation of $E(\Delta\psi^p)$ are given in Appendix A for the interested reader. The basic result is as follows:

$$\frac{E(\Delta\psi^p)}{E(H)} = \frac{b^2}{4(1-b)} [(1 + C_p^2)\ \Phi(C_p) + \frac{C_p}{\sqrt{2\pi}} \exp(-1/2C_p^2)] \ . \tag{5-18}$$

where: $$C_p = -\frac{1-b}{b} \sqrt{2(1-\rho)}\ \frac{\sqrt{V(H)}}{E(H)} \tag{5-19}$$

$\Phi(C_p)$ = probability that a normal random variable with coefficient of variation C_p is greater than zero

ρ = correlation coefficient between successive headways.

The quantity $\sqrt{V(H)}/E(H)$ in equation (5-19) is simply the coefficient of variation in the headway distribution. Thus, we have the result that the relative benefits of holding depend on three factors: 1) the coefficient of variation of headways, 2) the correlation coefficient between successive headways, and 3) the proportion of total passengers who must ride through the control point.

Control of headways will make the greatest reduction in total delay when headways alternate (i.e., short, long, short, long, etc.). This happens on routes where vehicles are influenced substantially by the operation of the vehicle in front of them. For example, this would tend to be the case where loading delays are relatively more important than traffic congestion in determining overall vehicle operating speed. Routes in which pairing or bunching is prevalent would be of this type. In such a situation, holding a vehicle to lengthen a short headway also serves to reduce the long one which follows. Thus, the variance of headways is reduced by a greater amount for a given delay to the held vehicle than if a short headway might be followed by another short headway.

The extreme case is when the observed sequence of headways alternates between two discrete values. In this case, the sum of any two consecutive headways is a constant. That is, if one headway is two minutes too short, the next one must be two minutes too long. By the same argument, if the second headway is two minutes too long, the third must be two minutes too short, etc. In a statistical sense, successive headways are perfectly correlated ($\rho = -1$), so that knowledge of one headway implies knowledge of the entire set. It is for this case that headway control will have maximum benefits.

The opposite case is one in which headways between successive vehicles are statistically independent ($\rho = 0$). This means that knowing a given headway is short gives us no additional information about the probable values for the next headway. Such a situation would arise, for example, when traffic conditions have a much greater effect on vehicle operations than does the loading time at stops. In this case control will be less effective, because we have no guarantee that by lengthening a short headway we are also reducing a long headway. We might be simply reducing another, already short, headway. This case of independent headways thus provides a lower bound on the effectiveness of control strategies.

Figure 5-2 illustrates sets of values of the headway coefficient of variation, headway correlation, and proportion of passengers delayed for which Prefol control could reduce average passenger delay by at least 10%. By analyzing the two extreme cases of independent headways and perfectly correlated headways in detail, we can bound the regions of effectiveness for a class of headway control strategies, as shown in Figure 5-2. For situations in which control produces benefits under the least favorable circumstances ($\rho = 0$), we can be fairly confident that it will be beneficial. On the other hand, there are situations in which it does not appear to be desirable to control under the best of circumstances ($\rho = -1$); hence, control in these situations is unlikely to be useful. There remains a middle region in which control would probably produce benefits on routes where vehicles are substantially influenced by the vehicles in front of them, but not on routes where vehicles move relatively independently of one another. For situations in this region, more detailed and specialized analysis is required.

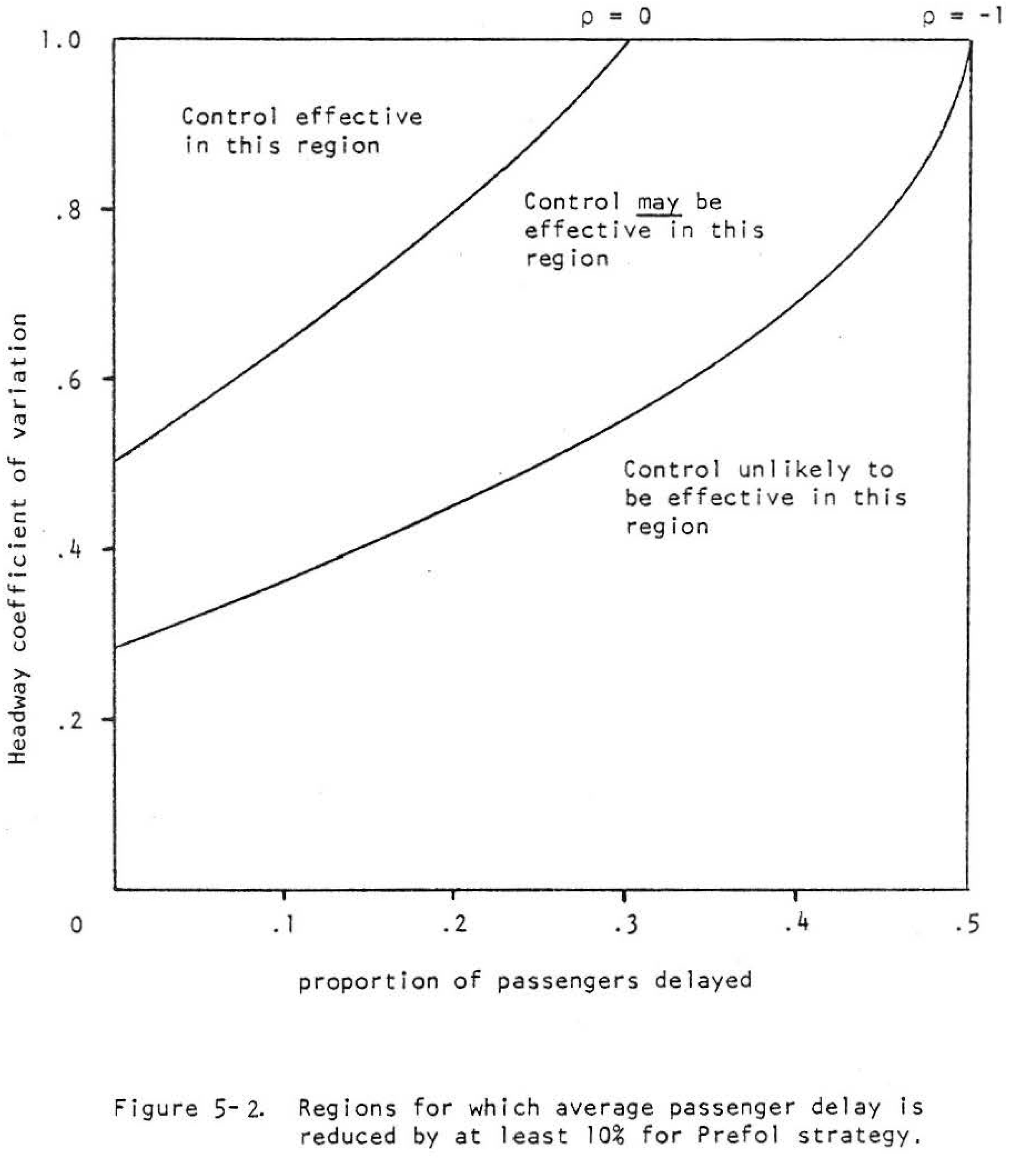


Figure 5-2. Regions for which average passenger delay is reduced by at least 10% for Prefol strategy.

As an example of using Figure 5-2, if headways are independent ($\rho = 0$), and $b = .2$, the coefficient of variation in the headway distribution would have to be at least .8 before Prefol control would yield a 10% reduction in average passenger delay. However, if headways are negatively correlated, control is more effective, and a 10% reduction in delay can be achieved at lower values of the coefficient of variation. At the extreme, when successive headways are perfectly correlated ($\rho = -1$), a 10% reduction in delay could be achieved with a coefficient of variation as low as .45 when $b = .2$.

A major implication of the result shown in Figure 5-2 is that it is wise to control a route at a point where there are relatively few people on the vehicle and relatively many waiting to board at subsequent stops, in order that the value of b be small. Generally, this means that the control point should be located as early along the vehicle's route as possible. However, it is also generally recognized that reliability problems worsen as one proceeds along a route. If dispatching at the route origin is effective, the headways will be reasonably regular at the early stops along the route, implying that the coefficient of variation will be small. At stops further along the route, however, the coefficient of variation in headways will tend to be larger. Thus, the decision of whether or not to implement a control strategy is tied to identification of a logical control point along the route.

Each stop along a route will have a particular headway distribution (with implied coefficient of variation) and value of b associated with it. Thus, each stop could be plotted as a point in the space defined by these two variables, as shown in Figure 5-3. Then, by looking at the "trajectory" of the route relative to the boundary values, the transit operator can make a decision about whether or not to control the route and, if so, where. For example, for the route illustrated by Figure 5-3, Prefol control at stop 3 might be worthwhile, but at stop 8 it is unlikely to be beneficial.

5.2.3.2 Improvement in Average Passenger Delay for the Single Headway Strategy

In the same way as we have analyzed potential effectiveness of the Prefol strategy, we can examine the effectiveness of the Single Headway strategy. Because the mathematics of the derivation are very similar to those for the Prefol strategy, we will omit the details, and simply present the result:

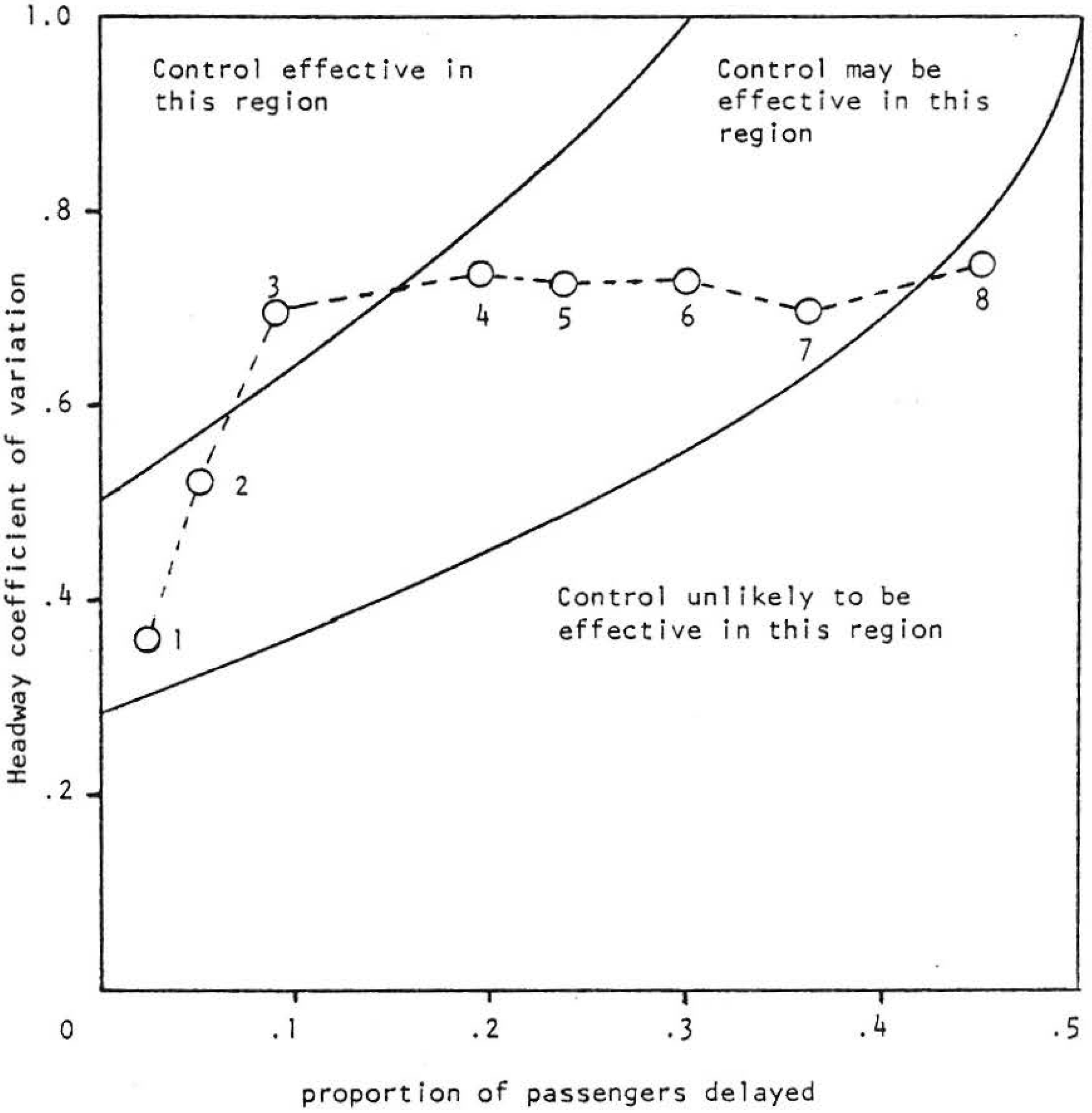


Figure 5-3. "Trajectory" of a bus route, showing headway coefficient of variation and proportion of passengers riding through each stop.

$$\frac{E(\Delta\psi^s)}{E(H)} = \frac{b^2}{4(1-b)} [(1+C_s^2)\ \Phi(C_s) + \frac{C_s}{\sqrt{2\pi}} \exp(-1/2C_s^2)] \qquad (5\text{-}20)$$

where:

$$C_s = -\frac{(1-b)(1-\rho)}{b} \frac{\sqrt{V(H)}}{E(H)} \ . \qquad (5\text{-}21)$$

Figure 5-4 shows the boundaries of the 10% benefit regions for the Single Headway strategy, along with those for the Prefol strategy. Note that the region for which the Single Headway strategy produces definite benefits is much smaller than that for the Prefol strategy. In general, the Single Headway strategy is less effective than the Prefol strategy, because it uses no direct information about the following headway. However, the difference between the strategies diminishes as $\rho \to -1$. As the correlation between successive headways becomes stronger, the predictability of the following headway is increasing. Thus, we need less actual information about it, because the current headway tells us more and more about what the following headway will be. In the limit, with $\rho = -1$, the strategies become identical.

5.3 Application of the Strategies

Implementation of vehicle holding strategies on any given route requires a sequence of steps:

1) collection of data on mean and variance of observed headways at various stops along the route, and passenger boarding/alighting counts by stop;

2) determination of the "trajectory" of the route in the manner illustrated in Figure 5-3;

3) preliminary evaluation of whether any control is desirable; if so, proceed to step 4;

4) selection of control (Prefol or Single Headway) and location at which control will be exercised;

5) test period with manual implementation and data collection for thorough evaluation;

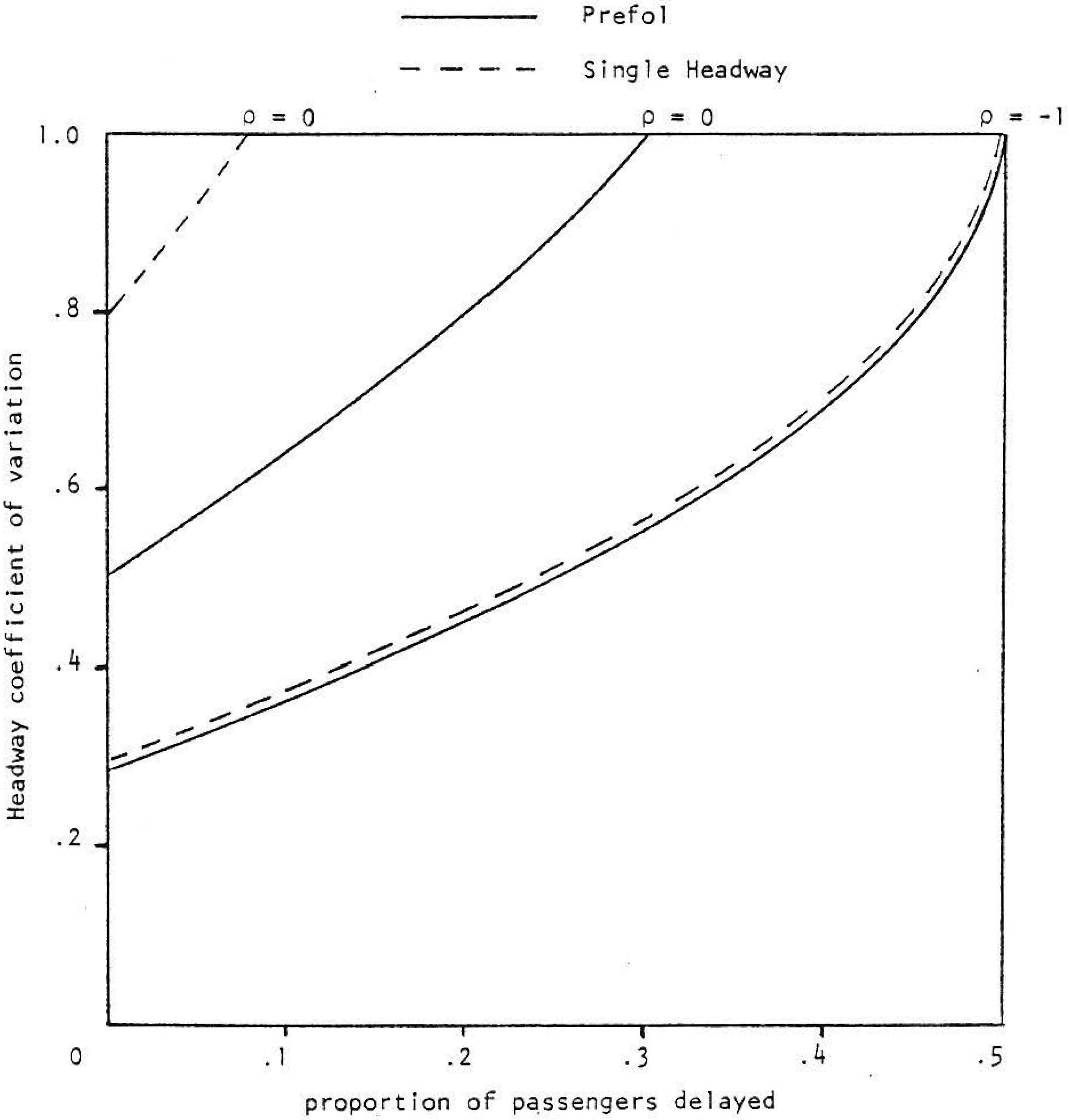


Figure 5-4. Comparison of regions of effectiveness for Prefol and Single Headway strategies.

6) acquisition of monitoring hardware (if any required) and assignment of personnel;

7) full scale implementation.

It should be emphasized that in many cases, data on correlations between successive headways will <u>not</u> be required. While the overall effectiveness of the Prefol headway control policy will depend to some extent on the degree of correlation present, the value of the correlation coefficient is not required for calculation of the optimal strategy. For the Single Headway policy, knowledge of the autocorrelation coefficient improves the prediction of the following headway, and thus is of value. In general, analysis of autocorrelation in the headway sequences is most important if the situation is a marginal one in which control could prove beneficial if headways are highly correlated, but would not be otherwise. However, such marginal situations are not prime candidates for control. It is advisable to begin such a program in a situation in which we are relatively confident of seeing significant results.

As an example of the analysis, let us consider the portion of the Cincinnati route network used for model validation in Chapter 3. The route layout with potential control points, is shown in Figure 5-5. Figure 5-6 shows the "trajectories" for each of the three inbound (toward Downtown) routes for the morning peak period. Since these trajectories lie well outside the boundary of the region for which headway-based holding is attractive, these routes do not appear to be good candidates for implementation of such strategies.

This is due to two principal reasons. First, the headways on these routes are relatively long, averaging about 15 minutes. This means that the standard deviation of the headway distribution would have to be quite large in order to make the coefficient of variation large enough to indicate that headway-based holding would be attractive. Furthermore, with such long headways, it is doubtful that passenger arrivals are random, as assumed in the model, so headway-based holding strategies are likely to be less effective than schedule-based holding.

