APPROXIMATE, ANALYTIC PERFORMANCE MODELS OF

INTEGRATED TRANSIT SYSTEM COMPONENTS

by

CHRIS T. HENDRICKSON

B.S., Stanford University
(1973)

M.S., Stanford University
(1973)

B.Phil., Oxford University
(1975)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(MAY, 1978)

Signature Redacted

Signature of Author ....
Department of Civil Engineering, May 1978

Signature Redacted

Certified by ....
Thesis Supervisor

Signature Redacted

Accepted by ....
Chairman, Department Committee

APPROXIMATE ANALYTIC PERFORMANCE MODELS

OF INTEGRATED TRANSIT SYSTEM COMPONENTS

by

CHRIS T. HENDRICKSON

Submitted to the Department of Civil Engineering on May 23, 1978, in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the Massachusetts Institute of Technology.

ABSTRACT

Explicit models of transit system performance are useful as an aid to system design and as a means to gain insight into the system's performance. This thesis develops performance models of scheduled, fixed and flexibly routed transit services. Flexibly routed transit services of this type use shared vehicles which are routed in response to the locations of individual patron demands. The models apply to transit services which operate in a service area or on a route with scheduled vehicle departures from a depot or transfer point. The use of these models is illustrated by comparisons of the performance of vehicle deployment, control and routing alternatives.

The models developed are approximate, analytic representations of steady state system performance. Applications may be performed manually or by computer. Models are developed using a deterministic queueing system framework with fluid approximations to the system's service and arrival processes. Corrections to account for stochastic phenomena are then superimposed upon the deterministic models. Such corrections are developed for the stochastic nature of the arrival and service processes using diffusion approximations and truncated probability distributions. The resulting models are relatively simple in structure, inexpensive to use, and relatively accurate, particularly in the case of flexibly routed systems, in which service becomes more efficient as patronage increases.

Thesis Supervisor: Nigel H.M. Wilson
                   Associate Professor
                   Department of Civil Engineering

ACKNOWLEDGEMENTS

-4-

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION AND SUMMARY

1.1 Introduction

During the review or design of transportation systems, it is often

useful to construct explicit models of a proposed system in order to

facilitate comparison of alternatives or to gain insight into the system

performance. This thesis develops models of feeder transit service per-

formance, concentrating upon flexibly routed transit services. The use

of these models is illustrated by applications to alternative vehicle de-

ployment, control and routing strategies for integrated transit systems.

The process of transit system design has some general characteristics

which suggest the requirements for models of system performance. Avail-

able resources of both time and money are limited in any analysis. Con-

sequently, models should be inexpensive to use and cannot have excessive

data requirements. Secondly, models must be relevant to the objectives

of the design process; model predictions should be reasonably accurate

and sensitive to relevant policy alternatives. Since new alternatives

may be generated during the design process, models should also be flexible,

with the potential for rapid modification. Finally, the formulation of

a model should be simple, so that designers may easily understand the li-

mitations of the model and, by understanding the causal relationships,

gain insights into the system operation. The approximate, analytic models

developed in this thesis are intended to satisfy these requirements.

Transit system design as an application area is a subject of con-

siderable intellectual and practical interest. Until the post war period,

public transit in the United States had developed as a regulated private industry. Analysis of the impact of service or efficient operation of systems was sporadic and often inconsistent. Under the influence of fixed investments and regulations, routes and operating policies were only rarely changed once service was inaugurated.

A number of trends in the past ten years have emphasized the need for changes in transit operations and the importance of system design. The most important of these trends is that of continuing and increasing budget deficits among systems. Declining patronage, expanded service and rapidly increasing costs are generally responsible for this worrisome trend. Faced with increasing deficits, operators have a strong incentive to investigate more efficient operating policies.

While deficits have been increasing, there are also factors which encourage the introduction of new services. Shifts in population to suburban or low density areas have led to a demand for service in areas for which traditional fixed route service is prohibitively expensive. There has also been an increasing public committment to the provision of transit services to the elderly, handicapped and other groups without ready access to automobiles. Finally, transit systems are often thought to have a significant impact on a variety of social concerns, including energy use, environmental quality and congestion. This impact may be directly due to vehicle operation or indirectly via an influence upon private automobile use and the pattern of social activity. Transit operators are expected to consider these wider public concerns in planning studies.

These trends have resulted in an environment in which new system designs and operations are imperative. At the same time, the range of

possible designs has expanded, especially with the consideration of para-transit modes such as taxicabs, jitneys, dial-a-ride and others. The performance of these alternatives is the principal subject area of this thesis.

This chapter is intended to provide a summary of the main results contained in the thesis. The next section describes the types of services and problems analyzed in this thesis. Following this discussion, Section 1.3 describes the modelling methodology used to construct performance models, including a discussion of alternative modelling methods. Section 1.4 lists some results from application of the performance models. Section 1.5 provides an introduction to the following chapters. Section 1.6 provides a brief summary of the models developed here. Finally, Section 1.7 contains a glossary of notation. These latter two sections (Secs. 1.6 and 1.7) are intended as reference sections, not as introductions for the general reader.

1.2 The Problem Addressed

This thesis develops performance models of transit services in which vehicles have scheduled departures from a depot or transfer point. Such services may be operated with fixed routes or be flexibly routed in response to the locations of patrons' origins and destinations within a service area. While most patrons in such services are expected to travel to or from the depot, some trips may both originate and end within the service area. In this thesis, _feeder_ services will denote transit services in which all patrons either originate at or are destined for a single depot. Services in which some intra-zonal trips are served will be called _zonal_ services.

One example of these transit services is a feeder service in which buses provide access to a line haul service such as a commuter rail line. Such feeder service may consitute the local circulation component of regional, integrated transit systems. In the Ann Arbor, Michigan transit system, flexibly routed vans are operated in zonal service, with most patrons transferring from the vans to fixed route bus lines. It is also possible to use zonal services as the only transit service in a region, with service oriented towards a central depot.

Performance models are intended to provide estimates of the important level of service attributes and resource requirements of a particular service. Thus, the problem addressed in this thesis is that of predicting the level of service provided to patrons and some other direct impacts of particular transit system components. The level of service attributes which are estimated include the expected waiting time, the expected riding time, and the variances of these times. For patrons to be delivered in the service area, waiting time is defined as the time between arriving at the depot and actually boarding a vehicle. For patrons requesting collection from the service area, the waiting time is defined as the time between the service dispatcher or operator becoming aware of the request and the patron boarding a vehicle.* Riding time is simply the time between a patron boarding and leaving a vehicle. The expected travel time on a service is defined as the sum of riding and waiting time of a random

*Due to congestion on telephone lines and other causes, delays may occur between the time a potential patron wishes to request a collection and the time the service dispatcher or operator is made aware of the request. This delay is usually minor; in the Ann Arbor transit system, this delay averaged 2.75 minutes during the busiest 4 hour period of the day [60]. Advanced reservation requests experience no delays of this type.

patron. Vehicle miles and hours of operation can also be estimated or calculated; these resource requirements may then be used to estimate system costs, fuel consumption and other impacts. However, analysis of the costs or benefits of particular service options is beyond the scope of this thesis.

To enable estimation of these level-of-service and impact measures, input parameters consisting of the patronage volume using the service, the system operating policies, the available vehicle fleet size and various service area characteristics are used. The models are based upon the assumption that these input parameters are constant during the period of analysis.* Thus, illustrative comparisons are made only between the estimated level of service of alternative systems with constant input parameters; generally, such comparisons will be made between services with identical vehicle fleets, patronage volumes and service area characteristics.

While fixed route services are discussed and modelled, this thesis concentrates upon models of flexibly routed transit services. While fixed route services are more common than flexibly routed services, there are a number of reasons for studying the flexibly routed service options in detail. Flexibly routed zonal services are a basic component of several existing integrated transit systems, such as the Ann Arbor, Michigan and the Regina, Ontario transit systems. Flexibly routed feeder services have not received a greate deal of attention in the professional literature. Moreover, the performance of flexibly routed service is somewhat more

_____

*With the use of the models in deterministic simulation, this assumption can be relaxed.

complicated than fixed route service, so the development of an adequate

model of flexibly routed service requires greater attention than does a

comparable model of fixed route service. Finally, an adequate performance

model of flexibly routed feeder services may be extended to van-pooling

(in which commuters regularly travel together in a van), shared ride taxi

service to airports, and other serivces.

A variety of options exist for operating scheduled, flexibly routed

services. Vehicles may be restricted to specific zones, so an operator

must decide how (and if) to district a service region: the models de-

veloped here estimate system performance in any particular zone. The num-

ber of vehicles to be operated in each zone is another service option. In

some situations, capacity of vehicles may be a planning option. The veh-

icle's schedule or time between visits to the depot is yet another operat-

ing option. Communication between a driver and a dispatcher concerning

patron demands may occur at the depot or continuously. Finally, service

may be offered in phases, in which vehicles first deliver from the depot

and then collect patrons in the service zone, or unphased, in which deli-

very and collection stops are interspersed. With phased service, an ope-

rator may also schedule some idle time between the delivery and collection

pahses. The performance model of flexibly routed transit service developed

here in a general analytic model which can be applied to any combination

of these operating options.

1.3 Approximate, Analytic Performance Models

The classic examples of performance models in transportation are

models of delays due to road traffic flow or at isolated signalized inter-

sections (see, for example, [12] or [70]). In all performance models, both

the demand for service and the system characteristics are constant. Performance models may be regarded as very short term models of transportation supply, since it is assumed that the system operators cannot or do not change the system immediately in response to changes in demand or the environment. Since making changes in a transportation system generally involves substantial costs, the performance of a given system over a period of time is of considerable interest to operators.

It is possible to use performance models alone or in conjuction with models of travel demand to find the equilibrium or expected level of demand and system attributes. An equilibrium solution satisfies the necessary condition that the level of demand attracted to the system is equal to the level of demand which causes that particular level of service or attractiveness. To find the equilibrium, an analyst must make some assumption about the response of travel demand to changes in the level of service provided. The problem of identifying equilibrium demand and level of service is especially important in the case of integrated transit systems since travel demand and the level of service depend upon one another.

It is also possible to use a performance model as part of the formulation of an optimization problem. For example, performance models of road links may be used in the mathematical programming formulation of the traffic assignment problem. Unfortunately, the performance of integrated transit systems is generally a non-linear function of demand and system characteristics. Moreover, transit systems are intended to achieve a multitude of social objectives. Consequently, the use of optimization or mathematical programming for transit system design involves multiple

-17-

objectives and non linear constraints. However, simplifying assumptions may be introduced to make design problems tractable in a mathematical programming framework, yielding good but not necessarily optimal designs as a preliminary screening technique. For example, least cost zone sizes i scf 'ules m··· be identified for flexibly routed feeder services, but only under the assumption of fixed demand and values of time.

Various techniques for constructing performance models are available, including Monte Carlo simulation, econometrics and the theory of queueing systems of Markov processes. Monte Carlo simulation can be used for any system. However, simulation models tend to be expensive to use, relatively inflexible, and typically yield little insight into the causal factors of system performance. Econometric or empirical models are based upon simple analytical relationships, with parameter values calibrated to fit observations of system performance or the results of simulation models. Econometric models tend to be convenient to use but are only valid within the range of calibration. Models based upon queueing theory provide exact predictions of system performance, given the accuracy of modelling assumptions. However, queueing models become intractable even when applied to fairly simple transportation systems. Consequently, exact queueing models tend to be difficult to solve and inflexible in applications.

Approximate analytic models are intended to offer an attractive intermidiate methodology between econometric or empirical models and the techniques of queueing theory. These models are constructed in two stages, first developing a deterministic model of system behavior and then superimposing stochastic corrections upon the deterministic model. The deterministic model may be based upon engineering relationships, deterministic

-18-

queueing theory or the expectations of the performance of system components. For example, our deterministic model of flexibly routed transit service is based upon the expected length of tours among patron origins and destinations. However, modifications to the deterministic model should be introduced to account for the most important stochastic effects in service operation. Such effects may arise from variations in the arrival process, vehicle speeds, patron boarding, or service area characteristics. For example, the number of patrons to arrive in a given time period is a random variable. Due to the resulting variability in tour length and to constraints such as the vehicle capacity, actual flexibly routed transit service deteriorates from the deterministic case. A modification may be superimposed upon the deterministic model to capture this effect.

This methodology offers several advantages. The resulting models are relatively simple in structure. Since the models are based upon an analytic formulation, the designer may obtain insights into system performance by examination of the model's equations. The models are also relatively inexpensive to use; the models developed in this thesis may be applied with only the aid of an electronic calculator or slide rule. Finally, quite complicated systems may be successfully modelled. For example, flexibly routed transit service is characterized by a very complicated service process which prohibits the solution of exact models of service performance. But the approximate, analytic model of such services is relatively easy to use.

This approach is not unique to this thesis. Much of traffic flow theory was developed by using approximation techniques. Newell [86] has emphasized the use of approximations in applications of queueing theory.

However, applications of this technique to transit systems have been rare.

The models developed in this thesis are based upon a deterministic queueing framework and continuum approximations, neither of which have been used extensively to study transit systems. The results of the modelling effort are encouraging. By comparison with simulation experiments, models of flexibly routed feeder service give quite similar predictions. In an application to an existing system, the model predictions were generally statistically indistinguishable from the observed level of service data.

1.4 Application Results

The basic product of this thesis consists of the models of transit service performance which are summarized in Section 1.6. These models may be applied in any specific case and illustrate the usefulness of the modelling methodology discussed earlier. An example of their application to the Ann Arbor Transit System is described in Section 5.2. In addition, examples and applications are presented during the course of this thesis which indicate some of the characteristics of integrated transit service. This section is intended to summarize the conclusions arising from these experiments. Observations concerning flexibly routed feeder service are presented first, followed by conclusions related to integrated transit service.

1.4.1 Flexibly Routed Feeder Services

Flexibly routed feeder service consists of public transit service in which all patrons originated at or are destined for a single depot and vehicles are routed among the specific patron origins and destinations.

- The average amount of vehicle travel required per patron or stop on a

vehicle tour decreases as demand or stop density increases.

The basic component of service in flexibly routed transit systems consists of vehicle tours among scattered origin or destination points. Vehicle travel time includes the wait for patrons to board or exit the vehicle and the time required for travel between vehicle stops. As the demand or stop density increases, successive stops are closer together, with the expected distance between stops decreasing approximately in proportion to the inverse of the square root of demand density. Consequently, the vehicle travel time required per patron on a tour decreases as demand density increases.

In the terminology of queueing theory, flexibly routed feeder service has a state dependent service process in which the efficiency of service (i.e. the number of patrons served divided by the length of vehicle tours) depends upon the number of patrons requesting service.

- The expected travel time on a system increases as patronage volume increases.

In common with classic queueing systems, the expected travel time on a flexibly routed feeder service increases as the volume on the system increases, even with a variable vehicle schedule. This phenomenon may be related to the characteristics of vehicle tours. As demand density increases, more stops are inserted on the vehicle tour and patrons must endure more detours. Eventually, it becomes advantageous to have patrons wait while delivering vehicle occupants, due to constraints on vehicle capacity or consideration of the vehicle occupants' desires for shorter rides. This phenomenon is in contrast to fixed route transit service, in which the level of service

is generally insensitive to the level of demand. Figure 5.5 illustrates the expected travel time (including wait time) through one flexibly routed feeder service

- In relatively uncongested services with fixed schedules, flexibly routed feeder services should be operated in phases, with vehicles first collecting and then delivering patrons.

  The phased service has the advantage of reducing expected riding time, at the expense of increasing the required amount of vehicle travel compared to unphased service. With more vehicle travel required, vehicles can serve fewer patrons in a given time period and, barring other changes, the efficiency of service or vehicle productivity would decline. However, for steady state operation, the number of patrons entering the system must equal the number of patrons leaving, so vehicle productivity must remain constant. As noted above, the vehicle travel required per patron on a tour declines as the density of stops increases. Thus, as the number of patrons waiting for service increases, the efficiency of tours may increase sufficiently to offset the decline in efficiency due to phased service operation. However, the expected waiting time increases in this situation. As patronage volume on the service increases (so that the service becomes more congested), the increase in waiting time exceeds the decrease in riding time due to phased service operation.

- With sufficient vehicle capacity available, a schedule for flexibly routed feeder service exists which minimizes patron travel time.

  Varying vehicle headways (ie. the time between vehicle departures

Figure 5.5 : Expected Travel Time in a Flexibly-Routed, Scheduled and Phased Feeder Service*

A = 7.07 sq. mi.
b_d = .2 min.
b_g = 1.0 min.
v^g = .25 mi./min.
R = 3.50 min.
r = 1.27
S = ∞
α = 1.0
$\lambda_g = \lambda_d$
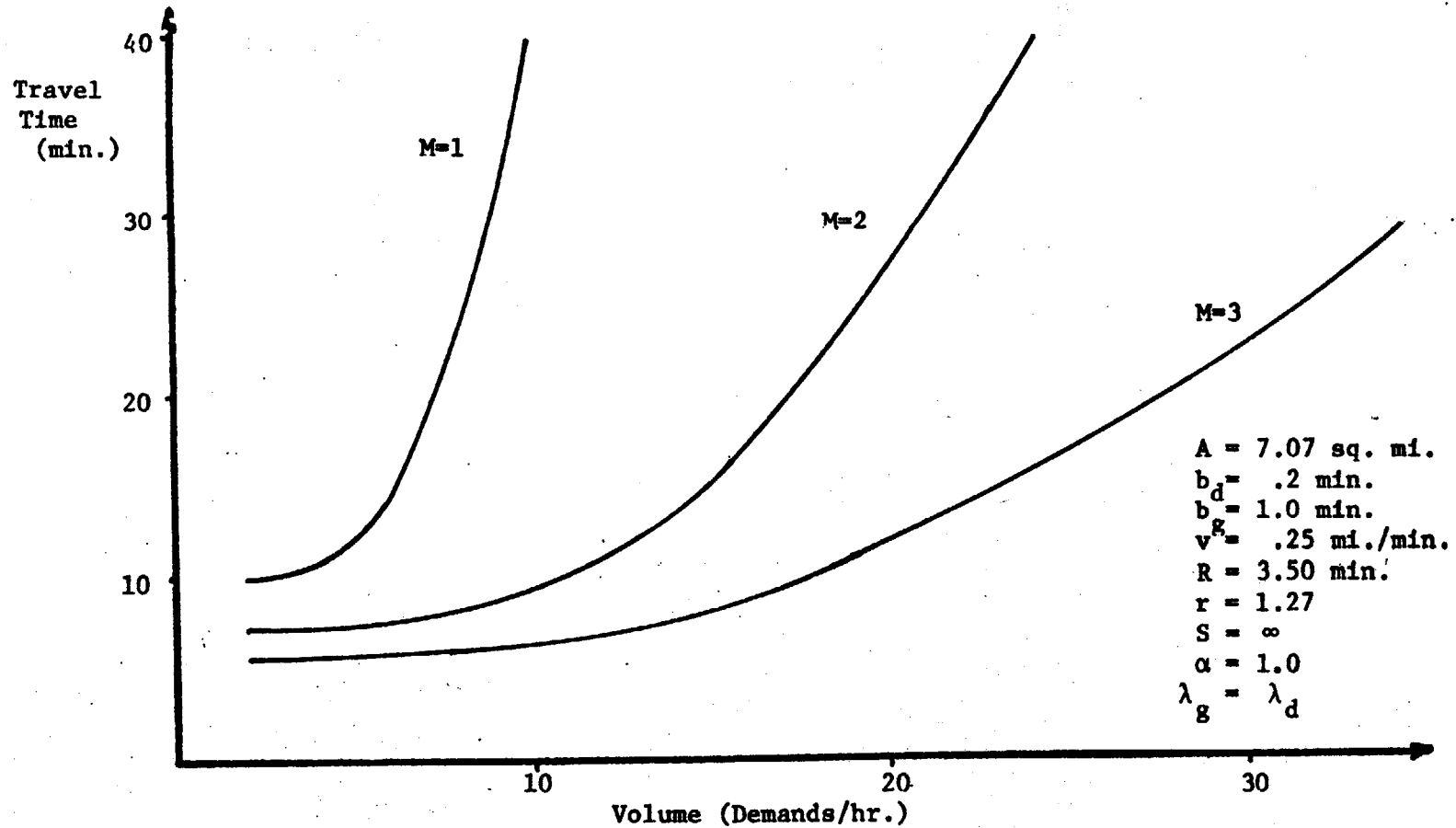
*Cycle Length Optimized to Nearest Minute, M= number of vehicles in service

from the depot) has a significant effect upon the expected travel time

of patrons in flexibly routed feeder services. Within the range of

feasible schedules, the expected travel time is everywhere convex with

respect to the time between a vehicle's visits to the depot (denoted

as the cycle time in this thesis). As a result, a single schedule

exists which will minimize the expected travel time.

- Vehicles in feeder service should be deployed into separate service
  zones or offset in time between visits to the depot.

    This is a rule to insure service with the lowest possible travel

  time, given the patronage volume to be served and the vehicle resources

  available. Offset feeder service is similar to conventional transit

  service in that a number of vehicles operate on one route (or in one

  area) with a scheduled headway between the vehicle arrivals at a stop.

  For flexibly routed services, an operator may offset vehicle stops at

  the depot in the same manner as in conventional fixed routed services.

  Alternatively, an operator may divide a large service area into smaller

  zones, with service operated independently in each zone. The choice be-

  tween offset service or area districting depends upon the particular

  characteristics of the area and the objectives of the operator.

- In constructing vehicle tours, it is desirable to consider both the veh-
  icle travel distance and the level of service provided users.

    In some situations, there exists a tradeoff between minimizing veh-

  icle travel distance and minimizing the user's riding time. In delivery

  tours, for example, patrons' riding time may often be reduced, at the ex-

  pense of greater vehicle travel, by changing the order in which stops

are made. In one series of experiments, vehicle travel distance could

be reduced 5% from manual transit routing, but at the expense of in-

creasing the patrons' travel distance by approximately 10% (Sec. 4.4).

1.4.2 Integrated Transit Service

Integrated transit services have a variety of service components,

which may include flexibly routed or special services such as express bus

service. This thesis does not treat the performance of such systems in

great detail, but it does suggest a means of analysis. From the discus-

sions in Chapter 6, a few conclusions may be drawn:

- At high patronage volumes, fixed route service results in lower expect-

  ed travel times than does flexibly routed service.

  Fixed route services can be more effective at high volumes because

  the time spent in access to the fixed route stops (generally by walking)

  is not incurred by the transit vehicle and its passengers, but only by

  individual patrons. With lower demand densities or higher access costs,

  flexibly routed service becomes more desirable. Fig. 6.5 illustrates

  one comparison between flexibly and fixed route feeder services in a

  region of eight square miles.

- Structured or zonal flexibly routed services can have expected travel

  times which are comparable with area-wide dial-a-ride services, even

  with randomly distributed origins and destinations. As origins or des-

  tinations become more concentrated, structured services become increas-

  ingly more desireable.

  Dial-a-ride service is a flexibly routed transit service in which a

v = .25 ml/min.
Cycle = 60
Offset = 2
A = 8 sq. mi.
b_d=.5 min.
b_g=1 min.
R =3 min.
αr= 1.27
I = C
S = 50

Travel Time (min.)

M=4

M=8

M=12

Flexible Routing
Fixed Route
(4 vehicles)

Volume (Demands/hr./sq. mi.)

Fig. 6.5: Expected Travel Times of Fixed and Flexibly Routed Services

-26-
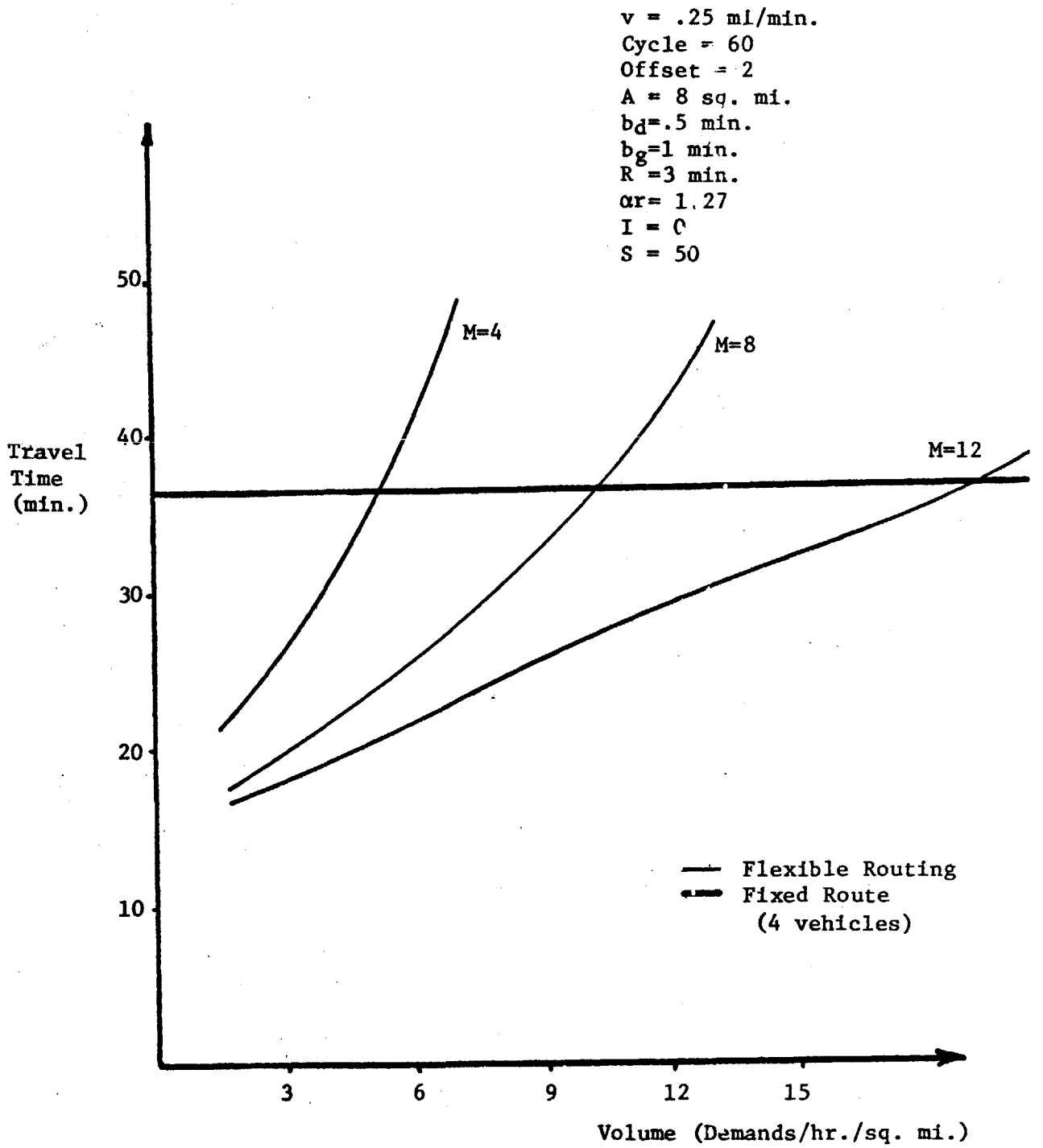
vehicle will transport a patron anywhere is a service area. A structured, flexibly routed transit system might consist of a series of zonal transit services in which patrons who wish to transfer between zones must go to a central transfer point. By restricting vehicles to particular zones, tours may be made more efficient and faster service offered with the same size fleet in some situations, but at the cost of increasing the number of transfers made on the system.

- With a system of line haul and feeder transit services, a line haul and feeder service schedule exists which minimizes patron travel time.

  This result follows from a similar result concerning the existence and uniqueness of a minimum travel time schedule for an isolated feeder service. Similarly, a minimum travel time schedule for a combined system must exist.

1.5 Organization of the Thesis

As noted above, the bulk of this thesis is concerned with the development of performance models of flexibly routed transit services. As an initial step, Chapter 2 is devoted to a review of existing models and a discussion of the use of performance models. Chapters 3 and 4 develop a performance model of flexibly routed transit service. Chapter 5 contains applications of this model. Chapter 6 discusses the performance of fixed route and integrated transit systems. Finally, Chapter 7 contains notes on the use of the models and discusses further research topics.

Following this summary chapter, Chapter 2, "Performance Models in the Design of Transit Systems" is concerned with the relevance of system

performance models to transit system design. Performance models may be used in comparative analyses, heuristic searches for good designs, evaluation studies or explicit optimization problems. In applications to transit systems, the effects of travel market equilibrium and the existence of multiple objectives should be considered. Chapter 2 formulates alternative frameworks for the use of performance models, defines various service alternatives which might be used in an integrated transit system, and reviews existing models of transit system performance and design.

In Chapter 3, an approximate, analytic performance model of scheduled, flexibly routed transit services is developed. The model uses a deterministic queuing framework with continuum approximations and corrections for important stochastic behavior. As an initial step, a simple expression for the expected tour length among a set or subset of randomly distributed points is derived in Section 3.2, based upon a next-nearest-point vehicle routing strategy. With this expression, a performance model of feeder services in circular service regions for any demand level, vehicle fleet size or area size is developed.

The model developed in Chapter 3 assumes uniformly distributed demands, next-nearest-point vehicle routing and a circular service area. In Chapter 4, these critical modelling assumptions are discussed and tested by means of Monte Carlo simulation experiments. In general, the feeder service model is found to be fairly robust, in that predictions based upon these assumptions are fairly accurate, even though the assumptions are not strictly correct. The effect of area shape on system performance is

found to be small, but can be significant. A technique for modifying the
feeder service model for cases of irregular service regions is presented.
Considering non-uniformity of demand, simulation experiments indicate that
even relatively extreme demand density gradients over the service area
do not substantially alter observed tour lengths. However, to enhance
the accuracy of the model, a more accurate means of calculating the dis-
tance from a depot to the first (or last) stop on a vehicle tour is sug-
gested. Finally, the next-nearest-point vehicle routing strategy is com-
pared with manual routing and minimum tour length routing algorithms.
The expression based upon next-nearest-point routing is found to be a
fairly good predictor of tour lengths even when vehicle routing is done
under these alternate strategies.

Chapter 5 presents a validation and some applications of the flex-
ibly routed feeder service model. The model is found to reproduce the
results of simulation experiments quite well and, in one test application,
gives relatively accurate predictions of average riding and waiting time
even without elaborate local calibration. Following these validation
experiments, the response of predicted system performance is discussed
as various input parameters are altered or as stochastic correction
terms are omitted.

Finally, Chapter 6 discusses the performance of fixed route and in-
tegrated transit service. In this chapter, an approximate, analytic model
of fixed route transit service performance is developed. Applications of
the fixed and flexibly routed models are made for isolated areas and
region-wide service. This chapter is primarily intended to indicate the
types of analysis which are possible, since a comprehensive treatment of

transit design is beyond the scope of this thesis.

A summary of the models developed and a glossary of notation appear in the following sections of the current chapter.

1.6 Model Summaries

The summaries appearing below are intended to provide a convenient reference of model assumptions, options and equations. Notation has been summarized in the following section (Section 1.7). In Section 1.6 1, the expression for expected tour lengths is presented which is then used in all the following flexibly routed models. Following this, models of feeder and zonal flexibly routed transit service are summarized. Finally, a model of fixed route service is presented. Equation and Section numbers provide references to derivations in the text.

1.6.1 Expected Tour Length

The expression below represents the expected tour length from a randomly located point through n of N points, without returning to the origin:

$$d \approx \alpha r \sqrt{A}(\sqrt{N+0.5} - \sqrt{N-n+0.5}) \qquad (3.6)$$

Assumptions of next-nearest-point vehicle routing, uniformly distributed points, regular areas, and continuum and other approximations are used in

its derivation (Sec. 3.2). Using these assumptions, simulation experiments indicate that the tour length expression is within 5% of observed average tour lengths with between 2 and 10 stops   (Sec. 3.2).   With tours among at least 3 non-uniformly distributed points in which the density gradient of points was not extreme, the tour length expression was also within 5% of observed tour lengths (Sec. 4.3). The factor α represents a correction term for irregular area shapes. Values of α are summarized in Table 4.4 for various shapes which may be used to approximate actual service regions.

The number of stops which can be made from a pool of N stops in a time t is:

$$n \approx \min\{N; \quad N+0.5-(-(\frac{vt}{\alpha r \sqrt{A}}) + \sqrt{N+0.5})^2\}$$

which may be derived algebraically from Eq. 4.6.

1.6.2 Flexibly Routed Feeder Services.

These transit services utilize shared ride vehicles and serve patrons who are either destined for or originate at a central depot or transfer point. Model options may be summarized by the vector H:

$$H = H(P, M, L, I, S, E)$$

where P is the option of operating in phased (with collection and distri-
    bution separated) or unphased (with collection and distribution
    interspersed in one tour) service,

 M is the number of vehicles operating in a service area and visiting
    the depot separately,

TABLE 4.4

Tour Length Expression Parameters*

	α	Grid Street Network ($r\alpha$)
Circle	1.01	1.29
Square	1.02	1.30
Rectangle		
3 x 2	1.04	1.31
2 x 1	1.07	1.33
3 x 1	1.14	1.37
Circular Sector[†]		
θ = 15°	1.3	1.6
30°	1.0	1.3
45°	1.0	1.3
60°	1.0	1.3
75°	1.0	1.3
90°	1.0	1.3
Isoceles Triangle[†]		
θ = 15°	1.4	1.5
30°	1.2	1.3
45	1.0	1.3
60	1.0	1.3
75	1.1	1.3
90	1.1	1.4

* from Eq. 4.5:

$$d \approx (.5r\sqrt{A} + E[d_r|A]) \, (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx (.5r + E[d_1|A=1])\sqrt{A}(\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx (.5 + E[s])r\sqrt{A} \, (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx \alpha r\sqrt{A} \, (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

where $\alpha = (.5r + E[d_1|A = 1])$

Values of α are calculated from the results in Table 4.1.

[†] θ is the angle between radii or equal length sides.

C is the vehicle cycle length or time between a vehicle's scheduled

   visits to the depot,

I is the (scheduled) idle time between collection and delivery,

S is the vehicle capacity, and

E is the option of assigning patrons for collection at the

   beginning of the collection period or during

   the stop at the depot.

These six option parameters may be used to describe a particular flex-

ibly routed feeder service.

In addition to the service options, input parameters are also re-

quired for the models, summarized by the vector P:

$$P = P(A,r,v,\lambda_d,\lambda_g,b_d,d_g,\gamma,R,\alpha) \qquad (3.60)$$

which includes area size (A), the route factor (r), average vehicle speed

(v), demand rates for delivery and collection (λ_d,λ_g), average boarding

times (b_d,b_g), average number of patrons travelling together (γ), ren-

dezvous time for the depot (R), and an area shape parameter (α).

The models result in a vector of level of service and impact measures:

$$J = J(T_g^w, T_d^w, T_d^r, T_g^r, \sigma, T\,T, VMT)$$

where the first four elements are the expected time of waiting at home,

waiting at the depot, riding for delivery, and riding for collection

$(T_g^w, T_d^w, T_d^r, T_g^r$ respectively). The vector σ is the variances of these

four travel time components. TT is the sum of the expected waiting and

riding time of a random patron. Finally, VMT is the vehicle miles of travel

per hour of system operation.

Rendezvous Time (R)

During the rendezvous period, R, the vehicle travels to the depot, spends a certain amount of layover time there (L) and then returns to the service region. The time spent travelling to or from the service region may be estimated as:

$$R-L = \frac{2(e + a/3)}{v} \qquad (3.11)$$

or more accurately as:

$$R-L = \frac{d^o_{n_d} + d^o_{n_g}}{v} \qquad$$

where d^o_i is the expected distance from the depot to the nearest of i points (Sec. 4.3). The layover time L should be set with reference to the variability of transfer vehicles and typically lies between 1 and 5 units (Section 6.3).

Phased Service

The model is based upon the tour length expression above. Vehicles are assumed to operate on a schedule from a depot or transfer point. Vehicles first deliver all patrons within the service area, then collect as many patrons as possible before returning to the depot.

Necessary conditions to insure the feasibility of steady state operation require sufficient vehicle capacity for both delivery and collection:

$$\lambda_d \cdot C/M\gamma \leqslant S \qquad (3.10)$$

$$\lambda_g \cdot C/M\gamma \leqslant S \qquad (3.15)$$

In addition, sufficient time must be available to at least board the number

of patrons who arrive during a cycle:

$$b_g \lambda_g . C/M\gamma \quad < \quad (C-D-R) \tag{3.16}$$

The model is solved in stages, first solving for the distribution time and the available collection period, then a calculation of the steady state pool pick-up size and finally, calculation of the output vector.

1. Distribution Time

$$n_d \sim \lambda_d . C/M\gamma \tag{3.8}$$

$$D \simeq n_d b_d + \frac{\alpha r \sqrt{A}}{v}(\Delta\sqrt{n_d+0.5} - \sqrt{0.5}) \tag{3.23}$$

with

$$\Delta = 1 - \frac{n_d}{8(n_d+0.5)^2}$$

$$u^* = n_d\gamma - S + (S - n_d\gamma)\Phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}}) + \tag{3.42}$$

$$\sqrt{n_d\gamma} \ \phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}})$$

2. Collection Phase

$$G \equiv C - R - D \tag{3.12}$$

$$n_g = \lambda_g . C/M \tag{3.13}$$

$$G' = \min\{G; \ n_g b_g + \frac{\alpha r \sqrt{A}}{v}(\sqrt{n_g+0.5} - \sqrt{0.5})\} \tag{3.19}$$

$$I = \max\{G - G' - 3\sqrt{G'} \ ; \ 0.\} \tag{3.43}$$

where we assume that a .99 probability of collecting all passengers is desired.

3. Steady State Pickup Pool Size

if $G' < G$ then $x^{*'} = n_g$, otherwise

$$x^{*'} = x^* + \max\{0; n_g - Y', n_g - Z'/\gamma\} \tag{3.37}$$

$$x^* = [\max\{0; \frac{0.5 + n_g - k^2}{2k}\}]^2 + n_g \tag{3.20}$$

$$k = \frac{v}{\alpha r \sqrt{A}} [G - n_g b_g]$$

$$Y' = x^* - (x^* - n_g)\Phi(\frac{x^* - n_g}{\sqrt{n_g}}) - \sqrt{n_g}\phi(\frac{x^* - n_g}{\sqrt{n_g}}) \tag{3.38}$$

$$Z' = S - (S - \gamma n_g)\Phi(\frac{S - \gamma n_g}{\sqrt{\gamma n_g}}) - \sqrt{\gamma n_g}\phi(\frac{S - \gamma n_g}{\sqrt{\gamma n_g}}) \tag{3.39}$$

4. Delays

$$E[T_g^w] = \frac{x^{*'}}{\lambda_g} - \frac{C}{2M} + \frac{G'}{2} \text{ for assignment when col-} \atop \text{lection begins} \tag{3.46}$$

$$\text{var}(T_g^w) = G^2/12 + C^2/M^2 (.08 + (\frac{x^{*'}}{n_g})^2 - \frac{x^{*'}}{n_g}) \tag{3.51c}$$

$$E[T_g^w] = \frac{x^{*'}}{\lambda_g} - \frac{C}{2M} + \frac{G'}{2} + C - G \text{ for assignment at the depot}$$

$$E[T_g^r] = \frac{R + G'}{2} + (G - G' - I) \tag{3.47}$$

$$\text{var} (T_g^r) = E[T_g^r]^2/12$$

$$E[T_d^r] = \max\{b_d + \frac{52\alpha r\sqrt{A}}{v} ; \frac{D}{2}\} + R/2 \qquad (3.49)$$

$$\text{var}(T_d^r) = E[T_d^r]^2/12$$

$$E[T_d^v] = \frac{u^*}{\lambda_d} + C/2M \quad \text{for random arrivals}$$

$$= u^*/\lambda_d \qquad \text{for transfers} \qquad (3.50)$$

$$E[ST] = \frac{\lambda_g}{\lambda_g + \lambda_d}(E[T_g^v] + E[T_g^r]) + \frac{\lambda_d}{\lambda_g + \lambda_d}(E[T_d^v] + E[T_d^r]).$$

5. Resources Consumed

$$\text{VMT} \approx \frac{C}{60} (\min \{\alpha r\sqrt{A}(\sqrt{n_d + 0.5} + \sqrt{n_g + 0.5} - 2\sqrt{0.5});$$

$$v(C-I-R-(b_d n_d + b_g n_g))\})$$

<u>Unphased Feeder Service</u>

This service does not separate the delivery and collection processes; stops are intermingled during one long tour.