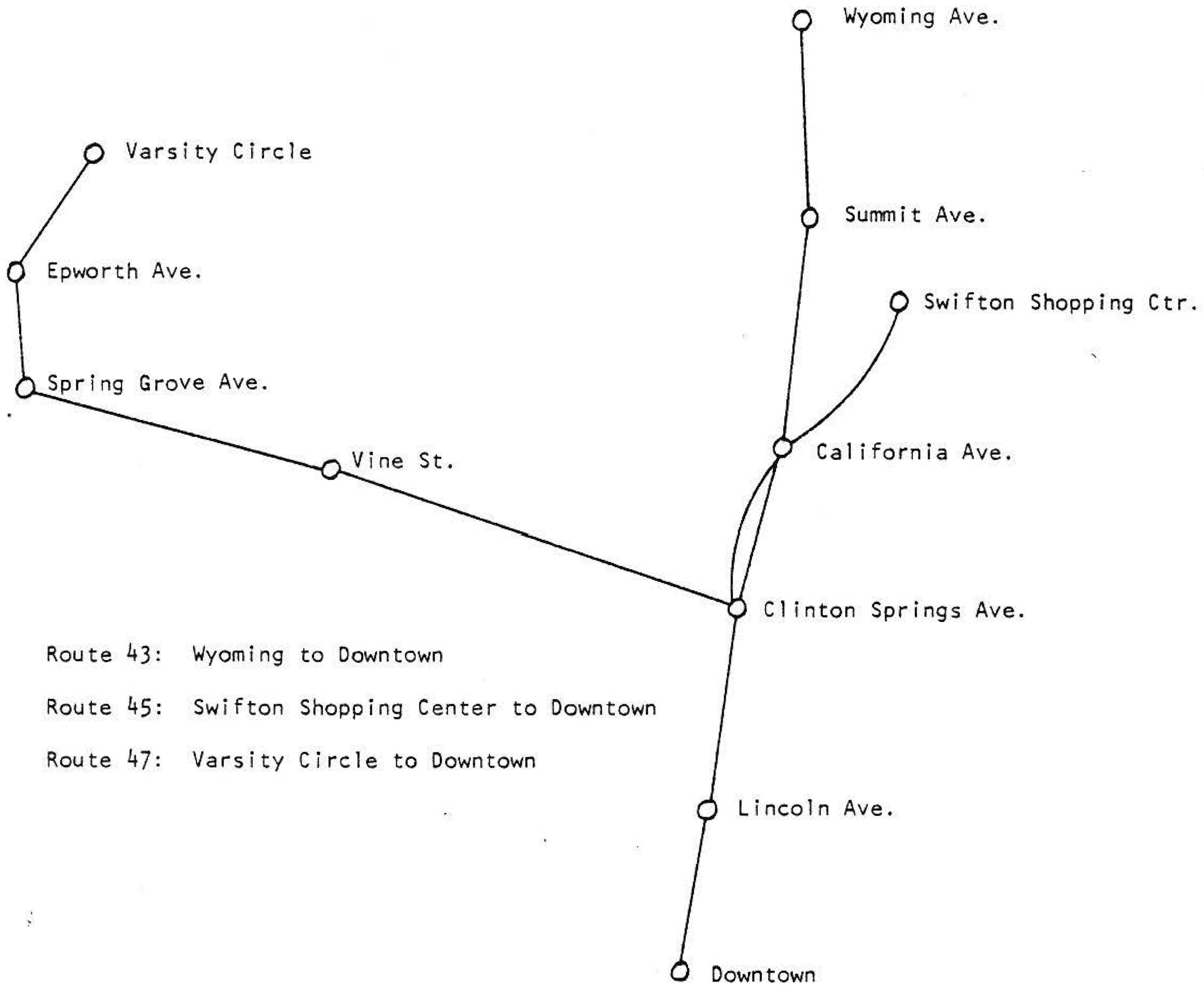


Figure 5-5. Route network from Cincinnati.

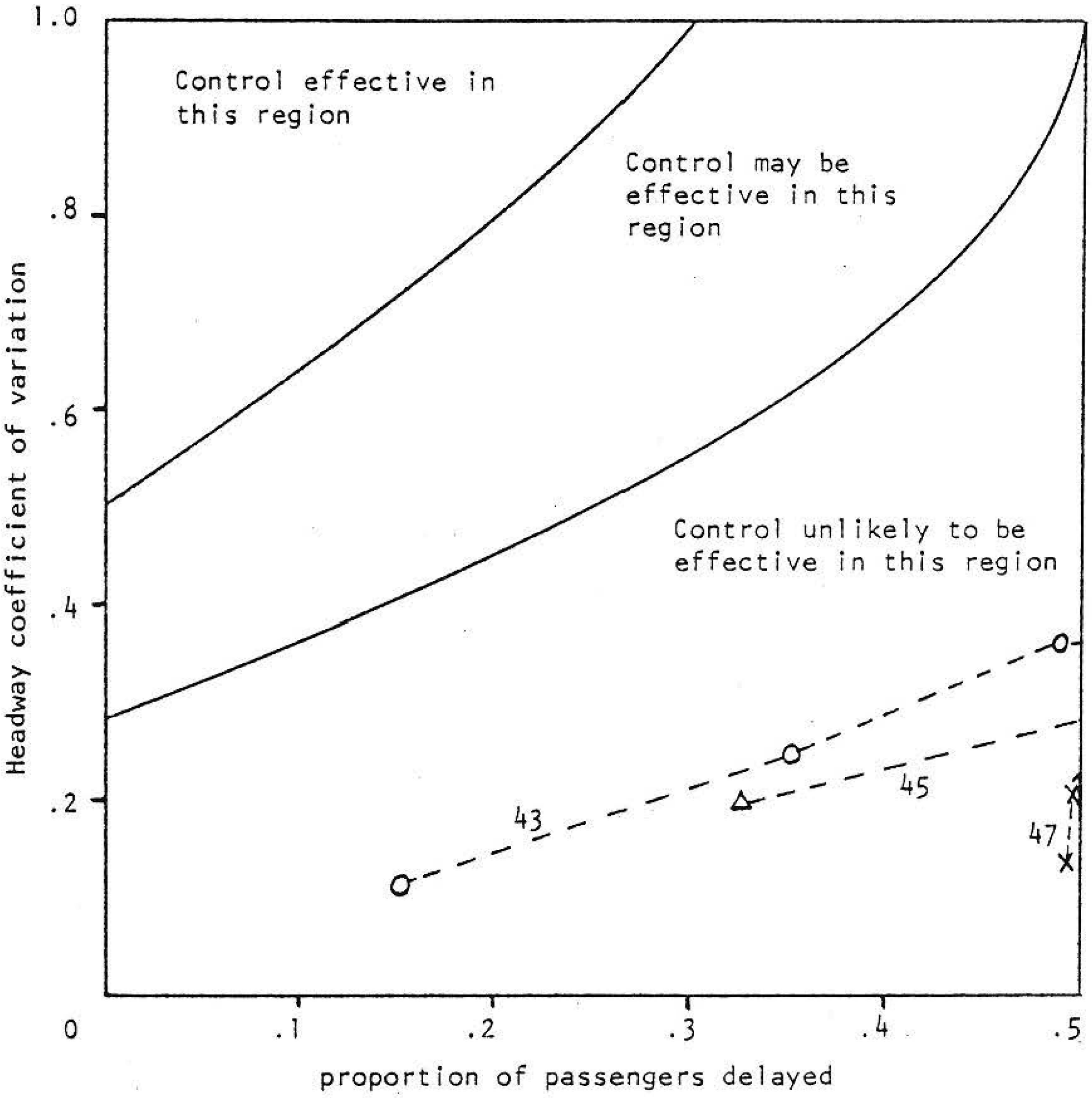


Figure 5-6. "Trajectories" of three inbound routes in Reading Road corridor.

The second reason for the unattractiveness of headway-based holding in this instance is that b tends to be relatively large for most stops. On each of these routes in the morning peak, a relatively large number of passengers board each bus near the beginning of the route and ride all the way downtown. Thus, the proportion of passengers delayed by a holding strategy tends to be relatively large, indicating that such strategies are unlikely to be beneficial in this instance.

5.4 Simulation of Schedule-Based Holding

Since the schedule-based strategy appears more appropriate for this situation, two variations of this strategy were simulated using the simulation model described in Chapter 2. In one case, holding was implemented for Route 43 only, at the Summit stop, early in the route. In the second case, holding was implemented at the Clinton Springs stop, and applied to all three inbound routes.

Each test comprised five simulation runs, with each individual run covering a 3-hour morning peak period (6-9 A.M.). The results of the five runs for each strategy were then averaged, and compared to a like average for a base case in which no holding was done.

The reason for conducting five replications in each test is that the simulation model is stochastic. Therefore, the result from any experiment is a sample statistic describing the performance of the system in that experiment. However, this result must be viewed as a random variable with a distribution, and we must be concerned about the statistical confidence of conclusions reached on the basis of these results. This requires that we be able to estimate the variance of the mean prediction value for any particular measure.

Tests performed on the results of earlier model runs indicated substantial variability in several measures of interest from the model. For example, with respect to mean deviation from schedule in the bus arrival times, for the Cincinnati network the average value is approximately 75 seconds. Multiple runs were made to estimate the standard deviation in this prediction. The resulting value was approximately 30 seconds. Thus, to reduce the coefficient of variation (standard deviation/mean) in a mean prediction of this value to .1 would require at least 16 replications. This number could be reduced somewhat by use of variance reduction techniques in the experimental design (see Fishman, 1973), but it would probably be prudent to plan on 10 replications for each experiment. Alternatively, we could accept a somewhat higher level of variability in the results, and require fewer replications. Accepting a coefficient of variation of .15, for example, could reduce the required replications to 7. With variance reduction techniques, this might be reduced to about 5. The increase in the coefficient of variation from .1 to .15 seems to be an acceptable price to pay for cutting the required replications in half. Thus, the experimental layout is based on 5 replications of each configuration.

The results of each test were similar, and indicated no significant benefits from schedule-based holding strategies for these routes. This is due in part to the relatively large number of passengers delayed, as discussed above, and in part to the fact that vehicles were seldom early, and hence holding was infrequent. With a modified schedule, the results might be somewhat better, but it is likely that alternate strategies are much more attractive for this particular network. These alternatives are discussed in subsequent Chapters.

These tests should not be construed to indicate that vehicle holding strategies are never attractive. A set of tests has been performed on a particular real network to provide a sense of realism to the experiments with alternative strategies. It is to be expected that some strategies will work well in a given situation, while others will not. An example of a route for which holding strategies would be very effective could be constructed quite easily.

5.5 Summary

The purposes of the analysis methods described in this chapter are first, to provide tools with which routes can be evaluated for potential implementation of holding strategies, and second, to describe the nature of the holding decision rules, should such strategies appear attractive. Two major subgroups of strategies have been discussed: schedule-based holding and headway-based holding.

The schedule-based "checkpoint" strategy is very simple to implement and may be useful on long-headway routes where the schedule is sufficiently slack so as to make holding to schedule a reasonable procedure. The key elements of implementing such a policy are constructing a reasonable schedule as a goal and enforcing adherence to that schedule. This enforcement requires both proper incentives for drivers and a mechanism for accurate monitoring of their performance.

For routes operating with shorter headways, two near-optimal headway-based control strategies have been developed. One strategy holds a vehicle until its preceding headway is as close as possible to its following headway, allowing for an adjustment in consideration of the people delayed on the vehicle. Referred to as "Prefol", it requires a prediction of the arrival time of the following vehicle. A similar strategy that is dependent only on the known magnitude of the current headway, called the "Single Headway" strategy, is also proposed. The strategies are simple in form, require limited data about the route, and are near-optimal over a wide range of situations.

Models of the effectiveness of the strategies indicate that they are sensitive to three important characteristics of a control point: (1) the current level of unreliability, as measured by the headway coefficient of variation; (2) the relationship between successive headways, measured by the correlation coefficient; and (3) the proportion of passengers who must ride through the control point.

The Single Headway strategy performs less well than the Prefol strategy when vehicles arrive relatively independent of each other. As passenger loading delays increase and successive headways become more dependent on each other, the Single Headway strategy prediction capability improves and it approaches the performance of the Prefol strategy. In any event, the Single Headway strategy requires less real-time information about the system, and could certainly be implemented without expensive AVM equipment.

However, it should be noted that such equipment can be very useful in the collection of the data required to evaluate whether or not headway-based control is likely to be effective, and in the design of the strategy if one is selected for implementation. The value of AVM equipment strictly for continual collection of data regarding system performance should not be underestimated.

CHAPTER 6
STOP REDUCTION AND ZONE SCHEDULING

Since a substantial portion of bus travel time is spent decelerating for stops, standing to allow boarding and alighting of passengers, waiting to re-enter the traffic stream and accelerating, reducing the number of stops made by each individual vehicle is one way of improving travel time. In addition, since the variability of stop dwell time is a major source of deviation from schedule, reducing the number of stops should improve reliability. This project has examined two different ways in which to accomplish a reduction in the number of stops each bus makes. The first is increasing stop spacing by eliminating some stops along a route, and the second is zone scheduling. These are discussed individually in the following sections.

6.1 Increased Stop Spacing

Increasing the spacing between stops is clearly one way to reduce the number of stops which must be made by each vehicle. The major disadvantage of increased stop spacing is that accessibility to the route is diminished. Passengers must walk further, on average, to get to a bus stop. This cost must be weighed against the improved travel time and reliability, in order to arrive at optimal stop spacing decisions.

Very little work has been done in this regard for bus operations. Vuchic (1969b) and Vuchic and Newell (1968) have considered such problems for rapid transit lines, but reliability improvements were not among their measures of performance. Mohring (1972) discusses optimal stop spacing for urban bus routes, but in the context of a very simple model, and with no attention to reliability of service.

Additional analysis is clearly necessary on this issue, since an important trade-off appears to exist between travel time and reliability on one hand, and accessibility on the other. This project has taken initial steps in this direction, by examining the degree of travel time and reliability improvements which might be attainable through an increase in stop spacing. We have not, however, addressed the demand-related issues regarding the relative importance of travel time, reliability and accessibility in traveler decision-making. This question, and the resulting implications for optimal service design, require further study.

Investigation of the potential service improvements which could result from increased stop spacing has included both analytic modeling and simulation tests. The analytic models demonstrate the general sensitivity of travel time and reliability to stop spacing, and the simulation experiments provide more detailed analysis of specific changes in the "Reading Road" network from Cincinnati.

6.1.1 Analytic Models of the Effects of Stop Spacing

Stop spacing has impacts on the mean and variance of both wait time and in-vehicle time for passengers. The basic reason for its effect on average in-vehicle time is obvious — more stops along the route increase the average time to traverse a given distance. Stop spacing also affects the variance of in-vehicle time, because the time required for each stop is variable; hence the total time a passenger spends on the vehicle becomes more variable as the number of stops increases.

The effects of stop spacing on wait time are more subtle. The more stops there are along a route, the higher the variance in headways between successive vehicles tends to be. If service is frequent enough so that we may assume random passenger arrivals, the impact of headway variance on expected wait time is indicated by equation (5-1), which is rewritten here.

$$E(W) = \frac{E(H)}{2} + \frac{V(H)}{2E(H)} \tag{5-1}$$

where:
$E(W)$ = expected wait time
$E(H)$ = expected headway between successive vehicles
$V(H)$ = variance of headways.

The variance in wait time at a stop has been derived by Friedman (1976):

$$V(W) = \frac{E(H^3)}{3E(H)} - [E(W)]^2 . \tag{6-1}$$

If the distribution of headways is symmetric, the third moment of the headway distribution, $E(H^3)$, is given by:

$$E(H^3) = 3E(H)V(H) + [E(H)]^3 \tag{6-2}$$

Since headways may be expressed as the difference in two identically distributed random variables (bus arrival times at a stop), the assumption of symmetry in the headway distribution is theoretically reasonable. Furthermore, empirical evidence (Sterman, 1974) tends to confirm this.

We see from (6-1) and (6-2) that the variance of wait time is also directly affected by the headway variance. Thus, the stop spacing on a route affects at least four important service quality measures: mean and variance of wait time, and mean and variance of in-vehicle time. In order to describe these effects more precisely, we need to characterize the distributions of in-motion times on links and dwell times at stops for buses moving along a route. These components can then be put together to describe the overall effects of stop spacing.

6.1.1.1 Vehicle and In-Motion and Dwell Times

A useful statistical model of vehicle in-motion times on links is the shifted gamma distribution described briefly in Section 2.3 of this report, and in greater detail by Turnquist and Bowman (1979). Using this model, the mean and variance of bus in-motion times are given by equation (2-1) and (2-2), reproduced here for convenience:

$$E(\gamma_j) = L_j(k_i z_i + 1/S_i) \qquad (2\text{-}1)$$

$$V(\gamma_j) = L_j \quad k_i(z_i)^2 \qquad (2\text{-}2)$$

where:

γ_j = in-motion time of bus on link j

L_j = length of link j

S_i = speed limit on links of type i

k_i, z_i = gamma distribution parameters for links of type i.

Dwell times at stops are a function of the numbers of boarding and alighting passengers served, and the service time per passenger. Basic models for dwell time in a variety of situations are given by Boardman and Kraft (1970). A more recent study, focusing on boarding passengers only, is reported in Jordan and Turnquist (1979). Let us consider a morning peak period on a radial route terminating in the central business district (CBD), for purposes of example. Most of the passengers are likely to be destined for the route terminus in the CBD. Because the number of alighting passengers at any non-CBD stop is likely to be small, boarding times will tend to determine bus dwell time. In this case, a simple, but relatively accurate, model of dwell times is:

$$Y = \beta_0 + \beta_1 N + e \qquad (6\text{-}3)$$

where: Y = dwell time

N = number of passengers boarding

β_0, β_1 = constants

e = random error term with zero mean and variance σ^2.

Models of the form shown in (6-3) have been calibrated by both Boardman and Kraft (1970) and Jordan and Turnquist (1979).

If β_0 and β_1 are assumed constant, the mean and variance of dwell time are given by:

$$E(Y) = \beta_0 + \beta_1 E(N) \qquad (6\text{-}4)$$

$$V(Y) = \beta_1^2 V(N) + \sigma^2 \qquad (6\text{-}5)$$

where: $E(N)$ = expected number of passengers boarding

$V(N)$ = variance of number of passengers boarding.

Under the assumption that passenger arrivals are Poisson, $E(N)$ and $V(N)$ are, respectively:

$$E(N) = P\,E(H) \qquad (6\text{-}6)$$

$$V(N) = P\,E(H) + P^2\,V(H) \qquad (6\text{-}7)$$

where: P = average passenger arrival rate

$E(H)$ = expected headway between successive vehicles

$V(H)$ = variance of headways.

Note that the largest component in the variance of dwell time is not the variance in boarding time for a given number of passengers (σ^2), but is due to the variance in the actual number of passengers boarding ($V(N)$).

Also note that the mean and variance of the headway distribution play a central role in determining the variance of dwell time at stops. This, in turn, reflects itself in the variance of wait and in-vehicle time. Thus, before proceeding further, it is important to develop equations to characterize the headway distribution.

6.1.1.2 The Distribution of Headways

For our purposes, it will be sufficient to characterize the headway distribution at each stop by its mean and variance. Since we assume that buses are dispatched regularly, and expected travel and dwell times are the same for each bus, the expected value of headway at stop i is simply $E(H)$, for all stops.

The variance of the headway distribution is somewhat more complicated. We begin by expressing the headway at stop i between buses j and j-1 as:

$$H_i = H_{i+1} + (Y^j_{i+1} - Y^{j-1}_{i+1}) + (\gamma^j_i - \gamma^{j-1}_i) \qquad (6\text{-}8)$$

where: Y^j_i = dwell time of bus j at stop i

γ^j_i = in-motion time of bus j from stop i+1 to stop i.

Our notation of stop numbers implies that the terminus of the route is stop 1, with numbers increasing as we progress toward the beginning of the route. Thus, a bus begins its run at stop n, and proceeds to n-1, n-2, etc.

If we assume that in-motion times of successive buses over a given link are independent random variables, and are independent of dwell times, we can write the variance of H_i, using the result from equation (6-5):

$$\begin{aligned} V(H_i) = {} & V(H_{i+1}) + \beta_1^2\, V(N^j_{i+1} - N^{j-1}_{i+1}) \\ & + 2\beta_1 \operatorname{cov}[H_{i+1}, (N^j_{i+1} - N^{j-1}_{i+1})] + 2\sigma^2 \\ & + V(\gamma^j_i) + V(\gamma^{j-1}_i). \end{aligned} \qquad (6\text{-}9)$$

where: β_1 = boarding time per passenger

N_i^j = number of passengers boarding bus j at stop i

σ^2 = variance of dwell time, given number of boarding passengers.