1. Steady State Depot Pool Size

$$n_d = \lambda_d \cdot C/M\gamma$$

$$u^* = n_d\gamma - S + (S - n_d\gamma)\phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}}) + \sqrt{n_d\gamma}\phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}})$$

2. Steady State Pickup Pool Size

$$n = n_d + n_g = (\lambda_d + \lambda_g)C/M\gamma$$

if $nb + \dfrac{\alpha r\sqrt{A}}{v}(\sqrt{n+0.5} - \sqrt{0.5}) < \ C - R$ then $x^{*\prime} = n$, otherwise:

$$x^* = [\max\{\ 0;\ (\frac{n+0.5-k^2}{2k})\}\ ]^2 + n$$

$$x^{*\prime} = x^* + \max\{\ 0;\ n_g - Y;\ n_g - Z/\gamma\}$$

$$k = \frac{v}{\alpha r\sqrt{A}}(\ C - R - nb)$$

with

$$Z = S - (S - n_g\gamma)\ \Phi(\frac{S-n_g\gamma}{\sqrt{n_g\gamma}}) - \sqrt{n_g\gamma}\phi(\frac{S-n_g\gamma}{\sqrt{n_g\gamma}})$$

$$Y' = (x^*-n_d) - (x^*-n)\Phi(\frac{x^*-n}{\sqrt{n}}) - \sqrt{n}\ \phi(\frac{x^*-n}{\sqrt{n}})$$

3. Delays

$$E[T_g^r] = E[T_d^r] = C/2$$

$$E[T_g^w] = \frac{x^{*\prime} - n_d}{\lambda_g} - \frac{C}{2M}$$

$$E[T_d^w] = \frac{u^*}{\lambda_d} + C/2M \qquad\qquad \text{for random arrivals}$$

$$= u^*/\lambda_d \qquad\qquad \text{for transfers}$$

4. Resources Consumed

$$VMT = \frac{C}{60}\ (\min\{\alpha r\sqrt{A}\ (\sqrt{n+0.5} - \sqrt{0.5});\ v(C-R-nb)\})$$

1.6.3 Zonal Service

Zonal Service is mainly feeder service with a small amount of intra-zonal trips. Such trips are generally served during the normal collection or delivery tours. To model this service, demand rates for collection or delivery are incremented:

$$\lambda'_g = \lambda_g + 2\lambda_m \qquad\qquad (3.57)$$

for intra-zonal service only during the collection phase, or

$$\lambda'_d = \lambda_d + 2\lambda_m$$

for intra zonal service only during the delivery phase, or

$$\lambda'_g = \lambda_g + \lambda_m \text{ and } \lambda'_d = \lambda_d + \lambda_m$$

where λ_m is the demand rate of many-to-many trips. The expected riding time is:

$$E[T^r_m] = \frac{G'}{3} \text{ or } \frac{D}{3} \text{ or } \frac{G' + D}{3} \qquad\qquad (3.58a)$$

in phased service or

$$E[T^r_m] \text{ unphased} = (C-R)/.3 \qquad\qquad (3.58b)$$

in unphased service. The other model outputs are found by applying the equations summarized above for feeder services.

1.6.4. Fixed Route Service

The fixed route transit service model is based upon expected travel time along a route. Input parameters are summarized by the vector P:

$$P = P(\tau_{ij}, b, \lambda_{ij}, h, \sigma_{ij}, \varepsilon_j, n_j, S)$$

where τ_{ij} is the expected travel time from i to j,

     b   is average boarding time,

     λ_{ij} is the demand for boardings or exiting between i and j,

     h   is the average headway on the route,

     ε_{ij} is scheduled slack time between i and j,

     σ_j is the standard deviation of headway

     n_j is the expected vehicle load after stop j and

     S   is vehicle capacity.

and output by the vector J:

$$J = J(T^r_{ij}, T^w_j)$$

where T^r_{ij} is ride time from i to j and T^w_j is wait time at j.

The expected ride time from i to j is:

$$E[T^r_{ij}] = \bar{\tau}_{ij} + b \cdot \lambda_{ij} \cdot h + \varepsilon_{ij}$$

and expected wait time at j is:

$$E[T^w_j] \approx \frac{h}{2} \left(\left(\frac{\sigma}{h}\right)^2 + 1.5 - \frac{(S-n_j)}{\sqrt{\lambda \cdot h}} \right)$$

Walking times must be calculated from the route density and patron distribution (see Section 6.2).

1.6 Glossary

d_i :   expected distance to the nearest of i points from a random point

A   :   area size

a   :   radius of service area; distance from service area center to depot or boundary point nearest the depot.

d : expected tour length (without return to the origin)

r : route factor; expected ratio of street network to straight line distance

α : shape factor; correction term to tour length expression for the effects of area shape.

n : number of stops on tour

N : number of eligible stops in service area

v : expected vehicle speed

e : distance from service region boundary to depot

D : period for delivery of patrons

G : period for collection of patrons

G' : time actually used for patron collection

R : rendezvous period for travel to and from the depot

I : idle period in between delivery and collection periods.

M : number of vehicle visits to depot per cycle; number of vehicles in offset service

S : vehicle capacity

λ : average arrival rate (λ_d for delivery, λ_g for collection)

b : average boarding time (b_d for delivery, b_g for collection)

γ : average group size boarding or exiting at a stop

$N_w(t)$: number of patrons waiting for service at time t

$N_r(t)$: number of patrons riding a vehicle at time t

u^* : steady state pool size of stops for delivery at the beginning of the delivery phase who cannot be carried in the vehicle

x^* : steady state pool size of potential stops for collection at the beginning of the collection phase

$x^{*'}$: value of x^* modified for stochastic variations

Y : number of stops made in a tour

Z : number of patrons served in a tour

$\phi(.)$: standardized normal density function

$\Phi(.)$: standardized comulative normal density function

$L(.)$: logistic curve approximation to the cumulative normal density function

T^r : riding time (T^r_g for collection, T^r_d for delivery)

T^w : waiting time

T^a : access time to fixed route

ST : total travel time of a random patron

VMT : vehicle miles of travel per hour

$\bar{\tau}_{ij}$: average travel time from i to j on a fixed route

h : average headway on a fixed route

σ_j : standard deviation of headway distribution at j

ε_{ij} : slack time in transit schedule between i and j.

CHAPTER 2

PERFORMANCE MODELS IN THE DESIGN OF TRANSIT SYSTEMS

2.1 Introduction

Transit system design is the process of preparing and analyzing plans for transit services. It is undertaken at major planning epochs -- particularly for the design of major new facilities such as a new rapid transit route -- and during periodic reviews of existing services. The process may be carried out solely by a staff technician or, more likely, it may involve a variety of professionals. Ideally, the design process reviews all relevant alternatives, conducts an analysis of the impacts of the various service alternatives, and eventually results in a concensus among policy makers concerning the best decision.

The design process as discussed here is only one element of a more general planning process. Design includes the tasks of analysis of alternatives and the preparation of evaluation reports for decision makers. The planning process itself involves additional considerations, such as public involvement, education and the structuring of decision making. This thesis will not discuss these elements of the planning process.*

Design models are representations of system performance or patrons' travel decisions which may be used to indicate the level of service, resources consumed, demand and other impacts of a particular system design. Such models may be useful at two levels of design effort. The first is

---

*This view of the planning process and the role of the technical analyst has been advocated by Manheim [44]. For a case study of a fully developed transportation planning process of this type, see Gakenheimer [24].

that of resource allocation, in which decisions concerning the desirabil-ility of particular types of services or funding provisions are made. The second is that of service design, in which detailed operational plans are formulated, such as service area boundaries and vehicle dispatching poli-cies.

The key issue involved in the development of such design models is the extent to which they are usable. At the present time models of sys-tem performance have had virtually no impact on service operators.* To be perceived as an improvement over planners' judgment and experience, models must be simple, accurate, and relevant to the local policy alterna-tives and issues. A goal of this thesis is to explore a particular method-ology for constructing performance models and to develop in detail a model of one important element of integrated transit systems, flexibly routed feeder service.

This chapter is concerned with the relevance of system performance models to transit system design. By the end of the chapter, we shall have:

- formulated alternative frameworks for the use of performance models.
- commented upon the importance of equilibrium and multiple objectives as factors in design studies.
- defined various service alternatives which might be used in an in-tegrated transit system, and
- reviewed existing models of transit system performance·

*A conclusion reached by Wilson and Hendrickson [76] for flexibly routed transit models, and Wilson [77] for fixed route transit planning.

2.2 Performance Models in Design Studies

The transit design process addresses questions of system performance and improvement. In the simplest terms, a designer asks the questions: "How will this proposed service perform?" and "How can I improve it?" Models of travel demand and system performance are intended to answer the former question directly, while the latter question involves techniques of optimization, evaluation and the generation of alternatives.

A performance model is intended to estimate the level of service provided to patrons, the resources consumed, and other external impacts caused by a particular system, given the level of demand and a specific service configuration. Level of service attributes which are of interest might include riding time in vehicles, waiting time, the number of transfers required, and the variability of these factors. Resources consumed by the system include fuel consumption, vehicle hours or miles of operation, and other factors. Other external impacts might include air pollutant emissions or congestion delays on auto users. In economic terms, a performance function may be characterized as a very short run supply function, in which the system characteristics are held constant during the period of analysis. Since changes in transportation system characteristics (such as changes in vehicle fleet sizes or infrastructure) may take a relatively long time to effect, the "short run" of the performance function may last several years in practice.

A designer may use performance models to compare alternatives, in a process of heuristic search for better designs, or in an explicit optimization problem. For example, a designer might be interested in comparing

the alternatives of fixed and flexibly routed service in a particular area and with a given fleet size; this would be a problem of comparison. Having decided upon one type of service, a designer might then investigate changes in the vehicle fleet to improve the contemplated service; this investigation is one of heuristic search. Finally, a designer might want to deploy a given vehicle fleet in order to maximize some objective(s); this problem is one of optimization.

Explicit models of system performance are not universally used in design studies. To be useful, models must be relatively accurate, sensitive to policy changes, and compatible with available data. In addition, it is desireable if models are flexible and inexpensive to use. Another desireable characteristic is a simple structure, so that users may understand the causal relationships in the model.

Unless models are relatively accurate, they offer no advantage over intuition. Critical elements of service behavior which influence the system performance must be considered in developing models. For example, random fluctuations have important implications for the level of service provided to users of a transit system, even with excess vehicular capacity available overall. While fixed route transit service has many deterministic features (or nearly so), transient and stochastic effects, with resulting congestion, are likely to be important and should be considered in performance models.

Of course, no model will be completely accurate in predicting the outcome of a particular design change. Indeed, increasing the accuracy of predictions may be quite costly and inhibit the flexibility of the performance models. At the present time, simulation is potentially the

most accurate modelling methodology, but it is quite expensive to develop a detailed simulation model for every design situation. In many applications, high accuracy of results is unnecessary since the same design decision will be the best over a broad range of conditions; for example, fixed route service may be preferable to flexibly routed service in all cases of relatively high demand. Even relatively crude models may give insights into design in these cases. Moreover, the expected demand is uncertain in all design studies, so costly increases in the accuracy of performance prediction may not be warranted since the overall prediction of the equilibrium system characteristics will be uncertain in any case.

Since models are only analytical tools, they are only useful to the extent to which they are applicable to particular areas and sensitive to policy changes. Thus, model outputs should distinguish between operating policies and include impact measures for relevant policy goals. To be applicable, models must only require data which are readily obtainable during the design process.

Due to time and financial constraints on the design process, it is desireable to have models which are flexible and inexpensive to use. Since alternatives are often generated during the design process, model flexibility can be very useful. The expense of using models is often related to the cost of calibration, so extensive calibration of models to local conditions should be avoided. The desireability of these two attributes is illustrated by the history of one large computer simulation model of a flexibly routed, shared ride transit service, the M.I.T. Dial-A-Ride model [72]. This model is relatively expensive to apply, which discourages its use in design studies, and its inflexibility has been one

factor in over-concentration of professional effort on a particular type of service, rather than more restricted and more effective alternatives [76].

Performance models which satisfy the criteria of accuracy, policy sensitivity, and ease of use could be of significant assistance in analyzing alternative transit designs. However, designers should be aware of the implications of equilibrium in the transportation market and of multiple objectives in the use of performance models.

'Equilibrium in the Transportation market.

Performance models estimate the level of service provided by a particular service, given an expected level of demand. However, the demand for travel is generally sensitive to the level of service which patrons experience. Consequently, the equilibrium level of demand which is attracted by the equilibrium level of service is the best estimate of the demand for service and should be used in comparison of alternatives. To illustrate this point, suppose that a designer is comparing an existing system ("old") with a new alternative ("new"). Average travel times via these services increase with increasing demand, as shown in Figure 2.1. Demand for travel decreases as the expected travel time increases (as shown by the curve "demand"). The equilibrium volumes are D_0 and D_N respectively for the old and new systems. Using the existing equilibrium volume level, D_0, and the expected performance of the new system (shown by the curve "new"), the expected travel time would be s_{ON}, a reduction of $s_{OO} - s_{ON}$ in average travel time. However, the actual impact of implementing the new alternative would be a reduction of average travel time from s_{OO} to s_{NN} and an increase in travel volume from D_0 to D_N. A numerical example of this
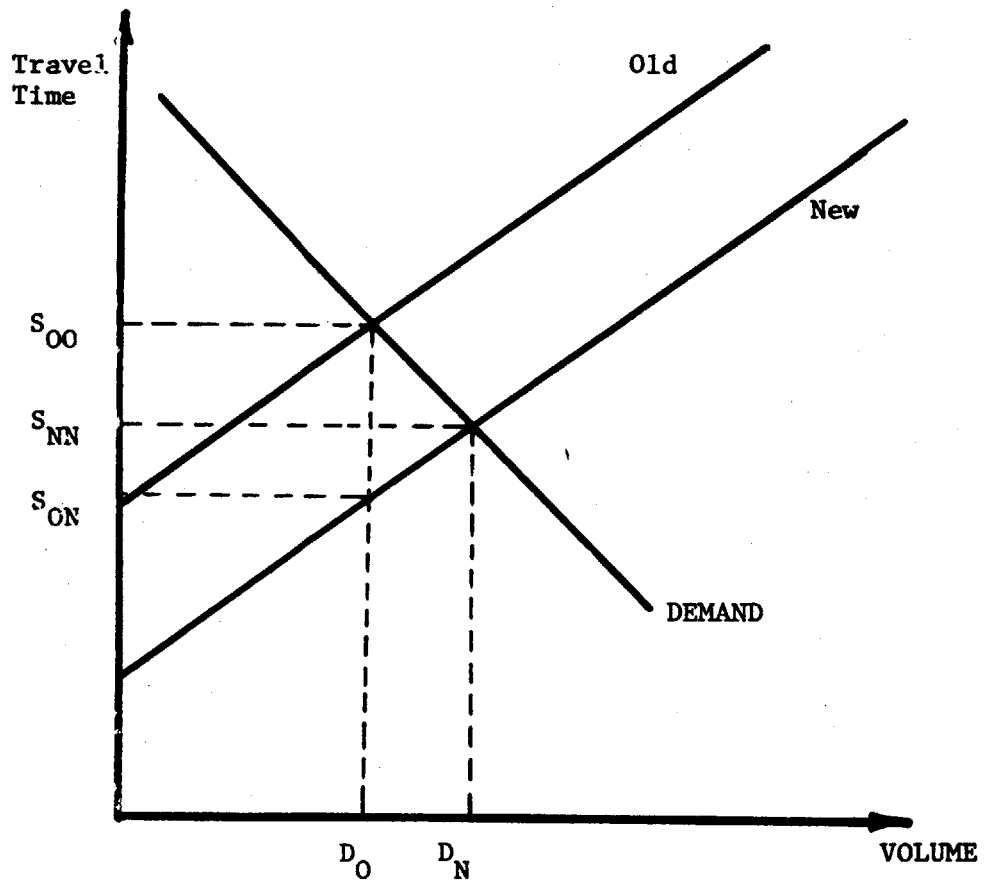
Figure 2.1: Comparison of Two System Designs.

phenomenon for a flexibly routed transit service appears in Section 5.6.

Many design studies have assumed that demand is inelastic with respect to the level of service provided or that the level of service is not a function of the demand level. In these cases, analysis of new alternatives need not contend with the problems involved with identifying equilibrium performance characteristics. However, numerous studies of demand indicate that individuals are responsive in their travel choices to the level of service provided [41]. Other studies have shown that travel times on fixed route transit service are not strongly dependent upon the level of patronage [54]. However, this assumption is not correct for integrated transit services with para-transit components, as will be shown in Chapters 5 and 6.

Global optimization of service in the situation of market equilibration may be difficult to achieve because objective functions such as maximizing consumer surplus or minimizing cost need not be convex (or concave) with respect to the systems' characteristics, so that multiple local optima are possible [22]. Of more immediate concern to operators, however, is the difficulty of defining one (or a few) objective functions for transit system design.

Multiple-Objectives

Transit systems are commonly operated to achieve a number of public objectives, including relieving congestion, reducing air pollution, and improving general mobility. The quality and cost of service are also of concern to patrons, public officials and operators. Consequently, a designer should consider a variety of impacts, some of which may be difficult to quantify or to measure on a commensurate scale. Even in the simple

example of Fig. 2.1, the reduction of travel time $(s_{00} - s_{NN})$ and the increase in patronage, $(D_{NN} - D_{00})$ may both be desireable, but quantifying the improvement may be difficult. One may measure the improvement as the change in consumer surplus, represented by the shaded area in Fig. 2.2. However, this measurement requires inter-personal comparisons of benefits and estimation of the demand function [55]. While issues of evaluation are beyond the scope of this thesis, the existence of multiple objectives suggests caution in the use of optimization techniques for anything but the identification of good initial designs.

2.3 Integrated Transit Alternatives

The bulk of existing public transit service consists of fixed vehicle routes which are generally oriented towards the Central Business District (CBD) (Fig. 2.3). Access to fixed routes is usually accomplished by private automobile or walking. Patrons travelling circumferentially in such networks (such as from point A to point B in Fig. 2.3) receive relatively poor service, often with a trip to the CBD and a transfer required. In addition to the fixed route transit service, a ubiquitous taxi service is available in most urban areas. Taxis charge higher fares and are usually operated and regulated separately from the fixed route transit system.

An integrated transit system would consist of a number of coordinated transit modes. In the existing system, integration might be achieved by facilitating transfers between transit and taxis and by accepting a joint fare structure. It may also be advantageous to substitute taxi service for some fixed route service at particular times or places. Beyond the possibility of using existing modes, there are a number of additional modes
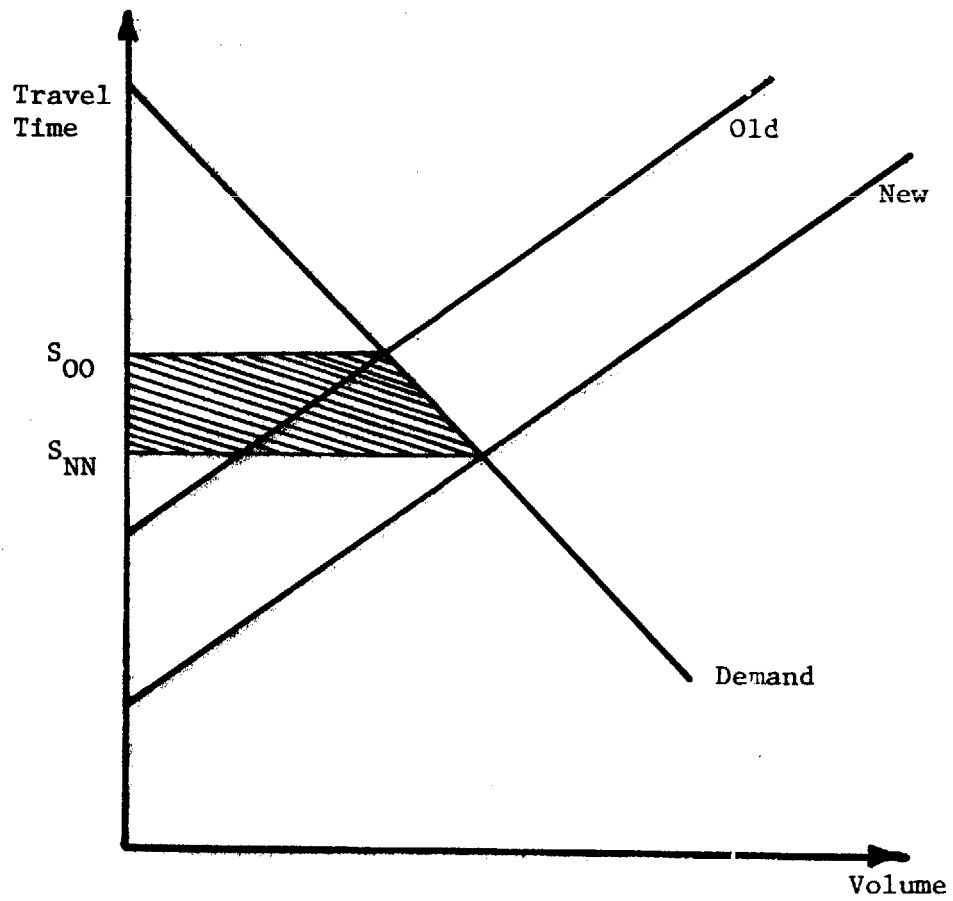
Figure 2.2: Difference in Consumer Surplus with Two System Designs
(Valued in Travel Time Savings).

— Transit Routes
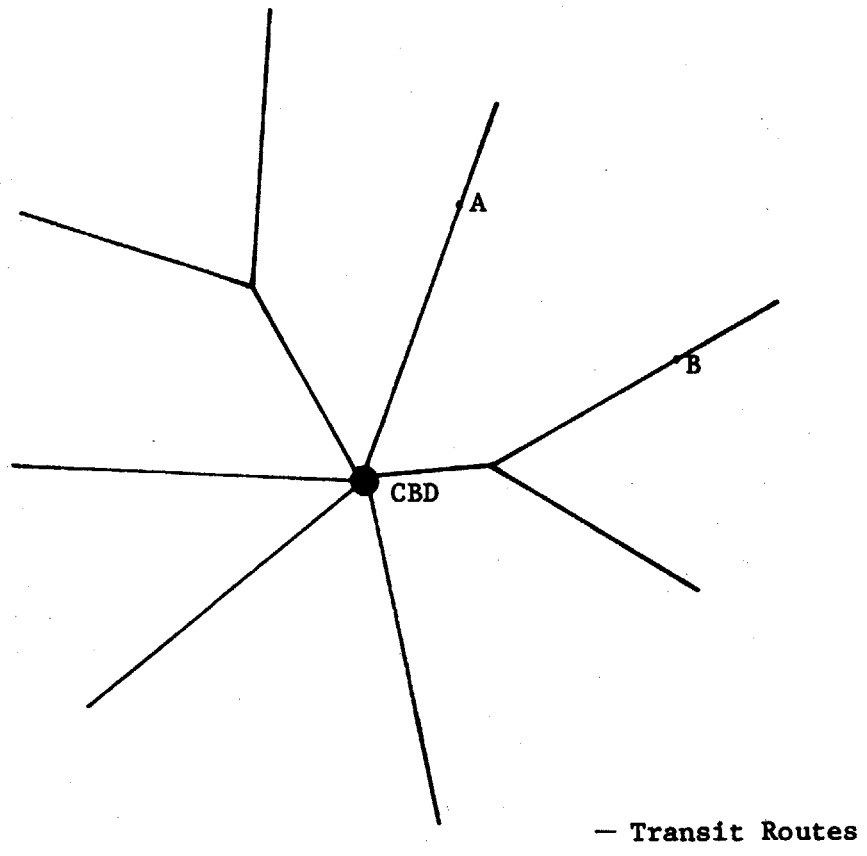
Figure 2.3:   Central Business District (CBD) Oriented Transit Network.

which may be useful components of integrated transit systems. These
other modes are commonly referred to under the category of para-transit
[37].

Integrated transit systems are of interest because different transit
modes and operating policies are most advantageous under different condi-
tions. By selecting the most advantageous designs, it is hoped that
transit service might be operated with increased efficiency, service and
patronage.

In this section, some of the more important components of integrated
transit service will be defined. In particular, we shall concentrate
upon the use and options for local feeder service. Table 2.1 summarizes
the various options discussed here.

At the regional level of operation, transit systems may have the op-
tion of operating fixed or flexibly routed services of various sorts. For
example, fixed route service may be provided by jitneys in the manner of taxi
service, i.e. with many private operators and no scheduled stops. In some
areas, it may be advantageous to operate fixed route service during peak
demand hours and flexibly routed service at other periods. Flexibly routed
service might also be restricted to particular geographic sectors or users,
such as the physically handicapped. With flexibly routed service, the sys-
tem may be structured into specific zones or offered on an area-wide basis.
The latter alternative is known variously as "dial-a-ride," "demand respons-
ive" or "demand-actuated" services, in which a shared vehicle provides door-
to-door service on demand to travellers with different origins and destina-
tions. In this thesis, this type of service will be referred to as many-to-
many, flexibly routed transit, where the "many-to-many" refers to the

-54-

Table 2.1 Options for Integrated Transit

Regional

    . Fixed Route

        conventional transit

        jitneys

    . Flexibly Routed

        exclusive-ride taxis

        shared-ride taxis

        many-to-many transit ("dial-a-ride")

        zonal many-to-many transit

Feeder Service

    . Fixed Route

        conventional transit

        jitneys

    . Flexibly Routed

        scheduled, phased

        scheduled, unphased

        unscheduled

Specific Services

    . Van-pools

    . Subscription Bus

multiple origins and destinations of patrons and "flexibly routed" refers to the modification of vehicle movements in response to patrons' origins and destination. A structured, flexibly routed service consists of a number of dial-a-ride services restricted to particular service areas. Patrons who wish to travel outside their origin zone are required to transfer. Depending upon the size of the zone and the pattern of demand, the majority of trips may require such inter-zonal and inter-vehicle transfer. This type of service will be called flexibly-routed zonal transit.

At the local level, it may be advantageous to operate feeder service to a line haul transit service. Feeder service has the effect of consolidating patronage on a few routes, thereby permitting the realization of the scale economies associated with high capacity line haul systems. Feeder services may be fixed or flexibly routed and may be characterized as many-to-one service, in which all (or nearly all) patrons originate at or are destined for a depot or transfer point. If advance requests for service are required, the service is called subscription. A variety of services are also possible which combine elements of both fixed and flexible routing [19]. Flexibly routed check point service involves stops at specific points in the service area. Route deviation checkpoint service involves a basically fixed route service from which route deviations may be made in response to patrons' demands.

In addition to the possibility of adding specific stops to a flexibly routed service ("checkpoint service"), there are a variety of alternatives for operating flexibly routed services. Vehicles may be scheduled in the sense that each vehicle returns to a depot or transfer point at a specified time. This type of service is particularly advantageous with feeder serv-

ices which are coordinated with line haul routes; schedules may be arranged so as to insure smooth inter-vehicle transfers. Consequently, scheduled feeder services are of particular interest in the design of integrated transit system.

Flexibly routed transit services may also be operated as phased or unphased service. In phased service, patron deliveries and collections are separated into distinct phases, so that all patrons are delivered from the depot before any new patron is collected. Thus, vehicle tours are made first among all the stops for delivery and then among the collection stops. In unphased service, collection and delivery stops are interspersed in one long vehicle tour.

For specific origin/destination pairs, dedicated services may be provided for patrons who travel together on a regular basis. Operating as a transit service (rather than as carpools) such services are called van-pools if a patron drives the vehicle and subscription bus service if a transit employee drives. These specific services may also be flexibly routed - in response to individual patrons' locations - or follow a fixed route.

In this thesis, we shall concentrate upon developing models of feeder services for three reasons. First, such services form a basic component for many integrated transit system designs. Secondly, feeder services, particularly when flexibly routed, have not received a great deal of attention in the literature. Finally, an adequate feeder service model may be easily generalized to the cases of zonal service and van-pooling .

2.4 Existing Performance Models of Transit Services.

The literature concerning modelling and the performance of transit services is quite large, particularly with respect to fixed route services. Much of this literature is concerned with the optimization of service or with identifying good vehicle schedules. In the review that follows, we shall concentrate upon performance models of service, rather than studies of transit supply or optimization. The review is divided into models of flexible, fixed, and integrated transit service.

2.4.1 Performance Models of Flexibly Routed Transit Service

Numerous performance models of flexibly routed transit services have been developed in the past ten years, corresponding to the growth in professional attention to this type of service. Most of these models were developed to demonstrate the general characteristics of such services or to experiment with different routing algorithms. Few have been applied in actual planning or design studies.

Simulation

Digital computer simulation was the first approach used to predict system performance of flexibly routed, demand responsive systems, and it remains the most generally accepted approach. Simulation models have been developed for feeder service [72] and area-wide, many-to-many service, and have embodied a variety of vehicle routing algorithms [8, 25, 31, 47]. These models operate at quite a fine level of detail, generating individual service requests from specified distributions, executing the specified assignment algorithm to select a vehicle and measuring the elapsed time between request generation, pickup, and delivery. Detailed informa-

tion is thus available on the service provided to the hypothetical set of demands, and is summarized in statistics indicating the overall system performance.

Simulation models tend to be quite large computer programs which are difficult to acquire, calibrate and use successfully in applications. Even when established, use of such models to find equilibrium solutions may be quite expensive. Simulation is also a technique which generally yields little insight into performance beyond the system actually simulated. Moreover, once a simulation program exists, it is difficult to modify so as to study alternative designs. So, while simulation models have been used to assist in the design of systems [52], it is not a particularly attractive tool for this purpose and requires skilled planners to be successful.

Apart from these general criticisms of the utility of simulation as a system design tool, initial use of simulation for design of dial-a-ride systems resulted in inaccurate predictions. Initial studies with one simulation model, for example, resulted in under-prediction of wait time by about 30% [73] due to the assumption of a constant number of vehicles in service. While this model was modified to reflect the effects of vehicles entering and leaving service, a designer must always be wary that a particular model may not be comprehensive or may be coded incorrectly for computer use.

Empirical Models

Empirical models of demand responsive transportation systems attempt to develop simple relationships between the key attributes of system per-

formance and design. These models lack a sound theoretical or analytical

basis and require calibration using data from operating systems and/or

from simulation model experiments. Neither approach to calibration is

perfect since there are significant differences among operating systems,

and as just discussed the simulation model is often not a good representa-

tion of the actual system.

Several empirical models of many-to-many dial-a-ride service have

been developed. Early models of this type were based upon linear rela-

tions between critical parameters [1] or on intuitive model forms [74].

An intuitive model [74] was used in one application to evaluate the

Santa Clara County Personal Transit Service [5]. The application in-

dicated that the planned number of vehicles would be unable to provide

an acceptable level of service at the anticipated demand level, and this

was indeed the case when the system was implemented. Another model of

many-to-many dial-a-ride service [23] was developed with separate wait

time and ride time relationships. Model parameters were calibrated from

simulation model results. The model was initially developed as part of

a supply-demand equilibrium model for many-to-many dial-a-ride and

shared-ride taxi systems, using the demand model developed by Lerman et.al.

[41]. In six trial applications of the model system, system demand was

predicted to within 30% of actual figures.

Empirical models are much easier to use in planning situations than

are simulation models. Some empirical models of many-to-many system per-

formance are currently being used in practice. However, due to the lack

of a sound theoretical base, they have occasionally neglected important

non-linear characteristics of system behavior. Careful studies of the

-60-

range of potential applications of existing models have not been performed, so application must be performed cautiously.

Deterministic Models

One series of models of flexibly routed transit service have been based on the following deterministic expression for the length of a vehicle tour among n points:

$$d_n^* = 2\sqrt{A} \ (0.8 + 0.08n) \tag{2.1}$$

where d_n^* = minimum length of a tour linking n points

    A   = area served by a single vehicle

    n   = number of stops

This linear relationship was derived from simulation experiments [47] and is inaccurate as analysis based upon geometric probability has shown (see Section 3.4). Nonetheless, this estimate of tour length has been used as the basis of models of many-to-one, subcription and many-to-many service, although its applicability in the latter case is highly questionable due to the calibration procedure for tour lengths [3,6]. However, none of these models have been validated with real or simulated data nor have they been applied in other than purely conceptual work.

One deterministic model system was developed to compare fixed and flexibly routed feeder services [69]. Vehicle collection tours were assumed to increase linearly in length with respect to the number of stops made. Comparisons between services were made only for square service areas with rectangular grid street systems. The models used were neither validated nor applied to actual data.

Stochastic Models

Most models which treat system operation as a stochastic queueing process have been of exclusive-ride taxi systems. Since the complication of constructing tours is avoided for exclusive ride systems, a stochastic treatment of this problem is considerably easier than for the shared-ride case.

In an early study, a model of many-to-one taxi service was developed assuming exponentially distributed passenger trip times--directly applying the well known results of classical queueing theory to obtain average passenger wait times [49]. This same model was subsequently used in an equilibrium study of the taxi market [46]. In neither case were the validation or applications of the model published.

Several studies of many-to-many taxi service have also been undertaken. McLeod used a single server queueing model for the purpose of estimating wait times [49]. He assumed exponential service times and a Poisson arrival process. Each vehicle is assumed to operate independently in defined sub-areas. The model predicted passenger wait time to within 10% of the observed wait in a single application. In all the taxi models mentioned, however, use of standard queueing models in which the service time is independent of the number of patrons waiting for service ignores the important characteristic that service time is actually a decreasing function of queue length, since the travel time to patrons

decreases as the density of patrons waiting for service increases (see Section 3.2).

Another model of taxi service assumed a general distribution of trip lengths and a limitless city. Comparison with simulation results revealed that the assumption of an unbounded city results in optimistic predictions, with significantly lower expected wait times resulting from the analytic model than from the simulation model [25].

Turning to multiple rider services, a model of one-to-many service [66] as a Markov process has appeared in the literature. Tour lengths were assumed to be negatively exponentially distributed. As each vehicle returned to the depot, as many people as possible boarded (up to the vehicle's capacity) and all these patrons were then delivered; a returning vehicle was assumed to depart even if the queue is empty. Numerical solution of the model was required to obtain the expected patron waiting time. While the model incorporated some important stochastic elements of operation (including random arrival and service processes) and a service constraint (the vehicle capacity), the operating policy is likely to be inefficient (since strategies which either hold a vehicle at the depot to wait for patrons or have fixed cycle lengths are more effective [58]) and has neither been used in practice nor validated.

A stochastic model of multiple rider, many-to-many service has also been proposed. A single server, exponentially distributed service time with the mean interstop time based on a linear function of trip length and productivity is assumed [42]. This linear function was calibrated using results from a series of simulation experiments. Wait time was based on the distance between the vehicle assigned and the passengers' origins, and

assumed to be a linear function of the density of vehicles and demand. Predictions of wait and ride time were felt to be valid only in relatively uncongested systems, and the assumptions of linearity in interstop distance certainly suggest that the model would at best be useful only within a narrow range. This model also fails to recognize that service time is a decreasing function of queue length. It is interesting to note that this supply model was part of the first attempt to model demand responsive systems in an equilibrium framework.

In general, models based primarily on stochastic processes have proved to be difficult to develop, somewhat inflexible, and fairly complicated to apply.

2.4.2 Performance Models of Fixed Route Transit

Fixed route transit service has been the subject of a great many modelling efforts. Most of these studies have been devoted to determining optimal or good designs with respect to vehicle routing or scheduling, such as the studies by Lampkin and Saalmans [38], Hauer [30], Scheele [64] and Hurdle [33]. Other studies have developed empirical supply functions of public transit, with system configuration and design as a dependent variable [54]. The most notable general performance model has been developed under the sponsorship of the U.S. Urban Mass Transportation Administration as part of their computerized transit network model [68]. Virtually all of these studies neglected stochastic effects in transit performance.

Deterministic models of service are generally based upon the expected or scheduled vehicle travel time between stops. Both simulation [e.g. 45] and analytical [e.g. 68] models have been developed in this framework.

Developing a deterministic model of an isolated bus route is fairly straightforward. The expected time elapsed in travel from i to j, $\bar{t}_{ij}$, is:

$$\bar{t}_{ij} = \bar{\tau}_{ij} + b \cdot \bar{p}_{ij} \qquad (2.2)$$

where $\bar{\tau}_{ij}$ is the expected travel time, without boardings,

$\bar{p}_{ij}$ is the expected number of patrons boarding or descending, and

b is the average boarding time.

The expected travel time, $\bar{\tau}_{ij}$, may be modified to reflect the probability that the bus need not stop at some potential stops or may be derived from observations [68]. The expected ride time from i to j is simply the vehicle travel time:

$$E[T^r_{ij}] = \bar{t}_{ij} \qquad (2.3.)$$

with the assumption of random arrivals, the expected wait time to board at stop j is simply half the headway*:

$$E[T^w] = h_j/2 = h/2 \qquad (2.4)$$

where h is the scheduled headway. This general framework has been used in a number of studies (such as [38]).

Another series of studies have concentrated upon the effect of random fluctuations in travel time. If a particular bus falls behind schedule due to such fluctuations, then some excess passengers will have arrived at stops during the schedule delay, and it will take the bus longer to

---

* It is possible to generalize this model to the case of non-random arrivals, but it is not pursued here because the desired departure times are likely to be random, even though patron arrivals are not.

load passengers. Alternatively, a particular driver may have a persistent tendency to travel slower than the norm. In either case, a given slow bus tends to fall further behind schedule. As a result, the level of service deteriorates on the bus route. Eventually, the phenomenon of buses bunching along a route may occur.

As a result of random fluctuations and the mechanism leading to bunching, both the expected wait time and ride time for patrons would deteriorate. In fact, the expected wait time depends upon the variability of headways [61]:

$$E[T^W] = (\frac{h_j}{2}) \ (1 + (\frac{\sigma_j}{h_j})^2) \tag{2.5}$$

where σ_j is the standard deviation of the headway distribution at stop j and h_j is the scheduled headway.

Since slower buses collect more individuals and tend to fall increasingly behind schedule, expected ride time also increases. However, bus operators are likely to institute controls to insure that such progressive deterioration does not occur.

Studies by Newell [59] and Potts [62] analyzed the phenomenon of vehicle bunching. Vehicle holding strategies at checkpoints were suggested by Osuna and Newell [61] and Barnett [2] to mitigate the effects of bunching. Analytical and simulation models which include the phenomenon of bunching are more complicated than deterministic models, due to the effects of vehicle interactions along a route, and more difficult to apply to specific cases without extensive calibration.

2.4.3 Performance Models of Integrated Transit Systems

The performance of integrated transit systems have received less

attention than the components discussed above. One series of papers considered the least cost allocation and scheduling of the line haul and fixed route feeder or express buses. Hurdle [33] used continuum approximations for route density and dispatch rates to find least cost configurations for fixed route feeder service. Clarens and Hurdle [11] developed a model of fixed route feeder services in order to find least cost zone sizes and route frequencies. Wirasinghe [78, 79] employed this feeder service model in an analysis of integrated transit service in an isolated corridor with rail rapid transit. He found that feeder service to the rapid transit system was preferred to separate bus service, with the desirable area for feeder service increasing with increasing distance from the CBD. Wirasinghe also showed that optimum line haul frequencies and interstation spacing depends upon the feeder service characteristics. All these papers assumed constant values for the cost of waiting and riding time, completely inelastic demand which is wholly destined for the CBD, no reuse of vehicles, a constant route density in service zones, uniformly distributed demand origins, and no stochastic effects on system performance. With these assumptions, analytic expressions for system performance and cost could be derived and good system designs identified.

Two other studies of the performance of regional integrated transit deserve mention [3,6]. These studies developed deterministic models of fixed and flexibly routed services to find the costs and travel time experienced by alternative designs. One of the studies [3] used the flexibly routed feeder service model developed in Chapter 3. The other study used the flexibly routed feeder service model developed by Ward [69] and discussed above. The intent of the studies was to explore the effects

of large variations in the level of demand. Again, these studies neglected

stochastic and congestion effect on transit service and have not been val-

idated.

CHAPTER 3

AN APPROXIMATE ANALYTIC MODEL OF SCHEDULED,
FLEXIBLY ROUTED FEEDER SERVICE PERFORMANCE

3.1 Introduction

A basic component of integrated para-transit systems consists of flexibly-routed, zonal feeder service to (or from) a line haul station or transfer point. In this feeder service, patrons either arrive at the transfer point for distribution within the service area or are collected in the service area and taken to the transfer point. Vehicles operate in the manner of a fixed cycle service, involving scheduled stops at the transfer point. This type of service has the advantages of facilitating inter-vehicle transfers in coordinated or integrated systems and of offering a regular scheduled service to patrons desiring service from the depot.