Evaluation of equation (6-9) requires that we have expressions for $\mathrm{cov}[H_{i+1}, (N_{i+1}^j - N_{i+1}^{j-1})]$ and $\mathrm{cov}(N_{i+1}^j, N_{i+1}^{j-1})$. These covariance terms are not zero, because of the "pairing" or "bunching" phenomenon. However, if we assume that the sums of two consecutive headways are essentially constant (that is, there is no long term delay propagation along the route), we can derive expressions for these covariance terms. The algebra is quite tedious, but produces the following results:

$$\mathrm{cov}[H_{i+1}, (N_{i+1}^j - N_{i+1}^{j-1})] = 2P_{i+1}V(H_{i+1}) \tag{6-10}$$

$$\mathrm{cov}(N_{i+1}^j, N_{i+1}^{j-1}) = P_{i+1}^2 V(H_{i+1}) \tag{6-11}$$

where: P_{i+1} = passenger arrival rate at stop i+1.

We may substitute (6-10) and (6-11) into (6-9), along with the variance of in-motion time given by (2-2) and the variance of the number of boarding passengers from (6-7). This produces the following result:

$$V(H_i) = V(H_{i+1})(1 + 4\beta_1 P_{i+1})^2 + 2[P_{i+1} E(H) \beta_1^2 + \sigma^2] + 2k L_i z^2 \tag{6-12}$$

Note that we have denoted link i (with length L_i) as connecting stops i and i+1.

Equation (6-12) gives the variance in headway between consecutively dispatched buses j and j+1 at stop i. However, if passing had occurred these buses may not arrive successively or in the dispatched order at stop i. In this case, the headway from bus j to bus j+1 will not be relevant to passenger wait time. Passing is to be expected when the standard deviation of headways approaches E(H) (i.e., when $V(H_i) = [E(H)]^2$). At this point, bus arrivals at stop i are Poisson. Empirical observations by Holroyd and Scaggs (1966) indicate that $[E(H)]^2$ is a probable upper bound on effective headway variance. Thus, the variance used in the model is as follows:

$$V'(H_i) = \min\{V(H_i, [E(H)]^2\}. \tag{6-13}$$

The result in equation (6-13) allows us to evaluate the variance of dwell times, and resultant effects on variance of in-vehicle time. It also allows evaluation of the mean and variance of wait times. We will first use this result to discuss the effects of stop spacing on in-vehicle time.

6.1.1.3 Mean and Variance of In-Vehicle Time

Travel time over a route is the sum of in-motion and dwell time. If a route (or the portion of a route over which a given traveler rides) has n links, with n-1 intermediate stops, we can use the results of the previous sections to write in-vehicle time as follows:

$$T_n = \sum_{i=1}^{n} \gamma_i + (n-1)\beta_0 + \sum_{i=1}^{n-1} (\beta_1 N_i + e_i) . \tag{6-14}$$

The expected value of T_n is then given by:

$$E(T_n) = (kz + \frac{1}{S})D + (n-1)\beta_0 + \beta_1 E(H) \sum_{i=1}^{n-1} P_i \tag{6-15}$$

where: D = length of the route (or a travelers trip).

Note that expected in-vehicle time, as given by (6-15), is linearly related to both distance traveled (D) and the number of stops (n). For a given n, average time is a linear function of distance, of the form:

$$E(T_n) = t_0 + r\ D \tag{6-16}$$

where: $t_0 = (n-1)\beta_0 + \beta_1 E(H) \sum_{i=1}^{n-1} P_i$

$r = kz + \frac{1}{S}$.

Rewriting (6-16) in the form of average speed, we have

$$\text{average speed} = \frac{D}{E(T_n)} = \frac{D}{t_0 + r\ D} \quad . \tag{6-17}$$

Average speed, for given n, clearly increases as D increases, reflecting the fact that the same number of stops are spread over a longer distance, or stop spacing has increased. An alternative way of seeing the same result is to consider the effect of changing n and keeping D fixed. Increasing n (reducing stop spacing) will increase t_0 (since $\beta_0 > 0$), and hence will decrease average speed.

The effects of stop spacing on the variance of in-vehicle time can be determined using equation (6-14), and substituting results from equations (2-2) and (6-7). The variance of T_n is:

$$V(T_n) = \sum_{i=1}^{n} V(\gamma_i) + \beta_1^2 \sum_{i=1}^{n-1} V(N_i) + (n-1)\sigma^2$$

$$+ 2\beta_1^2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \operatorname{cov}(N_i, N_j) + 2\beta_1 \sum_{i=1}^{n-1} \operatorname{cov}(N_i, \sum_{j=i}^{n} \gamma_j). \qquad (6\text{-}18)$$

The variances of γ_i and N_i are obtained from equations (2-2) and (6-7) respectively. Derivation of expressions for the covariance terms is tedious, but relatively straightforward. The details are omitted here, but the interested reader may refer to Jordan (1979) for the complete deviation. The resulting expression for $V(T_n)$ is:

$$V(T_n) = D\ kz^2 + \beta_1^2 \sum_{i=1}^{n-1} P_i^2\{V(H_i) + [E(H)]^2\}$$

$$+ 2\beta_1^2 \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} P_i P_j [1 + K(i,j)] V(H_j)$$

$$+ 2\beta_1\ kz^2 \sum_{i=1}^{n-1} P_i (\sum_{j=i}^{n} L_j) \qquad (6\text{-}19)$$

where: $$K(i,j) = \sum_{s=i+1}^{j} \sum_{r=s}^{j} [(2\beta_1)^{r-s+1} \prod_{q=s}^{r} P_q] . \qquad (6\text{-}20)$$

While equation (6-19) is complicated, evaluation for any particular value of n is straightforward. This expression allows us to trace the effects of stop spacing on the variance of in-vehicle time.

6.1.1.4 Mean and Variance of Wait Time

As indicated in equations (5-1), (6-1) and (6-2), the mean and variance of wait time depend very directly on the variance of the headway distribution. Thus, to trace the effects of stop spacing on wait time, we need to examine first the impacts on headway variance, as given by equations (6-12) and (6-13).

Note that equation (6-12) is of the form:

$$V(H_i) = \alpha V(H_{i+1}) + \nu \qquad (6\text{-}21)$$

where $\alpha > 1$ and $\nu > 0$. Thus, the headway variance grows monotonically as we proceed stop-by-stop along the route, until we reach the bound indicated by equation (6-13).

As stop spacing increases, ν increases since it is proportional to link length. However, since the number of stops in a given length of route decreases, the overall rate of variance increase is smaller. Reductions in headway variance as a result of increased stop spacing translate directly into reductions in the mean and variance of waiting time. Thus, there is reason to believe that reducing the number of stops made by buses through increasing the stop spacing can improve service reliability.

6.1.2 Simulation Experiments With Increased Stop Spacing

To explore this possibility in greater detail for a specific situation, a series of simulation experiments were run. These experiments changed the stop density along a 7.1 km (4.4 mile) section of the Reading Road corridor in Cincinnati, from Clinton Springs Ave. to Downtown. Over this section, all three routes in the test network run together. The reader is referred to Figure 5-6 for a route diagram showing this section.

In the base case (reflecting existing operations), there are 36 stops in this section, for an average stop spacing of .20 km (.12 mi.). For the tests, 17 of these stops were eliminated, resulting in an average stop spacing of .37 km (.23 mi.). This section of the network was simulated microscopically, with each individual stop and intersection represented explicitly. As in the previous simulation tests, five replications of each configuration were run, and the averages over these replications compared.

The results show that average passenger speed over the system increased from 14.1 km/h (8.8 mph) to 14.5 km/h (9.0 mph). This change, while in the right direction, is not statistically significant at any reasonable level of confidence, however. The standard deviation of passenger speed was unchanged, at 5.3 km/h (3.3 mph).

Reducing stop density also appears to have made small reductions in both the mean and standard deviation (or variance) of waiting time. Mean waiting time was reduced from 7.5 min. to 7.2 min., and the standard deviation from 7.7 min. to 7.0 min. However, as in the case of average passenger speed, these changes are not statistically significant.

Thus, the simulation results with respect to reduced stop density are not particularly encouraging. However, deeper inspection of the simulation output showed that a major reason why eliminating stops had such small effects was that buses were still being slowed by traffic signals. Because of the signal settings, they could not take advantage of the potential reductions in travel time along the route — they simply spent more time in queues at traffic lights. In an attempt to rectify this, changes in both stop density and signal operation were made simultaneously. These results were more encouraging, and are discussed in greater detail in Chapter 7, along with the results of other changes involving signalization.

6.2 Zone Scheduling

An alternative way of reducing the number of stops each vehicle must make, without increasing overall stop spacing, is to divide a route into "zones." Each zone is a set of consecutive stops with a subset of all the buses on the route allocated to it. An inbound bus, dispatched from the outermost stop in its zone, makes stops to pick up or let off passengers within its zone only, running non-stop to the route terminus after passing the inner zonal boundary. On its outward journey, the bus may provide local service all along the route, may travel express to the innermost stop of its zone, where it would again begin to offer local service, or may travel express all the way to the outer terminus of its zone and then begin another inbound run. Figure 6-1 depicts a zone scheduling scheme. Zones must overlap so that passengers bound from a stop in one zone to a stop in a different zone, other than the route terminus, can transfer.

Some obvious preconditions must exist for zone scheduling to be attractive. The benefits from zone scheduling accrue largely to passengers traveling to, or from, the route terminus - the stop at which all buses serving the route end their inbound express runs. Local passengers - those not starting or ending their trips at the route terminus - are not as well served by zone scheduling because they will frequently be forced to transfer and cannot take advantage of the express service offered. Zone scheduling will be useful when a large majority of bus route passengers from all stops are bound for (or originate at) one stop. Such situations (referred to as many-to-one or one-to-many) often exist during peak periods in urban areas when workers are bound for the central business district (CBD) in the morning, or are exiting the CBD in the evening. A bus route operating during the morning peak period will be examined in this analysis. The analysis could pertain to evening peak period operations just as easily, however.

Zone scheduling can improve both average bus speeds and reliability in two ways:

1. Average in-motion time and variability can be reduced by the non-stop service offered for a portion of each bus's run under a zone scheduling scheme. Making this express run on limited access roads reduces intersection and congestion delays relative to those experienced by buses operating locally along the entire route.
2. The number of stops each bus makes can be reduced by introducing zone scheduling. Reductions in the number of stops will lessen both average bus dwell time and variability in this time.

In this section, a model of an urban bus route is developed to test the effectiveness of zone scheduling as a means of improving service reliability. The model includes equations to express a bus route's service reliability in terms of the in-motion and dwell time distributions, bus route characteristics

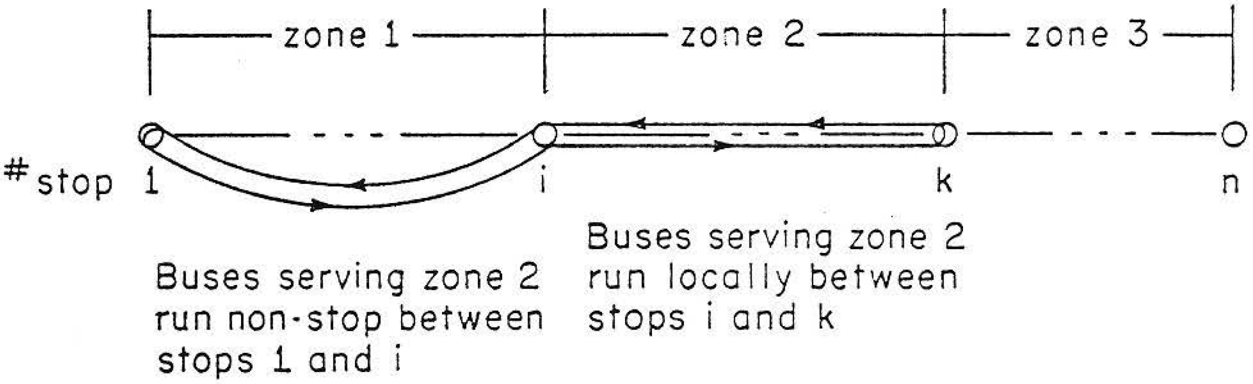


Figure 6-1. Structure of a zone scheduled route.

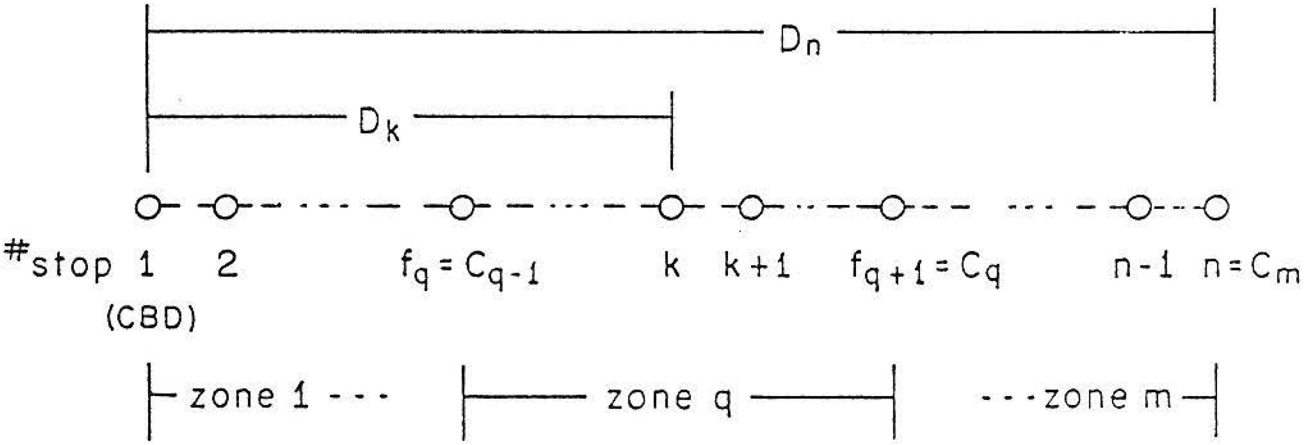


Figure 6-2. Bus route notation.

such as stop spacing and passenger arrival rates, and the zone structure imposed on the route (i.e., the manner in which the route is divided into zones and the allocation of buses among zones). A dynamic programming formulation is used to select the zone structure maximizing bus service reliability for a given route.

Hypothetically, service reliability can be improved at the expense of average level-of-service. For example, a zone scheduling strategy in which every stop is a separate zone could be very reliable, but average wait times would be excessive. Therefore, a measure of average level of service, expected wait plus travel time totaled over all passengers, is also computed. The tradeoff between this measure and the reliability measure can then be investigated.

6.2.1 Model Formulation

We assume that passengers arrive at each stop, i, as a Poisson process with rate P_i. The rate is assumed constant over the period of interest. Further, all passengers are assumed to be destined for the route terminus (e.g., the CBD), and board the first bus coming to their stop after they arrive. Finally, we assume that each bus stops at every stop in its local service area.

In reality, not all passengers will have a common destination. The presence of passengers with destinations other than the route terminus complicates the analysis, but they can be incorporated into the model. Since there are many routes whose ridership patterns are such that the vast majority of passengers have a common destination, the model is developed on that basis. Incorporation of local and interzonal passengers is discussed in Section 6.2.3.

Figure 6-2 shows a bus route labeled with the notation used in the model. The route has n stops and m zones numbered in ascending order from the inner route terminus. D_k is the distance from stop k to the route end. Although only CBD-bound passengers are being considered, neighboring zones have a one stop overlap to accommodate transfers so that the inclusion of non-CBD bound riders can be accomplished easily. That is, the innermost stop of zone q, f_q, for $q \neq 1$, overlaps with the outermost stop of zone q-1, C_{q-1}. It is assumed that all CBD-bound passengers arriving at this overlap stop board the bus from zone q, which immediately travels non-stop to the CBD.

Buses serving zone q are dispatched from stop C_q with a constant headway $\overline{h}_q$. It is assumed that buses are always dispatched on schedule. This dispatch interval is calculated as follows:

$$\overline{h}_q = RT_q/b_q \tag{6-22}$$

where: b_q = the number of buses serving zone q

RT_q = scheduled round trip time for buses in zone q.

RT_q must be sufficiently large so that the assumption that buses are dispatched on schedule will be violated with very small probability. Such a value for RT_q can be estimated by: 1) finding the in-motion time which would be exceeded only a small percentage of the time from travel time distributions; 2) adding to this a dwell time required to board and alight a number of passengers equal to the bus capacity (assuming few people on the bus get off at stops other than the CBD); and 3) adding a layover period for the bus driver to rest at the end of the run.

Several constraints are placed on the dispatch headway. First, a minimum level-of-service must be provided. This is represented by an upper bound on headway in each zone.

$$\overline{h}_q \leq H_{max} \quad \text{for all } q. \tag{6-23}$$

Second, headways must be adequate to meet zonal demand. Total capacity provided to zone q during the peak period is $\tau N/\overline{h}_q$ where τ is the length of the period modeled, and N is the capacity of a bus. Passenger arrivals are assumed Poisson so the mean and variance of total peak period zone q demand is $\tau \sum_{i \varepsilon q} P_i$. By approximating the cumulative Poisson distribution by a normal distribution, a capacity constraint can be imposed such that capacity should exceed demand with a given probability:

$$\tau N/\overline{h}_q \geq \tau \sum_{i=f_q}^{C_q} P_i + z_\alpha[\tau \sum_{i=f_q}^{C_q} P_i]^{.5} \quad \text{for all } q \tag{6-24}$$

where: z_α = value exceeded with probability α by a standard normal random variable.

Since demand is generally high during the morning peak period, the second constraint will usually dominate the first.

Another constraint affecting $\overline{h}_q$ results from the fixed number of buses, B, serving the route. This imposes a fleet size constraint:

$$\sum_{q=1}^{m} b_q = B. \tag{6-25}$$

Both mean trip time and reliability (measured by variance in trip time) are functions of route characteristics (such as stop spacing, average passenger arrival rates, etc.), parameters of the in-motion and dwell time distributions, and the manner in which the route is divided into zones. The variables of interest in determining zone structure are:

1) the number of zones into which the route is divided, m;
2) the stop number of the last stop in each zone, C_q, which determines the zone in which each stop is located; and
3) the number of buses serving each zone, b_q.

The zone structure, defined by these three variables, can be varied within the demand, capacity, and bus fleet size constraints on headway. Our objective is to find a zone structure which minimizes either average trip time or variance of trip time (or some combination of the two).

There are many potential zone structures for any bus route. Dynamic programming is an efficient way of searching through the combinations of m, C_q, and b_q (for q=1 to m) to determine the optimal one. The number of zones is the stage variable, and the last stop served and number of buses used are state variables.

The dynamic programming formulation can be summarized as follows. Define $F_r(C_r, B_r)$ to be the minimum total trip time (wait plus in-vehicle) for all passengers when the first C_r stops are divided into r zones and are served by B_r buses. B_r is defined as follows:

$$B_r = \sum_{q=1}^{r} b_q . \tag{6-26}$$

If we define wait time at stop i as W_i, and in-vehicle time from i to stop 1 as T_i, we can define total trip time, U_i, as their sum:

$$U_i = W_i + T_i , \tag{6-27}$$

with expectation:

$$E(U_i) = E(W_i) + E(T_i). \qquad (6\text{-}28)$$

Expected wait time will be dependent on both the number of buses allocated to each zone and the zone structure, since these together determine headways. Expected in-vehicle time will be dependent solely on zone structure.

For a single zone (r=1) we clearly have the result:

$$F_1(C_1, B_1) = \tau \sum_{i=1}^{C_1} P_i \, E(U_i | B_1) . \qquad (6\text{-}29)$$

For r>1, we can develop a recursion. If we have a total of C_r stops and B_r buses, we wish to find the first stop in zone r, denoted f_r, and the number of buses allocated to zone r, denoted b_r. Stops 1 through f_r-1 will be organized into r-1 zones and served by B_r-b_r buses, as best as possible. This leads to the following recursion:

$$F_r(C_r,B_r) = \min_{f_r,b_r} [F_{r-1}(f_r-1, B_r-b_r) + \tau \sum_{i=f_r}^{C_r} P_i \, E(U_i|b_r)]. \qquad (6\text{-}30)$$

Since in general, we might hypothesize that wait time is more onerous than in-vehicle time, we might choose to minimize a modified version of (6-30), using y_i defined as:

$$y_i = a \, W_i + T_i \ , \qquad (6\text{-}31)$$

where a is a weighting constant. This can be accomplished simply by replacing $E(U_i|b_r)$ with $E(y_i|b_r)$ in equation (6-30).

The problem of minimizing variance as a measure of reliability can also be addressed using a simple modification of equation (6-30). This involves replacing $E(U_i|b_r)$ with $Var(U_i|b_r)$, defined as follows:

$$Var(U_i) = Var(W_i) + Var(T_i) \ . \qquad (6\text{-}32)$$

In this case also, one might hypothesize that the criterion of interest is not so much variance of total trip time, but some weighted combination of the variances of wait time and travel time. This could be reflected by using $Var(y_i)$ instead of $Var(U_i)$.