To date this service has received little attention in the literature of supply or service models. Existing systems are described in several papers [27,9]. Sirbu [66] developed a variable cycle, Markov model of feeder service, with an expression for the expected tour length derived from simulation experiments; however this model was limited to one-way service and contained an inefficient dispatching policy assumption. Deneau [16] used an expression for inter-stop distance and a queuing framework which is similar to that used here, but neglected several important stochastic features of service. Neither of these models was generally validated nor applied.

In this chapter, we develop models of such feeder services using a

deterministic queueing framework with fluid approximations and corrections for important stochastic behavior. All the models rely upon a simple proceed-to-the-next-nearest-point dispatching strategy, which is embodied in a tour length approximation formula derived in Section 3.2. The performance model of fixed cycle, phased operation service is developed at length in Section 3.3 to 3.5; in this mode of operation, vehicles first deliver patrons from the transfer point, then collect patrons from the service area before returning to the depot. This mode of operation offers a fairly high level of service and is often used by parcel delivery services and integrated transit services such as the Ann Arbor Teltran service.

Section 3.6 extends the one vehicle model to multi-vehicle systems. In Section 3.7, a fixed schedule service with mixed collection and distribution of patrons is described. Section 3.8 discusses the introduction of some many-to-many service in the feeder system, in which some trips might not originate or end at the transfer point. Extensions of the model to irregular areas and non-uniform demands, comparisons with other dispatching algorithms, some applications of the model, and validation of the models are discussed in following chapters. Equations for the models developed and a glossary of notation appear in Section 3.9.

By the end of the chapter, we will have developed:

- a simple expression for the tour length among a set or subset of randomly distributed points, based upon a particular routing algorithm.

- a performance model of flexibly routed feeder service in circular areas for any demand level, vehicle fleet size or area size.

- expressions for patron delays in such feeder service systems. The performance model will then be extended in Chapter 4 and applications explored in Chapter 5.

3.2 Tour Lengths with Next-Nearest-Point Dispatching

An essential feature of the feeder service is the dispatching algorithm that is used to construct tours. Existing systems generally rely upon driver routing among the stops assigned by a dispatcher. In these systems, drivers are assigned a number of collection or delivery stops and may then choose a route among the various assigned stops. In this section, we consider a strategy in which vehicles always travel to the next nearest eligible point. This simple dispatching strategy is comparable in performance to other dispatching algorithms and permits the use of some simple results of geometrical probability to arrive at an approximate tour length expression.

It is a well-known result in the theory of geometrical probability (see Kendall and Moran [36] or Fairthorne [20]) that the expected distance between two random points in a circle is:

$$d_1 = \frac{128}{45\pi} a \approx .51\sqrt{A} \tag{3.1a}$$

where a and A are the circle's radius and area respectively.* The corresponding equation for a square is similar (see Daganzo [13]):

$$d_1^s \approx .52\sqrt{A} \tag{3.1b}$$

---

*"random" points is interpreted as points whose location probability density function is uniformly distributed over the service area.

In addition, the expected distance between a random point and the closest of a set of n random points distributed with an average density of (n/A) in an unbounded area is*

$$d_n = .5\sqrt{A/n} \qquad\qquad (3.2a)$$

In a given area, this expression is also correct as the density of points increases without bound [14]:

$$\lim_{n\to\infty} d_n = .5\sqrt{A/n} \qquad\qquad (3.2b)$$

For the case of two points, this asymptotic expression (Eq. 3.2a with n=1) is only 2% larger than the exact expression for the expected distance between two points in a circular service area (Eq. 3.1a). If we assume that the average of these expressions, Eqs. 3.1a and 3.2b, is approximately correct for small values of n, then:

$$d_n \approx .505\sqrt{A/n} \qquad\qquad (3.3a)$$

in a circular service area. A similar expression in a square, using Eq. 3.1.b, is:

$$d_n^s \approx .51\sqrt{A/n} \qquad\qquad (3.3b)$$

or a factor of 1.01 times the corresponding expression for a circular area.

An approximate expression for the tour length among n of N points in a circular service area is, then:

---

* The distance to the next-nearest-point is distributed as a Rayleigh dis-stribution with parameter $\sqrt{\dfrac{2n}{A}}$. See Borel [7] or Kendall and Moran [36].

$$d = \sum_{i=1+N-n}^{N} d_i \approx .505\sqrt{A} \sum_{i=1+N-n}^{N} \sqrt{1/i} \qquad (3.4)$$

To make Eq. 3.4 suitable for hand calculation, a continuous approximation may be used:

$$d \simeq .505\sqrt{A} \sum_{i=1+N-n}^{N} \sqrt{1/i}$$

$$\simeq .505\sqrt{A} \int_{N-n+0.5}^{N+0.5} x^{-1/2} dx$$

$$\simeq 1.01\sqrt{A}(\sqrt{N+0.5} - \sqrt{N+0.5-n})$$

$$\simeq \sqrt{A}(\sqrt{N+0.5} - \sqrt{N+0.5-n}) \qquad (3.5)$$

This expression assumes a circular service area, uniformly distributed points and the approximation to the expected distance to the nearest of n points contained in Eq. 3.3a. Section 4.2 will consider modification of this expression for application to differently shaped service areas. As suggested by the comparison to a square service area (Eq. 3.3b), this modification takes the form of a multiplicative constant.

Finally, to account for the circuity of the street network, this expression must be multiplied by the route factor, r, which is the expected ratio of the street network distance and the straight line or

air line distance between two points (see Haight [29]):

$$d \simeq r\sqrt{A}(\sqrt{N+0.5} - \sqrt{N+0.5-n}) \qquad (3.6)$$

The route factor must be inserted because vehicles travel on the existing street network, rather than travelling by straight lines from point to point. Section 4.2 discusses the estimation of the route factor in specific cases and describes a correction to account for the effect of area shape.

Unfortunately, the argument used to derive Eq. 3.6 is not entirely correct since the expression for the expected distance to the nearest of n points is only exact as A becomes very large and because the location of successive points on the tour are not random points in the service area.* To illustrate the latter point, suppose a vehicle begins at point 0 in Fig. 3.1 and makes its first stop at Point P. In this case, no other point on the tour can be within the dashed circle in Fig. 3.1, otherwise point P would not have been the closest stop to the original point 0. Thus, the length of the second leg of the tour is the average distance from a random point P to the closest of n-1 random points, given that no point falls in the dashed circle.

Since point P is on the edge of the area of feasible points, the expression for d_{n-1} (Eq. 3.3) may underpredict the actual distance. However,

---

*In discussing vehicle tours, we refer to the sequence of stops in delivery or collection, without requiring the vehicle to return to the starting point. In the literature of operations research, this sequence of stops would be called an "open" tour, while a "closed" tour requires a return to the starting point. The travel between stops will be referred to as links or legs in the tour.

-74-
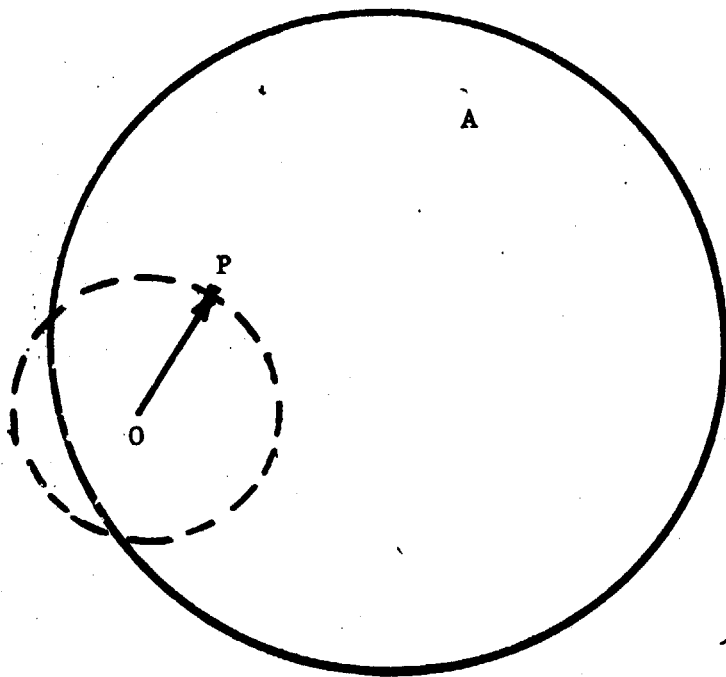
Figure 3.1: Illustration of Vehicle Movement

the remaining points are concentrated in an area smaller than the original service area, causing an overprediction of d_{n-1}. These effects tend to cancel each other. The same line of reasoning applies to all the legs of the tour, with the available area in which points may be located becoming increasingly smaller as the tour progresses. Consequently, one would expect the expression for d_n to be somewhat optimistic for segments at the beginning of a tour and slightly pessimistic for legs at the end of a tour.

A limited number of simulation runs have been performed to compare with the approximate expression for a tour length given above (Eq. 3.6). These simulations generated between 1 and 10 random points in a circular service area and then constructed tours among the points using two separate starting points.* The length of simulated tours with a starting point near the edge of the service area compare quite closely to those predicted by Eq. 3.6, with a maximum error of 10% occurring for two stops (Table 3.1). With more than two stops, starting at the center of the service area also compares quite closely with the predicted tour length. In an independent set of simulations to compare with Eq. 3.3., the expected distance to the closest of n random points, Daganzo [13] found that the equation was accurate to within 1% of simulated data. As a result, we conclude that Eq. 3.6 may be used for the feeder service tour lengths as long as the stops are randomly and uniformly distributed throughout the service area, vehicles start near the edge, and the number of stops is between 1 and 10.**

With the expression for the expected tour length, the expected time for n stops among N possible stops may be calculated from Eq. 3.6 as:

$$t(N,N-n) \simeq nb + \frac{r\sqrt{A}}{v}(\sqrt{N+0.5} - \sqrt{N+0.5-n}) \qquad (3.7)$$

---

*These tours are "open" in the sense that the vehicle does not return to the origin
**Between 10 and 20 stops, the expression should also be relatively accurate.

Table 3.1

Comparisons of Simulated and Predicted Tour Lengths*

Number of Stops	Predicted[†]	Simulation**			
		Start at Center	% Error	Start near Edge***	% Error
1	.64	.46	−39	.71	−10
2	1.13	.96	−18	1.16	3
3	1.50	1.42	− 6	1.51	1
4	1.83	1.80	− 2	1.83	0
5	2.07	2.12	2	2.12	2
6	2.38	2.38	0	2.38	0
7	2.63	2.61	− 1	2.60	− 1
8	2.85	2.81	− 1	2.78	− 3
9	3.07	3.02	− 2	3.01	− 2
10	3.21	3.31	3	3.22	0

* Results for a unit area Circle, radius= .56, points are uniformly distributed

** number of observations = 1000

***Starting Point coordinates (.5,0); circle center coordinates (0,0).

[†] Equation 3.6, r=1.27

where b is the average stop time and v is the <u>expected</u> vehicle speed.*

Note that the tour time <u>decreases</u> as the potential number of stops, N, <u>increases</u>, given a fixed number of stops, n. In the terminology of queuing theory, the potential number of stops, N, represents the number of people in the queue waiting for service, and the service rate of the system is state dependent, varying with the length of the queue. This phenomenon provides a mechanism by which the system may attain a steady state even with relatively high volumes.

We shall use Eq. 3.7 as a basic component of the feeder service models derived below.

3.3 Deterministic Model of One Vehicle Scheduled and Phased Service

We can now use the relationships just developed to build a deterministic model of scheduled feeder service with one vehicle. The variables of particular interest are the number of patrons waiting for service or riding the vehicle as a function of time, $N_w(t)$ and $N_r(t)$ respectively. In the deterministic model developed here, N_w and N_r are both periodic functions of t (with period equal to the scheduled time between visits to the transfer point, a cycle length of C) and so it will be sufficient to study the system in an arbitrary cycle.

We must consider the three phases of the vehicle's operation:

(1) distribution of patrons from the transfer point in period D.

(2) collection of patrons from the service area in period G.

(3) travel to and from the transfer point and a rendezvous time for transfer of patrons during period R.

---

*This expression for travel time, d/v, is less than the expected travel time whenever vehicle speed is variable. This conclusion follows from the fact that the inverse of velocity is a concave function, so E[1/velocity] > 1/v by Jensens's inequality. However, this effect is expected to be minor [39].

-78-

In deriving this model, we assume that the demand for the service is a deterministic arrival process to the system and that demand locations in the service area are uniformly distributed throughout the service area; these assumptions will be relaxed in following sections. We shall also assume in this section that the service area is circular with a homogeneous street network.

During the distribution phase, all patrons who transfer from the line haul service are delivered to their destinations in the service area. We assume that n_d stops are made in each delivery cycle. If groups of patrons travel together, then n_d is the number of stops made and the number of patrons delivered is γn_d, where γ is the average group size.[*] If λ_d is the arrival rate of patrons for distribution, then

$$\gamma n_d = C \cdot \lambda_d \qquad (3.8)$$

where C is the cycle time. In what follows, small n will refer to the number of stops, at which one or more people are delivered, and λ refers to the arrival or demand rate of patrons.

The length of the distribution period is given by Eq. 3.7, with N= $n=n_d$ and average stop time equal to average exiting time per stop, $b=b_d$, so the total time for the delivery phase is:

$$D = n_d b_d + \frac{r\sqrt{A}}{v} (\sqrt{n_d + 0.5} - \sqrt{0.5}) \qquad (3.9)$$

which is the sum of stopping time and the driving time for delivery.

A necessary condition for a steady state equilibrium arising from the delivery process is that the number of patrons arriving at the

---

[*] γ is the average number of patrons travelling together and is introduced as the proportionality constant between the number of stops made and the number of patrons served - with no patrons travelling together, $\gamma=1$.

transfer point during a cycle not exceed the capacity of the vehicle

$$\lambda_d \cdot C < S \qquad\qquad (3.10)$$

where S is the capacity of the vehicle. Otherwise, the queue waiting for delivery at the transfer point grows without bound.

During the rendezvous phase, the vehicle travels between the service area and the transfer point and stops at the transfer point for a given period of layover time, during which the patrons riding exit and the patrons for delivery board the vehicle. This layover time should be sufficiently long to allow passengers to transfer and to enable smooth connections, even with variation in the line haul vehicle's arrivals. For typical applications, layover time might fall between 1 and 5 minutes.

The length of the distribution tour (Eq. 3.9) is based upon the assumption that the initial point is randomly located in the service area. However, it is more efficient to make the first stop of the distribution tour at the destination nearest the transfer point. Eilon et.al [18] present expressions for finding the expected distance to the nearest of a set of random points in an area from a given point outside the area. However, the tour length is not particularly sensitive to the starting point as long as the number of stops exceeds two (Table 3.1). Consequently, travel time between a transfer point outside the service area and a starting point at the edge or two thirds of the way from the center to the edge of the service area should be sufficiently accurate for the design model. Using the latter assumption, then the rendezvous time may be calculated as:

$$R = L + \frac{2(e + a/3)}{v} \qquad\qquad (3.11)$$

where L is the layover or rendezvous time, e is the distance from the

transfer point to the edge of the service area, and a is the radius of the service area. The two thirds value is suggested by the difference between the expected distance to the service area center and the distance from the center to a random point (see Section 4.2.3 for a further discussion).

The period available for collection may now be calculated from the identity:

$$G = C - D - R \qquad (3.12)$$

which is the total cycle time less the time allocated for the trip to and from the transfer point and the distribution time. The vehicle is now assumed to collect patrons until either the period G ends, the vehicle's capacity is reached, or all patrons waiting have been collected.

The expected number of patrons arriving for collection during each cycle is $\lambda_g C$, representing n_g stops:

$$n_g = \frac{\lambda_g C}{\gamma} \qquad (3.13)$$

For steady state operation in the deterministic model, n_g stops must be made in every collection phase. If the time required to make n_g stops is less than the period available:

$$G \geq t(n_g, o) = n_g b_g + \frac{r\sqrt{A}}{v} (\sqrt{n_g + 0.5} - \sqrt{0.5}) \qquad (3.14)$$

then no queue of individuals waiting for collection remains at the end of the collection phase. Otherwise, some patrons must wait for the next cycle.

Two hypothetical vehicle cycles are plotted in Figures 3.2 and 3.3 to illustrate the variation in the number of patrons in the system over

Fig. 3.2 Collection phase shorter than required

Legend:
- ═══ Total $N(t)$
- ──── Total at home $N_{wh}(t)$
- ──── In bus $N_r(t)$
- ──·── Patrons at home assigned for collection
- $N=$ Number of patrons

N (vertical axis)

time (horizontal axis)

t_0 t_1 t_2 t_3

L D $G=G'$
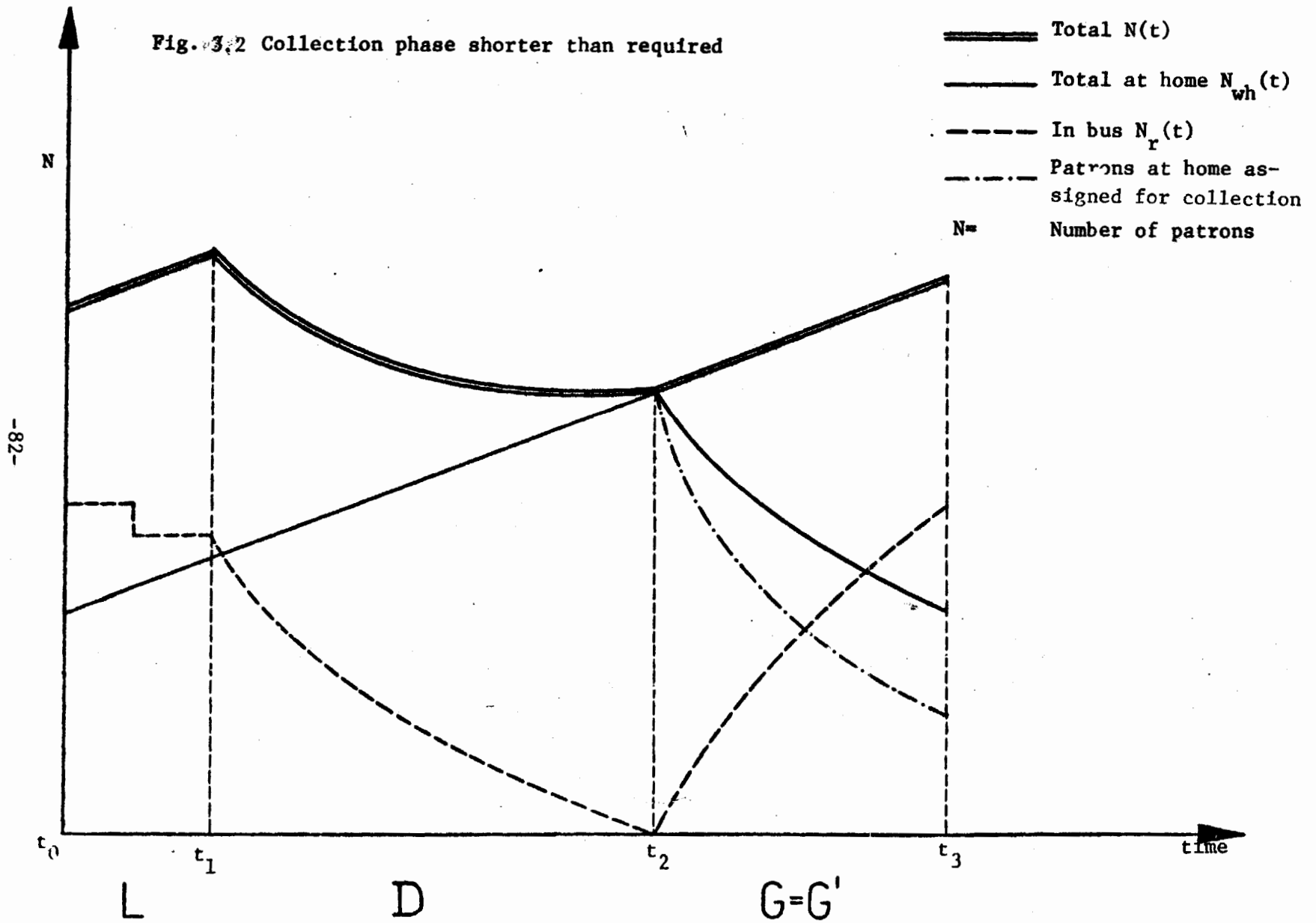
Fig. 3.3 Collection Phase longer than required to pick-up the requests available at time t_2

Total N(t)

At home $N_{wh}(t)$

In bus $N_r(t)$

Patrons at home
assigned for
collection

N

-83-

t_o t_1 t_2 t_3 time

L   D   G > G'
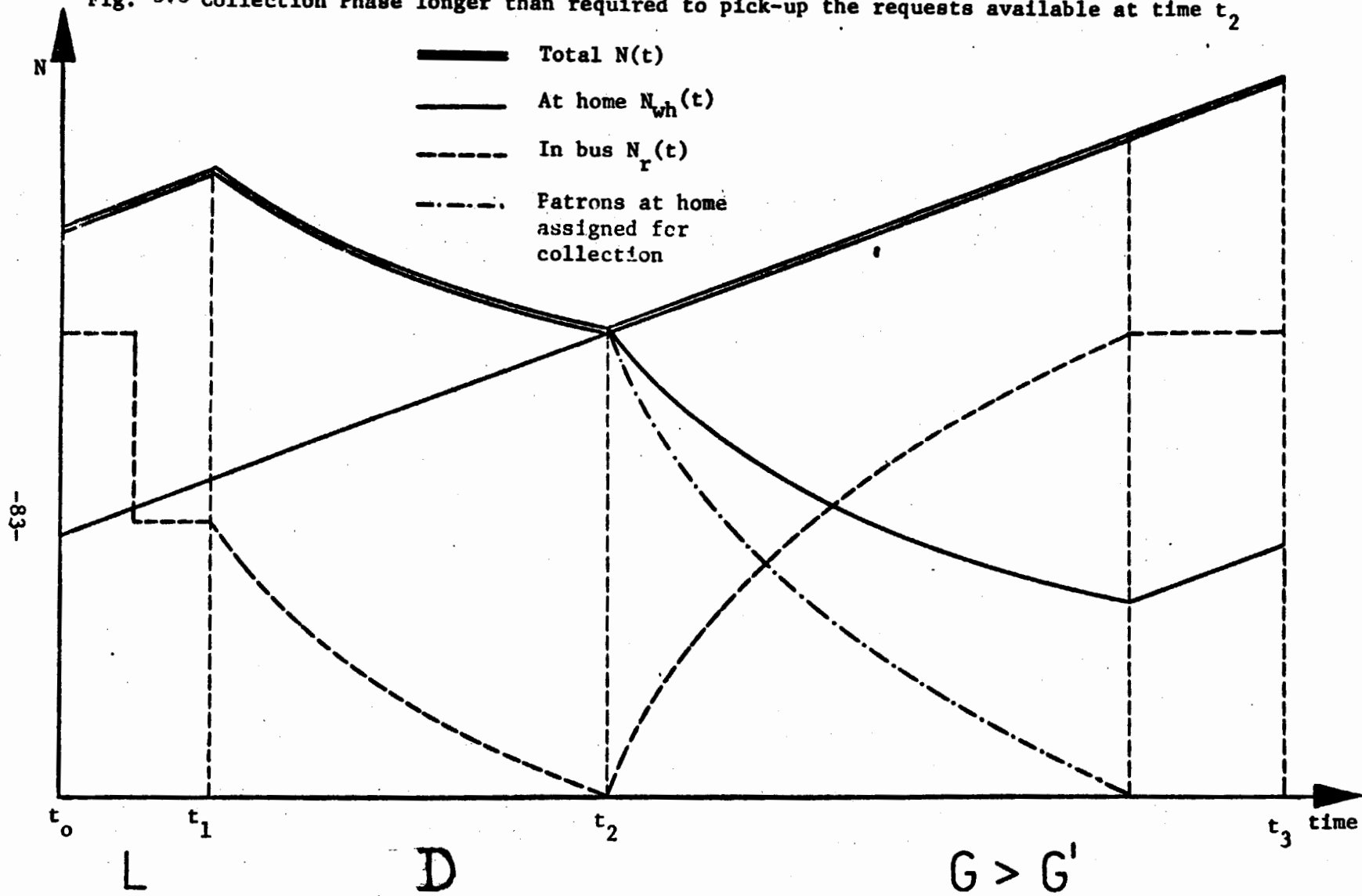
time. In Figure 3.2, the collection time available is insufficient to enable all waiting patrons to be collected. However the number of patrons collected during the cycle equals the number who arrived, so the system is in a steady state. Qualitatively, we see that during the lay-over period (t_0 to t_1) and the distribution period (t_1 to t_2), the number waiting at home steadily increases. The fall in the curve representing the number of patrons in the bus at the time $(t_0 + t_1)/2$ occurs during the rendezvous period, when patrons transfer to the line haul system. The vehicle completes all deliveries by time t_2, and then it begins to collect passengers. We assume that patrons requesting service during the collection period cannot be immediately scheduled, so there is a divergence in this period between the number of patrons waiting at home and the number of patrons assigned for collection. In figure 3.3 the collection time, G, is sufficient to enable all patrons waiting and assigned to be collected in time G'.

For the system to be in steady state, it is necessary and sufficient that the number of patrons collected not exceed the vehicle capacity:

$$S \geqslant \lambda_g . C \tag{3.15}$$

and that the number of patrons waiting for collection be equal at the start of each collection phase:

$$N_w(t_0) = N_w(t_0 + C) \tag{3.16}$$

where t_0 is the time at the start of a collection phase.

Since the collection tour becomes increasingly efficient as N_w increases, an equilibrium value of $N_w(t_0)$ may be calculated from the equations:

$$G = t_g(x; \; x-n_g) \qquad\qquad \text{if } x > n_g$$

$$\tag{3.17}$$

$$G \gtrless t_g(x,0) \qquad\qquad \text{if } x = n_g$$

where $N_w(t_o) = x$ and the time for the collection tour, $t_g(.,.)$, is given by Eq. 3.7 with b replaced by b_g, the average stop time for boarding patrons. A solution to Eq. 3.17 exists as long as:

$$D + R + b_g n_g \leqslant C \tag{3.18}$$

which simply requires that the sum of the layover, distribution, and total boarding times not exceed the cycle time. In the case that $x = n_g$, all assigned patrons are collected during the collection phase. Then the time actually spent collecting patrons, G', is:

$$G' = \min[G; \; t_g(x,0)]$$

With the equations defined above, it is possible to obtain a solution for the steady state value of the number of patrons waiting at home, x^*. If $x^* > n_g$ then:

$$t_g(x^*, x^*-n_g) = C-R-D \tag{3.19}$$

or

$$b_g n_g + \frac{r\sqrt{A}}{v} (\sqrt{x^*+0.5} - \sqrt{x^*-n_g+0.5}) = C - R - D$$

This may be reduced to a quadratic expression in x^* (as long as the necessary conditions relating to vehicle seating capacity and cycle time are satisfied, Eqs. 3.10, 3.15, and 3.18), taking into account Eq. 3.19, to obtain:

$$x^* = [\max\{\sqrt{0.5}; \; [\frac{n_g}{2k} - \frac{k}{2}] \; \}]^2 + n_g - 0.5 \tag{3.20}$$

where the factor k is:

$$k = \frac{v}{r\sqrt{A}}(C - R - D - n_g b_g)$$

which may be interpreted as the time available for travelling between stops on the collection tour, multiplied by the factor $v/r\sqrt{A}$.*

3.4 Fine Tuning Corrections

In order to enhance the accuracy of the model, four minor modifications are introduced to account for the "integerness" of customers and stops and to capture the most important stochastic effects of the system. In this section, we relax the assumption of deterministic arrival processes.

Since the arrival process is assumed to be random, the number of groups waiting for delivery from the transfer point is a random variable, Z. We assume that the distribution of the number of groups is Poisson, with mean and variance both equal to n_d. Then, the distribution time, D, is a random variable as well:

$$D = t_d(Z,0) \tag{3.21}$$

where $t_d(.,0)$ is given by Eq. 3.7, the expected tour time.
Since $t_g(.,0)$ is a concave function (see Eq. 3.7) Jensen's inequality implies that:

$$E[D] \leqslant t_g(E[Z],0) = t_g(n_d,0) \tag{3.22}$$

---

* The necessary conditions for this solution to be feasible are Eqs. 3.10, 3.15 and 3.18.

Consequently, the expression for the distribution time, Eq. 3.9, tends to overpredict the expected distribution time when the arrival rate is stochastic, due to the phenomenon that tours become more efficient with more patrons.* In order to better approximate the average distribution time, we take the first two terms of a Taylor expansion of Eq. 3.22:

$$E[D] \simeq t_d(n_d,0) + \frac{1}{2}[\frac{d^2 t_d(Z,0)}{dZ^2}\bigg|_{Z=n_d}] \, var(Z)$$

$$\approx t_d(n_d,0) - (\frac{1}{2}) \frac{r\sqrt{A}}{v} (\frac{1}{4})(n_d+0.5)^{-3/2} n_d$$

which reduces to:

$$E[D] \simeq n_d b_d + \frac{r\sqrt{A}}{v} \{\Delta\sqrt{n_d + 0.5} - \sqrt{0.5}\} \qquad (3.23)$$

with

$$\Delta = 1 - \frac{n_d}{8(n_d+0.5)^2}$$

The factor Δ is a correction term to account for the stochastic nature of the inbound (delivery) arrival distribution. It reaches a minimum of 15/16 for $n_d = 0.5$. Its effect is most noticeable when cycle lengths are very short, since in this case the small reduction in the average delivery period results in a small but relatively significant increase in the time

---

*This is one of the rare examples in which introducing randomness improves the system's performance, in this case due to the economies inherent in the state-dependent service rate.

available for the collection phase. Eq. 3.23 may thus be used instead of Eq. 3.9 to obtain the expected distribution period, D. In practice, Δ has a substantial effect only in the portion of the performance function at which service is deteriorating rapidly due to overly short cycle lengths. Since it is usually not advantageous to operate a system with such short cycle lengths, the correction term Δ may usually be set to one in manual application of the model without incurring a significant error (Sec. 5.3).

The next modification attempts to capture the effect of the indivisibility of customers during the collection phase. The dispatching algorithm used in the model restricts collection of patrons to a specified period of length G. In the fluid approximation of Eq. 3.14, however, the number of stops during the collection period is not constrained to be integer, as must happen in reality. If one assumes that the number of stops is integer, then in the ith cycle, the number of stops actually made, Z_i, will be the integer portion of n_g in Eq. 3.13:

$$Z_i = ||n_g||$$

(3.24)

For the system to be in steady state, it is necessary that:

$$E[Z] = E[n_g]$$

(3.25)

that is, the average number of stops made per cycle equals the number of stops required per cycle. However:

$$E[Z] = E[||n_g||] \simeq E[n_g] - 0.5$$

(3.26)

for any distribution of n_g with standard deviation much larger than

one*. It is then reasonable to use the steady state condition:

$$n_g = E[Z] + 0.5 \qquad (3.27)$$

or equivalently subtract 0.5 from the lower bound on the fluid approximation, Eq. 3.6, for the collection period, yielding:

$$t_g(x^*; x^*-n_g) = n_g b_g + \frac{r\sqrt{A}}{v}[\sqrt{x^*+0.5} - \sqrt{x^*-n_g}] \qquad (3.28)$$

so

$$x^* = [\max\{0, (0.5+n_g-k^2)/2k\}]^2 + n_g \qquad (3.29)$$

with k defined as in Eq. 3.20.

The next modification accounts for the stochastic nature of the collection process. We assume that the number of stops made, Y, and the number of passengers collected, Z, are random variables. Both of these random variables have upper limits consisting of the potential number of stops and the vehicle capacity respectively. Ignoring the capacity constraint for the moment, the true number of stops (assuming that we start with a pool of potential stops, x) is a random variable Y':

$$Y' = \min\{x,Y\}$$

and $E[Y'] \leq E[Y]$ in general. In the deterministic case, this was not a problem since $var(Y) = 0$.

If we assume that Y is approximately normally distributed with mean

---

*This result is intuitive and analogous to Poincare's roulette problem (see Feller [21], vol. II, pg. 62). Alternatively, one may assume that $Y + \theta = n_g$, with θ uniformly distributed [0,1] and Y,θ independent, to obtain the same result.

$\bar{Y}$ and variance σ^2, then we have:

$$E[Y'] = \int_{-\infty}^{x} \frac{w}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}[\frac{w-\bar{Y}}{\sigma}]^2\}dw +$$

$$+ \int_{x}^{\infty} \frac{x}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2}[\frac{w-\bar{Y}}{\sigma}]^2\}dw \qquad (3.30)$$

which after some algebraic manipulation reduces to:

$$E[Y'] = x - [x-\bar{Y}] \ \Phi(\frac{x-\bar{Y}}{\sigma}) - \sigma\phi(\frac{x-\bar{Y}}{\sigma}) \qquad (3.31)$$

where $\Phi(.)$ and $\phi(.)$ are the standardized normal cumulative distribution and density functions respectively.*

A similar line of reasoning may be applied to the vehicle capacity ignoring the pool size constraint:

$$Z' = \min\{Z,S\} \qquad (3.32)$$

so that

$$E[Z'] = S - (S-\bar{Z}) \ \Phi(\frac{S-\bar{Z}}{\sigma}) - \sigma\phi(\frac{S-\bar{Z}}{\sigma}) \qquad (3.33)$$

which is directly analogous to Eq. (3.31).

We shall assume that the number of stops made and the number of patrons collected is highly correlated. Consequently, we need only

---

*We use a Gaussian approximation rather than the Poisson in this case because the computations are somewhat simpler. Either approximation would give similar results.

consider the constraint with the lower upper bound. For example, if the expected number of stops made results in a larger number of patrons collected than the expected value of Z':

$$E[Y']\gamma > E[Z']$$

the vehicle capacity constraint is the binding constraint.

Returning to the case in which the potential pool of stops is of interest, we note that the formula reported earlier for the time necessary to reduce the pool of assigned stops from N to N-n (Eq. 3.7) must be corrected in order to capture the phenomenon described above. In the time $t(N,N-n)$, the vehicle will have stopped, on the average:

$$n'_g = N - (N-n) \ \Phi(\sqrt{N-n}) - \sqrt{N-n} \ \phi(\sqrt{N-n}) \qquad (3.34)$$

where we have assumed that $\sigma^2 = E(Y) = N-n$ since the underlying service process is likely to be Poisson in nature. Equation 3.34 is monotonic and therefore one may write that the reduction in pool size is some function of N and n_g':

$$(N-n) = f(N-n_g') \qquad (3.34)$$

and

$$t'_g(N;N-n) = t_g(N;f(N-n')) \qquad (3.35)$$

which is the time required to make n of N stops, corrected for variation in the service process. This expression $t'_g(.,.)$ should be used instead of $t_g(.,.)$ in order to obtain the corrected equilibrium value of x^*. Such analysis is computationally difficult, however, since $f(.,.)$ does not have a closed form. The following is a good approximation to the corrected value of x^*:

$$x^{*'} = x^* + \max\{0, n_g - E[Y']\} \tag{3.36}$$

where x^* is obtained as above (Eq. 3.20) and $E[Y']$ from Eq. 3.31.

A similar argument may be made with respect to the seating capacity, implying that:

$$x^{*'} = x^* + \max\{0, n_g - E[Y'], n_g - E[Z']/\gamma\} \tag{3.37}$$

So, in general, $x^{*'} \geqslant x^*$.

Calculation of $E[Y']$ and $E[Z']$ involves the evaluation of the cumulative normal function $\Phi(.)$. For ease of application, an approximation is quite useful. Using the normalized cumulative logistic function:

$$L(x) = \frac{1}{1 + \exp(-2x/\sqrt{2/\pi})} \simeq \frac{1}{1 + \exp(-2.5x)} \tag{3.38}$$

in place of the normal curve introduces a maximum error of 2% [17]. With this substitution, Eq. 3.34 becomes:

$$E[Y'] \simeq x - (x - E[Y]) \, L(\frac{x - E[Y]}{\sigma}) - \sigma\phi(\frac{x - E[Y]}{\sigma}) \tag{3.39}$$

We can now summarize the modification for the calculation of $x^{*'}$, the steady-state number of assigned stops to take account of the stochastic nature of the collection and arrival process. We have that:

$$x^{*'} = x^* + \max\{0, n_g - E[Y'], n_g - E[Z']/\gamma\} \tag{3.37}$$

$$E[Y'] = x^* - (x^* - n_g) \, L(\frac{x^* - n_g}{\sqrt{n_g}}) - \sqrt{n_g} \, \phi(\frac{x^* - n_g}{\sqrt{n_g}}) \tag{3.38}$$

$$E[Z'] = S - (S - n_g\gamma) \, L(\frac{S - n_g\gamma}{\sigma}) - \sigma\phi(\frac{S - n_g\gamma}{\sigma}) \tag{3.40}$$

where

$$\sigma = \sqrt{n_g E[\gamma^2]} \simeq \sqrt{n_g}\gamma \qquad (3.41)$$

and x* is given by Eq. 3.20. We have implicitly assumed that the number of passengers collected, Z, is a compound Poisson process yielding the standard deviation of Eq.(3.41). The secondary approximation of Eq. (3.41) is only intended for use when no other information is available.

A similar modification may be introduced to capture the effect of variations in the number of patrons arriving for distribution and the constraint of vehicle capacity. Even though the average number of patrons arriving must be less than the vehicle capacity (else no steady state exists), random fluctuations may result in a situation in which not all patrons arriving at the depot may be carried. As above, we assume that the number of patrons arriving is normally distributed with mean equal to the variance:

$$Z \approx N(\lambda_d C, \lambda_d C)$$

The number of patrons left after a vehicle leaves is:

$$u = \max(0, Z-S)$$

and the expected steady state value of u is:

$$u^* = \int_S^\infty \frac{w}{\sqrt{2\pi}} \exp\{-1/2(\frac{w-\lambda_d C}{\sqrt{\lambda_d C}})^{1/2}\}d_w$$

$$= \lambda_d C - S + (S - \lambda_d C)\,\Phi(\frac{S-\lambda_d C}{\sqrt{\lambda_d C}}) + \sqrt{\lambda_d C}\;\phi(\frac{S-\lambda_d C}{\sqrt{\lambda_d C}}) \qquad (3.42)$$

The effect of this modification is expected to have little or no effect on the service process unless patrons waiting for delivery are boarded by destination rather than on a first-come, first-served basis. However, the delay at the depot may be increased.

The final modification introduced is in effect a vehicle holding strategy. In very uncongested systems, it is often possible to insert a certain amount of idle time between the end of the delivery period and the beginning of the collection phase.* The duration of this idle period, I, should be set with the variability of the collection process in mind. A reasonable strategy would be to set I so that the probability of not collecting a patron is less than some parameter α. In this case, let the expected collection period be G' (given by Eq. 3.17 with $x^*=n_g$). As above, we assume that $Var(G') = G'$ and is normally distributed. Then we wish to set I such that:

$$P_r\{t_g(Y,0) < G-I\} < \alpha$$

$$\Phi(\frac{G-I-G'}{\sqrt{G'}}) < 1 - \alpha \qquad (3.43a)$$

If $\alpha=.01$, then I should be set such that:

$$\frac{G-I-G'}{\sqrt{G'}} = 3$$

$$I = \max\{G-G' - 3\sqrt{G'} \ ; \ 0\} \qquad (3.43b)$$

In section 3.3 we developed a deterministic model of feeder services with fluid approximations. At this point, we have developed stochastic corrections to account for the variability of the arrival process and tour lengths and for the "integerness" of patrons and stops. These are the major stochastic elements of the system's performance which were not captured in the original model. We now estimate the expected level of service of the system.

*The period I may be used for other vehicle operations; see Sec. 3.8.

3.5 Delays and the Level of Service

It is now possible to obtain the number of patrons waiting throughout the vehicle cycle, $N_w(t)$. The value of $N_w(t)$ at the beginning and end of the cycle process is:

$$N_w(t_3) = N_w(t_0) = x^* - n_g + \lambda_g G \qquad (3.44a)$$

with the assumption that patrons entering the system during the collection period are not immediately assigned for collection.

With assignment at the transfer point:

$$N_w(t_3) = x^* - n_g + \lambda_g (G+D+R/2) \qquad (3.44b)$$

In what follows, we use the assignment assumption of Eq. 3.44a, with assignment at the beginning of the collection period.

A linear approximation permits simple estimation of $N_w(t)$ during the collection process $(t_2$ to $t_3)$. The number of patrons riding may be obtained in a similar process:

$$N_r(t_2) = 0$$

$$N_r(t) = n_g \qquad t \in \{t_0, \frac{t_0+t_1}{2}\}$$

$$N_r(t) = n_d \qquad t \in \{\frac{t_0+t_1}{2}, t_1\} \qquad (3.45)$$

and a linear approximation may be made for (t_1, t_2) and (t_2, t_3)

Given the expressions for the number of patrons in the system, it is possible to determine average delays by using Little's formula [43]. Let T_g^w denote waiting time at home, T_d^r the riding time of an inbound (distributed)

passenger and T_g^r the riding time of an outbound (collected) passenger.
Then from figure 3.1:

$$E(T_g^w) = \frac{1}{n_g} \int_{t_0}^{t_3} N_w(t)d_t$$

$$= \frac{1}{n_g}[(x^* - \lambda_g(D+R) + n_g/2)C - n_gG/2 - \frac{(G-G') \cdot n_g}{2}]$$

$$\simeq \frac{x^*}{\lambda_g} - \frac{C}{2} + \frac{G'}{2} \tag{3.46}$$

$$E[T_g^r] = \frac{1}{n_g} \int_{t_1}^{t_3} N_r(t)dt + \frac{1}{n_g} \int_{t_0}^{\frac{t_0+t_1}{2}} N_r(t)dt$$

$$= \frac{R+G'}{2} + (G-G'-I) \tag{3.47}$$

The delivery time deserves more detailed attention. The actual delivery process ends with a discrete delivery event. Consequently, an estimate of the number of stops made for the purpose of calculating the expected riding time is:

$$E[T_d^r] = \frac{t_d(n_d+0.5,0) + R}{2} \tag{3.48}$$

For very uncongested systems, less than one stop may be scheduled, so a more accurate expression is:

$$E[T_d^r] = \max\{t_d(1,0), \frac{t_d(n_d+0.5,0)}{2}\} + \frac{D}{2} \tag{3.49}$$

A similar correction may be made to the collection riding time. However, since waiting time at home increases by the same amount that riding time

decreases, this correction is usually not important, especially since the available evidence is that waiting time at home is no more onerous than riding time [41].

The waiting time at the depot depends upon the pattern of patron arrivals and the expected number of patrons waiting who cannot be carried on the vehicle. Again, using Little's formula, the waiting time is:

$$E[T_d^W] = \int_{t_o}^{t_3} N_d^W(t)dt = \frac{C}{2} + \frac{u^*}{\lambda_d}$$

for randomly arriving patrons, and

$$E[T_d^W] = \frac{u^*}{\lambda_d}$$

for cases in which the feeder vehicle meets a line haul vehicle for transfer.

3.5.1 Variance of the Waiting Time.

Using some of the simple approximations described above, bounds on the variance of the waiting time at home may also be estimated. Our dispatching policy implies that stops are randomly chosen from among those assigned for collection. Waiting time at home may be divided into three components: waiting between arrival and the first collection period, w_1, (possible) waiting until subsequent collection periods, w_2, and waiting time during the collection period, w_3. Using the linear service assumptions, collections are randomly distributed over the collection period. Similarly, arrivals are randomly distributed over the period before the first collection phase. Thus

$$var(w_1) = c^2/12 \tag{3.51a}$$

$$var(w_3) = G^2/12 \tag{3.51b}$$

To calculate the variance of the (possible) wait until subsequent collection periods, we define g as the number of cycles until collection. Then:

$$var(w_2) = c^2 var(g)$$

The process of selecting an individual patron is a Bernoulli process resulting in a geometric distribution for the time until success (see Feller [21] pg. 303). Consequently,

$$var(w_2) = c^2(x^*/n_g)(x^*/n_g-1)$$

The variance of the wait time is then:

$$var(T_g^w) = G^2/12 + c^2(1/12 + (\frac{x^*}{n_g})^2 - \frac{x^*}{n_g}) \qquad (3.51c)$$

3.6 Several Vehicles Per Zone

When the line haul system headways are short, the service area large or the demand rate high, it may be advantageous to operate more than one vehicle per zone. In such cases, there are two efficient strategies for deploying vehicles. One might partition the service area into zones and operate just one bus per zone. Since inbound passengers will sort themselves at the transfer point and the vehicles may operate in a smaller area (resulting in more efficient tours), partition is more advantageous than operating all the vehicle simultaneously throughout the whole service area. The model described above can be applied separately to each of the subzones or to one representative zone for this case.

The other operating strategy consists of having vehicles operate out-of-phase or offset in time, in the sense that vehicles visit the transfer point at different times. This type of service is similar to common fixed route service: all vehicles provide the same service but at different times. In the fixed route situation, the operator usually schedules a

-98-

headway or gap between vehicle departures. For the feeder service discussed here, the headway is the time between vehicles' visits to the depot. With a cycle length of C and M vehicles in out-of-phase service, then the headway or time between visits to the depot is C/M. The out-of-phase operation tends to be more advantageous when the line haul system headways are short or if the transfer point is fairly distant from the service area (leading to long cycle lengths and consequent long delays at home).

A model for the multi-vehicle out-of-phase system operation may be developed in the same manner as for the one vehicle model. Figure 3.4 illustrates the fluctuations in the number of patrons waiting at home (N_{wh}) and riding in vehicles (N_r, N_r') over time for a two vehicle, out-of-phase service. In the figure, the delivery and collection demand rates are equal $(\lambda_g = \lambda_d)$ and the cycle length is not sufficiently long to permit all patrons to be collected before the end of any one collection period. An analogous situation was illustrated for the one vehicle case in Figure 3.2.

At time t_o in Fig. 3.4 , one vehicle has departed the service area for the depot, while the other vehicle is still in the midst of delivery. At time t_o', the same situation is observed but with the two vehicles' roles reversed. The first vehicle has now returned from the depot and is in the process of delivering patrons, while the second vehicle departs for the depot. Since the two vehicle cycles are identical (only displaced in time), the headway between the two vehicles is always $t_i' - t_i$, which is equal to half the cycle length. The number of patrons waiting at home declines during each collection phase, but the steady state pool size is attained every headway (at times t_2 and t_2' in the figure).

By symmetry, one can see that it suffices to study one bus during

-99-

Fig. 3.4 Two vehicles Operating Out-of-Phase ($\lambda_g = \lambda_d$)

Number in System: N(t)

Number at home: $N_{wh}(t)$

Number in Vehicle $N_r(t)$, $N'_r(t)$

N = Number of patrons

N

-100-

t_o t_1 t_o t_2 t_1 t_{o+3} t_2 t_{1+3} t_3
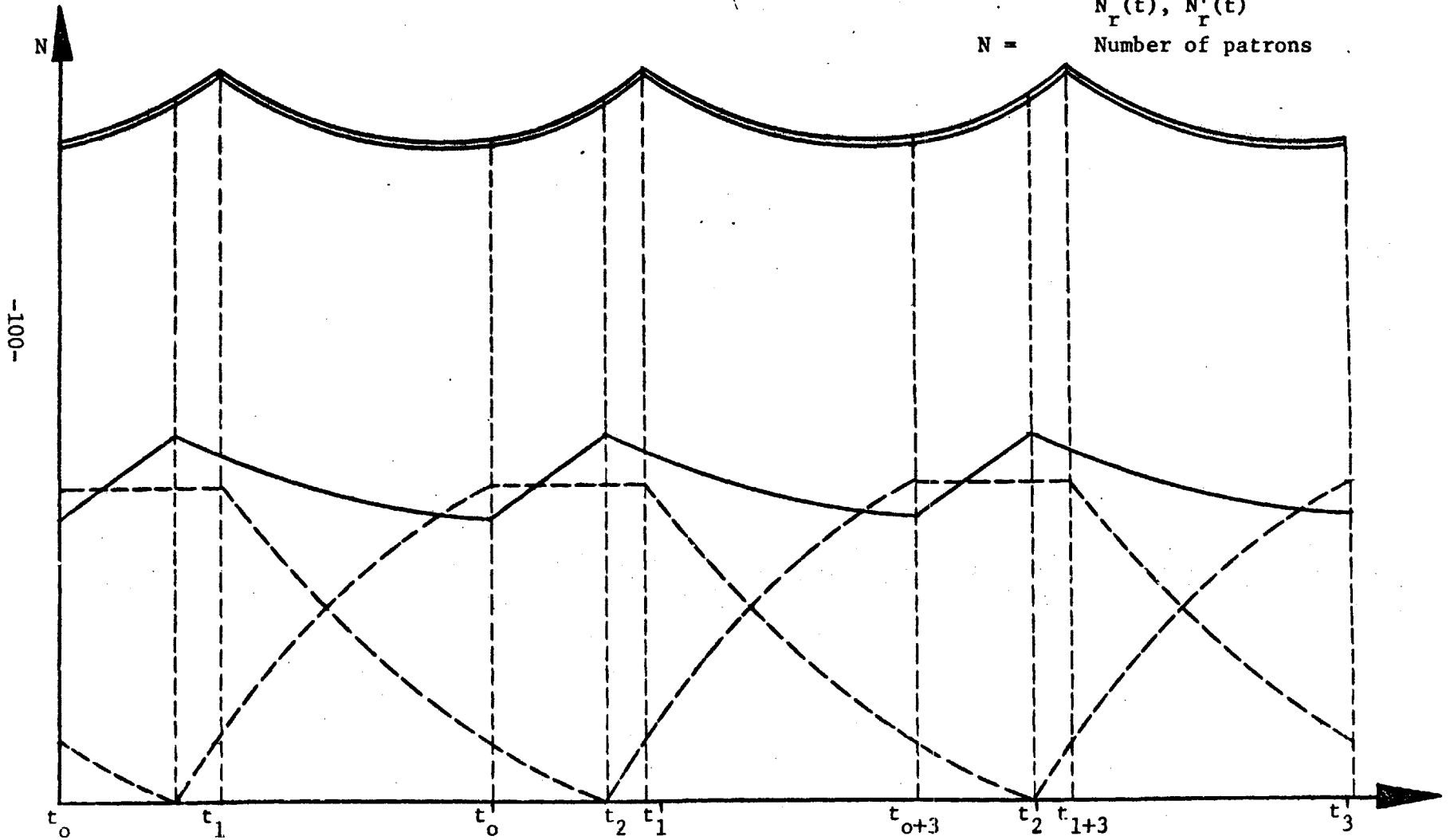
a cycle. The expected number of stops made during distribution is:

$$n_d = \lambda_d \cdot C/M\gamma$$

where M is the number of vehicles operating in the service area, which is only 1/M times the number of stops in the one vehicle case (Eq. 3.13). The distribution time may then be found using Eq. 3.8:

$$D \approx n_d b_d + \frac{r \sqrt{A}}{v} \ (\sqrt{n_d + 0.5} - \sqrt{0.5}) \tag{3.8}$$

The steady state conditions are analogous to the one vehicle case (Eq. 3.15):

$$\lambda_d C/M \quad < \quad S$$
$$\lambda_g C/M \quad < \quad S$$

to insure that vehicle capacity is not exceeded, and:

$$
\begin{aligned}
t_g'(x^*, x^* - n_g) &= G \qquad &\text{if } x^* > n_g \\
&\leqslant G \qquad &\text{if } x^* = n_g
\end{aligned}
\tag{3.52}
$$

That is, each bus must be able to serve its share of the total demand, $C \cdot \lambda_g/M$. The number of stops made and the patrons collected are:

$$Y = \lambda_g \cdot C/M$$

$$n_g = \lambda_g \cdot C/M\gamma = Y/\gamma \tag{3.53}$$

and the solution for the steady state delivery and pickup pool sizes is identical to the one vehicle case (Eqs. 3.37 to 3.39). The equations are summarized in Section 3.9.

A difficulty with the derivation of the model is that the tour length expressions do not capture the effect of competition of buses for the same

passenger when the collection or distribution phases of several buses overlap. This is expected to be only a minor effect unless the extent of the overlap is large and the system is relatively uncongested, however [16]. It is likely to be advantageous to partition the service zone under these conditions, so the problem is relatively minor.

3.7 Unphased Service

An alternative to the mode of operation described above consists of a system which intermingles delivery and collection stops. In contrast to the phased service in which all delivery stops precede all collection stops, the unphased service uses a single tour among the various stops. This service has the advantage of creating more efficient tours at the expense of increasing the delivery riding times. Fortunately, the tour length expression developed in Section 3.2 and the fine tuning corrections of Section 3.4 are generally applicable to this case, so constructing a performance function for this service is relatively simple.

To begin, we assume that patrons for collection are assigned to vehicles when the vehicle leaves the transfer point. The potential number of stops, N, includes the number of assigned collections and the number of deliveries, n_d. The time available for collection is simply the cycle time less the rendezvous time, so in a manner analogous to Eq. 3.17 we require:

$$C-R = t(x, x-n_g-n_d) \qquad \text{if } x > n_g+n_d$$

$$C-R \geqslant t(n_g+n_d, 0) \qquad \text{if } x = n_g+n_d$$

$$(3.54)$$

Similar conditions for solution exist as in the phased service case (Eqs.

-102-

3.10, 3.15, and 3.18). In particular, the vehicle capacity cannot be exceeded anywhere along the tour. If $x > n_g + n_d$ then:

$$C-R = b(n_g + n_g) = \frac{r\sqrt{A}}{v}(\sqrt{x^*+0.5} - \sqrt{x^*-n_g-n_d+0.5}) \qquad (3.55)$$

and we may solve for x^* to obtain:

$$x^* = [\max\{\sqrt{0.5}; \quad [\frac{n}{2k} - \frac{k}{2}]\}]^2 + n - 0.5 \qquad (3.56)$$

where

$$k = \frac{v}{r\sqrt{A}} (C-R-b_n)$$

$$n = n_g + n_b$$

which is quite similar in form to Eq. 3.20.

The fine tuning corrections made to the collection period of the phased service model - relating to the integerness of stops and the randomness of the tour time - should also be made for the unphased model. These corrections have the same form as those for the phased service model. Of course, the unphased service has the possibility of exceeding the vehicle capacity anywhere along the tour, whereas this could only occur at the beginning or end of the tour in the phased service.   Note however, that the variance of the unphased service process is larger than in the phased service case, since we assume a Poisson-type process in which the variance equals the number of stops made. As a result, of the higher variance, the tour is expected to be less efficient then it would otherwise be, which is what we expect.

Equations summarizing the unphased service appear in Section 3.9.

A comparison of phased and unphased service appears in Section 5.4.

3.8 Zonal Service

One strategy for operating an integrated transit system is to divide the metropolitan area into a number of zones. Flexibly routed vehicles then circulate in each zone, with scheduled stops at a transfer point to line haul services or for interzonal transfers. This mode of operation is similar to the Ann Arbor Teltran System (see Section 5.2).

In zonal service, the bulk of all trips are expected to have an origin or destination at the transfer point. However, a certain proportion of trips will have both origins and destination within the same zone. The only difference between the feeder service model described above and a zonal service model is that these intra-area trips are served by the same vehicles which are in cycled service to the transfer point. In the terminology of demand responsive models, there is a certain proportion of many-to-many trips (with both origin and destination in the zone) amid the majority of many-to-one trips, which originate or are destined for the transfer point.

There are three options for serving these many-to-many trips:

1) collect and distribute many-to-many trips during the idle period, I, between the delivery and collection phases.

2) include many-to-many collections during the collection phase and many-to-many deliveries during the delivery phase; in this case, all patrons visit the transfer point.

3) service many-to-many trips during either the collection or delivery period tours.

Strategy 1 is superior since it does not degrade the service of other patrons. However, a substantial idle period is required and this may be expensive. The length of a tour to serve these many-to-many trips may be developed from the tour expression Eq. 3.8.

The other two strategies simply require modification of the incoming and out-going demand rates and then solution of the steady state characteristics of the system as in Section 3.4. For strategy 2, for example:

$$\lambda'_g = \lambda_g + \lambda_m \tag{3.57}$$

$$\lambda'_d = \lambda_d + \lambda_m$$

where λ_m is the demand rate of many-to-many trips. The expected riding time for many-to-many trips is:

$$E[T_m^r] = \frac{G'+D+R}{2} \tag{3.58a}$$

Or in unphased service:

$$E[T_m^r] \text{ unphased} = (C-R)/2 \tag{3.58b}$$

Similarly, for the case in which many-to-many trips are served during delivery, the demand for delivery should be modified to be:

$$\lambda'_d = \lambda_d + 2\lambda_m$$

where the parameter 2 is introduced since both a collection and a delivery is required for many-to-many trips. The model may be optimistic in application to strategy 3 since the constraint that collections precede deliveries is not imposed.

A description of a zonal service and application of the model occurs in Section 5.2.

3.9 Model Summaries and Notation

The model summaries appearing below are intended to provide a convenient reference for equations and notation. Equation numbers provide a reference to derivations in the text.

3.9.1 Expected Tour Length

This expression represents the expected tour length from a randomly located point through n of N points, without returning to the origin. It assumes next-nearest-point vehicle routing, circular area shape, and uniformly distributed points. Using these assumptions, simulation experiments indicate that the tour length expression is within 5% of observed average tour lengths with tours of more than 2 stops:

$$d \approx r\sqrt{A}(\sqrt{N+0.5} - \sqrt{N-n+0.5}) \tag{3.6}$$

The number of stops which can be made from a pool of N in a time t is:

$$n = \min\{N; \quad N+0.5 - (-(\frac{vt}{r\sqrt{A}}) + \sqrt{N+0.5})^2\}$$

which may be derived algebraically from Eq. 3.6 (as long as stop time is zero).

3.9.2 Model Options for Pure Feeder Services

Within the chapter, a number of feeder service options are presented. These may be summarized by the vector H:

$$H = H(P,M,C,I,S) \tag{3.59}$$

where P is the option of operating in phased service (with collection and
distribution separated) or unphased (with collection and distribution interspersed) service,

-106-

M is the number of vehicles operating in a service area and visiting
the depot separately,

C is the vehicle cycle length,

I is the (scheduled) idle time between collection and delivery, and

S is the vehicle capacity.

These five input parameters may be used to describe a particular operating option.

In addition to the service options, input parameters are also required for the models, summarized by the vector P:

$$P = P(A,r,v,\lambda_d,\lambda_g,b_d,b_g,\gamma,R) \tag{3.60}$$

which includes area size (A), the route factor (r), average vehicle speed (v), demand rates for delivery and collection (λ_d,λ_g), average boarding times (b_d,b_g), average number of patrons travelling together (γ), and rendezvous time for the depot (R, given by Eq. 3.11).

The models result in a vector of level of service attributes J:

$$J = J(T_g^w, T_d^w, T_d^r, T_g^r, \sigma, TT)$$

where the first four elements are the expected time of waiting at home, waiting at the depot, riding for delivery, and riding for collection $(T_g^w, T_d^w, T_d^r, T_g^r$ respectively). The vector σ is the variances of these four travel time components. Finally, TT is the sum of the expected waiting and riding time of a random patron. In addition, resources consumed such as vehicle hours or miles of operation may be calculated.

3.9.3 Phased Feeder Service

This model is based upon the tour length expression in Section 3.9.1.

Vehicles are assumed to operate on a schedule from a depot or transfer point. Vehicles first deliver patrons in the service area, then collect patrons before returning to the depot.

Necessary conditions to insure the feasibility of steady state operation relate to the vehicle capacity for both delivery and collection:

$$\lambda_d \cdot C/M\gamma \quad \leqslant \quad S \qquad\qquad (3.10)$$

$$\lambda_g \cdot C/M\gamma \quad \leqslant \quad S \qquad\qquad (3.15)$$

In addition, sufficient time must be available to at least <u>board</u> the number of patrons who <u>arrive during</u> a cycle:

$$b_g \cdot \lambda_g C/M\gamma \quad \leqslant \quad (C-D-R) \qquad\qquad (3.16)$$

The model is solved in stages, first solving for the distribution time and the available collection period, then a calculation of the steady state pool pick-up size and finally calculation of the output vector.

1. Distribution Time

$$n_d = \lambda_d \cdot C/M\gamma \qquad\qquad (3.8)$$

$$D \simeq n_d b_d + \frac{r\sqrt{A}}{v}(\Delta\sqrt{n_d+0.5} - \sqrt{0.5}) \qquad\qquad (3.23)$$

with

$$\Delta = 1 - \frac{n_d}{8(n_d+0.5)^2}$$

$$u^* = n_d\gamma - S + (S - n_d\gamma)\Phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}}) + \qquad\qquad (3.42)$$

$$\sqrt{n_d\gamma} \; \phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}})$$

2. Collection Phase

$$G \equiv C - D - R - I \qquad (3.12)$$

$$n_g = \lambda_g \cdot C/M\gamma \qquad (3.13)$$

$$G' = \min\{G;\ n_g b_g + \frac{r\sqrt{A}}{v}\sqrt{n_g + 0.5}\} \qquad (3.19)$$

if $G' < G$ then:

$$I = \max\{0;\ G - G' - 3\sqrt{G'}\} \qquad (3.43)$$

otherwise $I = 0$*

3. Steady State Pickup Pool Size

if $G' < G$ then $x^{*'} = n_g$, otherwise

$$x^{*'} = x^* + \max\{0;\ n_g - Y';\ n_g - Z'/\gamma\} \qquad (3.37)$$

$$x^* = [\max\{0;\ \frac{0.5 + n_g - k^2}{2k}\}]^2 + n_g \qquad (3.20)$$

$$k = \frac{v}{r\sqrt{A}}[G - n_g b_g]$$

$$Y' = x^* - (x^* - n_g)\Phi(\frac{x^* - n_g}{\sqrt{n_g}}) - \sqrt{n_g}\phi(\frac{x^* - n_g}{\sqrt{n_g}}) \qquad (3.38)$$

$$Z' = S - (S - \gamma n_g)\Phi(\frac{S - \gamma n_g}{\sqrt{\gamma n_g}}) - \sqrt{\gamma n_g}\phi(\frac{S - \gamma n_g}{\sqrt{\gamma n_g}}) \qquad (3.39)$$

---

*Where we assume that a .99 probability of collecting all passengers is desired.

4. Delays

$$E[T_g^w] = \frac{x^{*'}}{\lambda_g} - \frac{C}{2M} + \frac{G'}{2} \quad \text{for assignment when collection begins} \tag{3.46}$$

$$\text{var}(T_g^w) \simeq G^2/12 + C^2/M^2 \left(.08 + \left(\frac{x^{*'}}{n_g}\right)^2 - \frac{x^{*'}}{n_g} \right) \tag{3.51c}$$

$$E[T_g^r] = \frac{R+G'}{2} + (G-I) \tag{3.47}$$

$$\text{var}(T_g^r) = E[T_g^r]^2/12$$

$$E[T_d^r] = \max\{b_d + \frac{.52r\sqrt{A}}{v} ; \frac{D}{2}\} + R/2 \tag{3.49}$$

$$\text{var}(T_d^r) = E[T_d^r]^2/12$$

$$E[T_d^w] = \frac{u^*}{\lambda_d} + C/2M \quad \text{for random arrivals}$$

$$\cong u^*/\lambda_d \quad \text{for transfers} \tag{3.50}$$

$$E[TT] = \frac{\lambda_g}{\lambda_g + \lambda_d}(E[T_g^w] + E[T_g^r]) + \frac{\lambda_d}{\lambda_g + \lambda_d}(E[T_d^w] + E[T_d^r]).$$

3.9.4 Unphased Feeder Service

This service does not separate the delivery and collection processes; stops are intermingled during one long tour.

1. Steady State Depot Pool Size

$$n_d = \lambda_d \cdot C/M\gamma$$

$$u^* = n_d\gamma - S + (S - n_d\gamma)\Phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}}) + \sqrt{n_d\gamma}\phi(\frac{S-n_d\gamma}{\sqrt{n_d\gamma}})$$

2. Steady State Pickup Pool Size

$$n = n_d + n_g = (\lambda_d + \lambda_g)C/M\gamma$$

if $nb + \dfrac{r\sqrt{A}}{v}(\sqrt{n+0.5} - \sqrt{0.5}) < C - L$ then $x^{*'} = n$, otherwise:

$$x^* = [\max\{ 0; (\frac{n + 0.5 - k^2}{2k})\}]^2 + n$$

$$x^{*'} = x^* + \max\{ 0; n_g - Y; n_g - Z/\gamma\}$$

$$k = \frac{v}{r\sqrt{A}}(C - R - nb)$$

with

$$Z = S - (S - n_g\gamma)\; \Phi(\frac{S-n_g\gamma}{\sqrt{n_g\gamma}}) - \sqrt{n_g\gamma}\,\phi(\frac{S-n_g\gamma}{\sqrt{n_g\gamma}})$$

$$Y' = (x^*-n_d) \doteq (x^*-n)\Phi(\frac{x^*-n}{\sqrt{n}}) - \sqrt{n}\;\phi(\frac{x^*-n}{\sqrt{n}})$$

3. Delays

$$E[T_g^r] = E[T_d^r] = C/2$$

$$E[T_g^w] = \frac{x^{*'} - n_d}{\lambda_g} - \frac{C}{2M}$$

$$E[T_d^w] = \frac{u^*}{\lambda_d} + C/2M \qquad \text{for random arrivals}$$

$$= u^*/\lambda_d \qquad \text{for transfers.}$$

3.9.5 Zonal Service

Zonal service is mainly feeder service with a small amount of intra-zone many-to-many service provided. Intra-zonal trips are generally

served during the normal collection or delivery tours. To model this service, demand rates for collection or delivery are incremented:

$$\lambda'_g = \lambda_g + 2\lambda_m \tag{3.57}$$

or

$$\lambda'_d = \lambda_d + 2\lambda_m$$

or

$$\lambda'_g = \lambda_g + \lambda_m \text{ and } \lambda'_d = \lambda_d + \lambda_m$$

where λ_m is the demand rate of many-to-many trips. The expected riding time is:

$$E[T^r_m] = \frac{G'}{3} \text{ or } \frac{D}{3} \text{ or } \frac{G' + D}{3} \tag{3.58a}$$

in phased service or

$$E[T^r_m] \text{ unphased} = (C-R)/3 \tag{3.58b}$$

in unphased service. The other model outputs are found by applying the equations summarized above for feeder services.

3.9.6 Glossary

d_i : expected distance to the nearest of i points from a random point

A : area size

d : expected tour length (without return to the origin)

r : route factor; expected ratio of street network to straight line distance

n : number of stops on tour

N : number of eligible stops in service area

-112-

v : expected vehicle speed

D : period for delivery of patrons

G : period for collection of patrons

G' : time actually used for patron collection

R : rendezvous period for travel to and from the depot

I : idle period in between delivery and collection periods

M : number of vehicle visits to depot per cycle; number of vehicles in offset service

S : vehicle capacity

λ : average arrival rate (λ_d for delivery, λ_g for collection)

b : average boarding time (b_d for delivery, b_g for collection)

γ : average group size boarding or exiting at a stop

$N_w(t)$: number of patrons waiting for service at time t

$N_r(t)$: number of patrons riding a vehicle at time t

u^* : steady state pool size of stops for delivery at the beginning of the delivery phase who cannot be carried in the vehicle

x^* : steady state pool size of potential stops for collection at the beginning of the collection phase

$x^{*'}$: value of x^* modified for stochastic variations

Y : number of stops made in a tour

Z : number of patrons served in a tour

$\phi(.)$: standardized normal density function

$\Phi(.)$: standardized cumulative normal density function

$L(.)$: logistic curve approximation to the cumulative normal density function

T^r : riding time (T_g^r for collection, T_d^r for delivery)

T^w : waiting time (T_g for collection, T_d for delivery)

TT : total travel time of a random patron.

CHAPTER 4

GENERALIZATION OF THE FEEDER SERVICE MODEL

4.1 Introduction

The basic development of the feeder service model was presented in

Chapter 3. This model assumed uniformly distributed demand in the serv-

ice area, next-nearest-point vehicle routing and a circular service area.

In this chapter, we investigate the effects on system performance if these

three assumptions are violated. In general, the feeder service model is

found to be fairly robust, in that predictions based upon these assumptions

are fairly accurate even though the assumptions are not strictly correct.

Heuristic methods to generalize the model to situations in which this con-

clusion does not hold are also discussed.

This chapter is divided into separate treatments of the three assump-

tions. Monte Carlo simulation experiments are used in each of the discus-

sions. Section 4.2 considers the effect of different service area shapes

and street network geometries. A heuristic procedure for modifying the

tour length expression is presented to correct the model for different

area characteristics. In Section 4.3, some experimental results concern-

ing tour lengths with spatially non-uniform demand are presented; the ef-

fect of all but extreme non-uniformity is found to be fairly minor. The

following section briefly considers the issues involved in selecting rout-

ing strategies and compares some alternatives. In the case of feeder serv-

ices, it appears that the tour length approximation formula (Eq. 3.8),

based upon the simple proceed-to-the-next-nearest point algorithm, gives

results comparable with the tours resulting from more complicated routing strategies, including manual routing of vehicles.

With this Chapter, the development of the feeder service model is completed. The experiments reported here indicate that:

- spatially uniform demand is not a critical model assumption, since even relatively extreme demand density gradients do not substantially alter observed tour lengths. With very large density gradients over the service area and a small number of stops on a tour, a correction term may be useful, however.

- the tour length expression based upon next-nearest-point vehicle dispatching is a fairly good model of tour lengths resulting from manual or minimum tour length vehicle routing.

- the effect of area shape on expected tour lengths is small but can be significant. A correction factor may be easily introduced into the feeder service model, however. Table 4.4. summarizes the factors by which to multiply the tour length expression to correct for the effect of area shape. Service areas may be approximated by one of the shapes listed and the appropriate factor, α, then used in applications.

4.2 Tours In Irregular Areas

The development of the feeder service model in Chapter 3 assumed a regular, circular service area. In this section, we shall retain the assumptions of uniformly distributed vehicle stops and next-nearest-point vehicle routing, but consider the effect of irregular or differently shaped areas.

A basic concept used in this discussion is that of the travel factor in a service area, defined as the ratio of the expected distances between two points along the street network or via a straight line:

$$r = E[d_1]/E[s] \tag{4.1}$$

where s is the straight line or airline distance and d_1 is the minimum street network distance between two points.

Regular homogeneous street networks are defined as street networks in which the network density or street length per unit area is a finite constant and the geometry with which streets intersect has a regular pattern. In such networks, the travel factor will be constant over a suitably defined area.

In practice, networks are rarely homogeneous, so the travel factor varies over a metropolitan area. We define the expected travel factor as the travel factor expected in the near vicinity of a random point in the service area. The expected travel factor may be estimated by the average of the travel factor observed around a relatively large number of randomly distributed points (i.e., distributed as a spatial Poisson process). Occasionally, it may be preferable to define the expected travel factor in relation to a spatial probability density function other than the uniform distribution. In what follows, however, the discussion is confined to uniformly distributed points.

In addition to the effect on trip lengths caused by the circuity of

---

*Note that this travel factor need not be exactly equivalent to the traditional route factor, defined as the expected ratio of the street network and straight line distances:

$$r_f = E[d_1/s]$$

the street network, the service area shape also influences the expected travel distance. As the service area becomes more irregular or elongated in shape, the expected straight line travel distance between random points increases and, thus, the expected tour lengths in the areas increase. Simulation experiments are reported below which indicate the extent to which distances increase in differently shaped areas.

The effects due to differently shaped service areas are relatively small, but a correction to the tour length expression (Eq. 3.8) may be desireable to improve accuracy. However, it is also the case that these effects become less significant as the number of points in the area becomes large. Indeed, as the number of points increases without bound, the distance to the nearest is unaffected by the area shape (see Section 3.2, Eq. 3.2b). Our strategy for introducing a correction term is to multiply the tour length expression by a factor which is the average of the expected distance between any two random points in the actual service area and the expected distance between the closest of many points in the service area. In the latter case, the effect of area shape may be disregarded.

In the remainder of the section, the distance between two random points in differently shaped areas and the suggested correction to the tour length expression are discussed. Section 4.2.3 presents an application and validation of the modification procedure.

4.2.1 Distance Between Random Points

It is possible to derive the expected distance between two uniformly and independently distributed points by means of the integral:

-118-

$$s = \int_A \int_A \frac{s(x,y)}{A^2} \, dx \, dy \qquad (4.2)$$

where $s(x,y)$ is the distance between the two points x and y and the probability density function of x or y is $1/A$ everywhere in the area.

Values of this distance have been reported for circular and various rectangular areas (see Borel [7] and Haight [29]). It is also possible to derive the expected distance between two randomly distributed points over a street network by a suitable redefinition of the distance function, $s(x,y)$. The network geometry and, to a lesser extent, the orientation of the street network with respect to the area boundaries influences this expected distance. For example a grid street network parallel to the boundaries of a square results in a 1% increase in the expected travel distance, compared to a grid network parallel to the area's diagonals [13;7].

Evaluation of the integral (Eq. 4.2) becomes quite laborious for all but simple areas, however. Consequently, simulation experiments have been used to find expected distances (see Eilon [18]). One notable exception to this observation is that of rectangular areas with grid networks, in which the expected travel distance is the sum of two independent distributions, corresponding to travel along the two axis of the street network. Analytic solutions are easily obtained in this case because of the independence of travel distance along the two street network axis [40]. The distance between two random points in a rectangular grid is [20]:

$$E[d_r] = 1/3(a+b) = \frac{(1+k)\sqrt{A}}{3\sqrt{k}} \qquad (4.3)$$

where a and b are the side lengths and $k = b/a$.

The expected distance between random points with various travel factors and area shapes is reported in Table 4.1. Included in the table are results based upon Monte Carlo simulation experiments using a standard pseudo random number generator.

The effect of area shape is surprisingly small on the expected street network distance between two points. The expected straight line distance between points becomes relatively large in area shapes that are relatively elongated. However, the ratio of street network travel to straight line travel distance goes down as the area becomes more elongated. This occurs because a greater percentage of travel may be served by streets along the long axis of the service area. For example, the expected straight line distance between two points in a 2x1 rectangle increases by 9% compared to the distance in a square, but the travel factor declines by 4%. As a result, the expected street network distance declines by only 6%. As a consequence of these offsetting factors, the street network travel distance remains relatively stable in comparing alternative shapes. An exception to this general observation is the case of rectangles with a high ratio between the side lengths (as in the 3x1 rectangle, with 18% greater travel distance than a circle), when the effect of area shape begins to dominate the decline in the travel factor.

Radial/circumferential street networks have shorter expected travel distances than traditional grid systems. However, we have not included any consideration of intersection congestion, which is more severe with radial/circumferential street networks.

TABLE 4.1

EXPECTED DISTANCES BETWEEN TWO RANDOM POINTS

Area	Straight Line	Along a Grid Network*	Along a Radial/ Circumferential Network**
Circle	.511	.651	.561
Square	.522	.667	.660
Rectangle			
3 x 2	.54	.680	
2 x 1	.57	.707	
3 x 1	.64	.770	
Circular Sector[+]			
θ = 15°	.83	1.0	.83
30°	.52	.62	.52
45°	.51	.61	.51
60°	.53	.64	.53
75°	.54	.65	.54
90°	.53	.68	.52
Isoceles Triangle[+]			
θ = 15°	.94	1.0	.94
30°	.66	.73	.66
45°	.62	.74	.62
60°	.52	.68	.52
75°	.57	.74	.57
90°	.58	.75	.58

Note: Unit areas are assumed; results with two significant figures are from simulation experiments (95% confidence interval +.04).

* Rectangular grid network with links parallel to boundaries or bisection of radial angle.

**Radial/circumferential network in the square is approximated by grid parallel to square diagonals.

+ Angles refer to angles between radii or equal length sides.

4.2.2. Trips from Boundary Points to a Random Point

The distance from a border point to a random point may also be of interest if the tour length approximation is modified for few stops or if an irregular area is approximated by two regular shapes. The expected distance may be calculated from the integral:

$$E[d_n] = \int_A s(t,x)f(x)dx \qquad (4.4)$$

where $s(t,x)$ is the distance from the transfer point, t, to the location x, and $f(x)$ is the density function of x.

Tables 4.2 and 4.3 summarize some analytical (see Haight [29] or Larson and Odoni [40]) and simulation experiment results for this expected distance.

4.2.3 Tour Lengths in Regular Areas

In Section 3.2, we developed an approximate expression for the tour length among n of N randomly distributed points in a circular area. In this discussion, we used the distance between random points in a circular area:

$$d_1 \approx .51 \ r\sqrt{A} \qquad (3.1)$$

and the distance to the nearest of n random points, as n becomes large:

$$\lim_{n \to \infty} d_n = .5 \ r\sqrt{A/n} \qquad (3.2)$$

irrespective of the area shape. We then assumed that the distance to the closest of n points was:

$$d_n \approx .505r\sqrt{A/n} \qquad (3.3)$$

TABLE 4.2

TRIPS FROM BOUNDARY POINTS

(Unit Areas)

Area Shape	Expected L_1 Grid Distance	Travel Factor*
Circle	.80	1.26
Square		
Corner	1.00	1.27
Midpoint	.75	1.27
3 x 2 Rectangle**		
Corner	1.02	1.2
Midpoint	.72	1.2
2 x 1 Rectangle**		
Corner	1.06	1.2
Midpoint	.88	1.2
3 x 1 Rectangle**		
Corner	1.15	1.1
Midpoint	1.01	1.1

* rectangular street grid is assumed.

**Travel factors from simulated data, number of observations = 45

TABLE 4.3

DISTANCE FROM A VERTEX TO RANDOM POINTS IN ISOCELES

TRIANGULAR AND CIRCULAR SECTORS OF UNIT AREAS

Vertex Angle*	Triangle**		Circular Sector**
	Straight Line Distance	Grid Travel Factor[+]	Straight Line Distance
15	1.8	1.1	1.8
30	1.3	1.1	1.3
45	1.1	1.2	1.1
60	1.0	1.2	.92
75	.83	1.3	.82
90	.69	1.3	.75

* the vertex angle is at the intersection of the sector radii or equal
length area legs of the isoceles triangle.

**simulation Results: # Observations = 45

[+] grid street network travel factor

Since the expected street network distance is relatively insensitive to area shape, it is still possible to use the tour length approximation formula of Section 3.2 with a substitution for the expected street network or straight line distance of the actual service area. Then the approximate distance from a random point to the nearest of n other points is:

$$d_n \approx .5(E[d_1|A] + .5r\sqrt{A})/\sqrt{n}$$

$$\approx .5(.5\sqrt{A} + E[s|A])r / \sqrt{n} \tag{4.5}$$

where $E[d_1|A]$ is the expected distance between two random points in the service area. By scaling units of measurement, we know that

$$E[s|A] = E[s|A=1]\sqrt{A}$$

and

$$E[s|A] = E[s|A=1]\sqrt{A}$$

where $E[s|A=1]$ and $E[d_1|A=1]$ are the straight line and street network distance between two random points in an area of unit area but of the same shape as the service region. Then, the approximate distance from a random point to the nearest of n other points is:

$$d_n \approx .5(.5\sqrt{A} + E[s|A])r/\sqrt{n}$$

$$\approx .5(.5 + E[s|A=1])r\sqrt{A/n}$$
$$\approx .5\alpha r/\sqrt{n}$$
where

$$\alpha = .5 + E[s|A=1]$$

that is, the parameter α is one half plus the expected straight line distance between two random points in a unit area which has the same shape as the service region. In the case of a circle, for example, $\alpha = .5 + .571 = 1.011$ using the result reported in Table 4.1. This factor α may be used to modify the tour length expression to account for area shape.

Following the derivation of Section 3.3, the expected tour length among n of N points is approximately:

$$d \approx \sum_{i=1+N-n}^{N} d_i$$

$$\approx \alpha r\sqrt{A}/2 \sum_{i=1+N-n}^{N} \sqrt{1/i}$$

and with a continuum approximation:

$$d \approx \frac{\alpha}{2} r\sqrt{A} \int_{N-n+0.5}^{N+0.5} x^{-1/2}dx$$

$$\approx \alpha\, r\sqrt{A}(\sqrt{N+0.5} - \sqrt{N+0.5-n}) \qquad (4.6)$$

with $\alpha = .5 + E[\mathbf{s}_i \mid A=1]$,

which is only α times the expression derived in Section 3.3.