This dynamic programming formulation provides a convenient framework for optimizing zone structures. It may be applied effectively using the representation of means and variances of both wait times and in-vehicle times developed in sections 6.1.1.3 and 6.1.1.4.

6.2.2 Model Application

The use of this model is illustrated by applying it to an idealized representation of the Sheridan Road bus route located in Chicago. Some details on route form must be sacrificed to make the model computationally feasible. For a long route, like the one studied, including all possible bus stops in the model would make an excessive number of computations necessary to determine the measures of reliability and level of service. Also, the model formulation considers every bus stop to be a potential zone boundary from which express service can be offered. However, only a few bus stops actually meet the requirement of being located at an access point to a limited access facility on which such an express run is made.

For these two reasons bus stops in the model are aggregations of neighboring real-world stops. The aggregate stops are located at access points to an expressway in the modeled route, shown in Figure 6-3. The data required to implement the model are shown in Tables 6-1 to 6-3.

Figure 6-4 shows the level of the reliability measure as a function of the number of buses serving the bus route. Using this same number of buses under the optimal zone scheduling strategy a 90% improvement in the reliability measure is achieved. However, as shown in Figure 6-4, reducing the number of buses serving the route results in only a slight increase in the measure relative to the all-local service value. That is, employing a zone scheduling strategy instead of all-local service can greatly improve reliability while simultaneously reducing the bus fleet requirements for the route. A parallel result for average trip time is illustrated in Figure 6-5, showing that vehicle productivity improvements from zone scheduling can simultaneously reduct fleet size and average travel time.

Figure 6-6 shows the optimal zone structures for different bus fleet sizes. The zone structure is very stable as the number of buses is decreased from 120 to 70, the only difference being in the allocation of buses among zones. This finding has important practical implications, as it indicates that transit operators can change the number of buses serving a zone-structured route without needing to change the entire bus operating pattern.

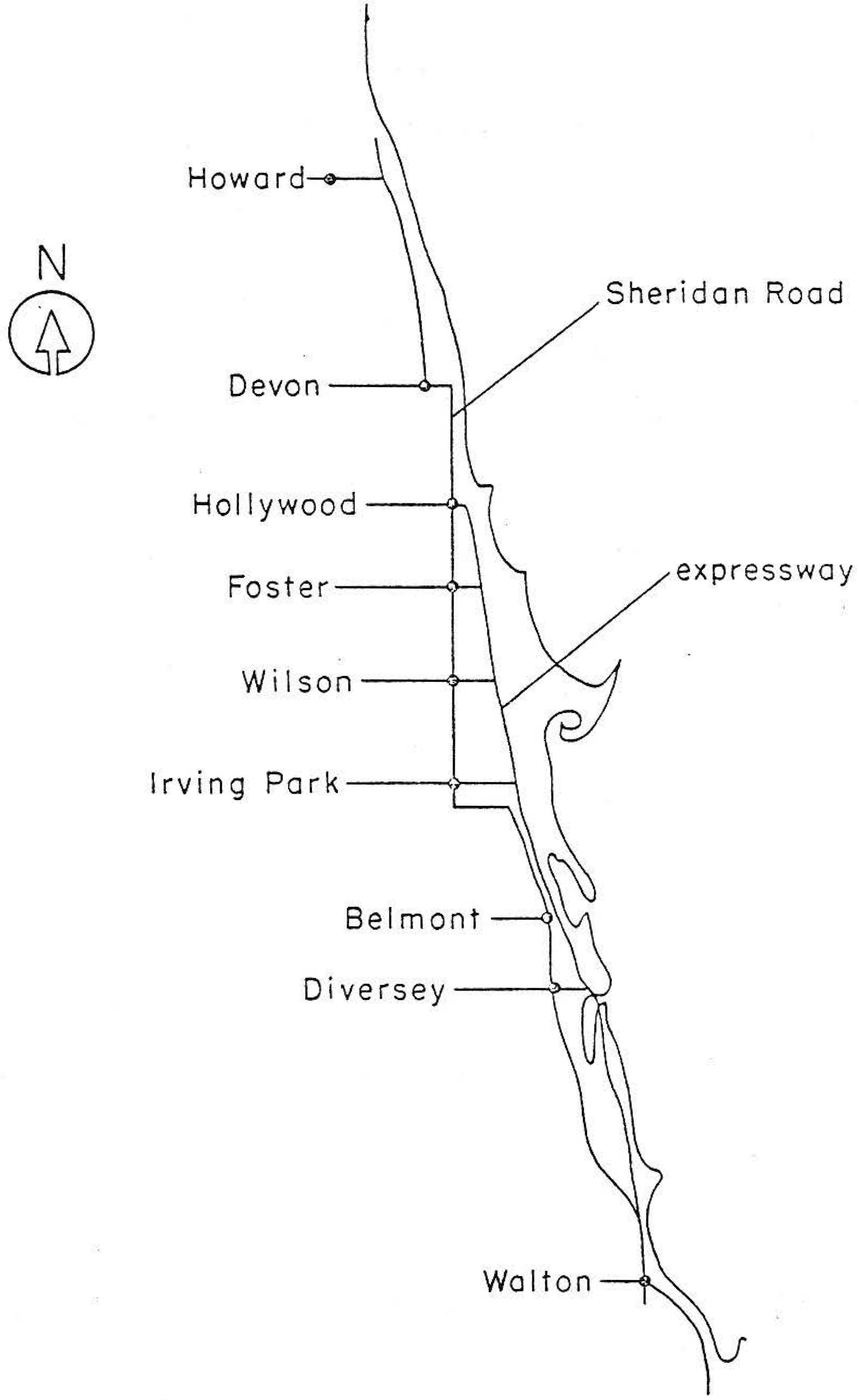


Figure 6-3. Route representation for model application.

TABLE 6-1

BUS STOP LOCATIONS

Stop location	Distance from CBD, D_k
Howard (H)	10.1 miles
Devon (D)	8.1
Hollywood (HW)	7.1
Foster (F)	6.3
Wilson (W)	5.6
Irving Park (IP)	4.8
Belmont (B)	3.1
Diversey (DY)	2.6
Walton (WN)	0.0

TABLE 6-2

MORNING PEAK ORIGINS AND DESTINATIONS (PASS./HOUR)

	Destination								
Origin	WN	DY	B	IP	W	F	HW	D	Total
H	500	2	2	3	8	5	0	13	533
D	534	9	9	12	12	20	0		596
HW	233	0	0	0	0	0			233
F	234	4	4	5	13				259
W	67	15	15	20					117
IP	300	20	20						340
B	700	17							717
DY	267								267
Total	2834	67	50	40	33	25	0	13	3062

TABLE 6-3

MODEL PARAMETER VALUES

Bus Dwell Time Constants:

β_0 = regression constant = 2.25 seconds

β_1 = boarding time per passenger = 2.71 seconds

σ^2 = variance in dwell time for given number of boarding passengers = 8.13 seconds^2

Bus In-Motion Time Gamma Distribution Parameters:

$k_{express}$ = 2/mile

k_{local} = 15/mile

z = .061 seconds^{-1}

N = bus capacity = 70 passengers

H_{max} = maximum headway = 10 minutes

τ = peak period length = 180 minutes

Free Flow Velocities:

$V_{express}$ = 55 miles/hour

V_{local} = 30 miles/hour

a = weight on expected wait time = 2

g = weight on variance of wait time = 1

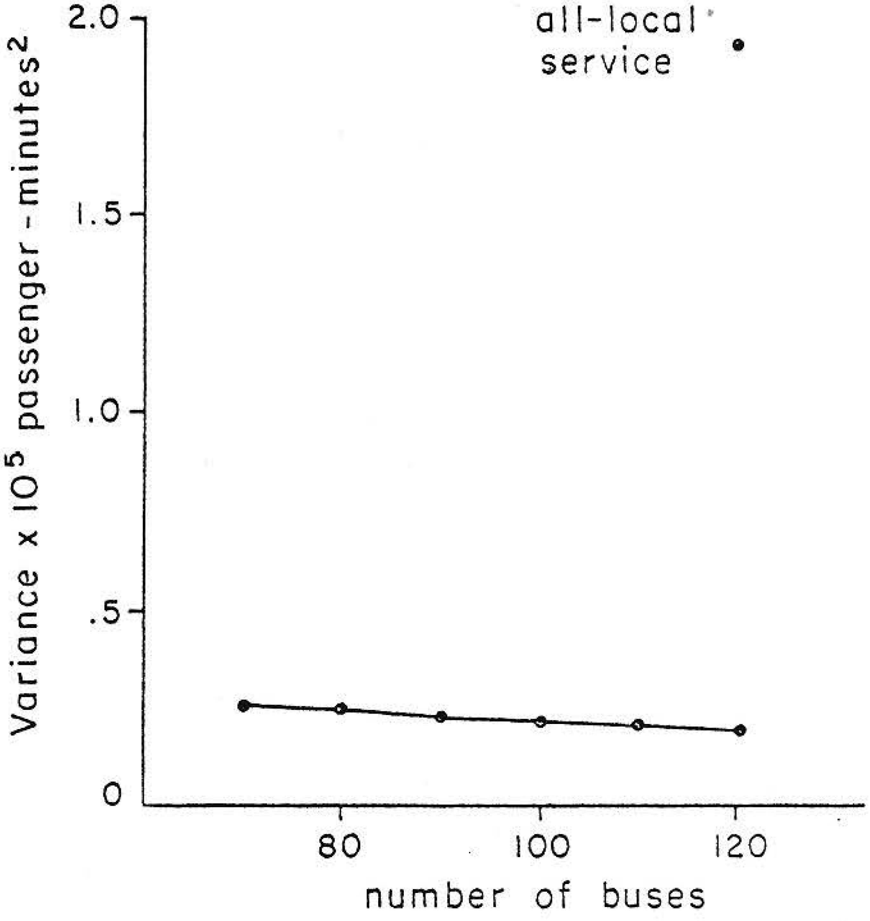


Figure 6-4. Service reliability as a function of fleet size.

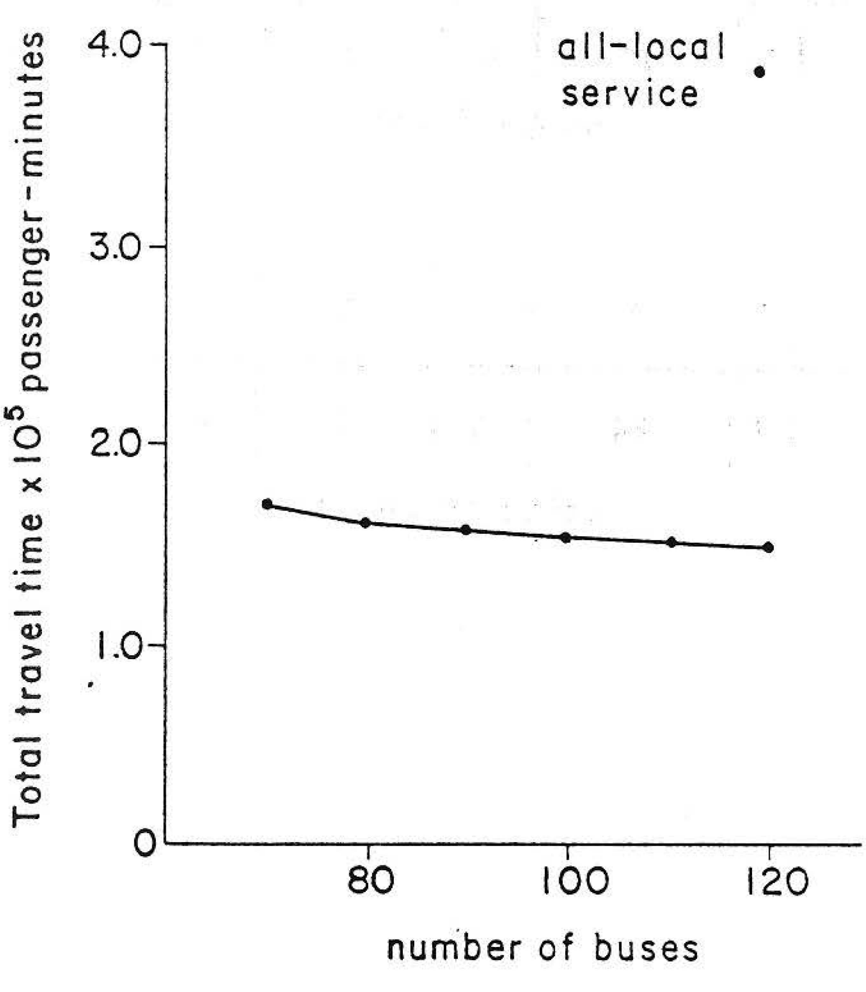
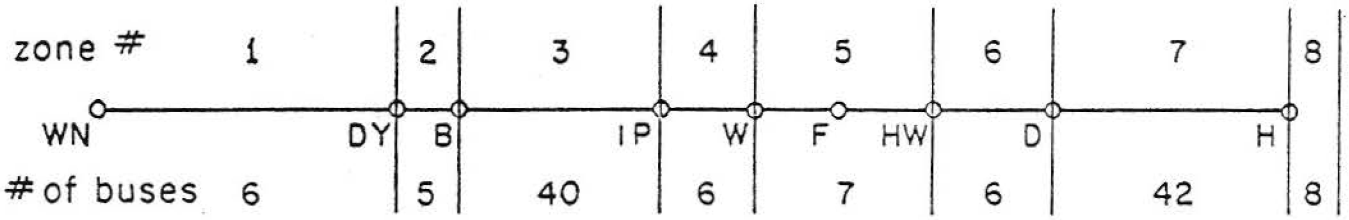
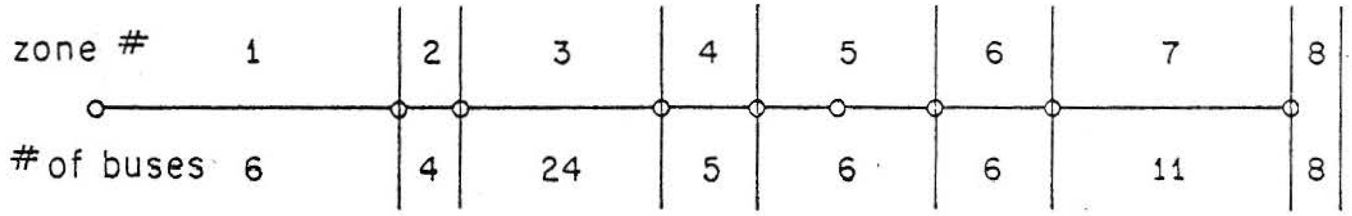


Figure 6-5. Total travel time as a function of fleet size.



using 120 buses



using 70 buses

Figure 6-6. Optimal zone structures for various fleet sizes.

Close to the maximum number of zones are used to create the optimal zone structures shown in Figure 6-6. Since this might appear to be impractical to implement, it is of interest to examine the effect on the reliability measure of constraining the number of zones allowed in any zone structure. Figure 6-7 shows the impact of applying this constraint at different levels. The graph shows that the most significant improvement caused by zone scheduling is achieved simply by moving from all-local to two-zone service. 70% of the overall improvement in the reliability measure is gained in this first step. This suggests that transit operators have considerable flexibility in determining the exact zone structure they should implement to improve service. This flexibility can allow them to respond to factors which are important in a local situation but which are not included in the model.

The sensitivity of model results to changes in important model parameters has also been analyzed. The in-motion time distribution is a key element of the model. The k/mile parameter of this gamma distribution is defined differently for local and express links and potentially has a great influence on the attractiveness of zone scheduling in the model. Values of 15 and 2 have been used for k/mile on local and express facilities, respectively. In the sensitivity analysis, the following combinations of values are used:

k/mile - local	k/mile - express
15	4
10	2
10	4

Table 6-4 shows the analysis results upon rerunning the model with different parameter values. As would be expected, the absolute magnitude of the reliability measure changes significantly. However, the relative improvement resulting from using the optimal zone scheduling strategy over all-local service is very insensitive to the parameter changes.

Additionally, the value passengers place on the variance in wait time relative to travel time variance could be argued to be greater than the value of 1 used in this model. Therefore, the model was rerun using weights of 2 and 4 for the variance of wait time, denoted by g. The effect on the model results is shown in Table 6-5 and Figure 6-8. The reliability improvements from zone scheduling are quite insensitive to this weighting factor, and the optimal zone structure almost unaffected.

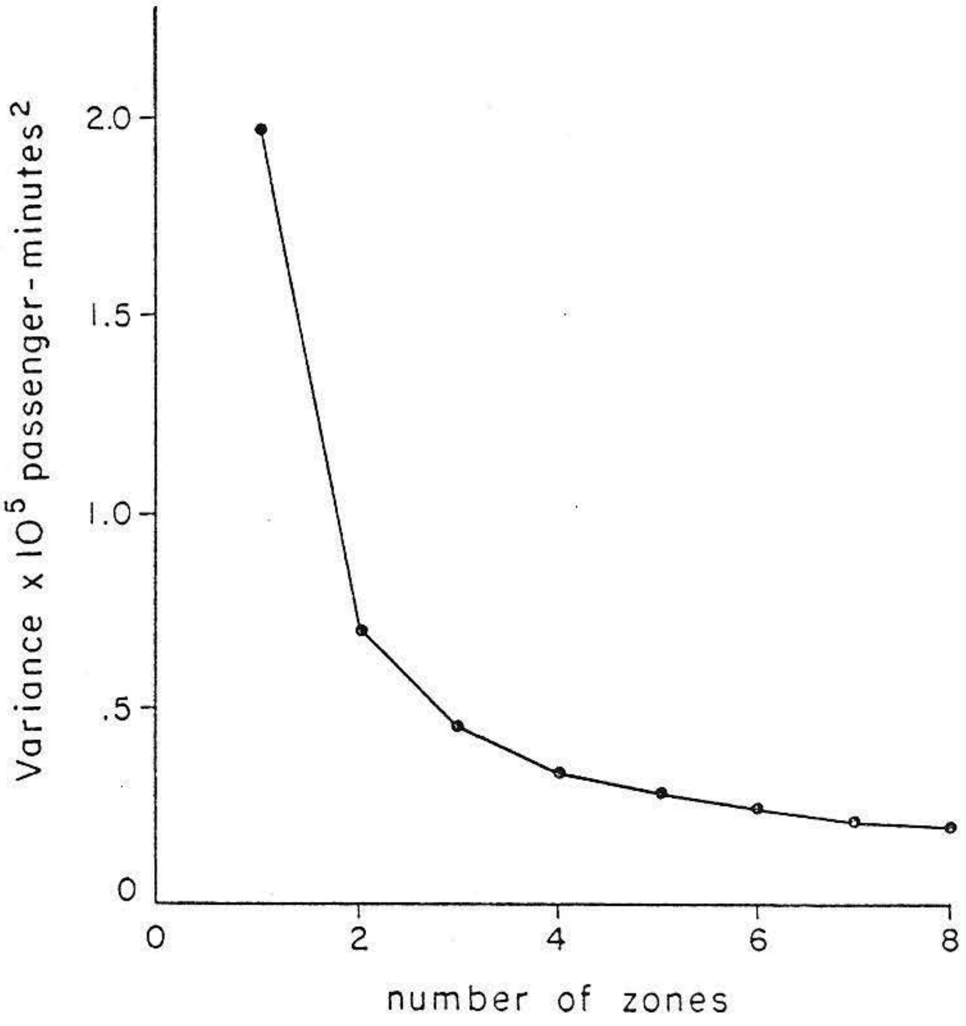


Figure 6-7. Service reliability as a function of number of zones.