The value of α for regularly shaped areas may be calculated from Table 4.1. For example, a tour in an area which was approximately a 3x2 rectangle would be:

$$d_n \approx (.5\sqrt{A} + .54\sqrt{A})r(\sqrt{N+0.5} - \sqrt{N-n+0.5}$$

$$\approx 1.04r\sqrt{A}\,(\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

or 4% longer than in the circular service area case, as long as r is un-
changed. With a grid street network, the tour in the 3 x 2 rectangle
would be:

$$d_n^g \approx (.5r\sqrt{A} + E[d_1|A]) \; (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx (.63 + .68)\sqrt{A} \; (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx 1.31\sqrt{A}(\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

which, due to the reduction in r from 1.28 to 1.26, implies that the
tour length is only 2% longer than in the circular service area case.

Values of the tour length constant $(.5 + E[s|A = 1])$ may be cal-
culated from Table 4.1 and are summarized in Table 4.4.

4.2.4 Tours in Irregularly Shaped Areas.

The preceding results indicate that the expected travel distance
between two random points is not particularly sensitive to area shape,
at least relative to other parameters estimated for planning purposes
(such as vehicle speed or demand for trips). As a result, irregularly
shaped service areas may be approximated by one of the regular shapes dis-
cussed previously. If the service area can be divided into two (or more)
regions which do not overlap, then the area may be approximated by two
(or more) regular regions, a and b. The expected distance between two
random points is, then:

$$E[s] = P_{aa} \cdot E[S_{aa}] + P_{ab}[S_{ab}] + P_{ba}[S_{ba}] + P_{bb}[S_{bb}] \qquad (4.7)$$

where P_{ij} is the probability of point 1 in area i and point 2 in

TABLE 4.4

Tour Length Expression Parameters*

	α	Grid Street Network $(r\alpha)$
Circle	1.01	1.29
Square	1.02	1.30
Rectangle		
3 x 2	1.04	1.31
2 x 1	1.07	1.33
3 x 1	1.14	1.37
Circular Sector[†]		
θ = 15°	1.3	1.6
30°	1.0	1.3
45°	1.0	1.3
60°	1.0	1.3
75°	1.0	1.3
90°	1.0	1.3
Isoceles Triangle[†]		
θ = 15°	1.4	1.5
30°	1.2	1.3
45	1.0	1.3
60	1.0	1.3
75	1.1	1.3
90	1.1	1.4

* from Eq. 4.5:

$$d \approx (.5r\sqrt{A} + E[d_1 | A]) \ (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx (.5r + E[d_1 | A=1])\sqrt{A}(\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx (.5 + s) \ r\sqrt{A} \ (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

$$\approx \alpha r\sqrt{A} \ (\sqrt{N+0.5} - \sqrt{N-n+0.5})$$

where $\alpha = (.5r + E[d_1 | A = 1])$
Values of α are calculated from the results in Table 4.1.

[†] θ is the angle between radii or equal length sides.

area j.

and $E[s_{ij}]$ is the expected distance from a random point in i to a

random point in j.

In calculating s_{ij}, it is usually simpler to find the grid or L_1 distance,

then convert to straight line or street distances by use of the appropriate travel factor.

The procedure suggested above was applied to six proposed dial-a-ride

districts in Marin County, California. All the districts were irregular,

with hills and coastline preventing a regular street network or boundaries.

One of the districts is illustrated in Figure 4.1,with the regular shapes

used for approximation of the area in the key. The expected straight

line distance between two points in this case is approximately:

$$E[s] = P_{aa} \cdot E[S_{aa}] + 2P_{ab} \cdot E[S_{ab}] + P_{bb} E[S_{bb}]$$

$$= (.64)(1.4) + (.32) \cdot (3.2) + (.04)(.70) = 1.96$$

The expected tour length among six random points for this area is, then:

$$d = (.5\sqrt{A} + E[s|A])r(\sqrt{N+0.5} - \sqrt{0.5})$$

$$= (1.48 + 1.96)1.4(\sqrt{6.5} - \sqrt{0.5})$$

$$= 8.9$$

using Eq. 4.5.

To test the approximation, the street network distance between 21

pairs of randomly and uniformly distributed points was calculated in each

service area (Table 4.5)*. Considering the extreme irregularity of the

---

*I am indebted to R. Shanteau of the University of California, Berkeley,
 for these simulation results.

$A = 8.8$ sq. mi
$r = 1.4$

Approximation to the
Region: 4x6 and 2x3
Rectangles

$P_{aa} = .64$
$P_{ab} = .16$
$P_{bb} = .04$
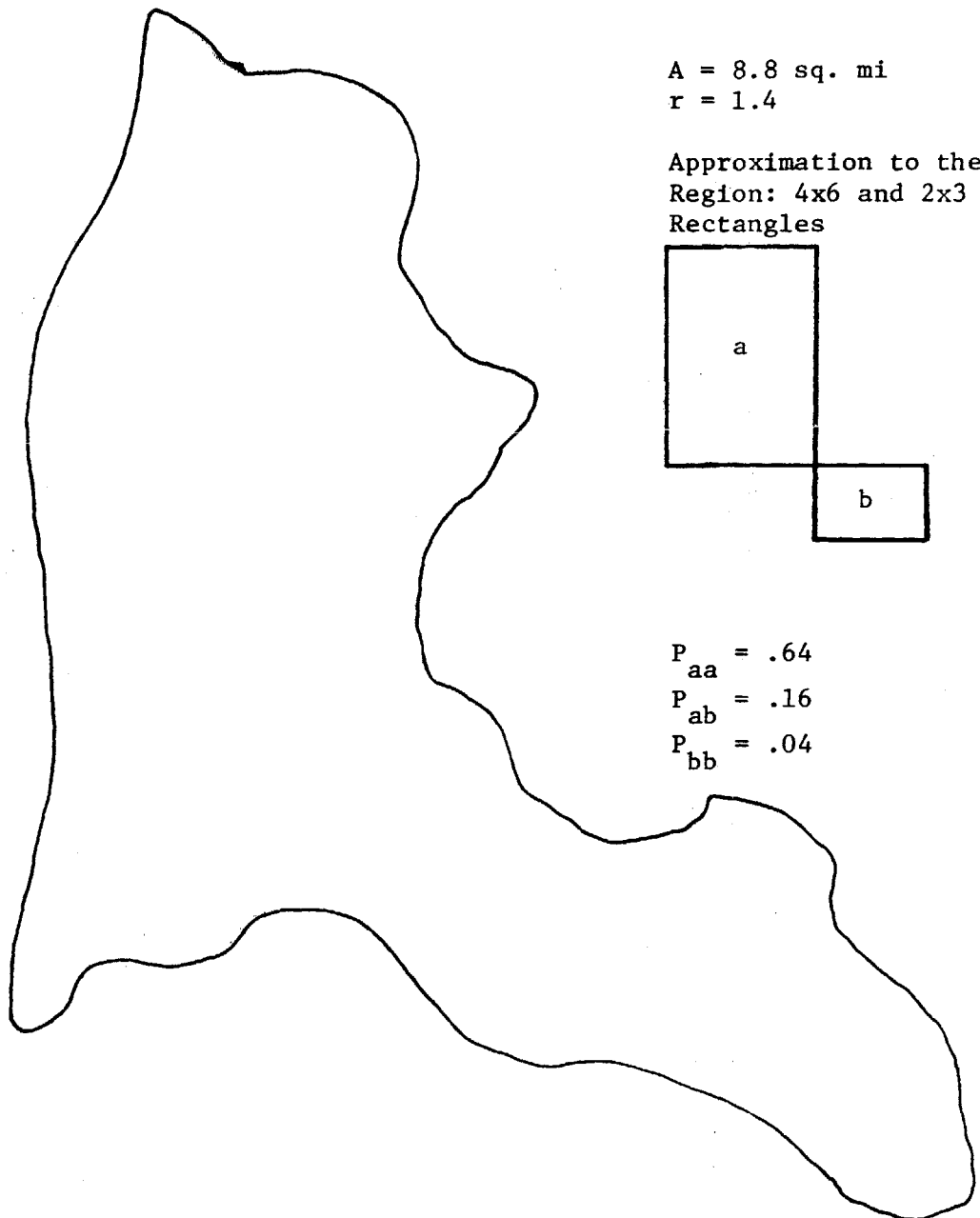
Fig. 4.1: San Rafael Basin Service Region

areas, the agreement between the simulation and approximation results

is quite good, with an average error of only 4%. The estimated tour length

among six stops was also calculated, and the estimated and simulated dis-

tances are shown in Table 4.6; again, the error is relatively small, with

the average error less than 1% and the absolute error 5%. This variance

is well within that expected from the simulation.

4.3 Spatially Non-Uniform Demand

All of the results developed in Chapter 3 rely upon the assumption

of uniform and randomly distributed demands. In this section, we ex-

amine the effect of typical non-uniform demand patterns on the expected

tour lengths. The approach adopted here is experimental, since analytic

solutions appear to be quite difficult to obtain. The encouraging result

of the experiments reported here is that non-uniformity of demand does

not significantly affect tour lengths as long as the distribution of points

is not extreme, such as exceeding, for example, a 7:1 decline in density

over the service area. Consequently, we can use the results of Chapter 3

fairly confidently in practice.

The possible types of non-uniformity of demand are quite many. In

practice, however, the most usual non-uniformity of spatial demand con-

sists of declining density as one moves away from the central city. More-

over, transfer points would generally be positioned in the middle or at

a corner of the high density portions of the service area. Our experi-

ments with non-uniform demand incorporated these assumptions, using a

square service area in which demand density declined in one direction but

was uniform in the vertical direction. The street network is assumed to

TABLE 4.5

DISTANCES BETWEEN PAIRS OF RANDOM POINTS
SIX DISTRICTS IN MARIN COUNTY

District	Estimated Straight Line Distance[1]	Observed Travel Factor	Estimated Street Distance[2]	Simulated Street Distance	Percent Difference
1	2.16	1.75	3.78	3.6	5
2	1.74	1.6	2.78	2.8	-1
3	1.96	1.4	2.74	2.4	14
4	1.60	1.35	2.16	2.2	-4
5	1.57	1.4	2.94	2.9	1
6	3.11	2.0	6.22	5.7	9
All			3.44	3.3	4

[1] based on Table 4.1 or Eq. 4.5

[2] $d = rs$

TABLE 4.6

COMPARISON OF ESTIMATED AND SIMULATED TOUR DISTANCES
IN SIX IRREGULAR AREAS*

District	Simulated	Estimated	% Difference
1	12.3	12.2	-1
2	8.8	9.0	2
3	7.9	8.9	13
4	7.6	7.0	-8
5	11.1	9.5	-5
6	19.9	20.1	1
Average	11.3	11.1	2

* the length of a vehicle tour among six randomly distributed points using next-nearest-point routing. The origin was chosen randomly.

**based upon Eq. 4.5: $d \approx (.5\sqrt{A} + E[s|A])r(\sqrt{N+0.5} - \sqrt{0.5})$
where N=6.

Note: The value of the travel factor, r, was calculated from the same points used in these tours.

be a rectangular grid. In each experiment, between one and ten points were distributed with assumed spatial probability density functions of the form:

$$p(x,y) = \frac{f(x)}{\sqrt{A}} \qquad 0 \leqslant x,y \leqslant \sqrt{A} \qquad (4.8)$$

where $f(x)$ is an assumed density function for demands. The area size, A, was always assumed to be 1 unit.

Figure 4.2 illustrates the four density functions, $f(x)$, which were tested. For density function 2, for example, the demand density is seven times higher at the high density side than at the low density side of the service area. Figure 4.3 shows a three dimensional representation of density function 2.

Results of the experiments are summarized in Tables 4.7 and 4.8. As one might expect, the estimated tour length with uniform demand tends to overpredict the length of tours, since demand density is concentrated near the transfer point. For tours with moderate non-uniformity of demand (that is, with a 7:1 density gradient over the service area or less) and more than three stops, the predicted tour length - based upon the assumption of uniform demand - is nearly always within 10% of the observed tour length.* Again, it should be noted that the expected error inherent in the simulation experiments is on the order of 5%. Starting at the corner of the service area rather than in the middle of an edge tends to reduce the error of prediction. Extreme demand density distributions, such as #3, are not expected to be common in practice; designers encountering such

---

*Density Function 1 has a gradient of 3:1 and function 2 a gradient of 7:1.
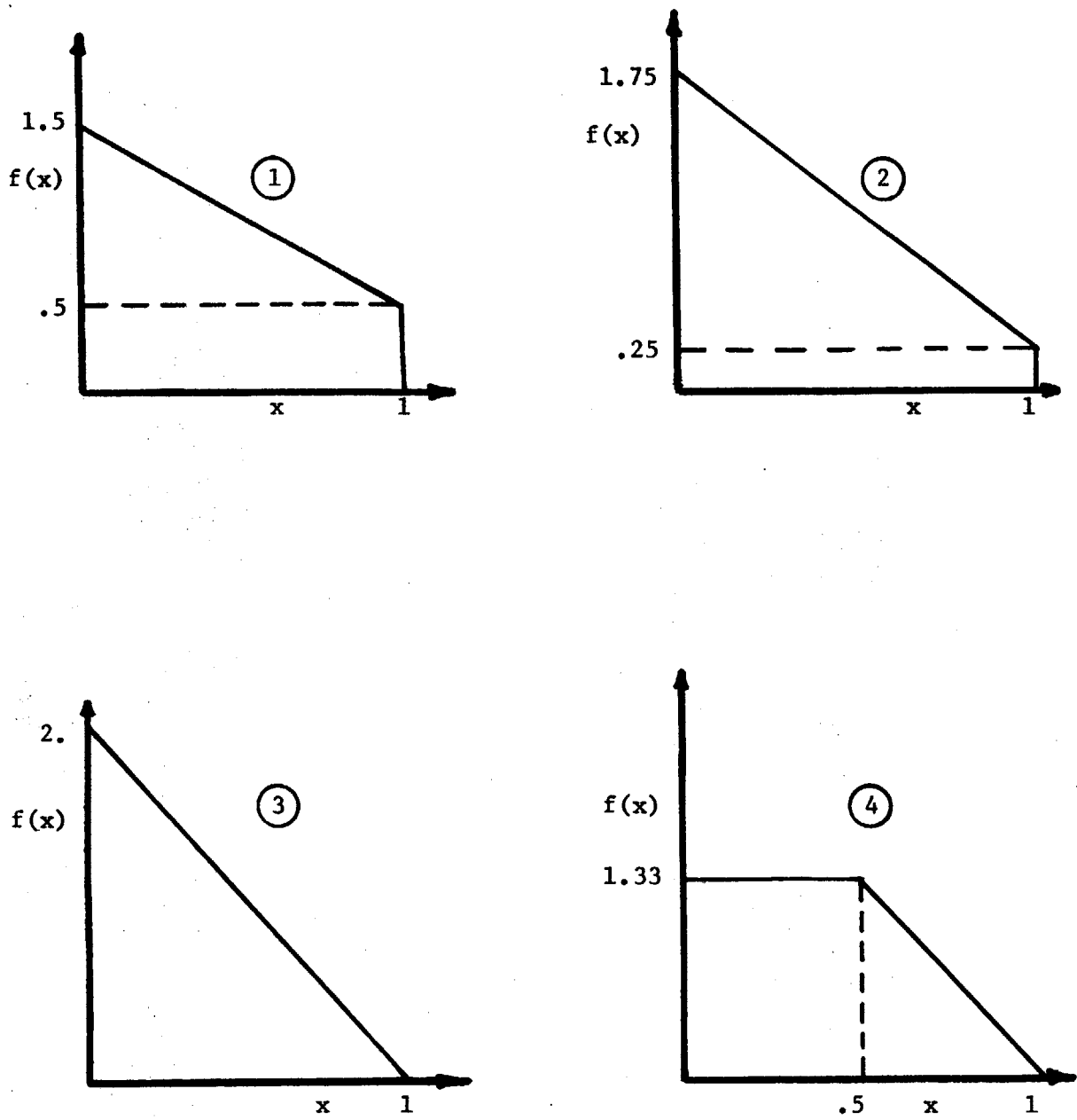
(Density Functions indicate the probability, f(x), that a stop
is located x units from the depot's side of the service region)

Figure 4.2: Density Distributions for Non-Uniform Distributions
Numbers 1-4.

Density
f(x)

x

◄ Possible Transfer
Point Location
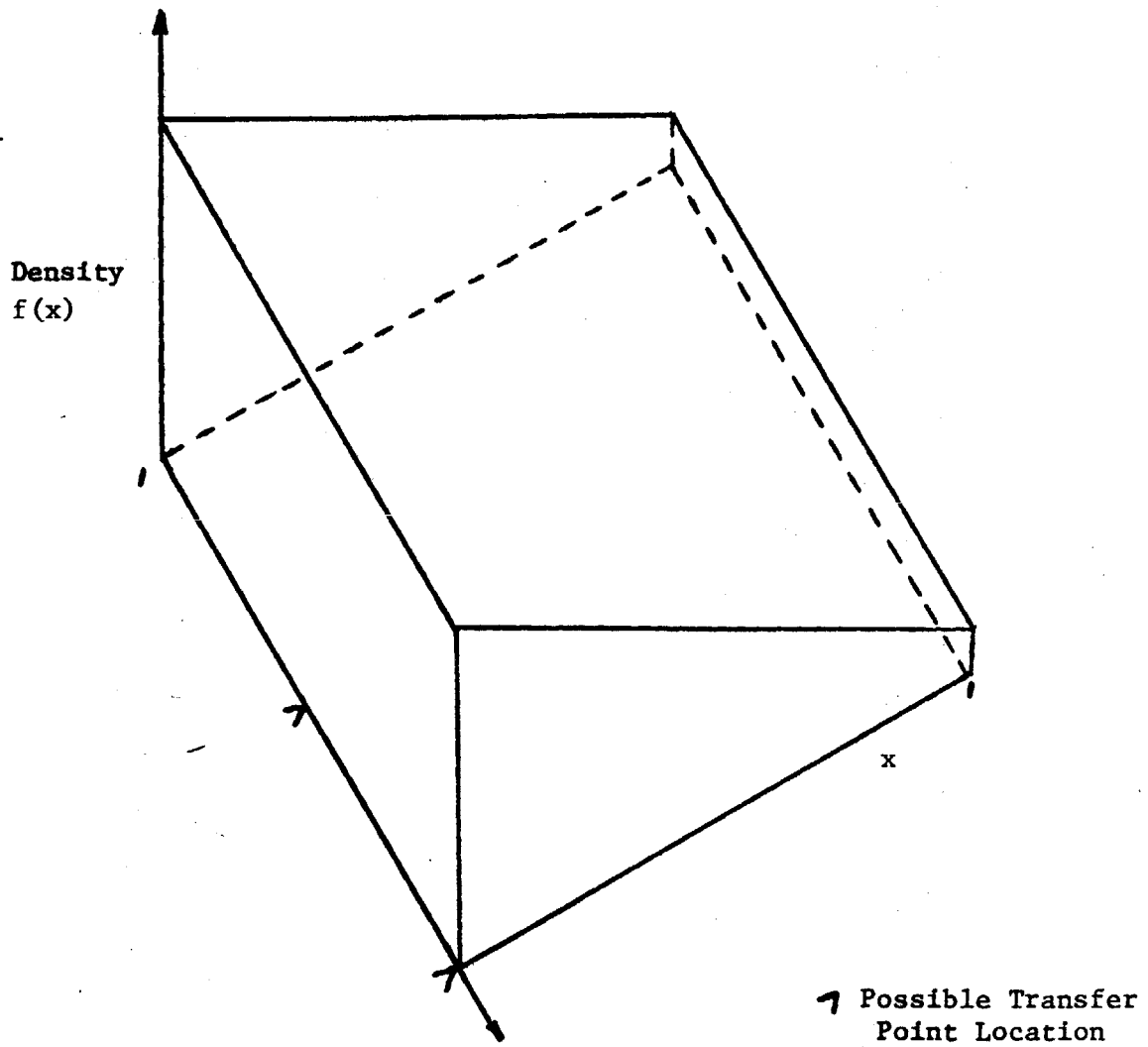
Figure 4.3: Example Service Area and Demand Density
(Distribution # 2)

distributions will probably have to resort to simulation experiments.

Tours with only a few stops deserve closer attention, since exclusive or shared ride taxis often have such tours. It is fairly simple to calculate the distance from the transfer point to a single demand point, given the spatial density function. For example, the expected travel distance from a transfer point located at the midpoint of a side of a rectangular service area, a x b, with a grid street network is:

$$d_1 = \int_0^b xf(x)dx + \int_\theta^{a/2} yf(y)dy + \int_0^{a/2} yf(y)dy \qquad (4.9)$$

One may use this value to modify the factor in the tour length which represents the expected distance from the transfer point to a random point, d_1 (see Section 4.2 for a discussion of this factor).

The predicted and observed tour lengths using this procedure are compared in Table 4.9. The extra work involved in calculating the value of d_1 has been beneficial, since the average percentage error is reduced by 4%. This is hardly surprising, since a similar correction for short tours enhanced the accuracy of prediction in the uniform demand case (Section 3.3).

4.4 Alternative Routing Strategies

Another important operational question for feeder service design is the type of vehicle routing strategy which is used. Different algorithms for routing have different impacts upon system productivity and the level of service. This section surveys some of the theoretical issues involved in algorithm choice and compares a number of algorithms with each other and with manual routing.

TABLE 4.7

TOUR LENGTHS WITH NON-UNIFORM DEMAND
(DEPOT ON THE SIDE OF THE SERVICE AREA)

# Of Stops	Predicted (Uniform demand)*	Uniform Dens. Demand		Density - Functions[†] 1		2		3		4	
1	.78	.74	- 5	.66	-18	.64	-22	.53	-47	.59	-32
2	1.25	1.24	- 1	1.15	-9	1.08	-16	1.05	-19	1.13	-11
3	1.64	1.65	- 1	1.60	-3	1.71	4	1.64	0	1.57	-4
4	1.97	1.96	- 1	2.06	4	1.84	-7	1.78	-11	1.84	-7
5	2.26	2.24	- 1	2.15	-5	2.06	-10	1.68	-35	2.08	-9
6	2.53	2.49	- 2	2.45	-3	2.49	-2	2.28	-11	2.48	-2
7	2.78	2.71	- 3	2.76	-1	2.54	-9	2.18	-28	2.51	-11
8	3.02	2.92	- 3	3.01	0	2.66	-14	2.68	-13	2.88	-5
9	3.24	3.14	- 3	3.35	3	2.83	-14	2.63	-23	2.98	-9
10	3.45	3.35	- 3	3.60	4	3.21	-7	3.00	-15	3.26	-6
Average Absolute % Error 1-10 stops		2		5		10		20		10	
Average Absolute % Error 3-10 stops		2		3		8		17		7	

Start at (0,.5); Tours in Square with center (.5,.5)
Number of Observations: 1000/tour.

*Eq. 4.6

[†] Percentage difference from the predicted tour length is shown after the simulated average.

TABLE 4.8

TOUR LENGTHS WITH NON-UNIFORM DEMAND
(DEPOT AT CORNER)

# of Stops	Predicted Length *	Uniform Dens. Demand		Density - Function † 1		2		3		4	
1	.83	.94	12	.96	14	.91	9	.79	-5	.84	1
2	1.30	1.42	8	1.35	4	1.21	7	1.20	-8	1.28	-2
3	1.68	1.77	5	1.69	1	1.81	7	1.67	-1	1.63	-3
4	2.01	2.06	2	2.02	0	1.92	-5	1.91	-5	1.76	-14
5	2.31	2.31	0	2.19	-5	1.91	-21	2.04	-13	2.12	-9
6	2.57	2.55	-1	2.49	-3	2.40	-7	2.17	-18	2.47	-4
7	2.82	2.76	-2	2.76	-2	2.60	-8	2.12	-33	2.64	-7
8	3.06	2.97	-3	2.99	-2	2.75	-11	2.65	-15	2.86	-7
9	3.28	3.19	-3	3.56	8	2.83	-16	2.62	-25	2.95	-11
10	3.49	3.38	-3	3.48	0	3.41	-2	3.08	-13	3.13	-12
Ave. % Error 1-10 stops		4		4		9		14		7	
Ave. % Error 3-10 stops		2		3		9		15		8	

Start at (0,0); Tours in Square with Center (.5,.5)
Number of Observations: 1,000/tour.

*Eq. 4.6

†Percentage difference from the predicted tour length is shown after the simulated average.

-139-

TABLE 4.9

OBSERVED AND SIMULATED TOUR LENGTHS WITH CORRECTIONS

	Number of Stops	Uncorrected Tour Length Estimate	Demand Distributions 1		2		3		4		Average Absolute Percentage Error
			Est.	Obs.	Est.	Obs.	Est.	Obs.	Est.	Obs.	
	1	.78	.67	.66	.63	.64	.58	.53	.64	.59	5
Start At	2	1.25	1.14	1.15	1.10	1.08	1.05	1.05	1.11	1.13	1
Edge	3	1.64	1.53	1.60	1.49	1.71	1.44	1.64	1.50	1.57	8
	4	1.97	1.86	2.06	1.82	1.84	1.77	1.78	1.83	1.84	3
	1	.83	.92	.96	.88	.91	.83	.79	.89	.84	5
Start At	2	1.30	1.38	1.35	1.29	1.21	1.30	1.20	1.36	1.28	6
Corner	3	1.68	1.75	1.69	1.67	1.81	1.68	1.67	1.74	1.63	5
	4	2.01	2.10	2.02	2.00	1.92	2.01	1.91	2.07	1.76	7

	1	2	3	4
Average Absolute % Error	4	5	5	6
Reduction in the % Error due to the correction	3	5	7	3

In the literature of operations research, attention to the problem of vehicle routing has focused upon the objective of minimizing the total vehicle travel distance. The classic problem in this area is that of the travelling salesman (TSP) in which a tour is to be constructed among a set of points such that each point is visited and the total distance travelled on the tour is minimized. An extensive literature has evolved concerning this problem and the extension to the case of multiple vehicles.* In contrast to the classic problem, tours for transit vehicles must be constructed with both the objectives of maximizing service to patrons and minimizing vehicle travel distance pursued. Minimizing vehicle travel is of interest in order to conserve system resources and to maximize the system capacity. Maximizing the level-of-service to patrons is desireable so as to provide a more desireable service and to attract a higher level of demand.

A simple example shows that the two objectives of minimum travel distance and maximum level-of-service may conflict. In Figure 4.4, one patron is to be delivered at each of the nodes A to E. The vehicle must begin and end at the depot and can travel at a constant speed of 1 unit per minute. The minimum length tour is ABCDE or, equivalently, EDCBA, which both have vehicle travel times of 8, but total patron riding times of 18 and 22 minutes respectively. Alternatively, for the tour CBADE, the vehicle's travel time is 9 minutes, but the total patron riding time is only 15 minutes. Although this tour incurs an additional minute of vehicle travel compared to the optimum travelling salesman tour, it has reduced total

---

*For an introduction to this literature, see Golden [26].
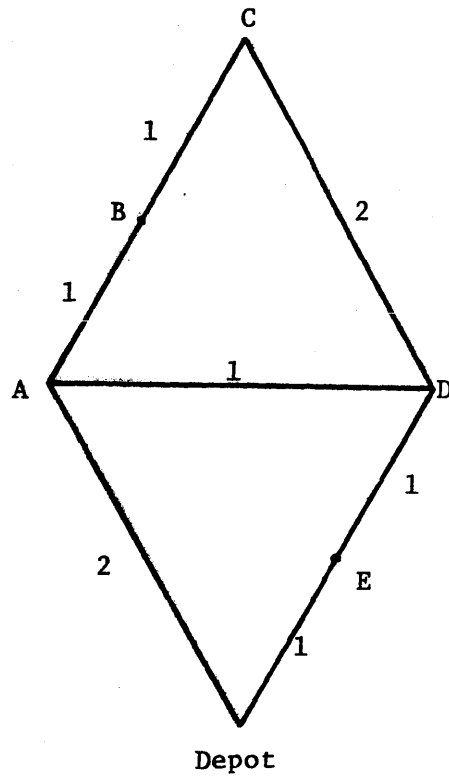
Numbers Indicate
Link Distances

Figure 4.4: A Five Patron Collection Problem

patron travel time by 3 minutes.

The example suggests that there is a trade-off between passenger travel time and vehicle travel time in at least some problems. Since the number of possible tours is finite, a finite number of tours which are not dominated in both tour length and service time exists. These undominated tours define a piece-wise linear envelopeof the best attainable tours (Fig. 4.5). Points marked on the graph in Fig. 4.5 represent particular tours among the given collection point; minimum passenger travel time and vehicle travel time tours are marked as points P and V in Figure 4.4. Since the two objectives of the transit bus routing problem conflict in practice, it is necessary to introduce a tradeoff or constraints on objectives in order to identify a single most desirable tour.*

4.4.1 Tour Length Approximations

Of particular interest for design modeling is the expected tour length arising from various routing algorithms. Beardwood, Halton and Hammersley [4] prove that the expected length of the optimum travelling salesman tour is proportional to the number of stops when the stops are uniformly distributed and the number of stops becomes very large:

---

*The bus tour with minimum patron travel time should converge to the optimum travelling salesman tour length as the number of stops becomes very large. To see this, consider a very small region m of the service area A. The expected number of patrons in the bus is very large relative to the number of stops in m (excluding the first zones visited). Consequently, any deviations from the optimal travelling salesman tour in m must incur penalties (to the patrons on the vehicle) much larger than the possible benefits to those delivered or collected in m. Therefore, the optimum travelling salesman tour should be used in m. Karp [35] proves that an algorithm which divides an area into equal regions and uses an optimum travelling salesman tour in each region must asymptotically converge to the optimum travelling salesman tour length. Thus, the optimum subscription bus tour should asymptotically converge to the travelling salesman tour length. However, the number of stops made on a transit vehicle tour is relatively small, so differences in the minimum length and maximum service tours are expected in practice.
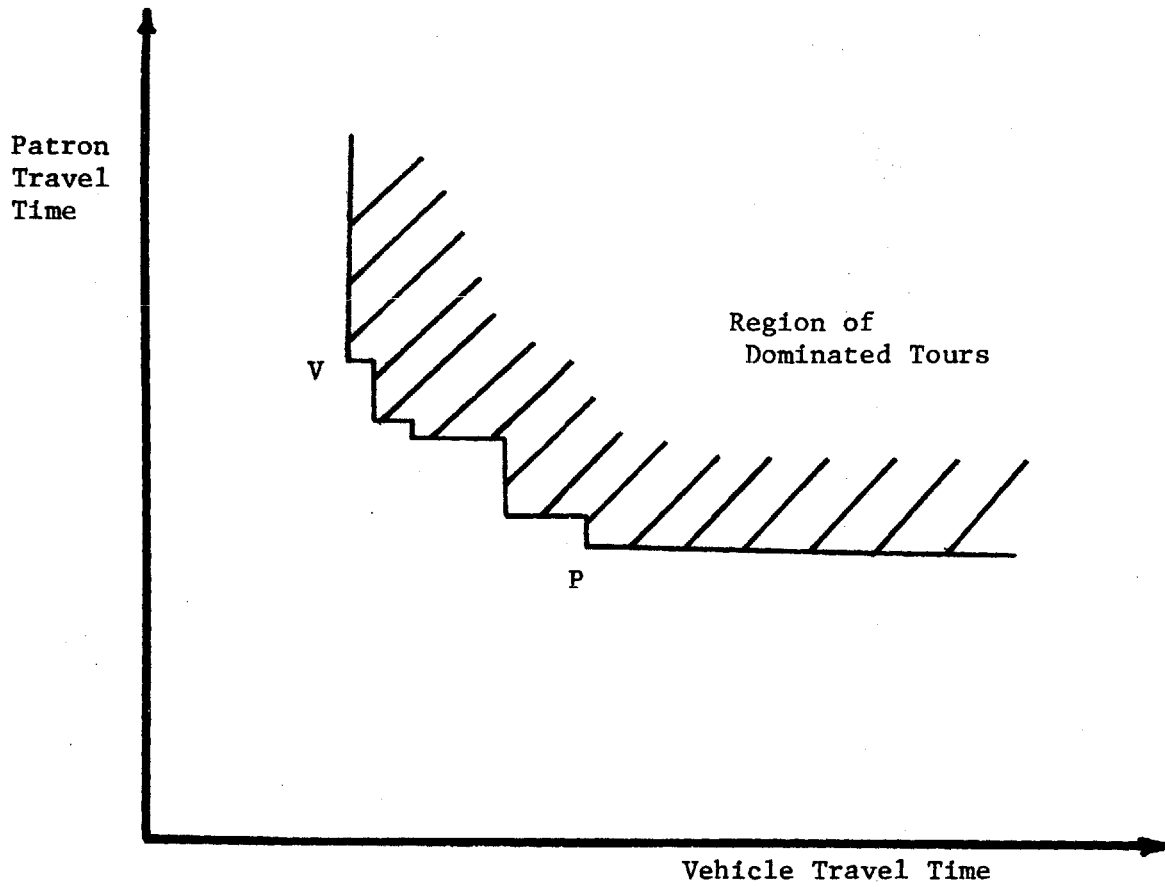
Figure 4.5: Passenger and Vehicle Travel Times of Alternate Tours

$$\lim_{n \to \infty} L_n^* = \alpha \sqrt{nA} \tag{4.10}$$

regardless of the area shape. With Monte Carlo simulation experiments, the constant parameter α was estimated to be .765 by Stein [67].

In Section 3.2 an approximate expression for the expected length of an open tour in a circular area based upon a next-nearest-point routing algorithm was developed:

$$d_n^o \approx \sqrt{A} \{\sqrt{n+0.5} - \sqrt{0.5}\} \tag{3.5}$$

Adding in the link between the starting point and the last stop results in an estimate of the closed tour length:

$$d_n = \sqrt{A} \{\sqrt{n+0.5} - \sqrt{0.5}\} + .5\sqrt{A}$$

$$= \sqrt{A} \{\sqrt{n+0.5} - .21\} \tag{4.11}$$

As n becomes very large, the next-nearest point tour length expression (Eq. 4.11) is 35% greater than the asymptotic travelling salesman expression (Eq. 4.10). For one stop, the difference is 25%. There are two major reasons for the difference between the two tour length expressions. First, the next-nearest-point algorithm will not generally result in the minimum length or optimum travelling salesman tour. Rosenkrantz et.al. [63] show that the worst case or maximum length of a tour based upon such an algorithm is:

$$d \leqslant d^*(1/2 \|\log_2(n)\| + 1/2)$$

where $\| . \|$ is the smallest integer greater than (.) and d^* is the minimum length tour. For a ten stop tour, the maximum length of the next-nearest-

point tour is 2.5 times the minimum length tour.

Also, the travelling salesman expression (Eq. 4.10) does not reflect the effects of randomness in patron location, since it is derived for the case in which demand density becomes quite large. We hope that the latter effect is more significant, so that use of the next-nearest-point algorithm would not result in substantial difference from other routing algorithms. Indeed, Stein found that Beardwood's expression underpredicts the observed minimum tour length by 24% with 10 stops in a square area, 13% with 30 stops, and 6% with 60 stops [67]. Eilon et.al. [18] present essential similar results. It appears that the asymptotic travelling salesman tour substantially and systematically underpredicts the expected tour length for small values of n.

4.4.2 Comparisons of Tour Characteristics

We shall compare four routing strategies with the asymptotic travelling salesman tour (TST) and the next-nearest-point (NNP) tour expressions discussed above. The first strategy employs an heuristic algorithm developed by Christofides which is intended to identify short tours [10]. The algorithm results in tours which are at most 50% longer than minimum length tours and generally finds tours within 5% of the minimum length. The second strategy uses next-nearest-point routing with vehicles always proceeding to the closest eligible stop. To be consistent with the Christofides TST algorithm, we shall impose a trip to the depot as the final leg in all cases, thereby finding closed tours among the patron's stops.

The other two routing strategies involve manually constructed

tours*

For one strategy, the dispatcher was asked to construct minimum length (TST) tours among the patrons. For the other strategy, the dispatcher was asked to route the vehicle as if he were delivering patrons from a transit vehicle. Tours were constructed separately by two dispatchers with the aid of maps of stop locations. Among thirty trials, two dispatchers constructed tours with only one small difference.

Ten tours were constructed for each strategy in each of three service areas: a square with the depot at the side, a square with the depot at the center, and a 2 x 1 rectangle with the depot on the side. A dense, rectangular street grid was assumed in each area.

Considering the resulting tour lengths (Table 4.10), next-nearest-point routing results in slightly longer tours than any of the other strategies. For this simple problem, manual routing to find minimum length tours is slightly better than the travelling salesman (TST) algorithm. Minimum length manual routing results in tours which are 4% shorter than with transit bus manual routing. However, the tour lengths from any of the strategies have little variation; there is only a 6% difference between the strategy producing the longest and shortest tours (NNP and minimum length manual routing respectively).

The next nearest point expression gives fairly good predictions of the tour lengths for all the strategies, with the exception of the case in which the depot is at the center of the service area. In the latter case, the maximum error was -12% (with respect to NNP routing) while for the other two situations the maximum error was 7% (for minimum length manual

---

*My thanks to Profs. C.F. Daganzo and N.M.H. Wilson for serving as expert dispatchers.

TABLE 4.10

COMPARISONS OF TOUR LENGTHS[1]

	Square: Depot At Side	Square: Depot At Center	2 x 1 Rectangle: Depot At Side
NNP Expression[2]	4.00	3.68	6.06
Asymtotic TST Expression[3]	3.09 (29)	3.09 (19)	4.37 (28)
TST Algorithms[4,5]	4.02 (-1)	3.96 (-8)	5.77 (5)
Next-Nearest-Point Routing[5]	4.10 (-3)	4.13 (-12)	6.25 (-3)
Manual: Minimum Length	3.94 (2)	3.93 (-7)	5.66 (7)
Manual: Transit	4.12 (3)	4.10 (-10)	5.97 (-1)

[1]Number in parenthesis are the percentage difference from next nearest point predicted tour length (NNP Expressions). Areas are 1 for the square and 2 for the 2x1 rectangle.

[2]$d = 1.31(\sqrt{A})(\sqrt{n+0.5} - \sqrt{0.5}) + E[d_1] + E[d_9]$ by Eq. 4.5

[3]$d* = .765r\sqrt{An} = 3.09\sqrt{A}$

[4]The TST algorithm is intended to find minimum or close to minimum length tours and is described in [10].

[5]Simulation uses uniformly distributed points, 10 observations per case, 9 patrons per tour.

routing). Since the expected error in the simulations is on the order

of 5% and the average error of prediction is 2%, the accuracy of the

next-nearest-point expression is relatively good. As expected, the asym-

totic TST expression underpredicted tour lengths.

Considering the average patron travel (Table 4.11), transit bus

manual routing results in the best level-of-service of the four strategies.

The trade-off between patron travel time and minimum length tours is il-

lustrated by the differences between the two manual routing strategies:

the minimum length tours result in a 4% reduction in travel distance and

an 11% increase in average patron travel compared to transit bus routing.

Next-nearest-point routing results in travel times statistically indist-

inguishable from the transit bus manual routing times. The travelling

salesman algorithm and minimum length manual routing result in travel

times which are similar and approximately 10% higher than for the other

two strategies.

With the linear and continuum assumptions concerning the delivery

process (Section 3.5), it is possible to estimate the expected patron

travel time as one half the estimated tour length plus a small amount of

travel from the depot into the service area (Section 3.5). Using the

NNP expression of tour lengths, this estimate gives fairly accurate pre-

dictions of travel time for the next-nearest-point and transit bus rout-

ing strategies, but tends to underpredict for the minimum tour length

strategies.

The results of the simulation experiments are fairly encouraging con-

cering the predictive ability of the next-nearest-point (NNP) tour length

expression, even in cases in which NNP routing is not used. In situations

-149-

TABLE 4.11

COMPARISONS OF AVERAGE PATRON TRAVEL TIME[1]

	Square: Depot at Side	Square: Depot At Center	2 x 1 Rectangle: Depot at Side
NNP Expression[2,3]	1.66	1.59	2.56
TST Algorithm	1.86 (-12)	1.88 (-18)	2.59 (-1)
Next-Nearest-Point Routing	1.74 (-7)	1.68 (-6)	2.43 (5)
Manual: Minimum Length	1.83 (-4)	1.93 (-21)	2.66 (-4)
Manual: Transit-Bus	1.70 (-2)	1.68 (-6)	2.44 (5)

[1] Simulation data as in Table 4.10.

[2] $E[T^r] = 1.31\sqrt{A}(\sqrt{n+0.5} - \sqrt{0.5})/2 + E[d_9] = 1.59\sqrt{A} + E[d_9]$ by Eq. 3.47.

[3] Numbers in Parenthesis are the Percentage Differences from the NNP Expression

in which the depot is not inside the service area, strategies designed to minimize tour lengths or maximize patron service have similar results, both with respect to each other and the NNP expression. Fortunately, the depot should not be at the center of a well designed feeder service area (see Section 3.5), so this is not a serious drawback to the tour length approximation expression.

From next-nearest-point or manual delivery tour routing, it is possible to reduce vehicle travel on the order of 5% by concentrating upon minimizing tour lengths, but at the expense of increasing patron travel by approximately 10%. Of course the gain in vehicle travel times may enable more frequent service, so one should not automatically reject the minimum tour length strategies. Clearly, however, the trade-off between vehicle travel and user's level-of-service should be considered in developing vehicle routes.

Finally, the approximate expression to predict average travel time (Eq. 3. 8) gives fairly accurate predictions for strategies which consider patron's level of service. However, it tends to underpredict the average travel time with minimum tour length routing strategies, particularly with the depot on the interior of the service area.

CHAPTER 5

VALIDATION AND APPLICATION OF THE

FLEXIBLY ROUTED FEEDER SERVICE MODEL

The previous two chapters have developed models of flexibly-routed feeder service in some detail; equations for these models were summarized in Section 3.9. In this chapter, the results from these models are compared with the results of a simulation model and with observations of the Ann Arbor transit system. Some applications of the models are also explored. These applications include sensitivity analysis of the model results as a single parameter is altered, comparison of alternative operating strategies in a particular case, and an example of the use of the model in a design study which considers the effects of demand sensitivity. In addition, the effects of the corrections for stochastic phenomena which were introduced in Section 3.4 are also discussed.

Section 5.2 begins the discussion with comparisons between the approximate, analytic models' results and both observations of the Ann Arbor, Michigan transit system and the results of a Monte Carlo simulation model. In addition to providing an indication of the accuracy of the models' predictions, the section describes application of the model to the Ann Arbor system.

In the following two sections, various characteristics of the models are explored. Section 5.3 presents the results of varying a single input or option parameter. For example, this section presents volume/delay performance functions of particular feeder services. Section 5.4 discusses

the effect of the fine tuning model correction terms which were introduced in Section 3.4.

Finally, two hypothetical applications of the models are discussed. In Section 5.5, alternative operating policies are compared in a particular case. Section 5.6 contains an example of a design study in an equilibrium framework which includes a demand function.

Throughout these applications, the expected travel time, $E[TT]$ is used for illustration. This level of service indicator may be thought of as the linear combination of waiting and riding time of a random patron:

$$E[TT] = E[T^r] + \beta E[T^w]$$

where $\beta=1$.

Waiting time has often been found to be more burdensome in empirical studies of demand, and is often weighted higher than riding time (with $\beta=2.5$ or 3) to form a composite service level indicator. However, available empirical evidence indicates that waiting time at home is no more burdensome to patrons than riding time. Hence, a simple sum ($\beta=1$) was used in the linear combination of travel time.

Specific conclusions arising from the discussion in the chapter include:

- the approximate, analytic performance model of feeder service can reproduce the results of simulation experiments quite well.

- the flexibly-routed, phased feeder service model gives relatively accurate prediction of expected travel times even without elaborate local calibration.

-154-

- a cycle length exists which minimizes the patrons' expected travel
  time.

- the stochastic corrections to the deterministic model may be quite
  significant for actual systems, so they should be included in appli-
  cations.

- phased and offset service is more advantageous at lower demand levels
  than is unphased or synchronous service.

- searches for least cost system designs without considering market
  equilibration may result in inappropriate designs.

5.2 Validation of the Model

Results from the approximate, analytic model of flexibly-routed
scheduled and phased feeder service developed in Chapters 3 and 4 have
been compared with the results of a simulation model which was implemented
on a digital computer. Next-nearest point dispatching and a regular, cir-
cular service area were assumed in both models. Vehicles could only travel
in parallel to a set of rectangular axes in the simulation, representing
a closely spaced grid street network. Poisson arrival processes for cus-
tomers and uniformly distributed origins and destinations were realized
with a pseudo random number generator, with arrival rates assumed to be
equal for inbound and outbound patrons. After patron demand locations
were generated, vehicle routing was accomplished by scanning the list of
available stops and proceeding to the nearest.

Figures 5.1 to 5.3 compare the travel times predicted by the feeder
service model and observed in the simulation model for various values of

Fig. 5.1 Travel Time Versus Cycle Length with Service Area=7.07 sq. mi. in the Phased, Flexibly Routed Feeder Service Model

-156-
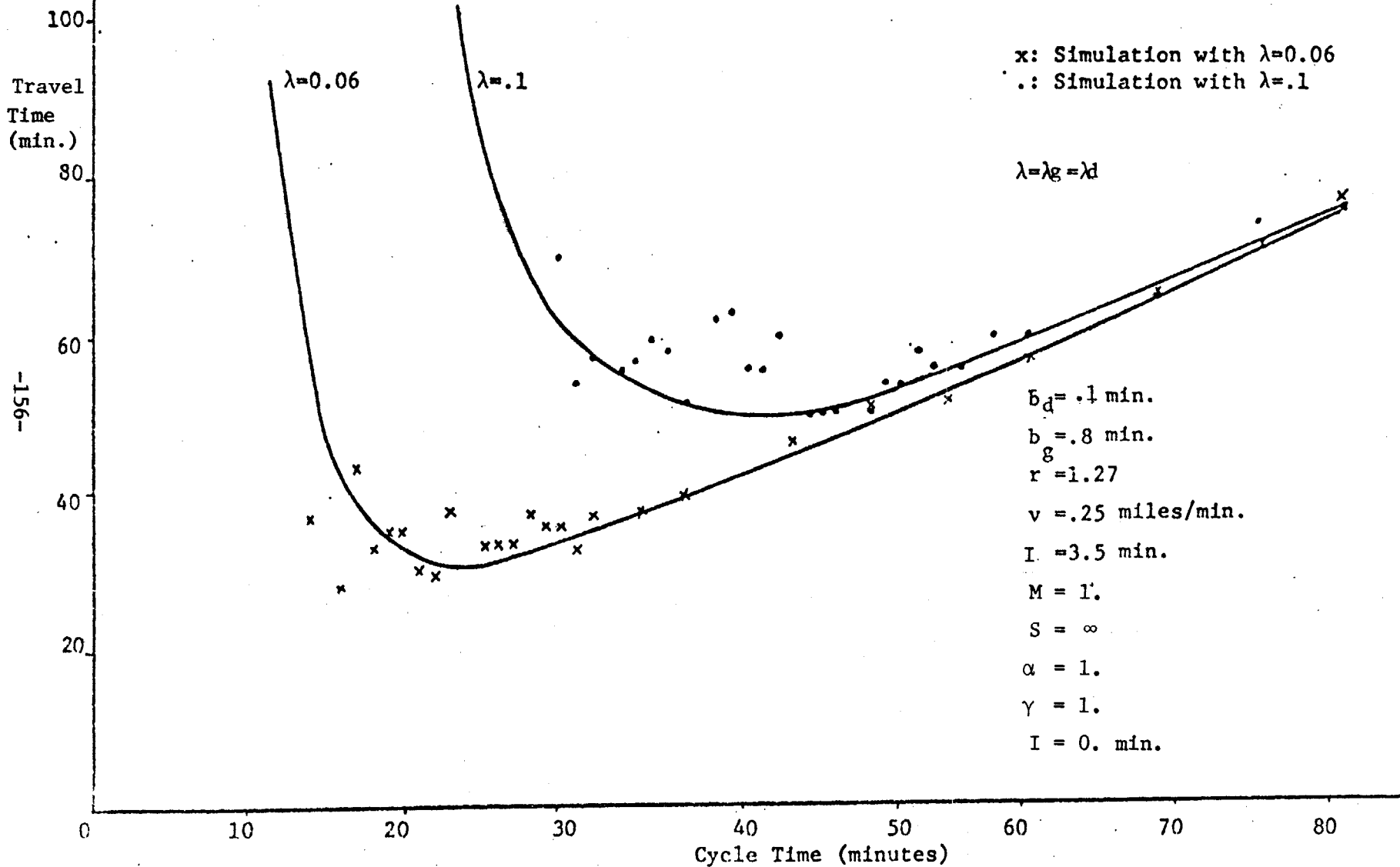
Fig. 5.2 Cycle Time Versus Travel   Time with Service Area=3.14sq. mi. in the Phased, Flexibly Routed Feeder Service Model
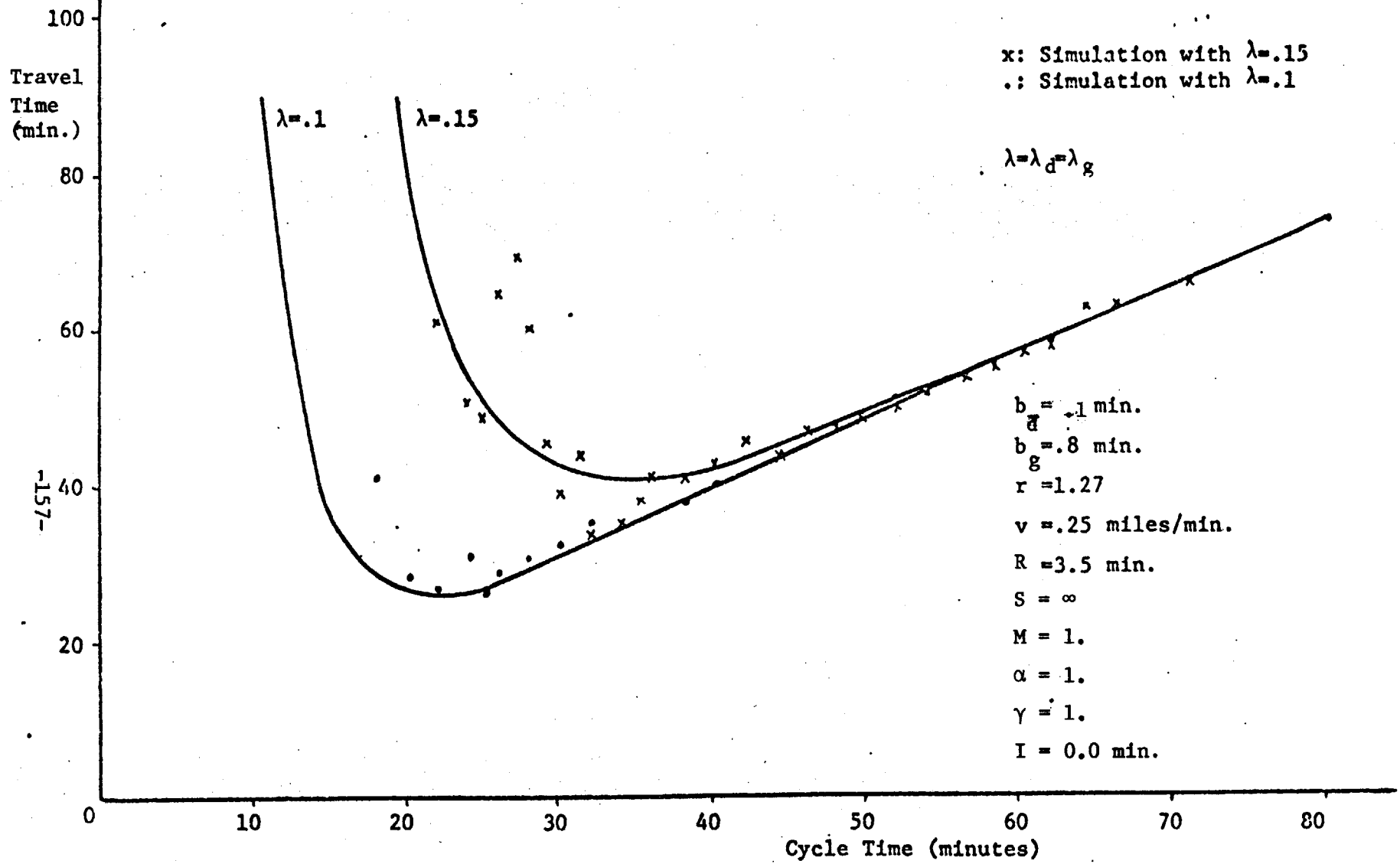
Figure 5.3: Expected Travel Time on a Phased, Flexibly Routed Feeder Service
with Various Cycle Lengths

. simulation results, S=4.

λd= .06 demands/min.
λg= .06 demands/min.
α = 1.
γ = 1.
b_g= .8 min.
b_d= .1 min.
r = 1.27
v = .25 mi./min.
R = 3.5 min.
I = 0. min.
M = 1.
A = 7.07 sq. mi.
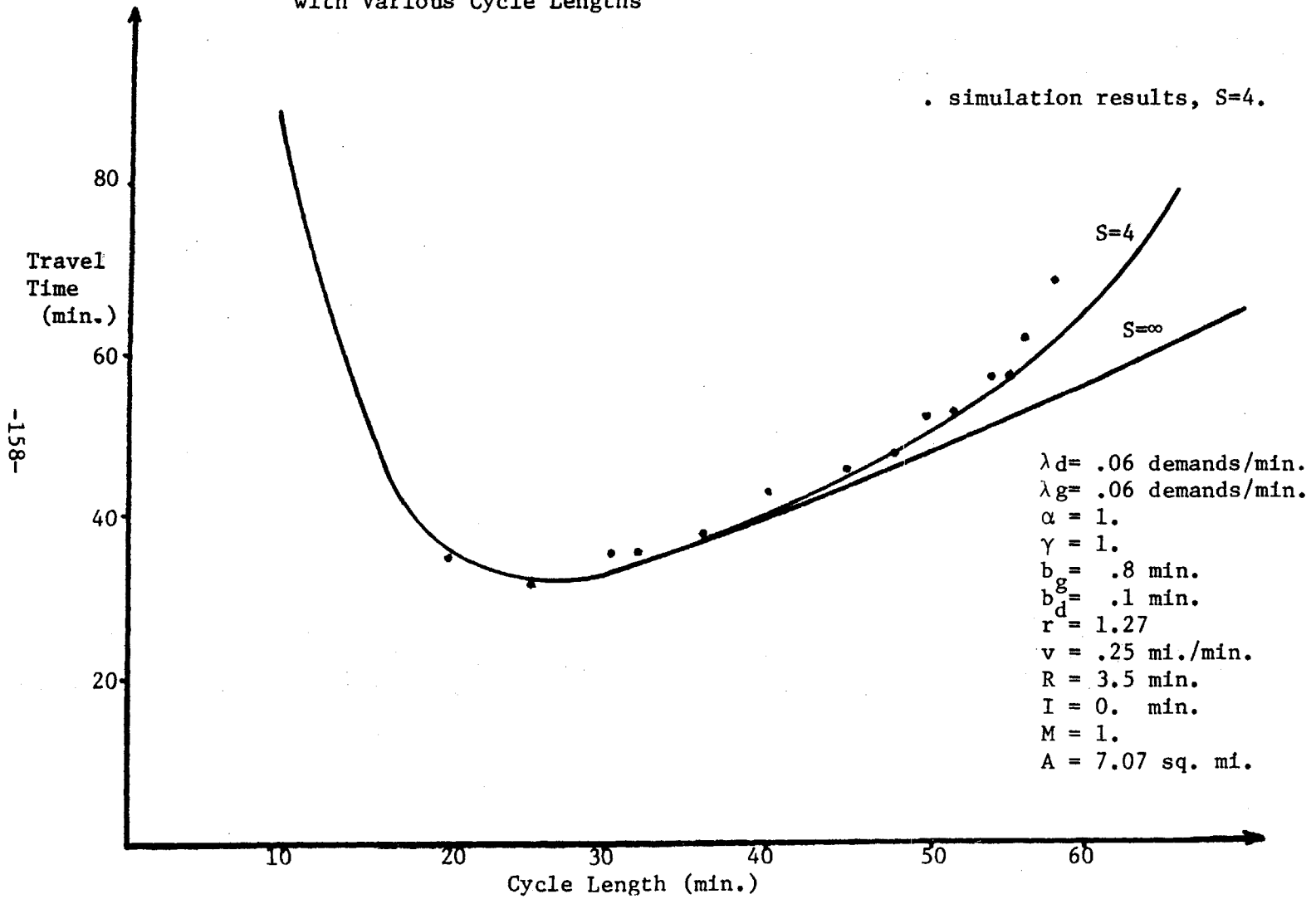
the demand level, cycle length, and area size.*   Travel time consists

of the expected waiting and riding time in the system for a random patron.

Each simulation result represents 100 or more vehicle cycles. Agreement

between the predicted and observed values is fairly close in all the

cases simulated, although slightly optimistic when vehicles are near capacity.

Predictions were also compared with actual service characteristics

of the Teltran feeder system in Ann Arbor, Michigan. This integrated

system consists of a number of feeder service zones and line haul transit

routes. The system is described by Guenther[28] · and Neumann [60]. Veh-

icles are routed manually, but with computer assistance in bookkeeping.

In applying the approximate, analytic models, all the zones were

assumed to be uniform square areas. Constant rendezvous time was also

assumed. Approximately 10% of patrons requested many-to-many service

within the individual zones; these patrons are assumed to be collected

and, if time is available, distributed during the collection phase. Thus,

the Ann Arbor feeder services are similar to the zonal service described

in Section 3.8.

One of the feeder service zones in which the model was applied is

called Pontiac Heights. In this and all other zones, feeder service was

flexibly routed, phased and scheduled so as to meet line haul vehicles.

Available data for the application is summarized in Table 5.1. A few as-

sumptions are required to enable application of the analytic model.

First, we assume that the percentage distribution of patrons between peak

and off peak periods is the same in Pontiac Heights as in the rest of the

---

*The feeder service model formulae are given in Section 3.9.

Table 5.1

Input Data For The Pontiac Heights Application

Volume = 230/weekday

Area = 1.50 sq. mi.

Effective Vehicle Speed = 14.6 miles per hour

Vehicles = 2 during the peak periods

        1 during the off peak period

Patronage Distribution* (system)	Collect	Deliver
6-9 a.m.	.27	.11
9a.m.-4p.m.	.11	.11
4-7 p.m.	.11	.27

Cycle Length: 30 minutes

Percentage of Many-to-Many Trips: 6%

Average Group Size: $Y = 1$

Vehicle Capacity: $S = 15$

*Collection statistics include many-to-many patrons.

system. We also assume steady-state operation during these periods. No

data on rendezvous time at the depot is available; we assume a time of

10 minutes for this value (which is used in all five districts).[+] Also,

we assume a regular square service area. Finally, only the effective

vehicle speed, including stops for boarding, is available. Thus, we

may set $b_g = b_d = 0$ and interpret v as the effective vehicle speed.[*]

Using the information in Table 5.1, the demand rates may be calcula-

ted as follows, in units of demands per minute:

	λ_g	λ_d	λ'_g
Morning Peak	.35	.15	.37
Midday	.15	.15	.16
Afternoon Peak	.15	.35	.16

where λ'_g includes the many-to-many delivery stops. We assume that many-

to-many trips are served during the collection phase because the potential

time available for collection is longer than the delivery phase.

Using the equations of Section 3.9.3, we find the number of stops

for delivery and collection during the morning peak:

$$n_g = \lambda'_g.C/M\gamma = 5.57$$

$$n_d = \lambda_d.C/M\gamma = 2.25$$

With these demands, vehicle capacity should virtually never be exceeded

and may be ignored. The time required for delivery is:

---

*This assumption does not accurately capture the variability of large
 tour times due to higher demand density, but should not be a large error.
[+]In the uncongested system studied here, a change in R of 1 min. changes
 expected ride time by .25 min. and has minimal effect on the wait time.

$$D = \frac{\alpha r \sqrt{A}}{v} (\Delta \sqrt{n_d + 0.5} - \sqrt{0.5})$$

$$= 6.6 (\sqrt{2.25 + 0.5} - \sqrt{0.5}) = 6.3 \text{ min.}$$

where we have assumed for simplicity that the correction factor, $\Delta = 1$. The time available for collection is:

$$G \equiv C - D - R$$

$$= 30 - 6.3 - 10 = 13.7 \text{ minutes}$$

Since a tour among the $n_g = 5.57$ stops requires:

$$G \leqslant \frac{\alpha r \sqrt{A}}{v} (\sqrt{n_g + 0.5} - \sqrt{0.5})$$

$$\leqslant 12. \text{ minutes}$$

which nearly exceedes the collection period, there must be a certain number of eligible patrons who cannot, on the average, be collected by a vehicle. To find the steady-state pool size, $x^{*'}$, we calculate:

$$k = \frac{v}{\alpha r \sqrt{A}} (G - n_g b_g) = 2.06$$

$$x^* = [\max\{0, (0.5 + n_g - k^2)/2k\}]^2 + n_g = 5.77$$

$$E[Y'] = x^* - (x^* - n_g) L(\frac{x^* - n_g}{\sqrt{n_g}}) - \sqrt{n_g} \phi(\frac{x^* - n_g}{\sqrt{n_g}})$$

$$= 5.77 - (.20)(1/(1 + \exp(-2.5 \times .08))) - 2.36 \phi(.08)$$

$$= 5.77 - (.11) - (.94) = 4.72$$

$$x*' = x* + \max\{0, n_g - E[Y']\}$$

$$= 5.77 + (5.57 - 4.72) = 6.6$$

Now we may calculate the expected riding times and the waiting time at home:

$$E[T_d^r] = \frac{D+R}{2} = 8.15$$

$$E[T_g^r] = \frac{G+R}{2} = 11.85$$

$$E[T_m^r] = \frac{G}{3} = 4.57$$

$$E[T_g^w] = \frac{x*'}{\lambda_g'} - \frac{C}{2M} + \frac{G'}{2}$$

$$= \frac{6.6}{.37} - \frac{30}{4} + \frac{13.7}{2} = 17.$$

The average riding time is, then:

$$E[T^r] = \frac{\lambda_d}{\lambda} [T_d^r] + \frac{\lambda_g}{\lambda} E[T_g^r] + \frac{\lambda_m}{\lambda} E[T_m^r]$$

$$= .3(8.15) + .66(11.85) + .04(4.57)$$

$$= 10.5$$

where $\lambda = \lambda_d + \lambda_g + \lambda_m = .5$

Similar calculations reveal that average riding time is 15.0 and 10.2

for the midday and evening peak periods in Pontiac Heights, so the average riding time during the day is:

$$E[T^r]= (.38)(10.5) + (.22)(15.0) + (.38)(10.2) = 11.2 \text{ minutes}$$

where the weights are the proportion of trips in each period. System wait times were calculated by averaging the wait time in the five zones studied, weighted by the demand in each zone.

Table 5.2 reports the observed and predicted service characteristics for the Ann Arbor system. In only one case could the hypothesis that the observed service times equaled the predicted times be rejected at the 5% level of confidence: for wait time during the midday period. During this period, the system was quite uncongested. As a result, many-to-many trips within the subzone could be served during the idle phase between the distribution and collection phases of the service cycle. Consequently, the observed wait time is lower than expected. With the assumption of many-to-many service during the idle period, the predicted waiting time is 22 minutes.

5.3 Peformance Functions

This section illustrates the variation in the expected service time and other level of service components as single input parameters to the feeder service model are varied. Of particular importance for design studies is the unmistakable sensitivity of the quality of service provided to the level of demand served by the system. Consequently, design studies must contend with the issues involving equilibrium demand and performance levels, a subject which was discussed in Section 2.2.

TABLE 5.2

COMPARISON OF MODEL PREDICTIONS WITH OBSERVED

ANN ARBOR SYSTEM CHARACTERISTICS

R I D E   T I M E

Zone	Predicted	Observed
Pontiac	11.2	11.6
Far SW	9.7	9.9
Far SE	10.9	11.1
Plymouth	9.9	9.7
Medford	6.9	6.5
Average	10.5	10.6

W A I T   T I M E

	Predicted	Observed
AM PEAK	22.0	25.4
MIDDAY	28.6	21.6*
AVERAGE	23.1	25.8
Average	23.9	24.7

_____

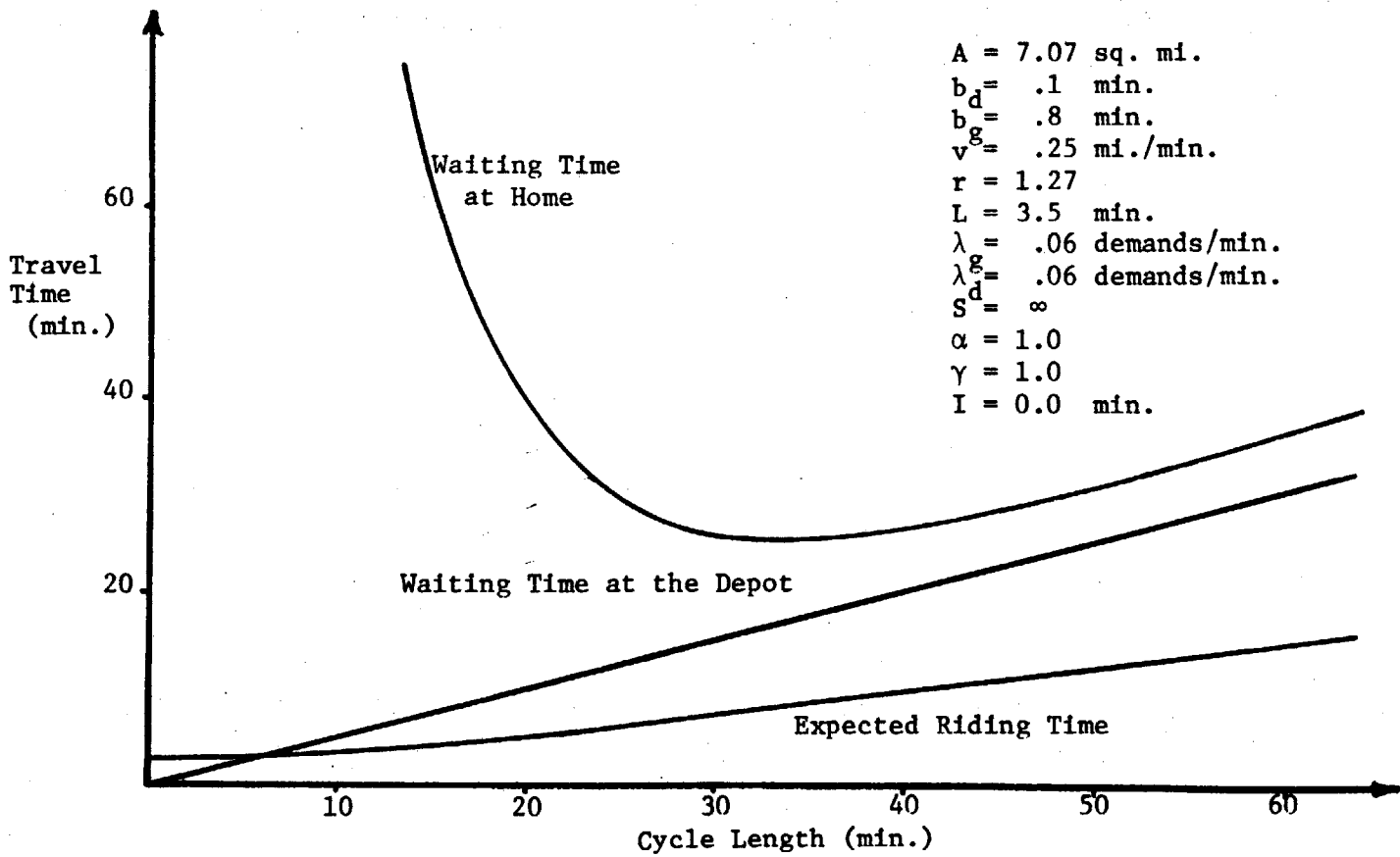*Hypothesis $\mu = \mu_0$ can be rejected at the 5% confidence level.

The effect of varying the scheduled depot visits or cycle length has already been illustrated in Figures 5.1 and 5.2. These figures indicate that there exists a cycle length which minimizes expected travel time. With short cycle lengths, the expected number of patrons which cannot be collected, $x^* - n_g$, tends to increase dramatically, thereby increasing the wait time at home and the expected travel time. With long cycle lengths, all the components of the expected travel time increase proportionally with the cycle length, with the expected wait time at home asymptotically approaching one half of C plus G. Figure 5.4 illustrates these various components of the travel time as the cycle length is varied. If the various components of the travel time are weighted to form a linear objective function, a cycle length will still exist which minimizes the objective function.

Since changing the cycle length does not require more vehicles or incur higher costs, the choice of a good cycle length may result in better service at no extra cost. For example, it may be desireable to operate with the cycle length which minimizes the patron travel time or some other benefit measure.* Due to uncertainty in demand predictions, other input parameters, and the model predictions, the optimum cycle length of an actual system cannot be found exactly. It can be seen that the increase in travel time is much faster at cycle lengths below the optimum than above the optimum (Figs. 5.1 and 5.2). Since the penalty of using a low cycle length is higher than that of using a cycle length

---

*As noted in Section 5.1, the empirically found weights on time spend riding and waiting at home were equal in one study, so travel time serves as a measure of net benefit. Other benefit measures are discussed in Williams [71].

Figure 5.4: Travel Time Components in a Flexibly-Routed, Scheduled and Phased Feeder Service

A = 7.07 sq. mi.
b_d = .1 min.
b_g = .8 min.
v = .25 mi./min.
r = 1.27
L = 3.5 min.
λ_g = .06 demands/min.
λ_d = .06 demands/min.
S^d = ∞
α = 1.0
γ = 1.0
I = 0.0 min.

Waiting Time at Home

Waiting Time at the Depot

Expected Riding Time

Travel Time (min.)

60

40

20

10   20   30   40   50   60

Cycle Length (min.)

-167-

above the optimum, it is desireable to operate with slightly longer

cycle lengths than the cycle length which is predicted to exactly min-

imize the travel time.

The response of the expected travel time to variations in the cycle

length has a very strong resemblance to the delay curves for pre-timed

traffic signals as a function of cycle length (Webster [70]). In fact,

the factors influencing the shape of the curves are also similar. For

very long cycle lengths, the system has excess capacity and the bus is al-

ways able to collect everyone in the pool of waiting calls. Consequent-

ly, the waiting time at home can never exceed one cycle plus one collec-

tion period, and it asymptotically approaches half a cycle. On the other

hand, for very small cycle lengths, collection times are shorter,

and larger pools of waiting requests will form at home. In this

case, customers will often be forced to wait more than one period.

The system under consideration here is somewhat more complex than

the traffic light problem because the departure rate from the queue is

not deterministic and constant. Consequently, exact analytical solutions

such as Darroch's solution for the pretimed traffic light delay problem

(Darroch [15]) appear to be out of the question. The approximate ap-

proach to the problem taken is similar to the fluid approximation model

of Clayton [12]. In this case however, the approximations are better be-

cause the feeder service has a state dependent service process. Thus,

with a variation in some system input (such as the number of patrons ar-

riving), there is a corresponding response in the service process efficien-

cy (i.e the efficiency of vehicle tours) which serves to "dampen" the

original variation. As a result, stochastic fluctuations are "damped out"

from cycle to cycle and the deterministic elements of service become more important.*
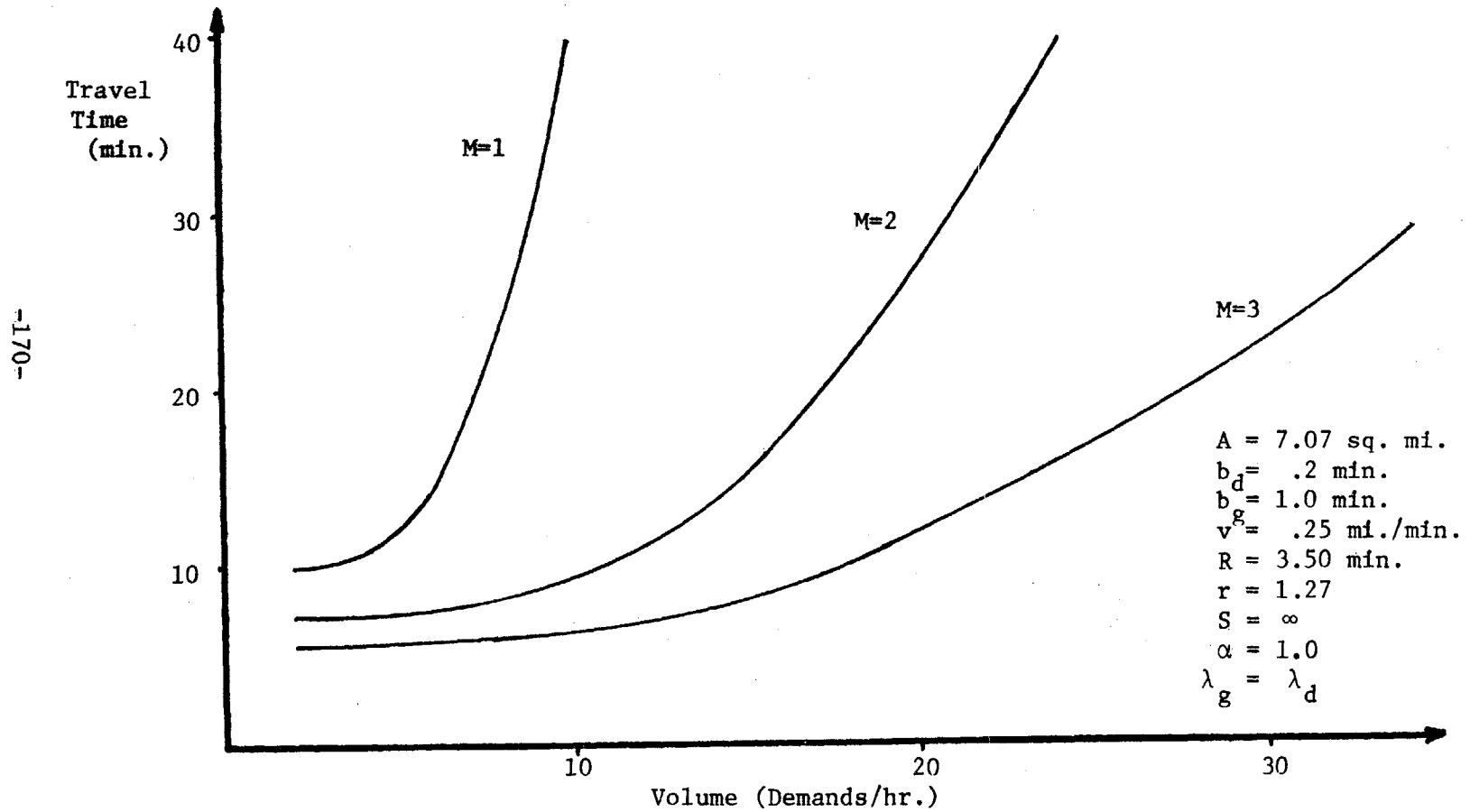
In figure 5.5, the expected travel time of a flexibly routed, scheduled and phased feeder service with various levels of demand is illustrated. In this figure, the cycle length is chosen at each demand level so as to minimize the expected travel time. The three curves represent the system performance with fleets of 1, 2 or 3 vehicles operating in an out-of-phase or offset policy (see Section 5.4 for a comparison with an in-phase or synchronous policy). The expected travel time increases rapidly as congestion in the system increases beyond a certain point. The rate of increase of the travel time is lower for systems with larger fleets, in a manner analogous to multiple server queueing systems. The general shape of the volume/delay curve is quite similar to typical queueing systems.

In developing this figure, it was found that the cycle length which minimized the expected travel time increased nearly exponentially with the demand level (Figure 5.6). While this result could not be obtained analytically, it is useful in searching for desirable cycle lengths.

In Fig. 5.7, the expected travel time of a system corresponding to that of Fig. 5.5 is illustrated, but with fixed rather than variable cycle lengths. Fixed cycle lengths are of interest in cases in which the feeder service is coordinated with a line haul system. In this situation, the time between successive visits of feeder service vehicles to the depot should equal the headway on the line haul system. To insure smooth inter-vehicle transfers, control mechanisms such as vehicle holding strategies may be helpful; such strategies are discussed in Section 6.3.

---

*this argument does not apply to services constrained by vehicle capacity, and as seen in Fig. 5.3, the model is less accurate when vehicle capacity is approached.

Figure 5.5 : Expected Travel Time in a Flexibly-Routed, Scheduled and Phased Feeder Service*

M=1

M=2

M=3

A = 7.07 sq. mi.
b_d= .2 min.
b_g = 1.0 min.
v^g= .25 mi./min.
R = 3.50 min.
r = 1.27
S = ∞
α = 1.0
$\lambda_g = \lambda_d$

Volume (Demands/hr.)

*Cycle Length Optimized to Nearest Minute

Figure 5.6 : Optimum Cycle Lengths in a
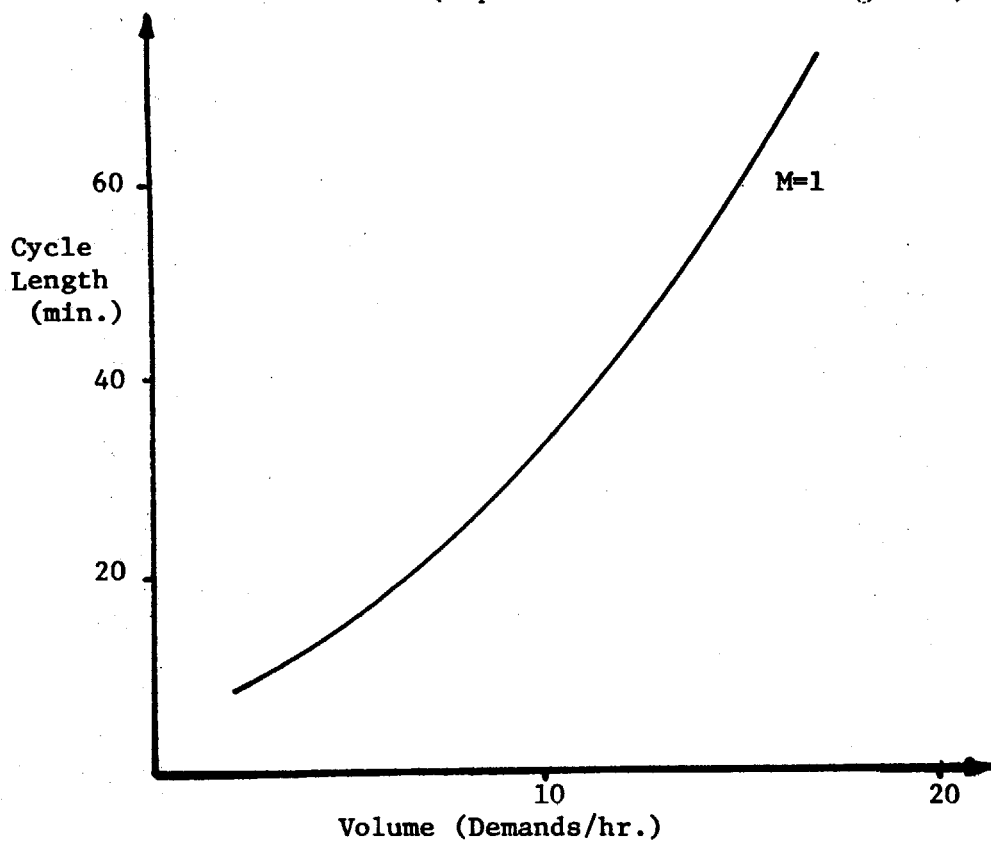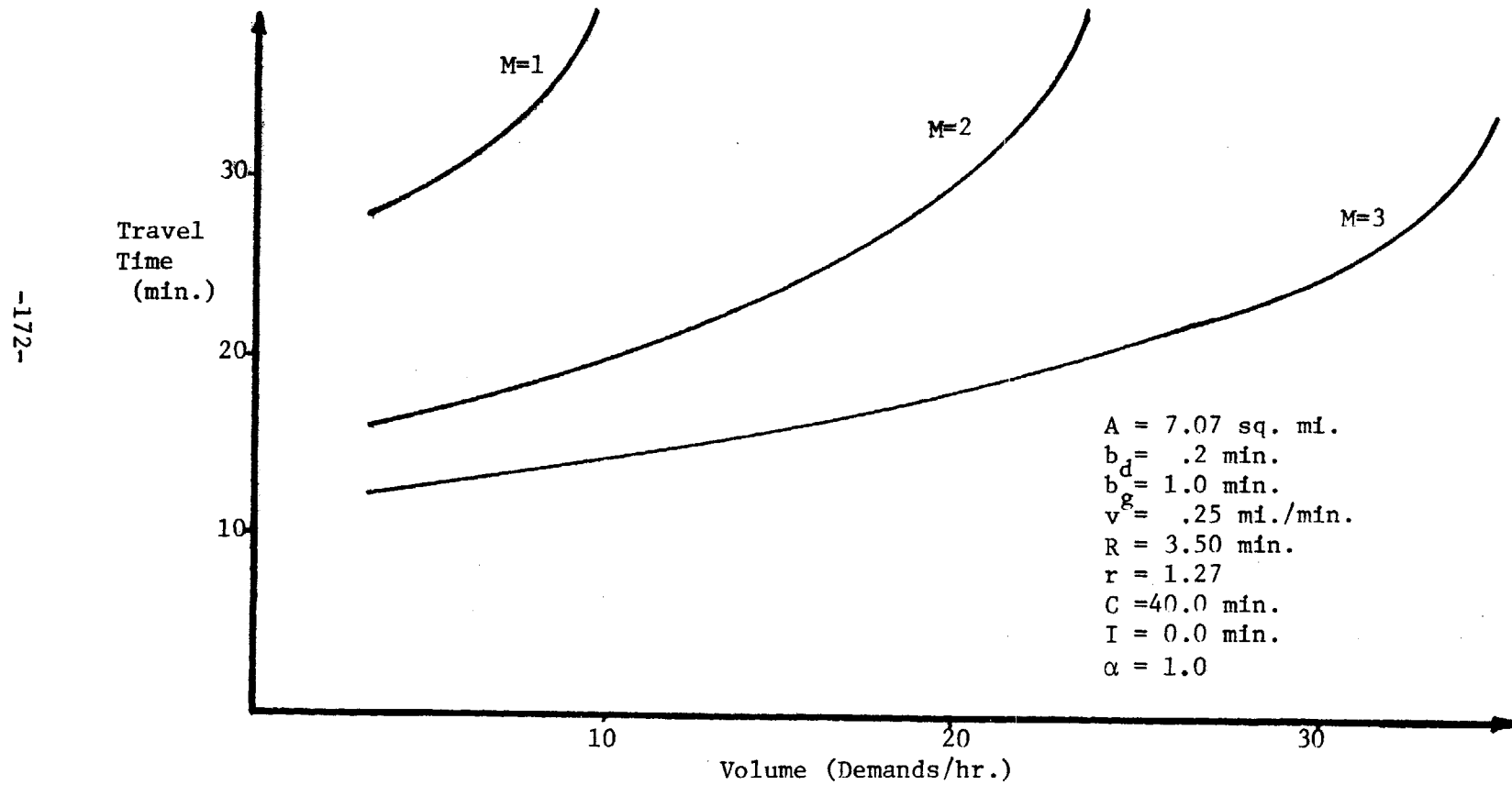   Flexibly-Routed Feeder Service
            (Input Parameters as in Fig. 5.5)

Figure 5.7: Expected Travel Time in a Flexibly-Routed,
Scheduled and Phased Feeder Service with
Constant Cycle Length

Travel
Time
(min.)