TABLE 6-4

SENSITIVITY ANALYSIS ON k PARAMETER

	Total Variance*	% reduc.	Total Trip Time**	% reduc.
k/mi.: local=15,express=2				
all-local service	192,500		350,700	
optimal zone ser.	19,080	90	127,600	64
k/mi.: local=15,express=4				
all-local service	192,500		350,700	
optimal zone ser.	28,260	85	160,100	54
k/mi.: local=10,express=2				
all-local service	126,300		273,100	
optimal zone ser.	16,740	87	122,600	55
k/mi.: local=10,express=4				
all-local service	126,300		273,100	
optimal zone ser.	25,250	80	154,300	44

* passenger-minutes^2
**passenger-minutes

TABLE 6-5

SENSITIVITY ANALYSIS ON WEIGHT OF Var(W)

	g=1	g=2	g=4
all-local service Total Variance*	192,500	196,500	204,500
optimal zone struc. Total Variance*	19,080	27,420	44,020
% reduction	90	86	78

* passenger-minutes^2

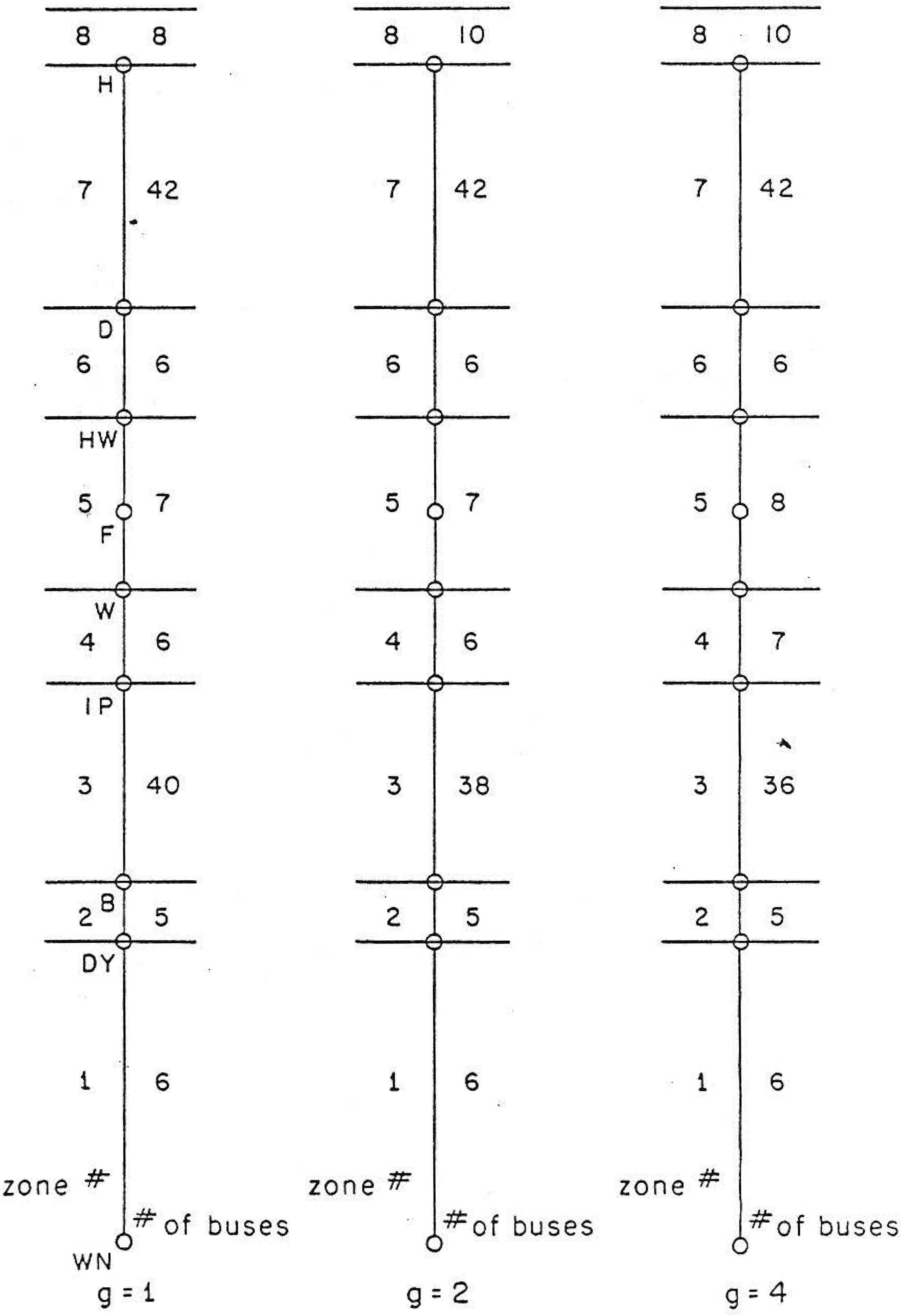


Figure 6-8. Effect on zone structure of changing weights on wait time variance.

The possibility that average trip time was sacrificed to improve reliability was examined in the following way. Two zone structures, one optimizing the reliability measure, the other optimizing average trip time, were determined. The measure which was not optimized was computed for each zone structure. Figure 6-9 shows a plot of the two solutions including also the all-local service situation. Indeed, it is seen that to achieve the optimal aggregate reliability (point B) the total expected wait plus in-vehicle time is increased by about 20% over the optimal average trip time solution (point A). However, the optimal aggregate reliability solution still provides about 90% of the improvement that the optimal average trip time solution does, compared to all-local service. In other words, the zone structure which minimized variance comes very close to minimizing average trip time, and vice versa, relative to all-local service.

The zone structures associated with points A and B on Figure 6-9 are shown in Figure 6-10. The number of zones is identical for the two cases, with zone boundaries being only slightly different. The most significant difference between the two solutions is the distribution of buses among the zones. The practical implication of this result is that transit operators may be able to make the fundamental decisions -- number of zones and zone boundaries -- without particular concern for the tradeoff between average trip time and reliability. In the example studied here, this tradeoff only reflects itself in the allocation of buses to zones. This decision is much less fundamental, and more easily changed in response to local concerns. It is also a decision which can be updated relatively easily as new information becomes available on the performance of the route.

6.2.3 Model Extensions

The model has been expanded beyond that presented here in two basic ways (see Jordan, 1979). First, non-CBD-bound (local) passengers have been incorporated into the model. Only slight modifications to the model are needed to include these passengers, but the notation becomes much more cumbersome. As would be expected, including local passengers makes zone scheduling somewhat less attractive, especially from an average trip time viewpoint, because these passengers are often forced to transfer. Referring to Table 6-2, about 7% of the Sheridan route users are not destined for the CBD during the

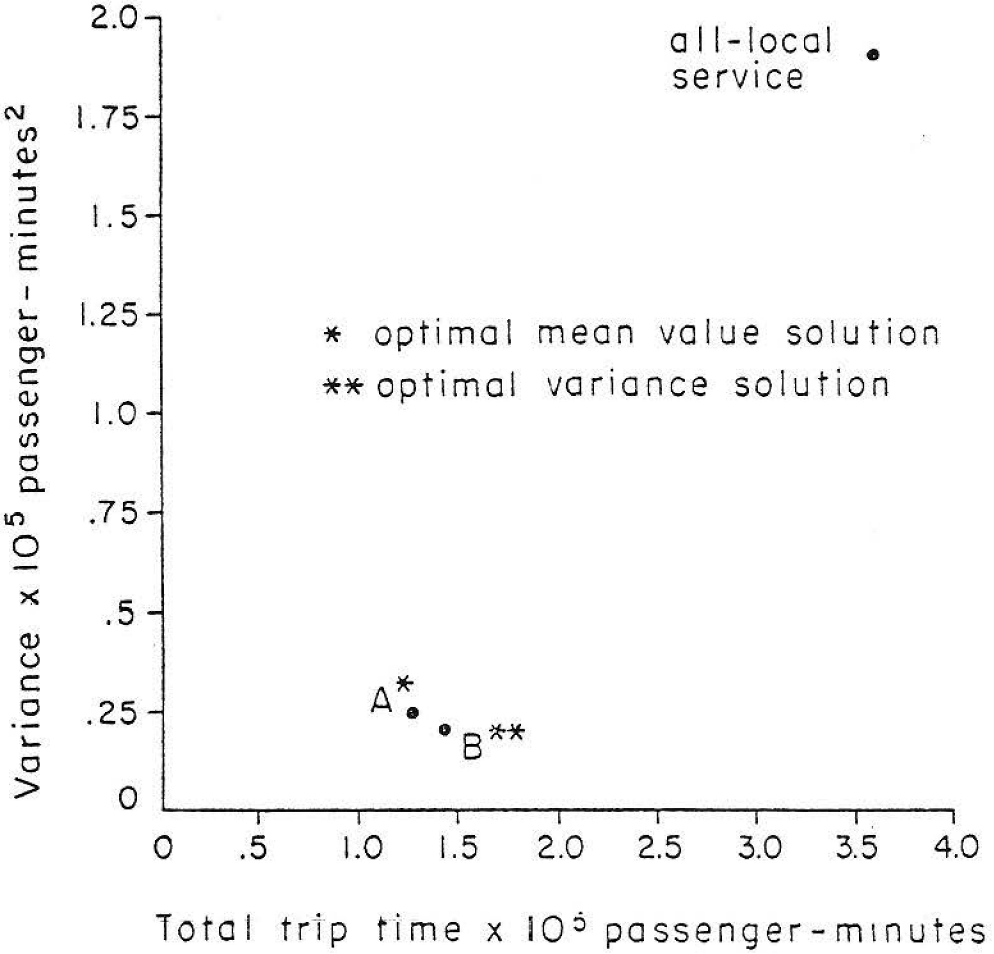
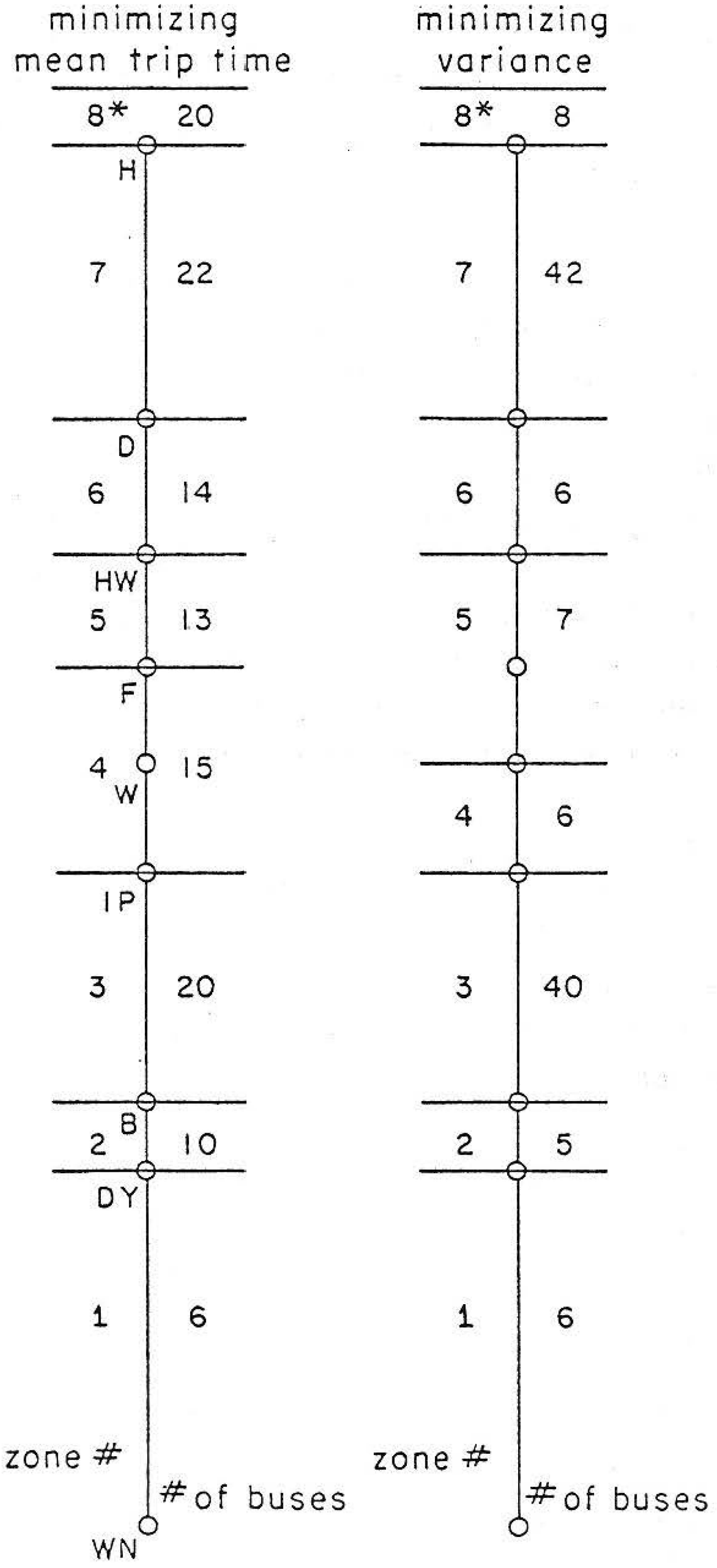


Figure 6-9. Trade-off of variance and average trip time, relative to all-local service.



*zone 8 serves stop 9 only, running express from stop 9 to the CBD

Figure 6-10. Zone structures for minimization of mean trip time and variance of trip time.

morning peak. Including these passengers in the model was found to have a very small impact (about 10%) on the values of the measures of reliability and average trip time for the best zone scheduling strategy relative to those for all-local service.

Also, a second set of measures has been used in the model. The average trip time and reliability measures for the entire route discussed above could indicate good overall service was being offered, while at the same time very poor service was being provided to a few low-demand stops. Therefore, extreme performance measures, the maximum average trip time, and maximum variance in trip time for any passenger, were included in the model. In applying the model it was found that the zone structure optimizing the aggregate measures also came very close to optimizing the extreme performance measures. Thus, while these extensions are conceptually interesting, and potentially valuable in some applications, they have little effect in the particular case study examined here. Additional work on other application sites should be useful in determining the real potential effectiveness of these extensions.

6.2.4 Summary and Conclusions

This research has developed a model of an urban bus route, operating during the morning peak period, to study the impact of zone scheduling on bus service reliability. Service reliability is measured as the variance in passengers' trip time, summed over all passengers using the route during the peak period.

At the core of the model are expressions for means and variances of wait and travel time for a bus route. These expressions reflect the dependence of wait and travel time on:

1) route and bus characteristics (stop spacing, bus fleet size, bus capacity);
2) the parameters of the in-motion time distribution and dwell time equations; and
3) the zone scheduling strategy or structure imposed on the route (the number of zones created, zone boundaries, and number of buses allocated to each zone).

A dynamic programming model is used to search efficiently for zone structures optimizing both average trip time and variance of trip times. The model is applied in a case study with the following results:

1) Very substantial average trip time and reliability improvements are attained through zone scheduline relative to all-local service.
2) While maintaining the average trip time and reliability improvements, substantial decreases in a route's bus fleet size can be made, as a result of the improved productivity of all vehicles.
3) Average trip time is improved simultaneously with reliability under a zone scheduling scheme.
4) The major portion of reliability improvements can be attained by a very simple zone structure.

These results appear to be relatively insensitive to changes in important model parameters, at least in the case examined. This tends to increase the level of confidence in these results.

Our analysis has demonstrated that zone scheduling can be a very effective way in which to improve the quality and productivity of urban transit service. The model developed in this section, appropriately calibrated, may be used in a variety of settings to help determine optimal zone structures and bus fleet allocations. It should be emphasized that the attractiveness of zone scheduling in a given situation will depend heavily on the relative express and local speeds attainable, the relative variability in travel times on express and local links, and the proportion of total route ridership which is destined for (or originates at) the route terminus.

The case analyzed as an example in this section is very favorable to zone scheduling in each of these respects. However, analysis of other situations has also demonstrated the effectiveness of the approach in cases somewhat less attractive to zone scheduling. In some practical situations, it may be advisable to retain some all-local service along the route to facilitate interzonal trips, or to otherwise modify the "optimal" solution in response to local needs. The most important implication of the research is that a model has been developed with which a transit operator can assess the potential impacts of zone scheduling on a given route, and can identify a "good" strategy as a basis for actual planning of the service.

6.2.5 Simulation Experiments With Zone Scheduling

A series of simulation experiments has been performed using the Cincinnati network, to test further the potential of zone scheduling for improving service reliability. This set of experiments accomplishes two objectives. First, it allows an investigation of particular zone scheduling schemes in the context of a more detailed model than used in the analytic optimization. Second, by testing zone scheduling strategies on the Cincinnati network, a basis for comparison with other strategies is established.

Two of the three inbound routes were zoned, routes 43 and 47. In each case, there is an Interstate Highway running into the CBD which can be accessed fairly easily from one stop on the route (see Figure 6-11). In the case of route 43, buses can be routed from California Ave. to I-71 for an express run to the CBD. For route 47, buses can be routed from Spring Grove Ave. to I-75, and run express to downtown. Thus, each of these routes is divided into two zones.

However, because of the large number of non-CBD-bound passengers handled by these routes, the operating strategy chosen is somewhat different from that described in the previous sections. In each case, all vehicles begin their inbound runs at the outer terminus of the route and run local to the stop at which access to the expressway is available. At this point, selected runs go express to the CBD, while the remainder continue local service along the rest of the route. In the outbound direction, all buses run local, as in the base case.

This operating strategy was chosen because it seemed realistic in the context of the system being examined. In fact, two "extra" trips each morning are currently being made over the 43-express route used in this test. Thus, the simulation experiments test the effects of expanding this service, and adding comparable service on route 47.

Because the nature of this zone scheduling strategy is limited, it is to be expected that the results may be less dramatic than those described in section 6.2.2. An additional consideration in this regard is that we are not necessarily testing an "optimal" strategy in the simulation. It is simply a strategy which seems likely to provide improvement relative to the base case.

Indeed, this is what occurs. As shown in Table 6-6, vehicle travel times on the express routes are reduced substantially. This translates into reduced in-vehicle time (or increased speed) for those passengers able to utilize the express runs. It also means that vehicle productivity is increasing, since the same number of passenger trips are being accommodated with fewer vehicle hours.

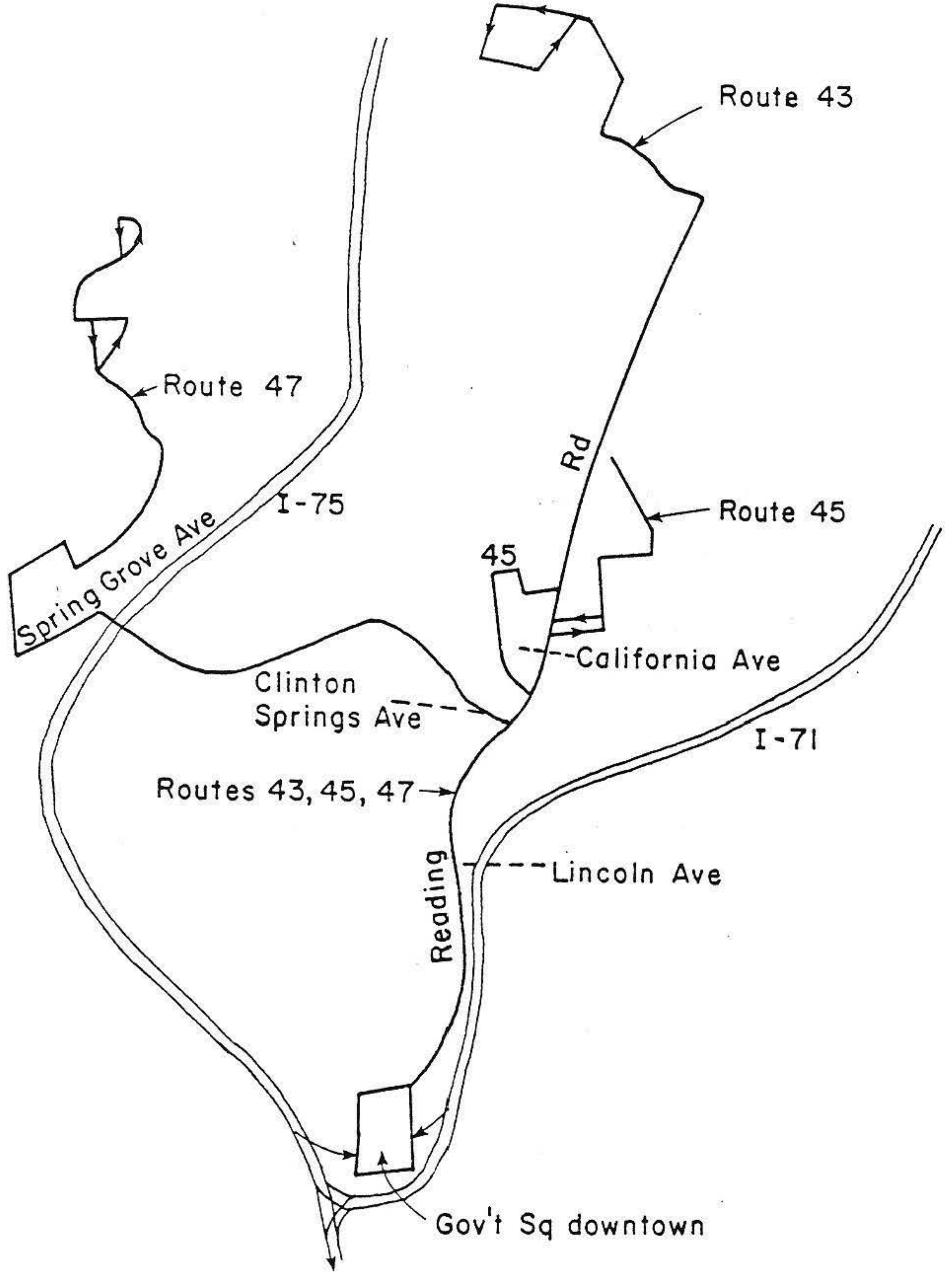


Figure 6-11. Route network from Cincinnati.