M=1

M=2

M=3

30

20

10

Volume (Demands/hr.)

A = 7.07 sq. mi.
b_d = .2 min.
b_g = 1.0 min.
v = .25 mi./min.
R = 3.50 min.
r = 1.27
C = 40.0 min.
I = 0.0 min.
α = 1.0

In comparing Figures 5.7 and 5.5, the fixed and variable cycle length situations, it is clear that the variable cycle situation always results in lower or equal average travel times than the fixed cycle case. This result is analogous to that of short and long run economic supply functions, in which the greater flexibility of the long run supply decision permits greater efficiency of production [32]*.

The response of the expected travel time to changes in vehicle speed and the square root of area size is generally linear. As can be seen in Figs. 5.5 and 5.7 the response of the expected travel time to increases in the vehicle fleet size is non-linear, with a diminishing marginal reduction in delay as the number of vehicle increases.

5.4 The Effect Of Fine Tuning Correction Terms.

In Section 3.3, a deterministic model of feeder service performance was derived. In Section 3.4, a series of fine tuning corrections were introduced to account for the variability of arrival and service processes and for the "integerness" of patrons. In this section, we shall consider the effect of the various correction terms derived in Section 3.5. The modifications are grouped in three categories:

- correction term for the variability of the delivery arrival process, Δ.

- corrections for the integerness of patrons and variability in vehcle tour lengths.

*As in the case of long and short run supply curves, the variable cycle performance function is the envelope of a series of fixed cycle curves.

- corrections due to the variability in the arrival and service pro-

cesses and the constraint imposed be the vehicle capacity.

In Figure 5.8 , the effects of the first two categories of corrections are illustrated. Initially, we shall assume that vehicle capacity is quite large. The "unmodified" model results are found using only the deterministic model of Section 3.3. The curve marked "Δ" shows the results of deterministic feeder service model in which only the factor Δ has been inserted in the delivery tour length expression (Eq. 3.23), where:

$$\Delta = \frac{n_d}{8(n_d+0.5)^2}$$

As can be seen, the effect of introducing Δ is quite small, except with relatively short cycle lengths. At the optimum cycle length (i.e. the cycle length with the minimum travel time), the inclusion of Δ results in approximately a 10% decrease in expected travel time. With longer cycle lengths (as would occur in practice), the effect is smaller.

The fully modified model results are also shown in Fig. 5.8 . The net effect of the correction terms is substantial, with the optimum cycle length increased by 5 minutes and the minimum expected travel time increased by nearly 50%. Considering a cycle length of 25 minutes (which minimizes travel time for the modified model), the corrections increase the prediction of travel times by approximately 36%. The net effect of the correction terms is to always increase the expected travel time.

Certainly, the effects of stochastic variations are not negligible in the situation illustrated in Fig. 5.8 . However, as the cycle length increases, the effect of the correction terms become proportionally smaller. Also, in uncongested systems, the effect of the correction terms become much

Figure 5.8: Effect of Correction Terms on the Expected Travel Time in a Flexibly-Routed, Phased and Scheduled Feeder Service

-175-
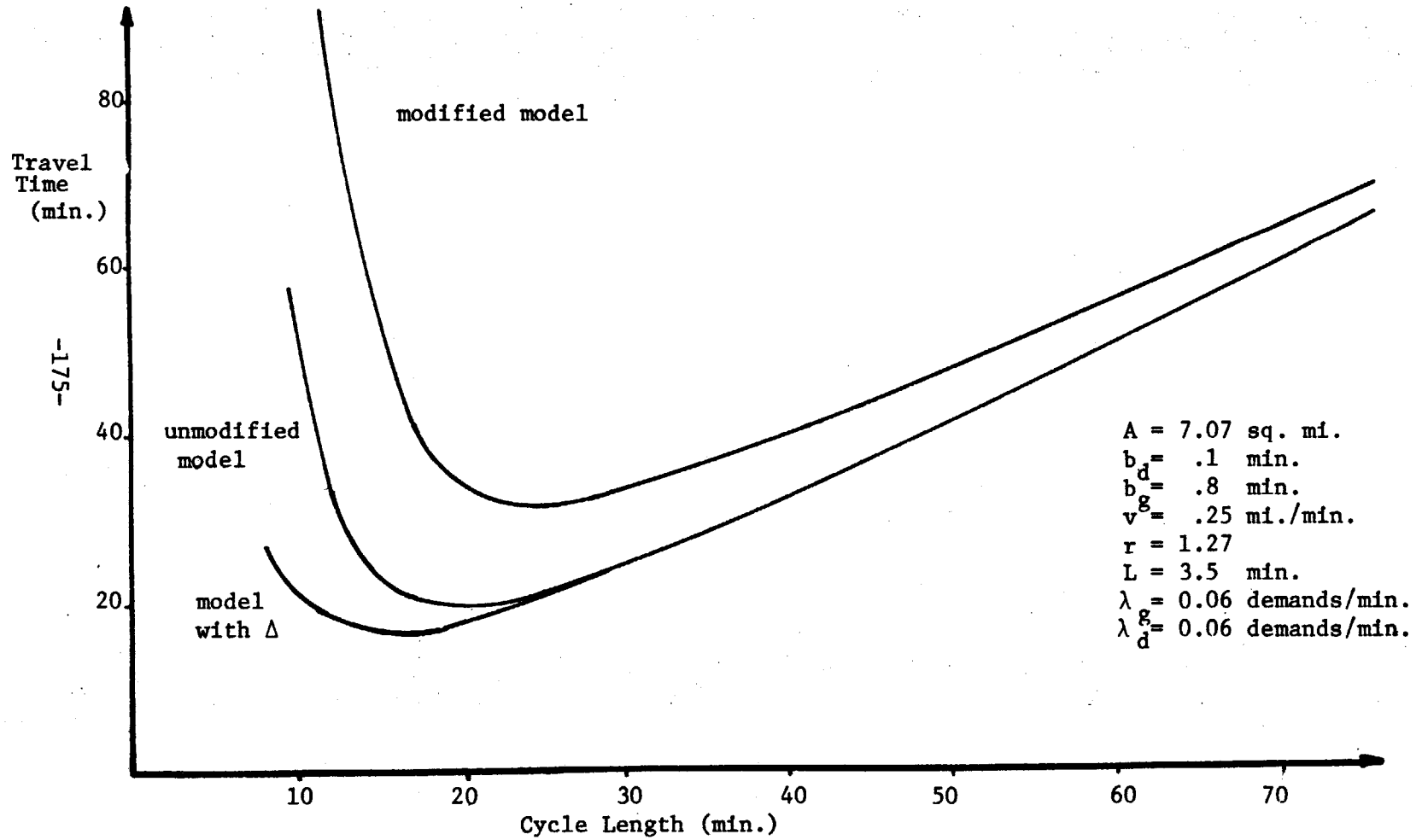
smaller.

The effect of vehicle capacity is illustrated in Fig. 5.3 for the same situation as in Fig. 5.7 but with a vehicle capacity of 4 . For the deterministic model, vehicle capacity only has an effect when the expected number of arrivals exceeds available capacity, at which time no steady state solution exists. In the modified model, service during occasional high demand periods may be hampered even though the vehicle capacity is not exceeded on the average. As a result, expected travel time increases, particularly as the average vehicle load approaches the vehicle capacity.

5.5 Alternative Operating Options

As discussed in Section 3.9, an operator has several options or alternatives for feeder services. We summarized these options by a vector H:

$$H = H(P, M, C, I, S)$$

where P indicated phased or unphased service,

M is the number of vehicles operating in a particular zone,

C is the cycle length,

I is the idle time scheduled between phases, and

S is vehicle capacity.

The effects of various C, I and S values have been illustrated earlier. In this section, we shall consider the options of phased and unphased service (P) and the number of vehicles per zone (M). In addition, we shall discuss the options of scheduled or unscheduled feeder service in a qualitative manner.

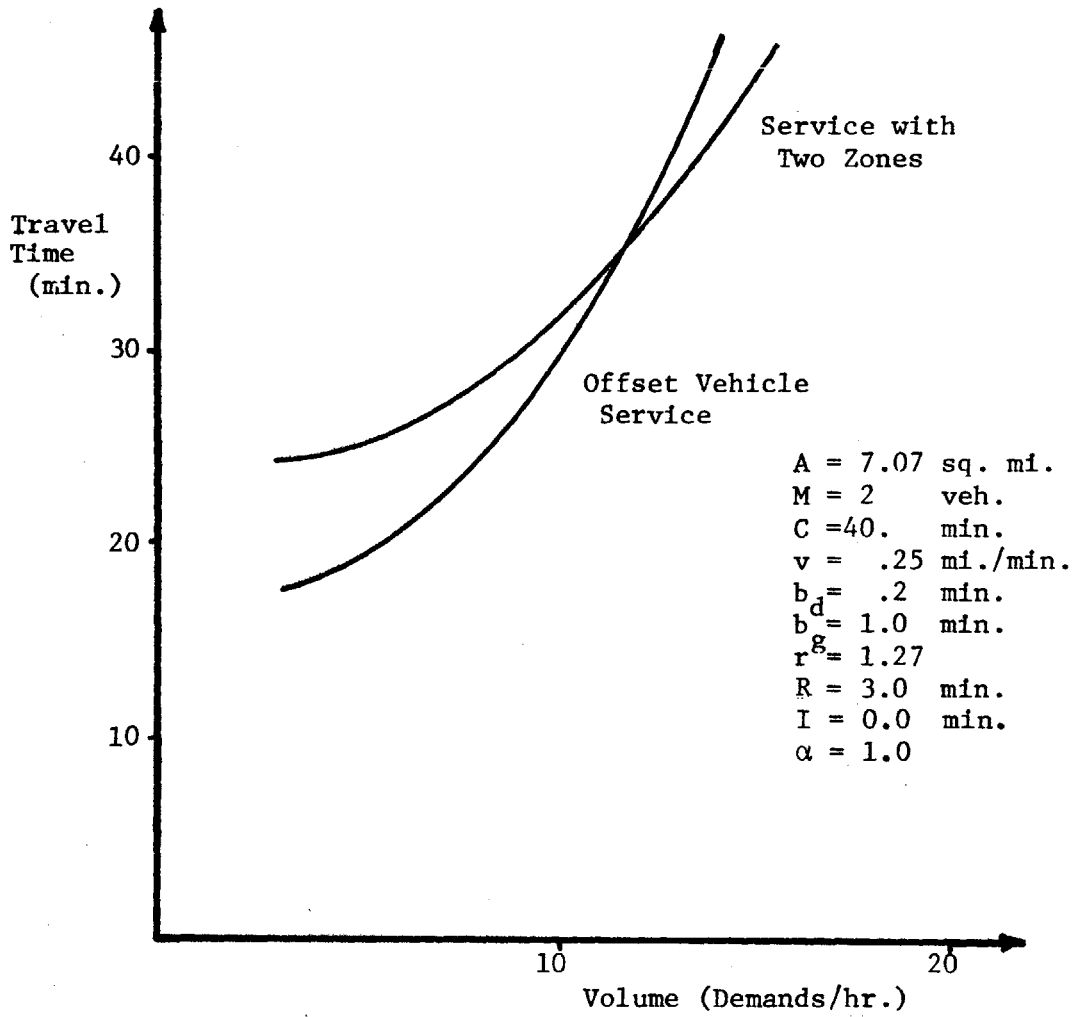With more than one vehicle available, an operator has the option of

operating the vehicles throughout the service area or of partitioning the service area into zones and operating just one vehicle per zone. If more than one vehicle is to be operated in the same zone, then the most advantageous mode of operation is to offset the vehicles' visits to the depot, just as it is more advantageous to operate fixed route service with headways between subsequent vehicles. To achieve offsets, different vehicles are simply scheduled to return to the depot at different times. Offset operation will tend to be more advantageous when cycle lengths are long or the depot is distant from the service area. Partitioning the service area results in more efficient vehicle tours, which is of importance in congested systems. Thus, in choosing between the two vehicle deployments, the operator is trading off reductions in waiting time for increased efficiency in routing.

To illustrate the differences in the options, consider a feeder service to be operated in a 4 sq. mi. service area with a 2 vehicle fleet. The operator may divide the service area into two zones and operate feeder services in each zone separately. Alternatively, the operator may operate the vehicles throughout the zone, so that M = 2. Figure 5.9 shows the resulting expected travel times with the two options for a variety of demand levels. The offset vehicle deployment strategy is more advantageous below 12 demands per hour.*

Turning to the option variable P, the choice between phased and unphased service may also be characterized as a tradeoff between service

_____

*The calculations for this illustration were made by Multisystems, Inc., and appear in Batchelder et.al. [3].

Figure 5.9: Alternative Vehicle Deployment Options
in a Flexibly-Routed Feeder Service

Travel
Time
(min.)

Service with
Two Zones

Offset Vehicle
Service

A = 7.07 sq. mi.
M = 2   veh.
C =40.   min.
v =   .25 mi./min.
b_d=   .2 min.
b_g= 1.0 min.
r = 1.27
R = 3.0 min.
I = 0.0 min.
α = 1.0

Volume (Demands/hr.)

quality and tour efficiency. Unphased service involves a single vehicle tour between visits to the depot in which deliveries and collections are interspersed. Unphased service results in more efficient tours, but riding time for deliveries increases, while riding time for collection increases by the same amount that waiting time at home decreases. The unphased service also results in tours in which patrons which are collected may be first driven directly _away_ from their destination, which may be psychologically disturbing. For exclusively collection or delivery services, the two types of operations are identical. The maximum differences occur for the case of balanced demand (that is, when the incoming and outgoing demand rates are equivalent).

The time required to collect and deliver patrons with the two types of operating policies may be derived from the tour length approximation formulae in Section 3.2. For the unphased service, the tour time is:

$$T_u = (b_g n_g + n_d b_d) + \frac{\alpha r \sqrt{A}}{v} (\sqrt{(n_g+n_d)+0.5} - \sqrt{0.5}) \qquad (5.1)$$

and for the phased service:

$$T_p = (b_g n_g + n_d b_d) + \frac{\alpha r \sqrt{A}}{v} (\sqrt{n_g+0.5} + \sqrt{n_d+0.5} - 2\sqrt{0.5}) \qquad (5.2)$$

with the difference of:

$$T_p - T_u = \frac{\alpha r \sqrt{A}}{v}(\sqrt{n_g+0.5} + \sqrt{n_d+0.5} - \sqrt{n_g+n_d+0.5} - \sqrt{0.5}) \qquad (5.3)$$

and with balanced service ($n_g = n_b = n$):

$$\max\ (T_p - T_u) = \frac{\alpha r \sqrt{A}}{v}\ (2\sqrt{n+0.5} - \sqrt{2n+0.5} - \sqrt{0.5}) \qquad (5.4)$$
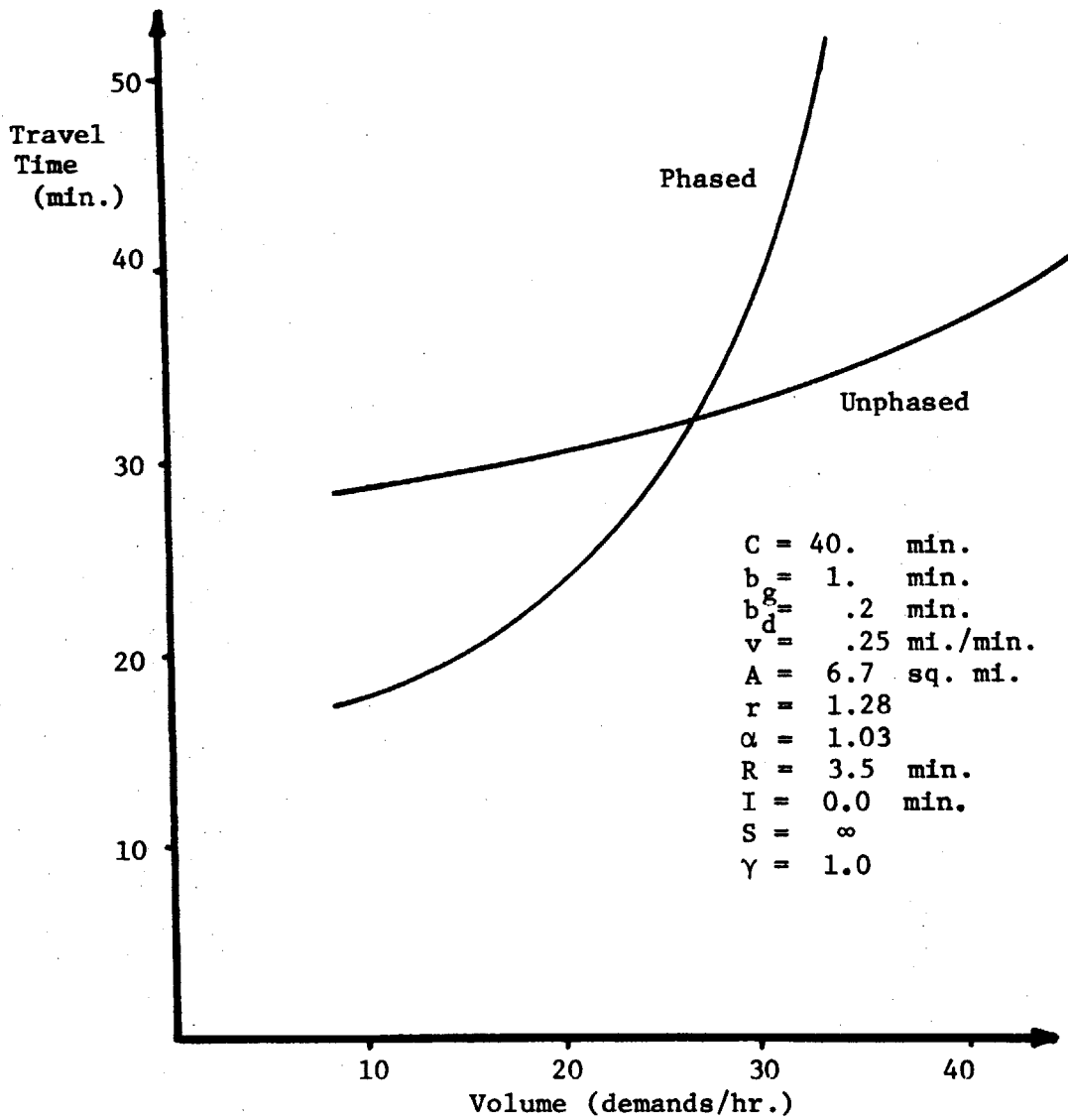
For the service illustrated in Figure 5.1, the maximum difference between the length of the phased and unphased tours is approximately 20% of the phased tour.

In Figure 5.10, the expected travel time with phased and unphased operating policies is illustrated for a specific case in which demand is balanced between deliveries and collections. This system is similar to that illustrated in Figure 5.2, including balanced demand, but with a constant cycle time of 40 minutes. At low demand levels, phased operation results in lower expected travel times. During high demand periods, such as the rush hour, the system demand would be expected to be unbalanced; inbound trips would predominate during the morning peak, for example. Consequently, the phased and unphased services would resemble one another during typical peak hours.

A final operating option consists of scheduled or unscheduled service from the depot. This problem has been studied as a topic in control theory, with major contributions by Osuna and Newell [61] and Barnett [2] in the case of fixed route transit service. Using fluid approximations to arrival processes, these authors found that it is desireable to hold vehicles a certain amount of time so as to improve the regularity of service. While the flexibly routed service has the additional complication of a state dependent service rate, the general conclusion concerning the desireability of regular service is applicable.

With some knowledge of the arrival pattern, however, it may be

Figure 5.10: Expected Travel Time in Phased and
Unphased Flexibly-Routed Feeder Services

Travel
Time
(min.)

Phased

Unphased

C = 40. min.
b_g = 1. min.
b_d = .2 min.
v = .25 mi./min.
A = 6.7 sq. mi.
r = 1.28
α = 1.03
R = 3.5 min.
I = 0.0 min.
S = ∞
γ = 1.0

Volume (demands/hr.)

advantageous to deviate from the regular cycle length. For example, if another collection may be made at the cost of slightly missing the deadline for return to the depot, it is often more desireable to make the collection. With a fairly uncongested system or low capacity vehicles, it may also be desireable to respond immediately to arrivals, as in the case of exclusive ride taxi systems.

5.6 An Equilibrium Application Example

To illustrate the use of the feeder service model in a design study, we consider the case in which a local social service or transit agency is planning a many-to-one feeder service to a shopping center. Patrons at home are expected to call a central dispatcher and be served as soon as possible, as in a taxi system. Since sufficient demand is expected to make regular service desireable (see Section 5.5) and to avoid the costs of dispatching from the shopping center, the system will be scheduled so that regular departures are made from the shopping center. The agency must decide how many vehicles to use, what schedule to operate, and what fare to charge for the service. It would also like to predict the system performance and patronage level.

The service is planned for a small suburban area and is intended to primarily serve elderly individuals. Consequently, the agency expects the service to attract a relatively low patronage and to operate at a deficit. As an initial design step, estimates of the service area characteristics are made. The area size (A) is 7.07 sq. miles, average vehicle speed (v) of 15 mph. is attainable, the travel factor in the area (r) is 1.27, and the area is roughly circular (so $\alpha=1$). From experience on a similar system, boarding time (b_g) is expected to be 1 min. and expected exiting

time (b_d) is .2 minutes. The shopping center is located slightly inside the service area, so the expected time to transfer passengers at the shopping center (3.5 minutes) is assumed as the rendezvous or layover time (R). These parameters are all that are necessary to apply the analytic performance models, with the exception of expected demand.

We shall assume that the task of predicting demand as a function of fare, level of service attributes, and socio-economic characteristics has been accomplished. This prediction may consist of estimating or updating a demand model for the service (see Lerman [47] for an example). Alternatively, a comparison to an existing system plus some estimate of the elasticity of demand permits the estimation of a straight line or constant elasticity curve approximation to the demand function. For simplicity in this example, we shall use a straight line approximation to the demand function. The base point for this approximation is 16 trips per hour (a demand density of 2.3 trips/sq. mi/hr.) at an average trip time of 20 minutes and a fare of $.40 per trip. The elasticities of demand at this point are assumed to be -.6 with respect to travel time and -.3 with respect to fare. Thus, the slope of the linear approximation to the demand function at the base point is:

$$\frac{\partial V}{\partial (ST)} = V/(ST) \cdot \varepsilon_{ST} = (16/20) \cdot (-.6) = -.48$$

where V is volume per hour,

ST is service or trip time,

ε_{ST} is elasticity with respect to travel time.

The expected travel time for a scheduled, phased feeder service

with fleets of 1, 2, or 3 vehicles is summarized in Fig. 5.11 as a function of volume (equations for this model are summarized in Section 3.9). In addition, Fig. 5.11 includes the approximate demand function at a fare of $.40 per trip which is discussed above. By inspection, the equilibrium demand and travel . time with fleets of 1, 2, and 3 vehicles are:

1 vehicle	9 patrons/hr.	34 min. travel time
2 vehicle	16 patrons/hr	19 min. travel time
3 vehicle	20 patrons/hr.	12 min. travel time

As expected, increased vehicle fleet sizes result in lower travel times and higher volumes in the system. If vehicles cost $12.00 per hour to operate, then the resulting system deficits are $8.40, $17.60 and $28.00 for the three vehicle fleet sizes.

By altering the fare charged, the demand for service may be increased. For example, Fig. 5.12 shows the performance curves and the approximate demand function at a fare of $.25. In this case the equilibrium predictions are:

1 vehicle	9.8 patrons/hr.	38 min. travel time
2 vehicle	18 patrons/hr.	24 min. travel time
3 vehicle	12.2 patrons/hr.	14 min. travel time

and the resulting deficits are $9.55, $19.50, and $30.50 respectively for the 1, 2, and 3 vehicle fleet sizes. With the fare reduction, expected travel time, system deficit and patronage all increased. By searching among various fare policies and fleet sizes, it is possible to identify the range of system performance and deficit characteristics. One may also

Figure 5.11: Expected Travel Time and Volume with a Flexibly-Routed Feeder Service with Three Fleet Sizes (Fare = $.25)

(Cycle Length chosen to minimize travel time)
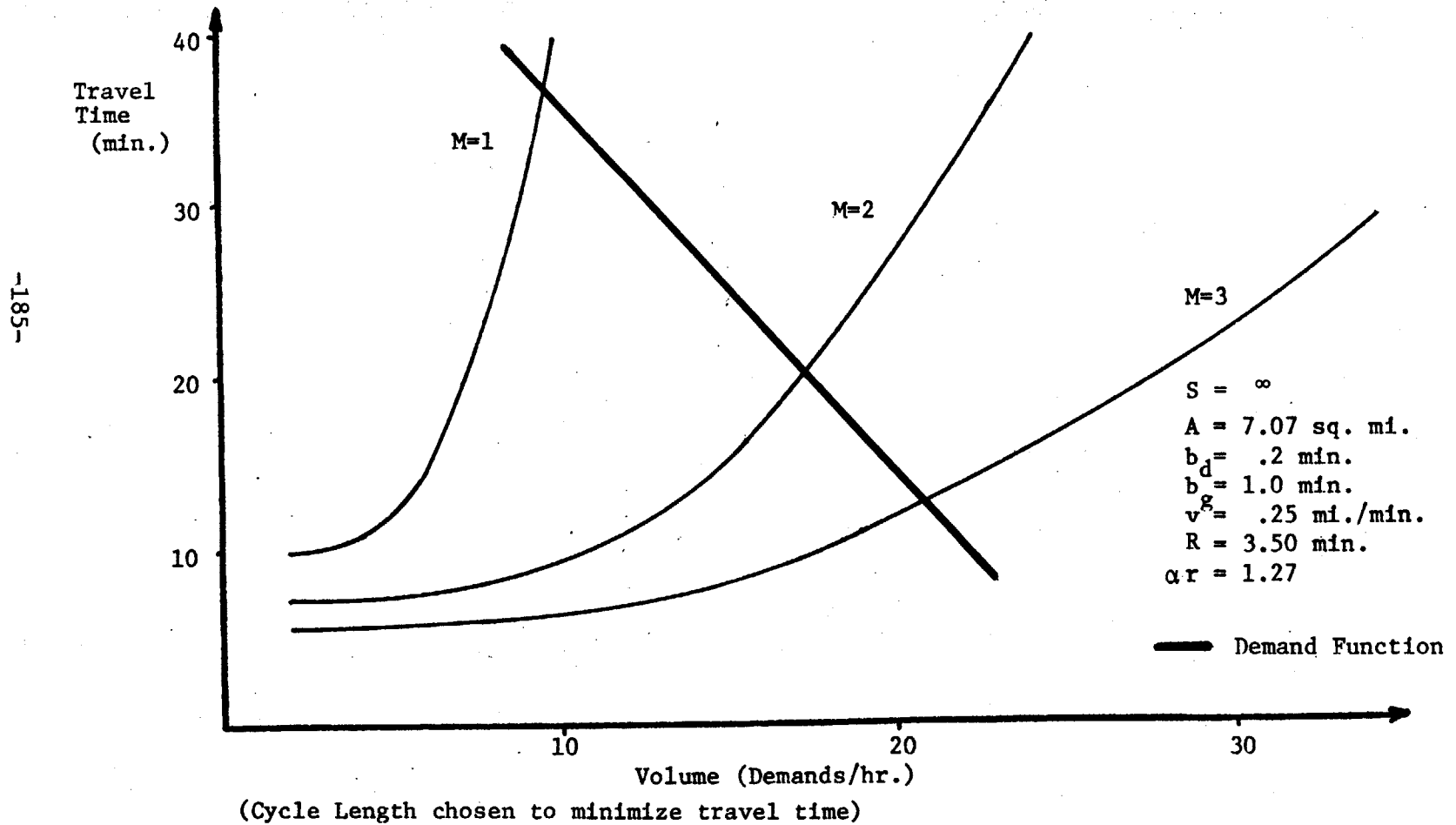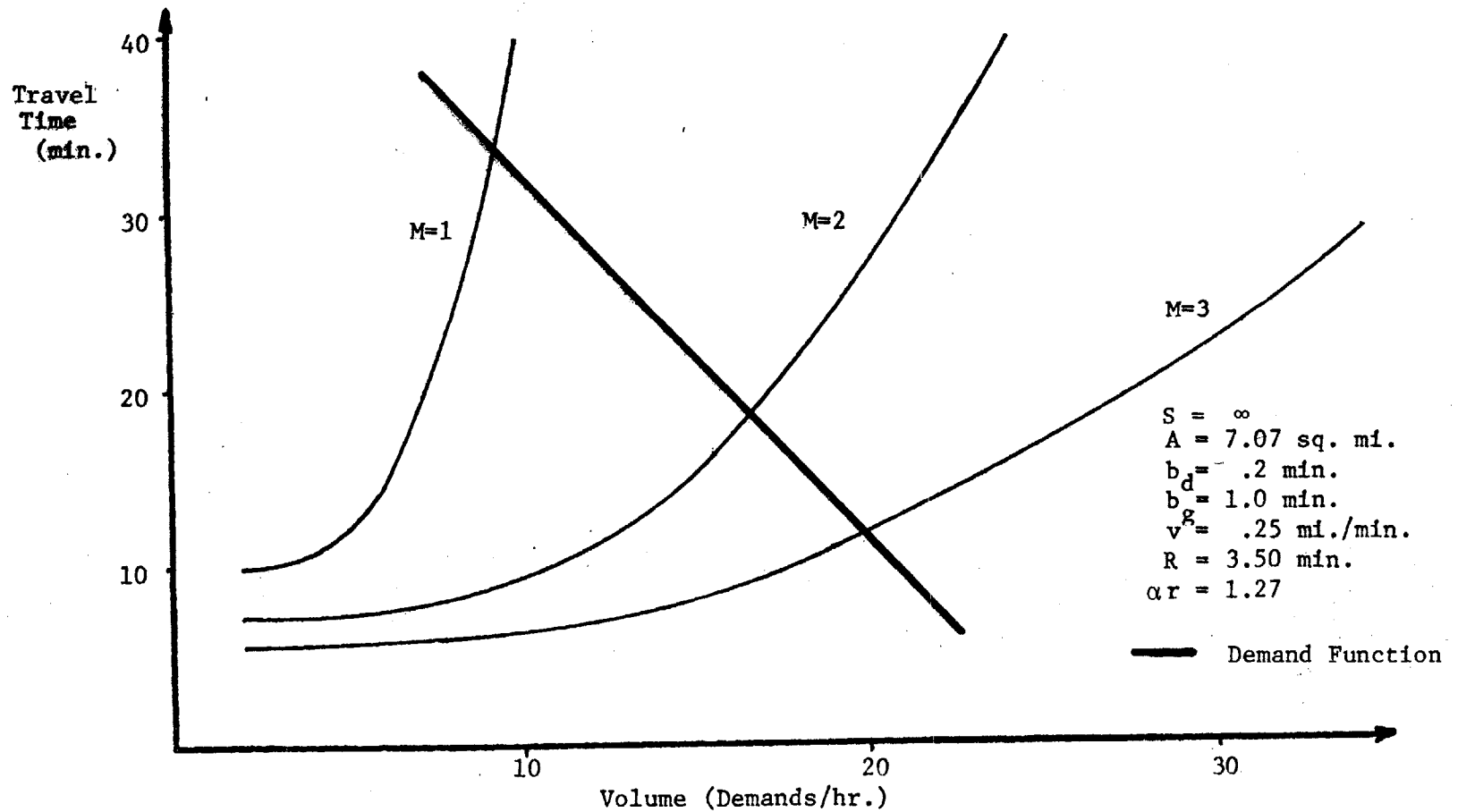
Figure 5.12: Expected Travel Time and Volume with a
Flexibly-Routed Feeder Service with Three Fleet Sizes

Travel
Time
(min.)

40

30

M=1

M=2

20

M=3

10

$S = \infty$
$A = 7.07$ sq. mi.
$b_d = .2$ min.
$b_g = 1.0$ min.
$v = .25$ mi./min.
$R = 3.50$ min.
$\alpha r = 1.27$

Demand Function

10   20   30

Volume (Demands/hr.)

(Cycle Length chosen to minimize travel time)

perform parametric tests on the sensitivity of the equilibrium character-
istics to cost, demand or model inputs. For example, Table 5.3 summarizes
the equilibrium characteristics if the base level of demand was 25% higher
or lower. In addition, it is possible to compare alternative operating
options such as phased or unphased service, as was done in Section 5.5.

An alternative design procedure consists of estimating an expected
demand level and then identifying the least cost (or maximum benefit)
design at that demand level. For example, suppose that patrons' time is
valued at $5.00 per hour while at home or shopping and $10.00 per hour
while waiting at the shopping center. Using the estimates of patron travel
time from the performance model (summarized in Fig. 5.11) and the vehicle
cost of $12.00 per hour, the total system costs (including patron travel
time) may be found at all demand levels, and least cost options identified
(Fig. 5.13). The least cost option envelope in Fig. 5.13 represents a
supply function, in the sense that this is the least cost (or maximum
benefit) operating policy at each possible volume level (Fig. 5.14).

It is incorrect to assume a fixed demand for service in this case,
however, and the assumption may lead to misleading results. For example,
if the least cost supply function is compared with the demand function
at a fare of $.40, the intersection occurs at a point at which three veh-
icles are desired. However, the equilibrium solutions identified previous-
ly (Fig. 5.11) indicate that the use of one vehicle results in lower costs.
This result is attributable to the sensitivity of demand to the level of
service and the structure of the least cost objective function. With one
vehicle in service, fewer patrons are attracted and this effect in combi-
nation with lower system operating costs results in a lower total cost

Table 5.3: Equilibrium Volumes, Level of Service and Deficits for Various
Demand Functions

Demand Function		1 Vehicle	2 Vehicles	3 Vehicles
Base*	Volume	9	16	20
	Service Time	34	19	12
	Deficit	8.40	17.60	28
Base + 25%	Volume	10	18	22
	Service Time	41	24	14
	Deficit	8	16.80	27.20
Base − 25%	Volume	8	14	17
	Service Time	28	15	10
	Deficit	8.80	18.40	29.20

*Base Case Demand is 16 trips/hr. at a service time of 20 min., a $.40
fare, and an elasticity of −.6.