Table 6-6. Vehicle travel times on routes

Route	43-local	43-express	47-local	47-express
Inbound Travel Time (min.)	54.	39.	50.	30.

An examination of the passenger speed distribution is illustrative of the effects of zone scheduling on both local and CBD-bound passengers. As shown in Figure 6-12, when zone scheduling is implemented, the local passengers are about as well off as they were before; no major changes in either the mean or variance of speed are observed. However, the CBD-bound passengers now receive much better service than previously. Average speed is increased from 14.3 km/h (8.9 mph) to 29.8 km/h (17.9 mph), and the standard deviation of speed is reduced from 4.8 km/h (3.0 mph) to 4.0 km/h (2.5 mph). Thus, CBD-bound passengers are receiving service which is both faster on the average, and more reliable as well.

It should be noted that these results are obtained with a reduction in total bus-hours operated. In this experiment, the same frequency of service was provided in both the base case and the zone scheduled runs. The reduced round trip travel time for buses was utilized to reduce the vehicle requirements for operating an equivalent frequency of service, rather than to provide more frequent service with the same number of vehicles. Thus, this service improvement is achieved while simultaneously reducing the cost of providing service.

The level-of-service improvements noted here provide a lower bound on what is achievable with this type of zone structure. If more of the bus-hours saved were used to increase service frequency, additional reductions in waiting time could be obtained, providing even greater improvements in the overall level-of-service.

6.3 Summary

Reducing the number of stops made by each individual vehicle can be a very effective method to improve service reliability, as well as average wait and in-vehicle time for passengers. Two major ways of accomplishing this reduction have been examined.

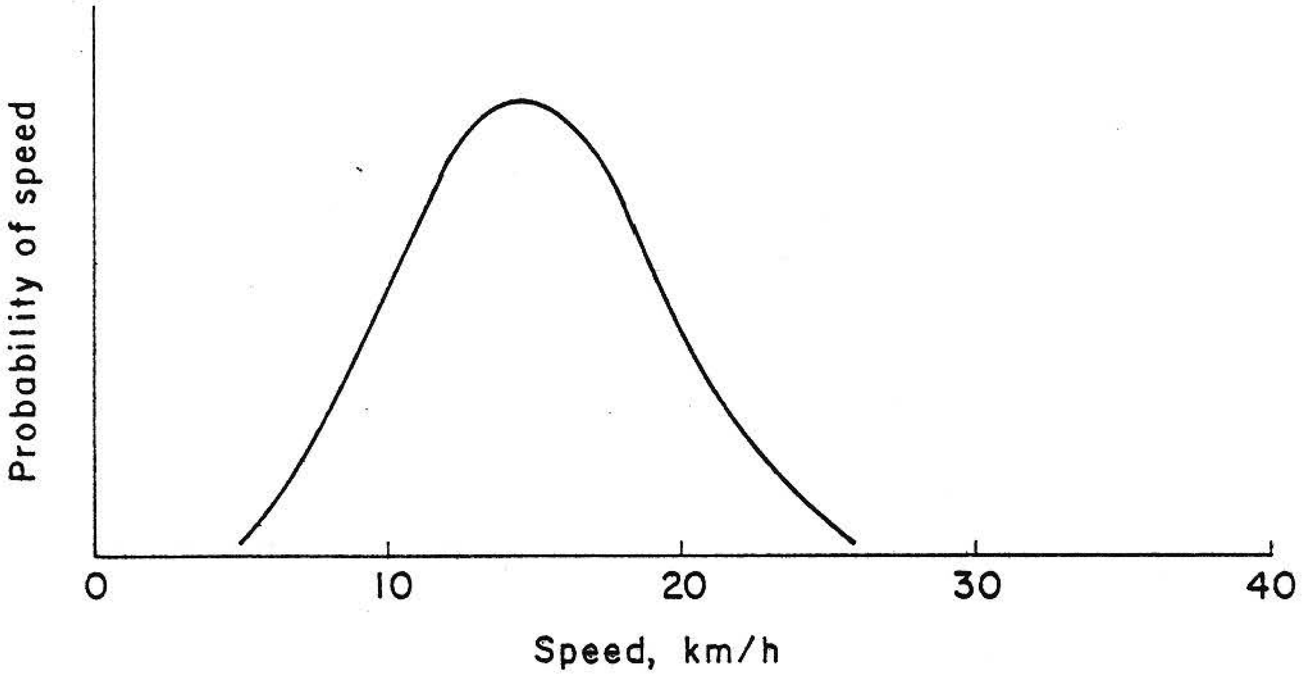


Figure 6-12(a). Distribution of passenger speeds in base case.

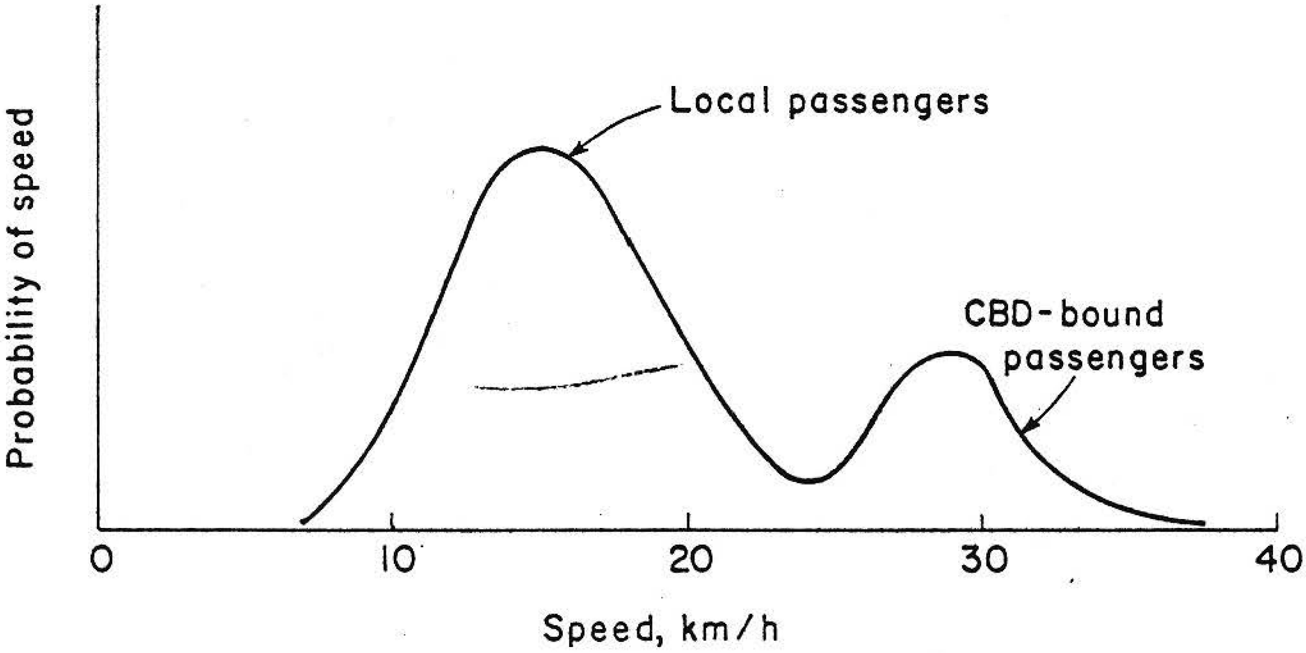


Figure 6-12(b). Distribution of passenger speeds with zone scheduling strategy.

Eliminating stops so as to increase average stop spacing can be useful if stop density is very high before reduction, and if traffic signal operation can also be changed to allow buses to take advantage of the potential for increased speed. This latter point will be discussed further in Chapter 7. It should be noted, however, that increasing stop spacing also has the effect of reducing the accessibility of the bus route to those who use it. Thus, there is a tradeoff of improvement in one (or more) dimension(s) of service quality and a degradation in another. The full implications of this can only be ascertained by including a demand analysis with the results of this work, in order to determine how travelers would react to this tradeoff. Such analysis remains for further study.

A more attractive way of reducing the number of stops made by each vehicle, without increasing overall stop spacing, is by zone scheduling. A model has been developed to design optimal zone structures and allocate buses to zones, using dynamic programming. Application of this model to a route in Chicago illustrates the potential effectiveness of zone scheduling as a service improvement strategy. Substantial improvements in reliability, as well as in other measures, appear possible. Simulation of a particular zone strategy for two routes in Cincinnati has indicated the attractiveness of such a method in an alternative context as well, using a more detailed model.

This chapter has focused on methods of reducing variability of dwell times at stops, as a means of improving reliability. The next two Chapters examine methods of reducing travel time variability between stops. Chapter 7 discusses signalization changes intended to reduce the variability in travel time resulting from intersection delays, and Chapter 8 discusses exclusive right-of-way provision as a means for reducing travel time variability by removing buses from mixed traffic.

CHAPTER 7

CHANGES IN TRAFFIC SIGNAL OPERATION

Several authors, including Welding (1957) and Jackson (1977), have emphasized the importance of variability in travel times between stops as a source of bus reliability problems. A major portion of this variability arises from delays at controlled intersections. This chapter examines two types of strategies aimed at reducing the impact of signalized intersections on average delay and the variability of delays. The first of these is signal preemption, and the second involves changes in signal timing and progression.

7.1 Prior Research on Signal Preemption

In preemption schemes, the traffic signal controller is made aware of the presence of a bus in some manner (optical or loop detectors, automatic or driver actuated bus transmitters), and then modifies the normal phasing of the signal. It may extend the green phase or cause early termination of the red, depending on when the bus arrives at the intersection. It could also activate a special phase to allow a bus to make a left turn or to permit a bus to re-enter normal traffic after the termination of a bus lane. Some schemes follow the preemption with a special cycle that rewards extra green time to the cross street to prevent long delays on that street.

Cottinet (1977) compared three different preemption strategies in an experiment in Nice, France. The first strategy allows an approaching bus to change the signal to green whenever it is detected. The second strategy shortened the red period only, while the third allowed only for extension of the green period. Not surprisingly, the first strategy was found to be superior. This strategy has thus formed the focus of the analysis in this research.

Several other studies have also reported experience with preemption strategies at isolated intersections[1] or on intersections along major streets that have light cross traffic. Such applications have shown mean transit time savings of 10-75% at insignificant costs to auto users. In Miami, mean travel time over a ten-mile section of an express bus route was reduced by 20% when 35

---

[1] "Isolated" in this context may mean problem intersections in a downtown area that are not timed to a network.

signals controlling minor street intersections were made subject to bus preemption (Courage, et al., 1977). A similar demonstration in Louisville involving eight intersections showed time savings of 9 to 19% on forty-minute routes (Mitre, 1975). A four phase, long cycle intersection in Bern, Switzerland was modified so that there were two small green phases per cycle, each subject to green extension upon detection of an on-street tram. Average intersection delay was reduced to 9 seconds from 37 seconds (Mitre, 1975). Quantitative documentation of corresponding improvements in bus reliability is given by the British, who have long experimented with priority measures and consider reliability a major attribute of signal preemption (TRRL, 1973). Two cities in England reported that the standard deviations of bus journey times through two intersections were halved after simple preemption schemes were implemented (Cooper and Layfield, 1977; Wood, 1976).

Extensive preemption has not been well tested in more congested conditions where disruption to traffic might be more substantial. A UTCS-1 model simulation of two streets in Washington, D.C. indicated that bus travel times are reduced by 15 to 20% when each signal is subject to bus preemption (Ludwick, 1975). The standard deviation of travel time is reduced by 20 to 50%. The delays to cross street traffic are serious and worsen with increasing bus frequency (from a 6% increase in vehicle travel times for 4 minute headways to a 30% increase for 30 second headways). When Washington, D.C. actually demonstrated a preemption scheme on thirty-four downtown intersections, bus travel time improvements of about 3% were shown on a typical mile-long route where 6 signals out of 11 could be preempted. On a 3.5 mile route where 10 out of 30 signals could be preempted, there was no improvement. Delays to overall traffic increased by .3 to 2.5% (JHK, 1975). An FHWA study concluded that all signals along a route should be subject to preemption, and that the scheme works best where the signal offsets are not critical and the bus flows are less than thirty per hour (Tarnoff, 1975).

In summary, consideration should be given to signal priority measures when: (a) bus flows are in the range of 10 to 30 per hour, and (b) the cross street bus flows are light. In addition, Jacobson and Sheffi (1980) have shown that other parameters of signal operation (cycle length and phase splits) should be adjusted when preemption is implemented, in order to achieve the greatest possible benefits.

It must be noted that most previous researchers have focused on preemption as a means of reducing average delay at intersections, and hence improving average speed of the buses. Except for the British studies, service reliability benefits from reducing the variability of delays have not been considered explicitly. In order to include this measure of effectiveness, simulation experiments with signal preemption have been performed in this project, again using the Reading Road corridor in Cincinnati as a test case.

7.2 Simulation Experiments With Preemption

The test case involves signal preemption over a 7.1 km (4.4 mi.) section of Reading Road, from Clinton Springs Ave. to Government Square in the CBD. Two of the 14 intersections between Clinton Springs and Lincoln, and 12 of 18 from Lincoln to Government Square, were made preemptable. The signals not made preemptable are generally ones with long cycle times and a very long main direction green phase. A typical example is a 90 second cycle length, with 70-75 seconds of green in the main direction. These signals do not produce serious delays for buses. The preemption tests were confined to intersections for which delays are likely to be more severe.

Over the northern 5.0 km (3.1 mi.) of the section tested, cycle lengths in the base case are all 90 seconds. In the southern 2.1 km (1.3 mi.), they are generally 70 seconds. This reflects current conditions along this section of Reading Road. Green phases in the main direction generally are 55-75 seconds in the northern segment, and 35-45 seconds in the southern segment. With a few notable exceptions, the signals are progressively timed for a speed of approximately 40 km/h (25 mph) over most of this section, but not in the downtown area.

Using the method described by Jacobson and Sheffi (1980), optimal cycle lengths and splits for the preemptable signals were found. These settings depend upon traffic parameters (main and cross direction flow rates, queue discharge rates and average vehicle occupancy), bus arrival rate and average occupancy, and preemption parameters (detection distance and minimum cross-direction green time in each cycle). For a range of parameter values typical of the Reading Road situation, a near-optimal solution is a 40 second cycle length, with 16 second main direction green. Since this appears to be a robust solution under several different sets of input values, it was chosen for all preemptable signals in the simulation tests.

Note that this signal setting is substantially different from that existing in the base case. The cycle is approximately one-half as long, and the split has changed from predominantly main direction green to predominantly cross direction green. The finding that a shorter cycle length is advantageous accords with the experimental results from Bern, Switzerland, cited in the previous section. The shift in phase splits reflects the fact that since the main direction will receive extra green every time the signal is preempted, the basic setting should favor the cross direction. Thus, the cross direction traffic will not be penalized heavily, relative to the base case, as a result of the preemption. In essence, the optimal solution represents a balancing of the marginal benefits of preemption to main direction travelers against the marginal delay costs imposed on cross direction travelers.

In addition to the signal changes, bus stops at intersections where the signal was made preemptable were moved to the far side of the intersection. This follows the recommendation of Ludwick (1975), and avoids the problem of having the detector cause the light to turn green as a bus approaches, only to have the bus stop at the near side of the intersection to load or unload passengers.

As in previous tests, five replications of the preemption strategy were made with the simulation model. The pooled average of these five runs is then compared with a similar average from the base case.

One of the major effects of the preemption strategy is that bus travel times over the section with signal preemption are reduced by approximately 3.5 minutes (out of a scheduled 24 minutes), implying an average speed increase from 17.7 km/h (11.0 mph) to 20.7 km/h (12.9 mph), or about 17%. The standard deviation of travel times was reduced by approximately .5 minutes out of a total of 2.7, a reduction of 18%.

A second major effect is on passenger wait time. Average wait time is reduced by .6 minutes, from 7.5 minutes to 6.9 minutes, a reduction of 8%. The standard deviation of wait time is also reduced, from 7.7 minutes to 7.0 minutes, or 9%. All of the changes to both travel times (or speeds) and waiting time are significant at the 90% confidence level.

Signal preemption thus appears to offer significant potential for improving both average speed and reliability, with concomitant effects on both mean and variance of waiting time. As a further test of this strategy, a second experiment was conducted, combining signal preemption with reduced stop density, as discussed in Chapter 6.

The results of this experiment were a small (but insignificant) further increase in average speed, and a small decrease in average wait (also insignificant), as compared to the use of preemption alone. However, the standard deviation of wait time decreased to 6.5 minutes, as compared to 7.0 minutes for preemption alone, and 7.7 minutes for the base case. This further reduction in the variability of wait time is statistically significant, and constitutes the major observed impact of combining signal preemption and reduced stop density.

7.3 Signal Timing

Signal preemption is not the only way in which to try to reduce intersection delays in order to improve service reliability. A different approach that has potential for congested areas is to time the signals of a network to match the overall travel speeds of a bus (including the time spent at stops). This strategy is less costly than the preemption methods which require bus detection equipment, and may be less disrupting to area traffic in situations in which there are many signals in a relatively small area. Glasgow, Scotland, has tried an extension of the computer package TRANSYT called BUS-TRANSYT to retime their signals using passenger delay as an objective criterion rather than vehicle delay. Bus travel times were reduced by 8% while auto times increased by only 1% (Robertson and Vincent, 1975).

Within the context of the current project, analysis of signal timing measures has involved simulation experiments, again using the Reading Road corridor in Cincinnati as a test case. There are 32 signalized intersections along the 7.1 km (4.4 mile) section from Clinton Springs Ave. to Government Square. Over the northern section (about 5.0 km; 3.1 miles) the signals all have a common cycle length of 90 seconds, and are generally progressively timed for a speed of about 40 km/h (25 mph). The progression breaks down over the southern segment, as the signals have different cycle lengths and are not timed progressively.

Two different timing strategies were tested. In the first, only minor modifications to existing timing were made. The second involved more substantial changes to try to match signal progression to a bus speed of 24 km/h (15 mph).

For the first test, the section was divided into two segments, essentially as at present, and an attempt was made to modify as few signals as possible while still providing a better progression for buses. Five of 21 signals on the northern segment were modified by simply changing their offsets. On the southern segment, the cycle lengths were made uniform (at 70 seconds) and three offsets were changed. The results of the simulation experiments showed no significant changes from the base case in any level-of-service measure.

A second test was then conducted, involving more substantial changes in offsets on the northern segment. Ten additional offsets were changed, in order to create a well defined progression at 24 km/h (15 mph). While this is somewhat faster than buses usually travel, one of the main reasons for slower speeds may be the traffic signal settings themselves. By timing the signals for a speed which might be attainable under favorable conditions, it was hoped that we would in fact create those conditions. However, once again the simulation experiments showed no significant changes from the base case.

Detailed examination of the runs indicated that one reason for the poor performance of the strategies was that buses were simply not able to maintain the speed of the progression because of the need to make frequent stops for boarding and alighting passengers. As a result, it was decided to do additional experiments, combining the changes in signal timing with reduced stop density. The results of the combination of reducing the number of stops to approximately quarter-mile spacing and the first timing strategy produced results similar to those achieved with signal preemption.

Bus travel times over the entire section were reduced by approximately 2 minutes, or about 8%. Average passenger wait time was reduced from 7.5 minutes to 6.9 minutes, and the standard deviation of waiting time from 7.7 minutes to 7.0 minutes. Thus, the combination of reduced stop density and changes in signal timing is slightly less effective than signal preemption in increasing vehicle speeds, but very nearly the same with respect to improvements in waiting time, at least for this network.

Somewhat surprisingly, the combination of reduced stop density and the more extensive signal timing changes was less effective. As with the signal changes alone, this combination produced no significant changes from the base case. This is probably attributable to the selection of 24 km/h (15 mph) as the progression speed. Even with the reduced number of bus stops, it is not possible for buses to maintain this average speed, and hence there is little change from the base case. An alternative timing strategy at a somewhat slower speed may be more effective.

7.4 Summary

The signal preemption strategies tested resulted in 17-18% increases in both mean and standard deviation of vehicle speed, and 8-9% reductions in mean and standard deviation of waiting time. These are significant improvements, and indicate that signal preemption can be an effective means to improve reliability of service, as well as average travel time.

The experimental results here have not measured carefully the effects of signal preemption on cross traffic. To do so requires a more detailed traffic simulation model. However, it should be emphasized that the revisions to signal cycle length and phases, determined using the model of Jacobson and Sheffi (1980), include the average effects on both main and cross direction traffic and find the solution which minimizes overall expected person-delay. Thus, the results cited here should be attainable without imposing large penalties on cross direction traffic.

The results of the experiments on signal timing indicate the important interactions between stop spacing and signal timing. Both characteristics must be considered jointly in order to make changes in either effective. The limited experimentation done in this project indicates that this may be difficult to do, but that if the stop spacing and signal timing are appropriately "matched," the effects may be similar to those from signal preemption.

CHAPTER 8

RESERVED BUS LANE STRATEGIES

In congested downtown areas, traffic stream delays may account for 15 to 30% of transit travel time (Vanselow and Sinclair, 1972). Measures that remove the bus from these random delays will reduce travel time and improve reliability. In extreme cases, buses that are allowed to bypass a severe bottleneck can travel significantly faster than the autos.

Freeway priority lanes for buses can be grade-separated roadways, contraflow lanes, shoulder lanes, reserved lanes, or lanes that improve bus access to the highway by bypassing various types of queues. On urban streets, similar measures are curb lanes for buses (and perhaps for right turning vehicles), contraflow lanes, bus-only streets, short reserved lanes preceding traffic signals, and automobile entry restriction combined with bus entry privileges.

There are several examples in U.S. cities of effective use of bus lanes. In Minneapolis, buses bypass vehicles queued at metered on-ramps to I-35W. Travel from the freeway exit to downtown stops is improved with contraflow lanes on two parallel streets in the CBD. The Shirley Highway in Washington, D.C., has two reversible lanes for buses and carpools in the median of the expressway. These lanes carry approximately 6000 persons/lane/hour at 50 mph in the peak period, compared with 2300 persons/lane/hour in each of the regular lanes, operating at much lower speeds (McQueen and Waksman, 1977). In Miami, buses that were already receiving signal priority on an express route were given an exclusive lane. Average travel time improved by about 8%, and the travel times across the peak hour were more uniform (Michaelopoulos, 1976).

Bus lanes have also been used effectively in Europe and Australia. One lane of a four-mile inbound street in Sydney, Australia, was designated as a transit lane, and bus travel times were cut nearly in half, from 24.2 minutes to 13 minutes. The standard deviation was cut by 80%, from 8.3 minutes to 1.8 minutes. Transit ridership and the number of carpools increased so much that congestion was reduced and auto travel times also decreased dramatically (Hallam, 1977).

A one-mile section of a CBD street in Melbourne, Australia, was modified to provide an exclusive right of way for its street tram system, resulting in a reduction of one auto lane. On the 6.5 minute route, the standard deviation of tram travel time was reduced from 66 seconds to 30 seconds (Vanselow and Sinclair, 1972). Three one-half to one-mile bus lanes in Paris and one in Dublin have also halved the standard deviation of bus travel times (Wilbur Smith and Associates, 1975).

Thus, there is considerable evidence that reserved lanes can improve both average transit speeds and reliability. Additional simulation experiments conducted in this project have been designed to examine the effectiveness of combinations of reserved lanes together with signal preemption and changes in signal timing.

As a test case, the Reading Road corridor was again used as a basis, but with substantial modifications. At present, bus operations are not heavy enough to justify a reserved lane, numbering only about 12 per hour. This would not be a very effective test of reserved lane strategies intended for areas of much higher activity. To obtain a test case, the bus frequencies and passenger arrival rates were multiplied by a factor of 5, resulting in average headways of about 1 minute with loadings comparable to the present case. Other elements of the corridor were left unchanged. Thus, our test is over a corridor 7.1 km (4.4 miles) in length, with a total of 36 stops and 32 signalized intersections. The reserved lane was specified as a curb lane (with-flow).

Table 8-1 summarizes the major results of testing the reserved lane alone, and in combination with signal preemption and signal timing changes. The addition of the lane itself results in a small reduction in average travel time, but it is not statistically significant. The reduction in the standard deviation of travel time, from 5.1 minutes to 4.4 minutes (14%), is statistically significant at the 95% level. Reductions in mean and standard deviation of waiting time are also statistically insignificant. Thus, the major impact of adding the reserved lane in this case appears to be a reduction in the variability of travel times, a direct improvement to service reliability.

Combining the reserved lane with changes in signal timing produced no further significant benefits, but the combination of the reserved lane and signal preemption is noticeably more effective. The changes in passenger wait time are still insignificant, but the changes in both mean and standard

Table 8-1. Summary of Test Results for Reserved Lane Strategies

Measures	Base Case (no lane)	Reserved Lane Only	Lane + Timing[1]	Lane + Preemption
Average bus travel time (min.)	25.2	23.8	23.6	20.8
Standard deviation of bus travel time (min.)	5.1	4.4	4.4	4.2
Average wait time (min.)	0.9	0.8	0.8	0.7
Standard deviation of wait time (min.)	1.3	1.2	1.2	1.1

[1] Timing changes correspond to the first strategy discussed in section 7.3.

deviation of travel time along the corridor are highly significant. The estimated reduction in average travel time is 17%, and the reduction in the standard deviation of travel time is 18%.

Two additional aspects of these experiments should be noted in order to aid interpretation of the results. First, the effects of the signal preemption on cross-traffic have not been analyzed in detail. As described in Chapter 7, the expected value of delay to cross-traffic has been included in the settings of the signals used in the preemption study. However, more thorough analysis would require a more detailed traffic simulation model. The second point is that the vehicle traffic levels (both main-direction and cross-direction) assumed for these experiments are relatively light. Main direction volumes are the heaviest, and are in the range of 350-400 vehicles/lane/hour. Most situations in which reserved lanes would be considered are likely to have heavier traffic volumes, as well as heavy bus volumes. Thus, the benefits of removing the buses from the mixed traffic scheme are likely to be greater than measured in these experiments. In this sense, these results are likely to be conservative.

Thus, the combination of a reserved lane for buses and signal preemption capability appears to be a potentially effective method for improving both average travel time and reliability in situations involving very heavy bus movements.

CHAPTER 9

CONCLUSIONS AND IMPLICATIONS OF THE RESEARCH

In the previous chapters, we have examined the effects on service reliability of various network characteristics, as a means for identifying the primary causes of unreliability in bus transit networks. We then proceeded to identify and test several potential strategies for improving service reliability. In many cases, these strategies result in improvements in other level-of-service measures as well. It is worthwhile at this point to summarize the major findings of these tests, as the basis for reaching general conclusions regarding the causes of unreliability and effective strategies for improving the situation.

9.1 Sources of Unreliability

Understanding the network relationships which influence the reliability of service provides insight into potential approaches for service improvement. Reliability of service, as affected by vehicle bunching, may be improved either by preventing bunches from forming, or by breaking them up after they form. The experiments performed in this research have indicated how vehicle bunching is related to frequency of service, level of demand and the variability of link travel times. In particular, these results illustrate the importance of reducing link travel time variability in an effort to prevent bunches from forming. This represents an extension to the results of Vuchic (1969a), which placed primary emphasis on the demand/capacity ratio and boarding times.

The importance of transferring to overall trip reliability focuses attention on the trade-off between the length of the scheduled wait time at transfer points and the risk of missing the intended connecting bus. Where arrivals can be scheduled to coincide, as in radially structured networks, the application of controls to the operation of transit service has the potential to permit more closely scheduled arrivals on connecting services while maintaining a reasonable assurance that the intended connection will be successful.

Finally, it is clear from the experimental results that service reliability is much more sensitive to frequency of service than to route density. This implies that there are substantial reliability impacts of the trade-off between operating fewer routes at higher frequency or more routes at lower frequency, given a limited amount of vehicle resources. Traditionally, this trade-off has been evaluated using simplistic models of expected passenger wait time and the accessibility of transit service to users. However, the present work has shown that service reliability is also an important factor in this trade-off and should be included in the evaluation.

9.2 Strategies for Improving Reliability

In light of the findings regarding sources of unreliability in passenger trips, the major objectives of control strategies are to keep bunches from forming (or to break them up after they have formed) and to ensure that scheduled arrival times at transfer points are met. At a more detailed level, deviations from schedule, which lead to bunching and poor transfer connections, can be traced to excessive variability in either link travel times between stops, or dwell times at stops. Therefore, potential control strategies should be focused on reducing one or both of these sources of variability.

This investigation has concentrated on four general classes of strategies:

1) vehicle holding;
2) reductions in number of stops served by each vehicle;
3) modifications to traffic signal settings and operation; and
4) provision of exclusive right-of-way for transit vehicles.

Such a classification provides a useful framework for discussion of many individual strategies, and a comparison of their relative effectiveness in particular situations.

Within the class of vehicle holding strategies, two subgroups have been analyzed: schedule-based holding and headway-based holding.

The schedule-based "checkpoint" strategy is very simple to implement and offers promise of significant benefits on long-headway routes where the schedule is sufficiently slack so as to make holding to schedule a reasonable procedure. The key elements of implementing such a policy are constructing a reasonable schedule as a goal and enforcing adherence to that schedule. This enforcement requires both proper incentives for drivers and a mechanism for accurate monitoring of their performance.

For routes operating with shorter headeays, two near-optimal headway-based control strategies have been developed. One strategy holds a vehicle until its preceding headway is as close as possible to its following headway, allowing for an adjustment in consideration of the people delayed on the vehicle. Referred to as "Prefol", it requires a prediction of the arrival time of the following vehicle. A similar strategy that is dependent only on the known magnitude of the current headway, called the "Single Headway" strategy, is also proposed. The strategies are simple in form, require limited data about the route, and are near-optimal over a wide range of situations.

Models of the effectiveness of the strategies indicate that they are sensitive to three important characteristics of a control point: (1) the current level of unreliability, as measured by the headway coefficient of variation; (2) the relationship between successive headways, measured by the correlation coefficient; and (3) the proportion of passengers who must ride through the control point.

The Single Headway strategy performs less well than the Prefol strategy when vehicles arrive relatively independent of each other. As passenger loading delays increase and successive headways become more dependent on each other, the Single Headway strategy prediction capability improves and it approaches the performance of the Prefol strategy. In any event, however, the Single Headway strategy requires less real-time information about the system, and could certainly be implemented without expensive AVM equipment, although such equipment could be useful in developing the data on which to base the design of appropriate strategies.

Reducing the number of stops made by each individual vehicle is the second class of strategies examined. This can be a very effective method to improve service reliability, as well as average wait and in-vehicle time for passengers. Two major ways of accomplishing this reduction have been analyzed.

Eliminating stops so as to increase average stop spacing can be useful if stop density is very high before reduction, and if traffic signal operation can also be changed to allow buses to take advantage of the potential for increased speed. It should be noted, however, that increasing stop spacing also has the effect of reducing the accessibility of the bus route to those who use it. Thus, there is a tradeoff of improvement in one (or more) dimension(s) of service quality and a degradation in another. The full implications of this can only be ascertained by including a demand analysis with the results of this work, in order to determine how travelers would react to this tradeoff. Such analysis remains for further study.

An alternative way of reducing the number of stops made by each vehicle, without increasing overall stop spacing, is by zone scheduling. A model has been developed to design optimal zone structures and allocate buses to zones, using dynamic programming. Application of this model to a route in Chicago illustrates the potential effectiveness of zone scheduling as a service improvement strategy. Substantial improvements in reliability, as well as in other measures, appear possible. Simulation of a particular zone strategy for two routes in Cincinnati has indicated the attractiveness of such a method in an alternative context as well, using a more detailed model.

The third class of strategies examined includes changes in signal timing and signal preemption as two methods of reducing both the average value and variability of delays to buses at signalized intersections. The results of simulation experiments indicate the important interactions between stop spacing and signal timing. Both characteristics must be considered jointly in order to make changes in either effective. The limited experimentation done in this project indicates that this may be difficult to do, but that if the stop spacing and signal timing are appropriately "matched," the effects may be similar to those from signal preemption.

The signal preemption strategies tested resulted in 17-18% increases in both mean and standard deviation of vehicle speed, and 8-9% reductions in mean and standard deviation of waiting time. These are significant improvements, and indicate that signal preemption can be an effective means to improve reliability of service, as well as average travel time.

The fourth class of strategies is the provision of exclusive right-of-way for transit vehicles. This action is likely to be considered only when there is heavy transit use of a facility, probably in excess of 30 buses per hour. In such cases, bus lanes can be an effective way of improving both average travel time and reliability. Based on limited simulation tests, the combination of a bus lane and signal preemption appears particularly effective.

9.3 Relative Effectiveness of Alternative Strategies

The most important determinant of the appropriate strategy for reliability improvement in a given situation seems to be frequency of service on the route(s) in question. In some cases, the focus of attention will be a corridor used by several routes, in which case the important variable is total bus volume.

For low frequency situations (less than 10 buses per hour), checkpoint control (schedule-based holding) is likely to be the most effective strategy, providing an appropriate schedule is constructed and adherence to schedule at checkpoints is enforced. In some low frequency situations, zone scheduling may also be effective, if most passengers are destined for (or originate at) one terminus of the route. The presence of an expressway roughly parallel to the route also makes this strategy more effective.

In medium frequency situations (10 to 30 buses per hour), the most effective strategies are likely to be zone scheduling and signal preemption. If the origin-destination pattern of passengers is suitable and an express facility is available, zone scheduling is likely to be the best choice. If these conditions are not met, signal preemption on the local facility used by the buses should be considered. Headway-based holding can also be useful, if an appropriate control point can be found along the route.

For high frequency situations (more than 30 buses per hour) an exclusive lane, together with signal preemption if an arterial, should be considered. Experience from several demonstrations of bus lanes, as well as modeling results from this study and others, indicates the effectiveness of such a strategy in improving both average travel time and reliability for buses.

While these recommendations provide general guidelines for transit operators and planners in selecting service improvement strategies, the most valuable product of this research is the battery of models developed for analyzing a number of strategies in any particular situation. These models include the analytic formulations for developing holding strategies, the dynamic programming model for designing zone scheduled systems, and the computer simulation model for detailed analysis of many possible strategies. With these tools, the transit operator or planner can design and test a service improvement strategy appropriate for his/her own particular situation.

9.4 Practical Implications of the Research

Quite clearly, the findings of this research regarding the relative effectiveness of various service improvement strategies, and the recommendations listed in the previous section, have substantial practical implications. In addition, there are several other implications of this research for the transit operator or planner interested in providing more cost-effective service to passengers.

First, the presence of large variability in link travel times can reduce substantially the benefits resulting from increasing frequency of service, due to the tendency of vehicles to bunch together along the route. In such cases, it is well worthwhile to investigate techniques for reducing this travel time variability, through the sorts of strategies described in this report, rather than simply allocating resources to providing more buses.

The second implication is also directly related to this same argument. Since level-of-service measures (primarily wait time) are less sensitive to increases in frequency of service than previously believed, they are also less sensitive to decreases in frequency of service. The implication of this is that reducing frequency of service may not increase wait time as much as previously believed. Hence, fewer resources allocated to running buses and more allocated to controlling the ones that are running may be beneficial. The results of several of the tests conducted in this research indicate that in many cases better service can be provided with fewer vehicles, if those vehicles are better organized and controlled. This is clearly an opportunity to provide more cost-effective service.

Third, even if no direct control strategies are to be implemented, it may be advantageous to operate a denser route network at lower frequency of service than to provide higher frequency on fewer routes. Such a shift would improve accessibility of the system, probably without major penalties in terms of wait time increases.

The fourth implication is that the influence of transfers on the level of service points out the need to pay special attention to the on-time arrival of vehicles at major transfer stations. This is especially true for radially oriented network structures. As a rule, providing excess slack time in the route schedule is to be avoided, due to its effect on slowing travel speed and reducing vehicle productivity. However, where a large number of passenger transfers can be aided by creating enough slack time to assure successful connections, allowing a short delay may be highly beneficial.

Fifth, it should be noted that effective implementation of service improvement strategies need not imply the installation of expensive AVM equipment. While such equipment is clearly beneficial in implementing headway-based holding strategies, for example, there are many other potential strategies which are likely to be just as effective (or perhaps more so) with substantially less required investment in hardware.

Finally, it is important to emphasize the need for cooperation between transit operating authorities and municipal departments responsible for streets and traffic signals. Many of the strategies for service improvement described in this report would require agreement and joint action on the part of both agencies for effective implementation. In order to reach the point of <u>acting</u> together, it is important that they begin to <u>plan</u> together. Communication and agreement on overall goals at an early stage is vital to the success of many of the strategies which seem to be most effective in improving service reliability in transit systems.

REFERENCES

Abkowitz, M., H. Slavin, R. Waksman, L. Englisher and N. Wilson (1978). Transit Service Reliability, Transportation Systems Center, U.S. Department of Transportation, report UMTA-MA-06-0049-78-1.

Barnett, Arnold (1974). "On Controlling Randomness in Transit Operations," Transportation Science, Vol. 8, pp. 102-116.

Bisbee, E.F., J.W. Devaney, D.E. Ward, R.J. vom Saal and S. Kuroda (1968). "Dispatching Policies for Controlled Systems," Fourth International Symposium on the Theory of Traffic Flow, Karlsruhe, pp. 255-263.

Boardman, T. and W. Kraft (1970). "Predicting Bus Passenger Service Time - Part II," Traffic Engineering, Vol. 40, No. 5, pp. 36-40.

CACI, Inc. (1976). SIMSCRIPT II.5 Reference Handbook, Los Angeles.

Cooper, B.R. and R.E. Layfield (1977). Bus Priority at Traffic Signals Terminating a Contraflow Bus Lane in Walsall, Transport Road Research Laboratory, Supplementary Report 252, Crowthorne, Berkshire, England.

Cottinet, M. (1977). Prise en Compte des Autobus sur un Itineraire Coordonné, prepared for the Institut de Recherche des Transports, Arcueil, France.

Courage, K.C., C.E. Wallace and J.A. Wattleworth (1977). Effect of Bus Priority System Operation on Performance of Traffic Signal Control Equipment on NW 7th Ave., prepared for the U.S. Department of Transportation, Report UMTA-FL-06-0006.

Fishman, G.S. (1973). Concepts and Methods in Discrete-Event Digital Simulation, Wiley, New York.

Friedman, R.R. (1976). Statistical Models of the Mean and Standard Deviation of Passenger Wait Time in Urban Bus Transit, M.S. Thesis, Northwestern University Transportation Center.

Golob, T.F., E.T. Canty, R.L. Gustafson and J.E. Vit (1972). "An Analysis of Consumer Preferences for a Public Transport System," Transportation Research, Vol. 6, No. 1, pp. 81-102.

Hallam, C.E. (1977). "The Sydney Transit Lane," Traffic Engineering and Control, Vol. 18, No. 2, pp. 70-71.

Holroyd, E.M. and D.A. Scraggs (1966). "Waiting Times for Buses in Central London," Traffic Engineering and Control, Vol. 8, No. 3, pp. 158-160.

Jackson, R.L. (1974). An Evaluation of Bus Control Systems by Simulation, Transport and Road Research Laboratory, Laboratory Report 712, Crowthorne, Berkshire, England.

Jackson, R.L. (1977). Evaluation by Simulation of Control Strategies for a High Frequency Bus Service, Transport and Road Research Laboratory, Laboratory Report 807, Crowthorne, Berkshire, England.

Jacobson, J. and Y. Sheffi (1980). Analytic Model of Traffic Delays Under Bus Signal Preemption: Theory and Application, presented at the 59th Annual Meeting of the Transportation Research Board, Washington, D.C.

JHK and Associates (1975). Evaluation of First Generation UTCS/BPS Control Strategy, Volume 1, Technical Report, prepared for U.S. Department of Transportation, Federal Highway Administration.

Jolliffe, J.K. and T.P. Hutchinson (1975). "A Behavioral Explanation of the Association Between Bus and Passenger Arrivals at a Bus Stop," Transportation Science, Vol. 9, pp. 248-282.

Jordan, W.C. (1979). The Effects of Zone Scheduling on Bus Service Reliability, M.S. Thesis, Northwestern University.

Jordan, W.C. and M.A. Turnquist (1979). "Zone Scheduling of Bus Routes to Improve Service Reliability," Transportation Science, Vol. 13, pp. 242-268.

Jordan, W.C. and M.A. Turnquist (1980). Control of Service Reliability in Bus Transit Networks: Simulation Model User's Manual, Version 2.0, prepared for the Office of University Research, U.S. Department of Transportation.

Koffman, D.,(1978). "A Simulation Study of Alternative Real-Time Bus Headway Control Strategies," Transportation Research Record 663, pp. 41-46.

Kulash, D.J. (1971). Routing and Scheduling in Public Transportation Systems, MIT Urban Systems Laboratory.

Ludwick, J.S. (1975). "Simulation of an Unconditional Pre-emption Bus Priority System," Transportation Research Record 526, pp. 1-10.

McQueen, J.T. and R. Waksman (1977). "Increasing the People-Moving Capability of Shirley Highway," Transportation Research Record 626, pp. 21-26.

Michalopoulos, P.G. (1976). "Bus Priority System Studies," Traffic Engineering, Vol. 46, No. 7, pp. 46-55.

Mitre Corporation (1975). Overview of Experimental Bus Priority Systems, prepared for U.S. Department of Transportation, Urban Mass Transportation Administration.

Mohring, H. (1972). "Optimization and Scale Economies in Urban Bus Transportation," American Economic Review, Vo. 62, No. 4, pp. 591-604.

Newell, G.F. (1971). "Dispatching Policies for a Transportation Route," Transportation Science, Vol. 5, pp. 91-105.

Newell, G.F. and R.B. Potts (1964). "Maintaining a Bus Schedule," Proceedings of the Second Conference, Australian Road Research Board, pp. 388-393.

Osuna, E.E. and G.F. Newell (1972). "Control Strategies for an Idealized Public Transportation System," Transportation Science, Vol. 6, pp. 52-72.

Paine, T.F., A.N. Nashe, S.J. Hille and G.A. Brunner (1967). Consumer Conceived Attributes of Transportation, An Attitude Study, University of Maryland, Department of Business Administration.

Radelat, G. (1973). Simulation of the Operation of a Mass Transportation System: Urban Buses on an Arterial Street, Ph.D. Dissertation, Catholic University of America.

Robertson, D.I. and R.A. Vincent (1975). Bus Priority in a Network of Fixed Timed Signals, Transport and Road Research Laboratory, Laboratory Report 666, Crowthorne, Berkshire, England.

Wilbur Smith and Associates (1975). Bus Rapid Transit Options for Densely Developed Areas, prepared for U.S. Department of Transportation, Federal Highway Administration.

Sterman, B.P. (1974). Factors Affecting Reliability of Urban Bus Service, M.S. Thesis, Northwestern University.

Tarnoff, P.J. (1975). "Results of FHWA Urban Traffic Control Research: An Interim Report," Traffic Engineering, Vo. 45, No. 4, pp. 27-35.

Transport and Road Research Laboratory (1973). Bus Priority: Proceedings of a Symposium Held at TRRL, 1972, Laboratory Report 570, Crowthorne, Berkshire, England.

Turnquist, M.A. (1978). Network Reliability in Urban Transit Systems, final report prepared for the National Science Foundation under grant ENG-76-09607.

Turnquist, M.A. (1979). "Zone Scheduling of Urban Bus Routes," Transportation Engineering Journal of ASCE, Vol. 105, No. TE1, pp. 1-13.

Turnquist, M.A. and L.Bowman (1979). Control of Service Reliability in Transit Networks, Office of University Research, U.S. Department of Transportation, report DOT/RSPA/DPB-50/79/5.

Vanselow, R.G. and R.B. Sinclair (1972). "Improving Street Public Transport -- Some Possible Solutions," Proceedings of the Australian Road Research Board, Vol. 8, pp. 33-46.

Vuchic, V.R. (1969a). "Propagation of Schedule Disturbances in Line-Haul Passenger Transportation," Revue de L'UITP, Vol. 28, pp. 281-284.

Vuchic, V.R. (1969b). "Rapid Transit Interstation Spacing for Maximum Number of Passengers," Transportation Science, Vol. 3, pp. 214-232.

Vuchic, V.R. and G.F. Newell (1968). "Rapid Transit Interstation Spacings for Minimum Travel Time," Transportation Science, Vol. 2, pp. 303-339.

Wallin, R.J. and P.H. Wright (1974). "Factors Which Influence Modal Choice," Traffic Quarterly, Vol. 28, pp. 271-290.

Welding, P.I. (1957). "The Instability of Close Interval Service," Operational Research Quarterly, Vol. 8, pp. 133-148.

Wood, K. (1976). Survey of Bus Priority at Traffic Light Signals in Woking, Transportation and Road Research Laboratory, Supplementary Report 228, Crowthorne, Berkshire, England.

APPENDIX A

DERIVATION OF EXPECTED BENEFITS FROM HEADWAY-BASED HOLDING STRATEGIES

As defined in Chapter 5, the expected reduction in overall delay to all passengers from holding vehicle i for X_i minutes is:

$$E(\Delta\psi) = \frac{1}{E(H)\cdot Q} E(\xi_i) \qquad \text{(A-1)}$$

where $$E(\xi_i) = (1-b)\ QE[X_i(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1} - X_i)]\ . \qquad \text{(A-2)}$$

The quantity ξ_i is the net savings in total passenger-minutes achieved by holding the i^{th} bus X_i minutes. It is the reduction in total wait, minus the delay to passengers already on the vehicle.

We want to find $E(\Delta\psi)$ for the Prefol policy of $X_i = X_i^P$, where X_i^P is given by:

$$X_i^P = \begin{cases} \max[0,\ .5(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}^P)], & i = 1,2,\ldots,n-1 \\ 0 & , \ i = 0,n. \end{cases} \qquad \text{(A-3)}$$

We will denote $E(\Delta\psi)$ for the Prefol policy as $E(\Delta\psi^P)$, and $E(\xi_i)$ as $E(\xi_i^P)$, with X_i^P substituted for X_i, as shown in equation (A-4).

$$E(\xi_i^P) = (1 - b)\ QE[X_i^P(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}^P - X_i^P)]. \qquad \text{(A-4)}$$

We assume that we have limited our period of analysis to a steady-state situation in which the underlying distribution of headways does not change with i. Thus the expected benefits do not change with i; $E(\xi_i^P) = E(\xi^P)$ for all i. The i in this equation and subsequent ones is used only as an index and is retained as a way to identify the previous and following headways, or the previous and current holds.

Define d_i^P as the quantity that is computed at the time of each control decision:

$$d_i^P = .5(H_{i+1} - H_i - \frac{b}{1-b} H_i + X_{i-1}^P). \qquad \text{(A-5)}$$

The control rule (A-3) can then be rewritten as:

$$X_i^P = \max[0,\ d_i^P]\ , \qquad \text{(A-6)}$$

and the benefits from an individual hold as:

$$\xi_i^P = (1 - b)\ Q[X_i^P(2d_i^P - X_i^P)]. \tag{A-7}$$

The quantity ξ_i^P can also be written as $(1 - b)\ Q(X_i^P)^2$. To see this, first consider ξ_i^P when d_i^P is greater than zero (i.e., when a decision is made to hold vehicle i). It is held an amount $X_i^P = d_i^P$ and by (A-7), $\xi_i^P = (1 - b)\ Q[X_i^P(2X_i^P - X_i^P)] = (1 - b)\ Q(X_i^P)^2$.

When $d_i^P \leq 0$, $X_i^P = 0$ and $\xi_i^P = (1 - b)\ Q(0\cdot(2d_i^P - 0)) = 0$. For this case, the quantity $(X_i^P)^2$ is also equal to zero; thus, we can write:

$$\xi_i^P = (1 - b)\ Q(X_i^P)^2\ . \tag{A-8}$$

As a result, we can also rewrite the expectations as follows:

$$E(\xi_i^P) = (1 - b)\ QE[(X_i^P)^2] \tag{A-9}$$

and

$$E(\Delta\psi^P) = \frac{1}{E(H)}(1 - b)\ E[(X_i^P)^2]\ . \tag{A-10}$$

To compute $E[(X_i^P)^2]$, the approach is to first make reasonable simplifying assumptions about the probability distribution of d_i^P, $f(d^P)$. Since the distribution of X_i^P, denoted $h(X^P)$, is closely related to $f(d^P)$, it is then a matter of integration to find $E(X_i^P)^2$, and hence $E(\Delta\psi^P)$. Since

$$X_i^P = \begin{cases} d_i^P\ , & \text{if } d_i^P > 0 \\ 0\ , & \text{if } d_i^P \leq 0 \end{cases}\ , \tag{A-11}$$

we can write the density function for X^P as follows:

$$h(X^P) = \begin{cases} f(X^P) & , \text{if } X^P > 0 \\ p(d^P \leq 0) & , \text{if } X^P = 0 \end{cases} \tag{A-12}$$

where $p(\cdot)$ denotes the probability of the event $(\cdot)$. See Figure A-1 for a graphical representation of $f(d^P)$ and $h(X^P)$.

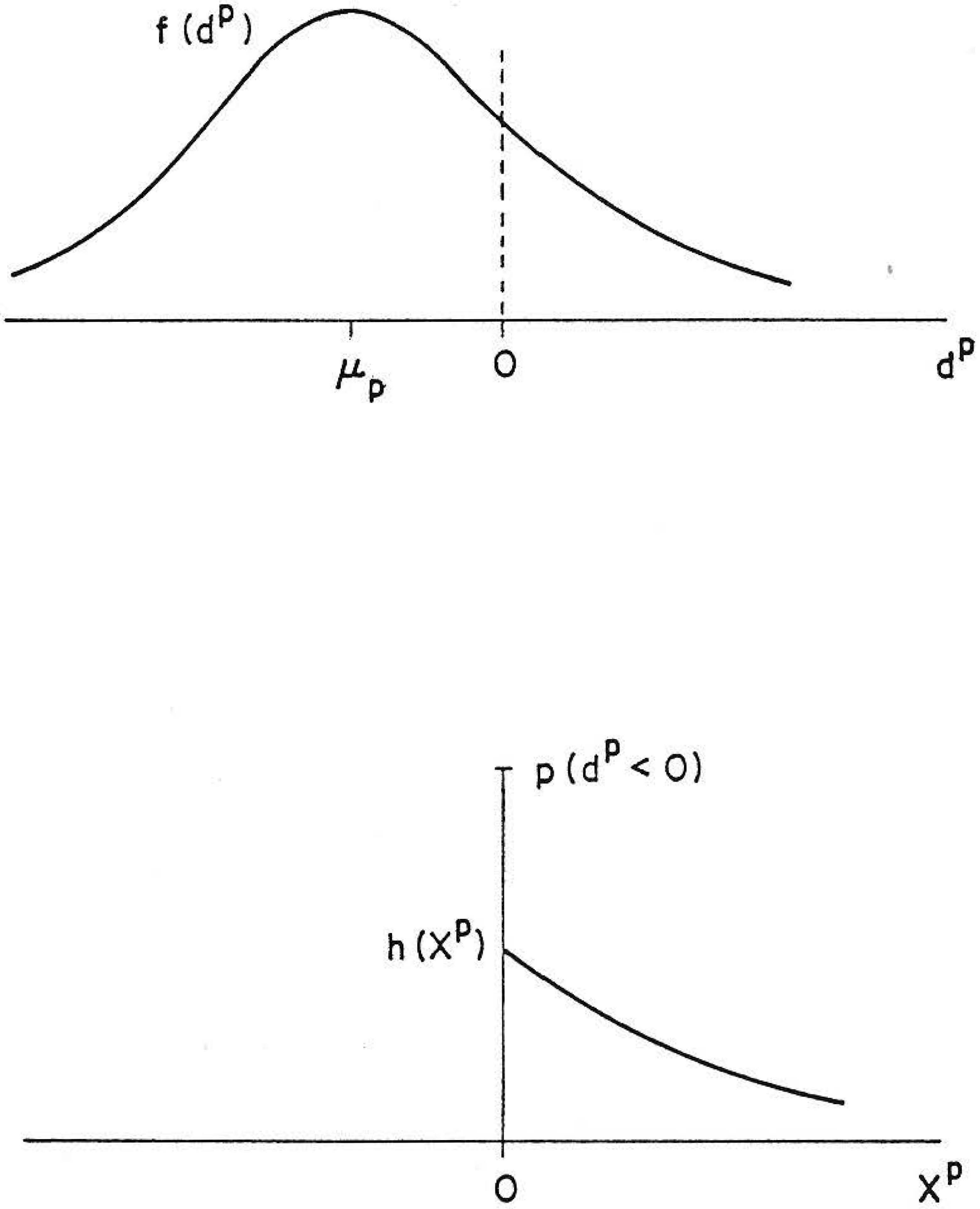


Figure A-1. Distributions of d^P and X^P.

An approximate form for the probability distribution $f(d^p)$ can be derived by making three simplifying assumptions. First, we will assume that the effect of the previous hold (vehicle i-1) on the distribution of d_i^p is negligible. Effectively, this implies an assumption that $X_{i-1}^p = 0$.

Second, we will assume that the number of passengers on the bus at the control point is always the same. This means that we can replace the term $\frac{b}{1-b} H_i$ by $\frac{b}{1-b} E(H)$.

Finally, we will assume that H_i and H_{i+1} are identically distributed random variables, and that the difference, $H_{i+1} - H_i$, is approximately normal. If H_i and H_{i+1} are identically distributed, their difference should be symmetric about zero, and Sterman (1974) presents empirical evidence that headways themselves are normally distributed. Hence, assuming that the difference between two headways is approximately normal appears quite reasonable.

Under these assumptions, we can write d_i^p as:

$$d_i^p = .5[(H_{i+1} - H_i) - \frac{b}{1-b} E(H)] \ . \tag{A-13}$$

Because the last term in (A-13) is just a constant, if $H_{i+1} - H_i$ is normally distributed, d_i^p will be also. The expected value of d_i^p is then approximated by:

$$\mu_p = .5[E(H_{i+1} - H_i) - \frac{b}{1-b} E(H)]$$

$$= - \frac{b}{2(1-b)} E(H) \ . \tag{A-14}$$

Since H_i and H_{i+1} are identically distributed, $E(H_{i+1} - H_i) \equiv 0$. Also, since (A-14) does not depend on i, but is the same for all vehicles, the subscript i in the expected value has been dropped. The distribution will be the same for all i, so the subscript is unnecessary.

The variance of d^p is:

$$\sigma_p^2 = .25 \ V(H_{i+1} - H_i)$$

$$= .25[V(H_{i+1}) + V(H_i) - 2 \ \mathrm{cov}(H_{i+1}, H_i)] \ . \tag{A-15}$$

If H_i and H_{i+1} are identically distributed, $V(H_i) = V(H_{i+1}) = V(H)$. Also we can write the covariance term as:

$$\text{cov}(H_{i+1}, H_i) = \rho\, V(H) \tag{A-16}$$

where ρ is the correlation coefficient between successive headways. Thus, we may rewrite (A-15) as follows:

$$\sigma_p^2 = .5(1-\rho)\, V(H)\ . \tag{A-17}$$

To summarize, d^P is distributed approximately normal, with mean μ_p and variance σ_p^2, given by equations (A-14) and (A-17), respectively.

We can then find $E[(X^P)^2]$ as follows:

$$E[(X^P)^2] = \int_{-\infty}^{\infty} (X^P)^2\, h(X^P)\, dX^P$$

$$= \int_{-\infty}^{0} 0 \cdot p(d^P \leq 0)\, dX^P + \int_{0}^{\infty} (X^P)^2\, f(X^P)\, dX^P$$

$$= \int_{0}^{\infty} (X^P)^2 \frac{1}{\sigma_p\sqrt{2\pi}} \exp\left(-\frac{(X^P - \mu_p)^2}{2\sigma_p^2}\right) dX^P$$

$$= (\mu_p^2 + \sigma_p^2)\, p(d^P > 0) + \frac{\mu_p \sigma_p}{\sqrt{2\pi}} \exp\left(-\frac{\mu_p^2}{2\sigma_p^2}\right)\ . \tag{A-18}$$

If we define the coefficient of variation, C_p, as:

$$C_p = \sigma_p/\mu_p \tag{A-19}$$

we may rewrite (A-18) as follows:

$$E[(X^P)^2] = \mu_p^2\left[(1 + C_p^2)\, \Phi(C_p) + \frac{C_p}{\sqrt{2\pi}} \exp(-1/2C_p^2)\right] \tag{A-20}$$

where: $\Phi(C_p)$ = probability that a normal random variable with coefficient of variation C_p is greater than zero.

Substituting equations (A-14) and (A-20) into (A-10), we obtain the following result:

$$\frac{E(\Delta\psi^{p})}{E(H)} = \frac{b^{2}}{4(1-b)} \left[(1 + C_{p}^{2})\ \Phi(C_{p}) + \frac{C_{p}}{\sqrt{2\pi}} \exp(-1/2C_{p}^{2})\right]. \tag{A-21}$$

In addition, we can substitute equations (A-14) and (A-17) into (A-19), and rewrite C_p as follows:

$$C_{p} = -\ \frac{1-b}{b} \sqrt{2(1-\ \ )}\ \frac{\sqrt{V(H)}}{E(H)} \tag{A-22}$$

Equations (A-21) and (A-22) are the results quoted in Chapter 5, as equations (5-18) and (5-19).

U.S. GOVERMENT PRINTING OFFICE 1982-0- 522-046/8461

Cut Out Along This Line

DOT/RSPA/DPB-50/91/27

"Strategies for Improving Bus Transit Service Reliability"

REQUEST FOR FEEDBACK TO
The DOT Program Of University Research

This booklet is intended to serve as a reference source for transportation analysts, planners and operators. Your comments will be reviewed by the persons responsible for writing and publishing this material. Feedback is extremely important in improving the quality of research results, the transfer of research information, and the communication link between the researcher and the user.

YES	NO	
☐	☐	Did you find the report useful for your particular needs? If so, how?
☐	☐	Did you find the research to be of high quality?
☐	☐	Were the results of the research communicated effectively by this report?
☐	☐	Do you think this report will be valuable to workers in the field of transportation represented by the subject area of the research?
☐	☐	Are there one or more areas of the report which need strengthening? Which areas?
☐	☐	Would you be interested in receiving further reports in this area of research? If so, fill out form on other side.

Please furnish in the space below any comments you may have concerning the report. We are particularly interested in further elaboration of the above questions.

COMMENTS

Thank you for your cooperation. No postage necessary if mailed in the U.S.A.

RESEARCH INTEREST

PLEASE CHECK YOUR AREAS OF INTEREST. YOU WILL BE NOTIFIED AS NEW REPORTS ARE PUBLISHED.

TRANSPORTATION

- ☐ Air
- ☐ Highway
- ☐ Maritime & Ports
- ☐ Mass Transit
- ☐ Pipeline
- ☐ Rail

ECONOMIC REGULATION

- ☐ Air Carrier
- ☐ Motor Carrier
- ☐ Railroads

TRANSPORTATION STUDIES IN SPECIAL AREAS

- ☐ Advanced Technology
- ☐ Energy
- ☐ Environment
- ☐ Facilities & Terminals
- ☐ Hazardous Materials
- ☐ Safety
- ☐ Tunnelling
- ☐ Construction, Rehabilitation & Maintenance
- ☐ Efficiency, Productivity & Innovation
- ☐ Elderly, Handicapped & Disadvantaged
- ☐ Freight, Goods Movement & Commodity Flow
- ☐ Modeling - Demand, Supply & Traffic
- ☐ Rural Development, Planning & Policy
- ☐ Urban Development, Planning & Policy

☐ OTHER (PLEASE LIST) ______________________________

____

____

U.S. DEPARTMENT OF TRANSPORTATION
Research and Special Programs Administration
Washington, D.C. 20590

NO POSTAGE
NECESSARY
IF MAILED
IN THE
UNITED STATES

BUSINESS REPLY MAIL

FIRST CLASS PERMIT NO. G-126

POSTAGE WILL BE PAID BY ADDRESSEE

OFFICE OF UNIVERSITY RESEARCH
Attn: HELEN L. ENGRUM (DPB-50)
U.S. DEPARTMENT OF TRANSPORTATION
400 Seventh Street, S.W.
Washington, D.C. 20590

PLEASE LIST YOUR NAME AND ADDRESS.

NAME ______________________________

TITLE ______________________________

ORGANIZATION ______________________________

STREET ______________________________

CITY, STATE, ZIP CODE ______________________________