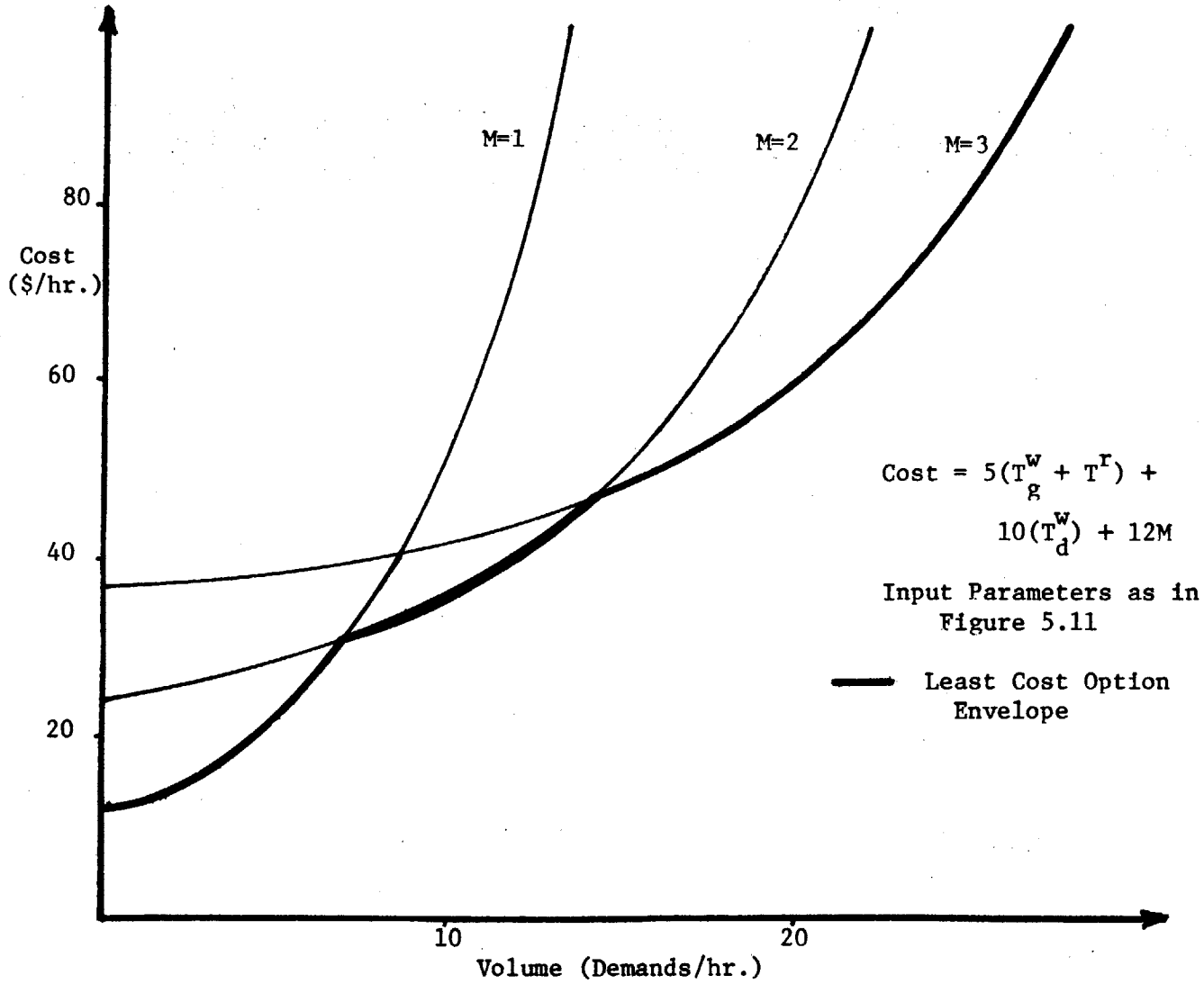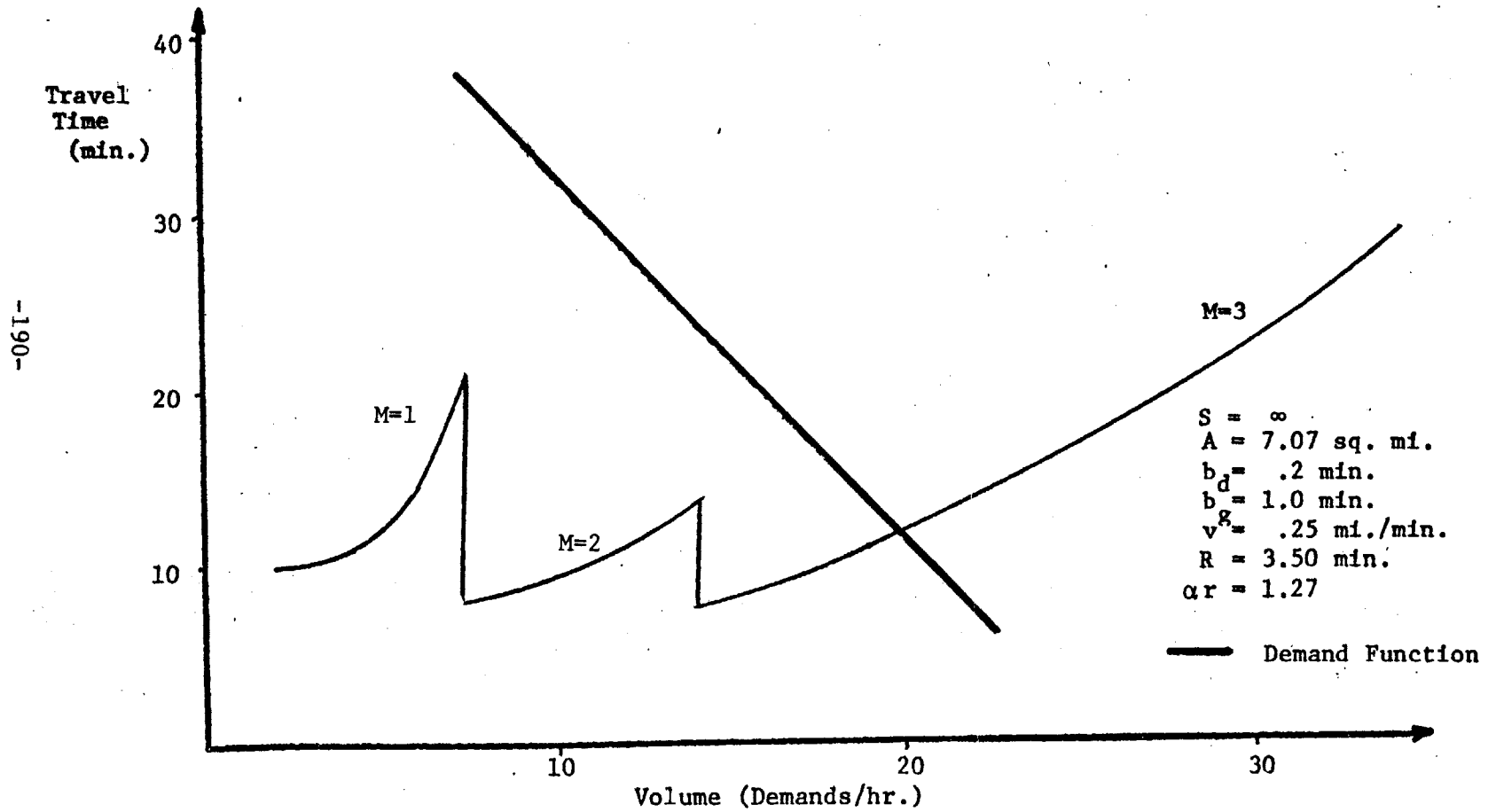Figure 5.13: Expected Patron and System Cost of Flexibly-Routed Feeder Services with Three Fleet Sizes

Figure 5.14: Expected Travel Time and Volume with a Least Cost Supply Function of a
Flexibly-Routed Feeder Service with Three Fleet Sizes

$S = \infty$
$A = 7.07$ sq. mi.
$b_d = .2$ min.
$b_g = 1.0$ min.
$v^g = .25$ mi./min.
$R = 3.50$ min.
$\alpha r = 1.27$

—— Demand Function

(Cycle Length chosen to minimize travel time)

(of user cost plus system cost). The supply curve of Figure 5.13 is de-
rived with the implicit assumption that demand is unaffected by the level
of service provided (which is the classic economic assumption for a per-
fectly competitive market). Clearly, the expected elasticity of demand
is a critical parameter in the design process and a careful analyst would
perform a sensitivity analysis around the expected value of the elasticity.

This example illustrates the importance of congestion effects and
demand sensitivity to the analysis of alternatives. In addition, it
serves to illustrate the use of a performance model in a design process.
The analyst worked in an environment in which demand was assumed to be
quite uncertain. A range of alternatives was analyzed. The output of
a simple analytic performance model could be rapidly employed to analyze
the various alternatives.

CHAPTER 6

PERFORMANCE OF FIXED ROUTE AND

INTEGRATED TRANSIT SERVICE

6.1 Introduction

The past three chapters developed performance models of flexibly-routed feeder services in detail. This model can be used to predict the level of service and resources consumed by an isolated feeder service operating in a particular area and with a variety of operating strategies. In the present chapter, we shall develop a performance model of fixed route service and then consider the performance of transit services in which routes are coordinated and in which both flexibly and fixed route services may be provided. Application of the models to isolated areas and to region-wide service will also be presented. This chapter is primarily intended to indicate the types of analysis which are possible, since a comprehensive treatment of transit design is beyond the scope of this thesis. In particular, the usefulness of the modelling methodology applied in Chapters 3-5 will be discussed.

Fortunately, development of an adequate performance model of fixed route service is somewhat easier than in the case of flexibly routed services. Since routes are known, a fairly accurate estimate of travel time between stops and the variability of travel time may be obtained without great effort. Moreover, the level of service provided by a fixed route system is generally insensitive to the level of demand, except for cases

when substantial increases in patronage occur.* The primary difficulty

in applying such a model is in accounting for the substantial number of

routes, links and constraints on service which occur in an actual transit
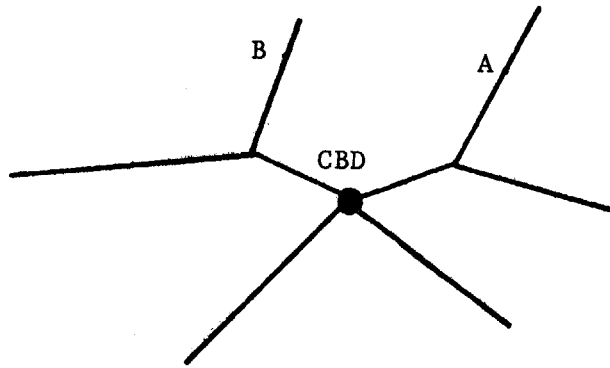
system.

In designing transit services, operators may have the option of

operating traditional fixed route services or various types of flexibly

routed services, possibly under contract to private carriers. This choice

is a central issue for the design of integrated transit systems. Some of

the service features which influence the choice between flexibly or fixed

route service may be mentioned qualitatively as an introduction to the

discussion below. Consider a distribution service in which patrons are

delivered within a service area from a central station. In general, flex-

ible routing results in greater vehicle travel - as patrons are delivered

nearer their homes - and longer in-vehicle or riding time for patrons com-

pared to fixed route service. Consequently, as the cost of vehicle ope-

rations, vehicle occupancy, or the value given to patron riding time in-

creases, flexibly routed service become less desirable. However, patrons

have shorter walks with flexible routing of vehicles. As route density

declines or the cost or disutility of walking increases, flexible routing

becomes relatively more desirable. It is also possible that at very low

demand levels, flexibly routing results in less vehicle travel, since an

entire vehicle route need not be traversed. In this case, the vehicle

travel, amount of walking, and patron riding time are all lower with flexible

routing.

---

*Note, however, that as a result of operator responses via a design process,
the system may be altered in response to a change in demand. While the
performance of a given system is insensitive to the level of demand, the
supply of service might be quite sensitive to the level of demand.

With these factors in mind, flexibly routed feeder services are like-
ly to be more desireable in situations with relatively low demand, such as
late night, midday and weekend transit service, or in situations in which
the access cost to a fixed route system is relatively high, as with systems
serving patrons carrying baggage or the physically handicapped, or in si-
tuations in which flexibly routed service is less expensive, as when van-
pools may be substituted for fixed route service.

Another central issue in transit network design is that of the net-
work shape and connectivity. The bulk of existing transit service is
oriented toward the Central Business District (CBD), resulting in a primar-
ily hub-and-spoke design (Figure 6.1a). Patrons travelling circumferential-
ly (such as A to B in Figure 6.1a) receive quite poor service, often with
a trip to the CBD and a transfer required. Service for circumferential
trips may be enhanced by inserting more direct services, such as the cir-
cumferential line of Figure 6.1b. With a fixed amount of resources avail-
able, introducing additional direct service or transfer points into a net-
work tends to reduce the variance in the level of service patrons ex-
perience, but tends to increases the average travel time [27].

An additional consideration is the structure of routes within the
overall network. It is common to operate all routes to the central hub
area, even if some lines travel over the same street or rail links. An
alternative arrangement is to operate feeder services to a central hub or
line haul system, thereby consolidating trips on a few routes (Figure 6.1c).
With economies of scale, such consolidation may be desirable. A second-
ary benefit of introducing feeder routes is the opportunity for better

(6.1a) Conventional
CBD-Oriented
System
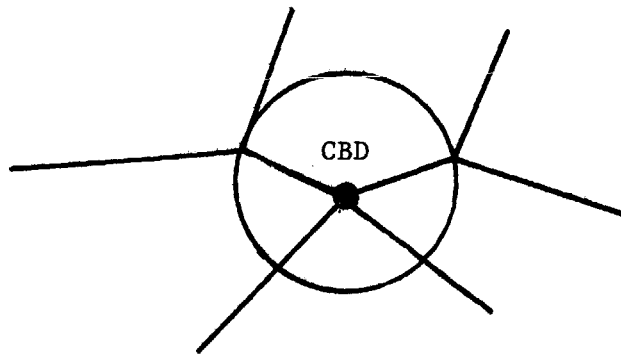
(6.1b) System with Cir-
cumferential
Line
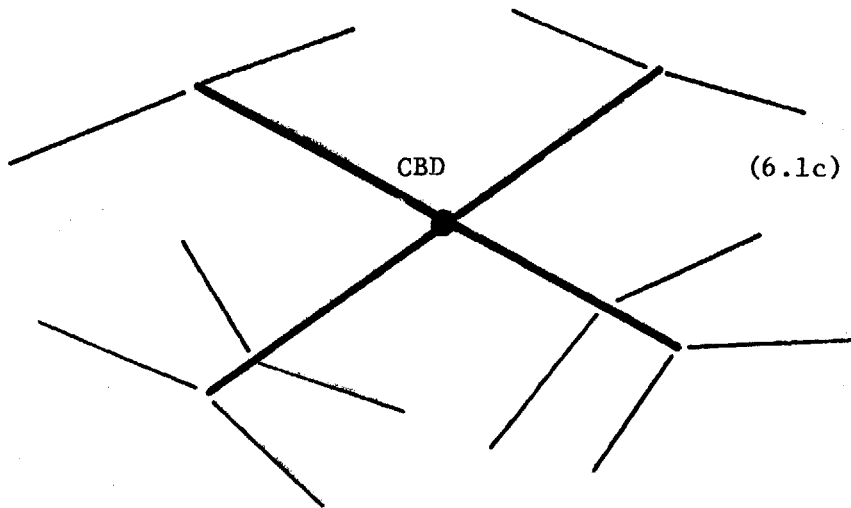
(6.1c) Feeder Service

Figure 6.1: Alternative Transit Network Designs.

control of headways and optimization of route headway in a smaller area. The disadvantage of introducing feeder services is the need for more transfers. Thus, a designer must consider the tradeoff between more efficient service - due to consolidation - and the number of transfers required. With a fixed amount of resources available, more efficient operation implies that more frequent or cheaper service may be provided, at the cost of the larger number of transfers required.

The next section develops a simple, approximate, analytic model of fixed route transit service. Following this, the effects of coordinating transfers are considered. In Section 6.4, a comparison is made between fixed and flexibly routed feeder service in the same area. Finally, Section 6.5 presents a comparison between two types of region-wide, flexibly routed transit systems. One of these systems divides the region into zones, while the other permits area-wide vehicle travel.

6.2 An Approximate, Analytic Model of Fixed Route Transit Service

In this section, a simple model of fixed route transit service is developed which uses the modelling methodology discussed in Chapter 1. First, a deterministic model of performance is developed. Then, consideration of stochastic effects, capacity limitations and control strategies lead to modifications of the basic model. The final result is a simple analytic model which offers a fairly good approximation of route performance.

Developing a deterministic model of an isolated bus route is fairly straightforward. The expected time elapsed in travel from i to j, $\bar{t}_{ij}$, is:*

---

*This model may be extended to time dependent travel times or demand levels, but this extension is beyond the scope of this thesis.

$$\bar{t}_{ij} = \bar{\tau}_{ij} + b \cdot \bar{p}_{ij} \tag{6.1}$$

where $\bar{\tau}_{ij}$ is the expected travel time, without boardings,

$\bar{p}_{ij}$ is the expected number of patrons boarding or descending, and

b is the average boarding time.

The expected travel time, $\bar{t}_{ij}$, may be modified to reflect the probability that the bus need not stop at some potential stops or may be estimated from observations. The expected ride time from i to j is the vehicle travel time:

$$E[T^r_{ij}] = \bar{t}_{ij} \tag{6.2}$$

With the assumption of random arrivals, the expected wait time to board at stop j is simply half the headway*.

$$E[T^w] = h_j/2 = h/2 \tag{6.3}$$

where h is the scheduled headway.

Random fluctuations in vehicle speed, passenger arrivals and the boarding process cause the system performance to deteriorate. In particular, equal spacing of buses may not be maintained. As discussed in Section 2.3, waiting time is a function of the variability of headways [61]:

$$E[T^w] = (\frac{h_j}{2})(1 + (\frac{\sigma_j}{h_j})^2) \tag{6.4}$$

where σ_j is the standard deviation of the headway distribution at stop j.

---

*It is possible to generalize this model to the case of non-random arrivals, but it is not pursued here because the desired departure times are likely to be random, even though patron arrivals are not.

Jolliffe and Hutchinson [34] report one series of observations on a route in which no buses were cancelled. In this case, waiting time was 12% higher than would be the case with perfectly spaced buses (so that $\sigma_j^2/h_j^2 = .12$). Barnett [2] found an increase of 19% due to headway variance during observation of a rail rapid transit line in Boston, Massachusetts.

It is difficult to derive the standard deviation of headway distributions, σ_j, as a function of route characteristics. Osuna and Newell [61] developed a model of a single vehicle, loop transit route which permits calculation of σ_j. Barnett [2] suggests empirical techniques for estimating σ_j and identifying good vehicle control strategies.

One such control mechanism is to insert a certain amount of slack time into schedules and to introduce a simple vehicle holding strategy. Typically, a series of stops with scheduled departure times are included in each bus line. If a bus can leave such a stop ahead of schedule, it must wait until the scheduled departure time; otherwise, the bus leaves immediately. In this case, the scheduled time between successive departures from stop i and stop j, H_{ij}, is:

$$H_{ij} = \bar{\tau}_{ij} + b \cdot \bar{p}_{ij} + \varepsilon_{ij} \tag{6.5}$$

where ε_{ij} is the slack time scheduled between i and j, and thus the expected ride time from i to j is:

$$E[T_r] = \bar{t}_{ij} + \varepsilon_{ij} \tag{6.6}$$

The value ε_{ij} should be chosen such that the vehicle has a good chance, on the average, of returning to the schedule if a random fluctuation slows the vehicle.

In one analysis, Newell [59] concludes that ε should be on the order of 10 to 30 seconds per mile for an urban route. More explicitly, in order to insure that buses will have reasonably high probability of maintaining a schedule, the value of ε should be chosen such that:

$$\varepsilon_{ij} > \sigma_{ij} \sqrt{2\lambda_{ij}\tau} \qquad (6.7)$$

where σ_{ij} is the standard deviation of trip time from i to j,

λ_{ij} is the arrival rate of passengers, and

τ is the expected time to serve one additional passenger.

Another consideration in transit system performance is the effect of vehicle capacity on expected waiting time. Suppose all patrons are destined for the CBD or a depot. Let j be the peak load stop along the route. As in the flexibly routed feeder service (Section 3.4), it is a necessary condition for steady state operation that available vehicle capacity not be exceeded:

$$\max_{j}(n_j) = \lambda_j h < S \qquad (6.8)$$

where n_j is the number of patrons aboard the vehicle at j,

λ_j is the demand rate of patrons along the route before the stop at j,

h is the route headway, and

S is the vehicle capacity.

Due to a fluctuation in demand or headways, the vehicle capacity may be exceeded for one (or more) trips, even through sufficient capacity is available overall. We assume that the variability in the number of patrons wishing to board, n_j, is normally distributed with variance equal to n_j.

-200-

We also assume that the number of patrons who wish to board a vehicle is independent of the number of patrons who could not board a previous vehicle. Since patrons turned away from one bus will likely board the next, this assumption is not strictly correct. However, the phenomena of vehicle bunching and transit control strategies make the assumption reasonable. As discussed earlier, a bus which becomes late is likely to encounter more passengers, while the following bus will likely have fewer passengers than normal. Hence, queues of patrons are likely to be able to board the next vehicle. Secondly, transit operators typically insert vehicle or schedule more runs whenever capacity on a particular run is regularly exceeded.

With the assumption of independent, normally distributed demands, the expected proportion of patrons who cannot board a particular vehicle is the probability that n_j exceeds S:

$$P_r\{\max(n) > S\} = 1 - \Phi(\frac{S-\max(n)}{\sqrt{\max(n)}}) \qquad (6.9)$$

Expected waiting time is then:

$$E[T^w] \approx (\frac{h_j}{2})(1 + (\frac{\sigma_j}{h_j})^2) + h_j(1 - \Phi(x))$$

$$\approx \frac{h_j}{2}(3 + (\frac{\sigma_j}{h_j})^2 - 2\Phi(x)) \qquad (6.10)$$

where $x = \dfrac{S-\max(n)}{\sqrt{\max(n)}}$

In addition to the waiting and riding time, patrons of fixed route bus service also incur the cost and time associated with travelling to the bus

route. The expected walking distance to a stop j may be calculated as:

$$E[T_j^a] = \int_0^\infty xf(x)dx \qquad (6.11)$$

where $f(x)$ is the probability of patrons originating a distance of x from the stop and using the stop. For example, in the case of uniformly distributed patron origins, the expected walking distance in a transit system with parallel routes a distance of h units apart is $h/4$ and the variance of this distance is $h^2/48$.

We have now developed a simple model of fixed route transit performance between two points. To summarize, the expected ride time is:

$$E[T_{ij}^r] = \bar{t}_{ij} + \varepsilon_{ij} \qquad (6.6)$$

The wait time is:

$$E[T^w] = \frac{h}{2}\left(3 + (\sigma/h)^2 - 2\Phi\left(\frac{S - \max(n)}{\sqrt{\max(n)}}\right)\right) \qquad (6.10)$$

where $\bar{t}_{ij}$ is expected vehicle travel time between i and j

ε_{ij} is the scheduled slack time between i and j

h is the average inter-vehicle headway,

σ is the standard deviation of headways,

S is vehicle capacity, and

$\max(n)$ is the expected number of patrons at the route's maximum load

point.

The walking time must be calculated from the route density and patron distribution, using Eq. 6.11.

This model may be used as a link performance model between two stops or as a model of an entire line. As will be seen in the next section, consideration of interlink transfers may also be added to the model.

6.3 Interline Transfers

As discussed above, the arrival of a vehicle at a particular stop is subject to random fluctuations, which may be magnified by the process of passengers' arrival and boarding along a line for fixed route service. Consequently, interline transfers are difficult to coordinate.

Without coordination, vehicles would be expected to arrive randomly at the transfer point with respect to one another. The expected wait for a transfer vehicle is then identical to that of a randomly arriving passenger:

$$E[T_j^w] = (h_j/2)(3 + (\sigma_j/h_j)^2 - 2\Phi(\frac{S-n_j}{\sqrt{n_j}}))$$ (6.12)

where h_j is average headway at stop j,

σ_j^2 is the variance of headways at stop j,

S is the vehicle capacity, and

n_j is the expected vehicle load at j

For coordinated transfers, there are a number of potential strategies. First, with regard to waiting rules, scheduled vehicles may be required to wait either until the arrival of the transfer vehicle or until a given scheduled departure time. The latter strategy insures that at least one route operates on schedule, but prohibits transfer from a late vehicle. Secondly, a certain amount of slack time may be inserted in the schedule to facilitate vehicle encounters. For example, in the development of the

model of flexibly-routed feeder transit, a certain amount of slack time at the transfer point was recommended for this purpose (Section 3.3). Since the feeder service is expected to be carrying fewer passengers than the line haul service, inserting more slack time in the feeder service than in the line haul schedule is a reasonable strategy.

The amount of slack time to include in the schedule depends upon the variability of the vehicle arrival process and the relative magnitude of the costs of increased travel time to all patrons versus the cost of missing a transfer.

For example, consider a bus coming to a transfer point. We assume that if any bus arrives at the transfer point, it will wait until the scheduled departure time. Thus, the probability that a bus will not make a transfer connection is simply the probability that it will be late. By inserting a certain amount of slack time, the scheduled departure time is made later and the probability of being late is reduced. Mathematically, the probability of missing a transfer is the probability of arriving after the scheduled departure time a:

$$\Pr\{\text{no vehicle encounter}\} = \Pr\{t > a\} \qquad (6.13)$$

where t is the actual arrival time, and this must exceed the probability of arriving after time a plus a slack time ε :.

$$\Pr\{t > a\} > \Pr\{t > (a+\varepsilon)\} \; ; \; \varepsilon > 0.$$

As in the isolated fixed route model above, the advantage of adding slack time is a reduction in wait time, in this case of the wait time of patrons who miss a transfer connection. The cost of adding slack time is increased

travel time for patrons remaining on the bus and the increased time required for a vehicle to traverse a route.

6.4 Feeder Service in a Specific Area

It is often possible to operate fixed or flexibly routed transit service in a particular area. With the performance models developed previously, it is possible to compare these operating options on the basis of level of service attributes and the number of vehicles required.

As a numerical example, we consider a feeder service to a railroad in a service area of 8 square miles. We shall compare fixed route and flexibly routed, phased feeder services in this area. Balanced demand between inbound and outbound patrons is assumed; unbalanced demand would improve the performance of the flexibly routed service. Headways on the line haul service are fixed at 30 minutes, and all feeder services are operated so as to insure transfers to the line haul system.[*]

For the fixed route service, we assume two routes are operated in the service area, each of length 4.5 miles. Fig. 6.2 presents one such route configuration and service area. With an average vehicle speed of .25 mi/min., travel time on the route is:

$$\bar{\tau}_{od} = 4.5(4) = 18 \text{ min.}$$

We assume that average boarding time is 6 min. and include 3 min. of slack time in the fixed route schedule at both the depot and the routes' ends. The time required to traverse a route is then:

$$\bar{t}_{od} = \bar{\tau}_{od} + b.n + \varepsilon_{od}$$
$$= 18 + 6 + 6 = 30 \text{ minutes}$$

---

[*]This comparison is one of non-optimal services, since area size and headways are fixed.

Depot

—— Transit Lines

Figure 6.2: Two Fixed Route Transit Lines in a Service Region.

and the time between visits to the depot is 60 minutes. With two veh-
icles on each route, the headway between vehicles would be 30 minutes.
Assuming uniformly distributed patrons, the expected walking distance is
approximately .25 miles   (Fig. 6.2). With such short routes and the
insertion of slack time, the variance of headways will be quite small;
we shall assume that it is 1 min^2. The expected riding time on the fixed
route system is then:

$$E[T^r] \approx (\bar{t}_{ij} + bn + \varepsilon_d)/2 = 13.5$$

The waiting time is:

$$E[T^w] = (h/2)(3 + (\sigma/n)^2 - 2\Phi(\frac{S - \max(n)}{\sqrt{\max(n)}}))$$

$$= 45 - 30\Phi(\frac{S - \max(n)}{\sqrt{\max(n)}}))$$

using 6.10, and the expected walk time is:

$$E[T^a_n] = E[d_{walk}]/v_w = .25/2 = 7.5$$

where v_w is the average walking speed, assumed to be 2 miles per hour.
Then, the expected travel time is the sum of these components:

$$E[ST] = E[T^r] + E[T^w] + E[T^a]$$

$$= 13.5 + 7.5 + 45 - 30\Phi(\frac{S - \max(n)}{\sqrt{\max(n)}})$$

$$= 66.0 - 30.0 \, \Phi(\frac{S - \max(n)}{\sqrt{\max(n)}})$$

which lies between 36 and 66 minutes, depending upon the probability of

exceeding vehicular capacity. In what follows, we shall assume the vehicle capacity is 50 patrons. Also, the maximum load always occurs at the depot, so max(n) is the volume served by a single vehicle:

$$max(n) = \lambda/4$$

An alternative to the fixed route service is a flexibly-routed, phased feeder service. The performance model of this service is summarized in Section 3.9.2 and a sample application appears in Section 5.2. To insure meeting the line haul system vehicles every 30 minutes, the flexible routed service will be operated with a cycle length of 60 minutes and with an offset of 2 vehicles per zone. A rendezvous time of 3 minutes at the depot is also assumed. Figure 6.3 presents volume/delay curves for flexibly routed services with fleet sizes of 4,8 and 12 vehicles for this situation.

As the demand for service increases, the fixed route service eventually becomes more desireable. Above a level of 5 demands per hour per square mile, for example, the fixed route service with four vehicles provides a lower travel time than does a flexibly routed service with four vehicles.

Considering the sensitivity of demand to the level of service should not substantially affect the results of the example. The elimination of walking and outside waiting might make the flexibly routed service more attractive to patrons. Patrons further from the fixed route line would be more likely to choose an alternate mode, thereby reducing the average walk time; however, walking time is a relatively minor component of the average

*I am indebted to R. Menhard [50] for performing the flexibly routed service calculations.

v = .25 mi/min.
Cycle = 60
Offset = 2
A = 8 sq. mi.
b_d=.5 min.
b_g=1 min.
R =3 min.
αr= 1.27
I = 0
S = 50

Travel
Time
(min.)

M=4

M=8

M=12

—— Flexible Routing
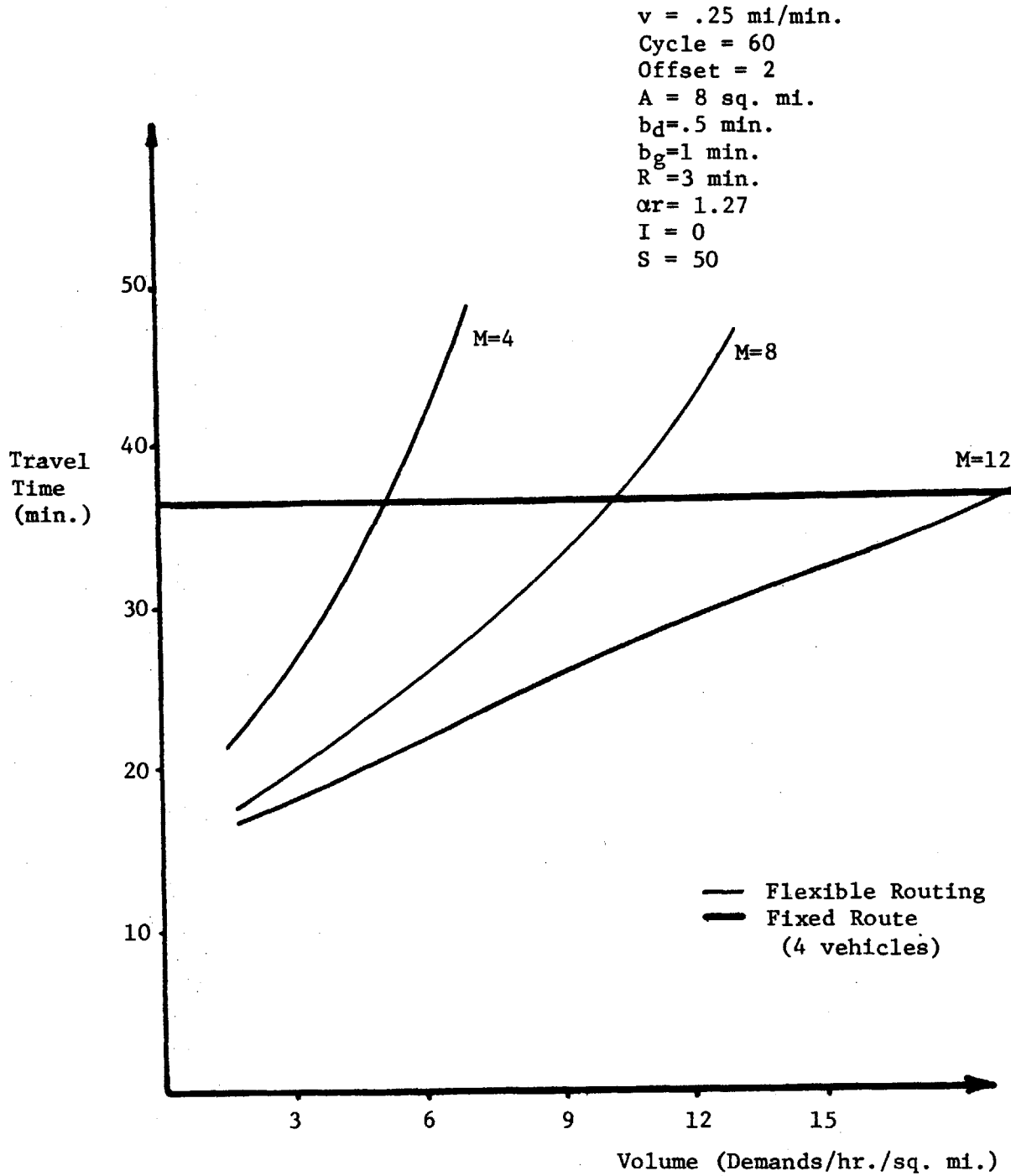━━ Fixed Route
    (4 vehicles)

Volume (Demands/hr./sq. mi.)

Fig. 6.5: Expected Travel Times of Fixed and Flexibly Routed Services

service time. Moreover, due to the possibility of being assigned to a following vehicle, the variance of travel time is larger for the flexibly routed feeder service than for the fixed route service. These modifications have offsetting effects so the demand for the two types of service at equal travel times should be comparable.

Of more concern are the costs of operating the services envisioned. With the 8 sq. mi. service area, the expected productivity of the buses with 6 demands/hr/sq.mi. is 1.3 patrons/revenue mile. This may be compared with the 1976 Boston area transit system average of 3.4 patrons/revenue mile [48]. The services examined here are relatively expensive due to relatively low demand densities. However, these densities are not exceptional for late night or weekend services.

6.5 A Comparison of Structured and Unstructured Flexibly-Routed Regional Transit.

As a final comparison of service options, we shall compare two types of area-wide, flexibly-routed transit services. One option consists of area-wide, many-to-many dial-a-ride service, in which vehicles travel everywhere in the region in response to patron demands. In this option, no patron is required to transfer between vehicles. The alternative design also employs flexibly-routed transit service, but divides the region into zones. Vehicles operate only within their assigned zones. In addition to the structure provided by zonal boundaries, this service will be scheduled, operated in phases of collection and distribution (as in Sections 3.3 to 3.6), and coordinated so that transfers may be rapidly accomplished at a central depot. The zonal model of feeder service

performance which was developed in Section 3.8 will be applied to this structured transit service.

Figure 6.4 summarizes the relevant characteristics of the region in which the transit service is to be provided. We assume that the region is square and has a Central Business District (CBD) located at its center. The service area size is 24 square miles, which is comparable to that of Ann Arbor, Michigan (described in Section 6.2). The street network is assumed to be a closely spaced rectangular grid. Patron origins and destination not in the CBD are assumed to be uniformly distributed throughout the region. An available vehicle fleet size of 12 is assumed.

An existing model has been used to estimate the performance of the many-to-many, regional dial-a-ride service [23]. This model yields estimates of wait and ride times in such systems from the equations:

$$T^W = \frac{r}{2v(1 - \frac{(b_g + b_d)\lambda}{M})} \sqrt{A/M} \, \exp(.22\sqrt{(A+4)/(M+12)} \, (\lambda/M)^{.9}) \qquad (6.14)$$

$$T^r = \frac{rF}{2v(1 - \frac{(b_g + b_d)\lambda}{M})} \, \exp(.084 \, (\frac{A\lambda}{M^2})^{.7}) \qquad (6.15)$$

where the variable not used previously, F, is the average direct trip length on the system and λ is in units of demands per hour.

These relationships were selected on the basis of best fit to simulation results and were calibrated using log-linear regression, producing root mean square errors of 10% or less. The simulation model used a particular computer control procedure, so we shall assume that the regional
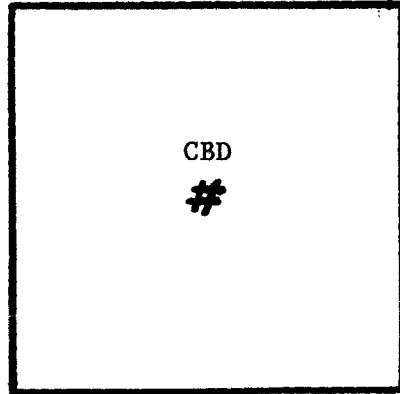
```
Area                      =  24 sq. mi.
Attainable Vehicle Speed  =  15 mph.
Boarding Time             =  1 min.
Exiting Time              =  .2 min.
Travel Factor             =  1.28
   (Rectangular Grid Street Network)
Fleet Size available      =  12 vehicles
```

Figure 6.4: Service Region Characteristics

dial-a-ride service has computer dispatching available.

Two structured transit services will be compared with the regional dial-a-ride service. The first divides the area into four zones and has three vehicles operating in each zone. The second option divides the region into six zones and has two vehicles operating in each zone. A depot for transferring patrons is assumed to be located in the Central Business District. Both configurations are operated as feeder service to this depot, with a rendezvous time of 15 minutes. Trips both originating and destined within the same zone, however, are served on the vehicle tour between visits to the depot. Equations summarizing the performance model for this zonal service appear in Section 3.9.5.

Figure 6.5 illustrates the expected service time with the unstructured and the two structured system alternatives for the case in which both patron origins and destinations are uniformly distributed throughout the region. Due to errors in the model predictions and inaccuracy in fitting the curves shown, one must be cautious in interpreting the results of this example*. However, it appears that the area-wide dial-a-ride service results in lower travel times throughout the range of demand

from 18 to 80 demands per hour, representing vehicle productivities of 1.5 to 7 per vehicle per hour. In this range, the dial-a-ride service has expected service times which are approximately 4 minutes less than the structured system with six zones. Above a demand level of approximately 24 demands per hour, the structured service with 6 zones is more desireable than the 4 zone option, as one would expect from the discussion of vehicle

---

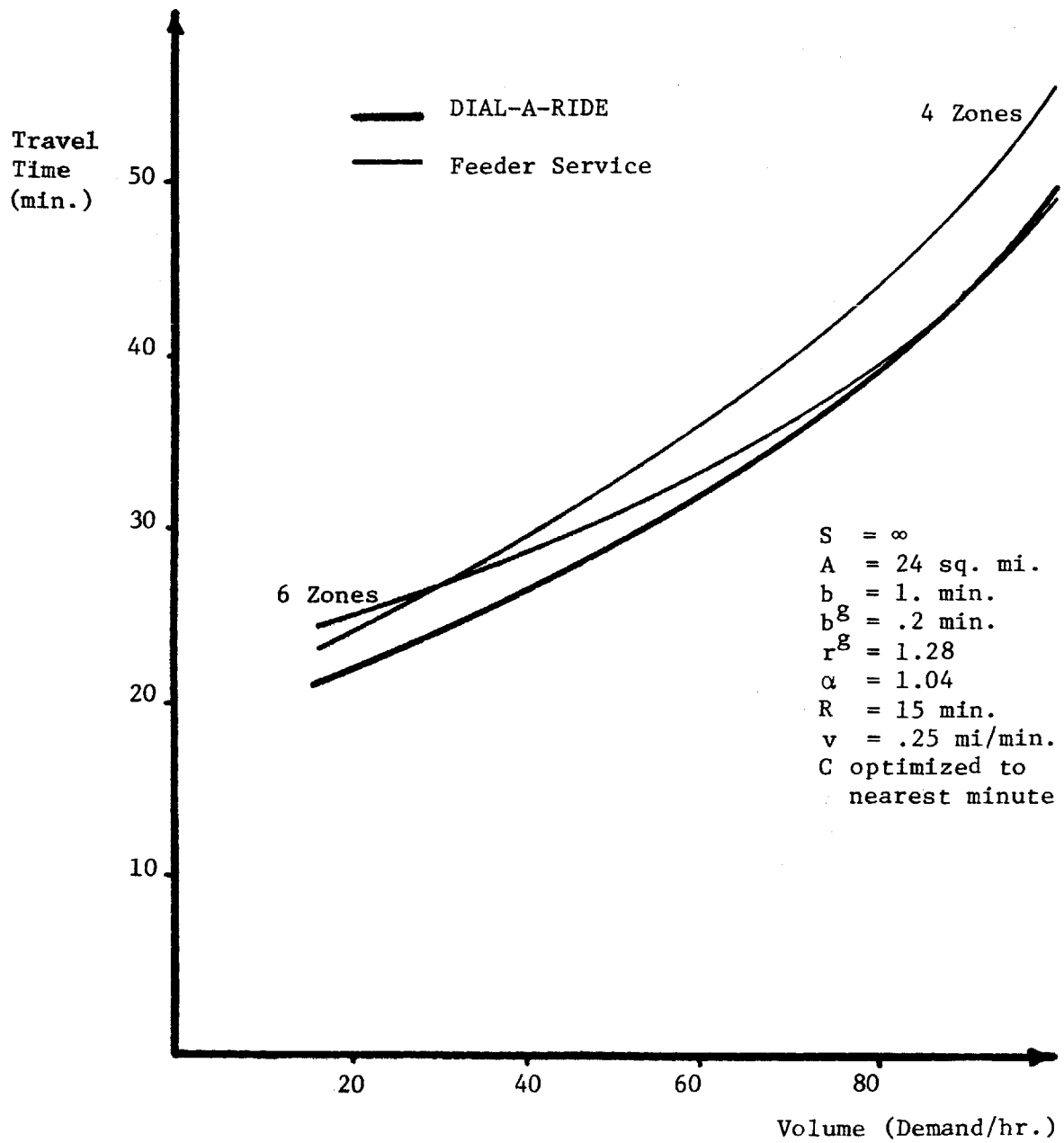*Curves were fit to approximately 10 observations by hand.

Figure 6.5: Expected Travel Times with Structured and Unstructured, Flexibly Routed Transit Systems (Patrons' Origins and Destinations Uniformly Distributed throughout Region).

deployment strategies in Section 5.4.

As the proportion of patrons travelling to the CBD increases, however, the structured transit system becomes more desirable. Figure 6.6 illustrates the situation in which 50 % of all trips either originate at or are destined for the CBD. In this case, the structured alternatives result in significantly lower travel times than does the dial-a-ride option, particularly as the system becomes more congested. Again, the six zone system is slightly more advantageous than the four zone system, so only the six zone system is illustrated in the graph.

In specific situations, it may be useful to investigate hybrid designs in which some vehicles provide regional dial-a-ride service while other vehicles operate as feeder services. However, investigation of such options is beyond the scope of this thesis. Also, even in situations in which the structured transit system results in lower expected travel times, the dial-a-ride system may be preferred because patrons do not have to make inter-vehicle transfers.

Travel
Time
(min.)

DIAL-A-RIDE

Feeder Service

50

40

30

20

6 Zones

10

A = 24 sq. mi.
b_g = 1. min.
b_d = .2 min.
r = 1.28
α = 1.04
R = 15 min.
r = .25 mi/min.
S = ∞
C optimized to
   nearest minute

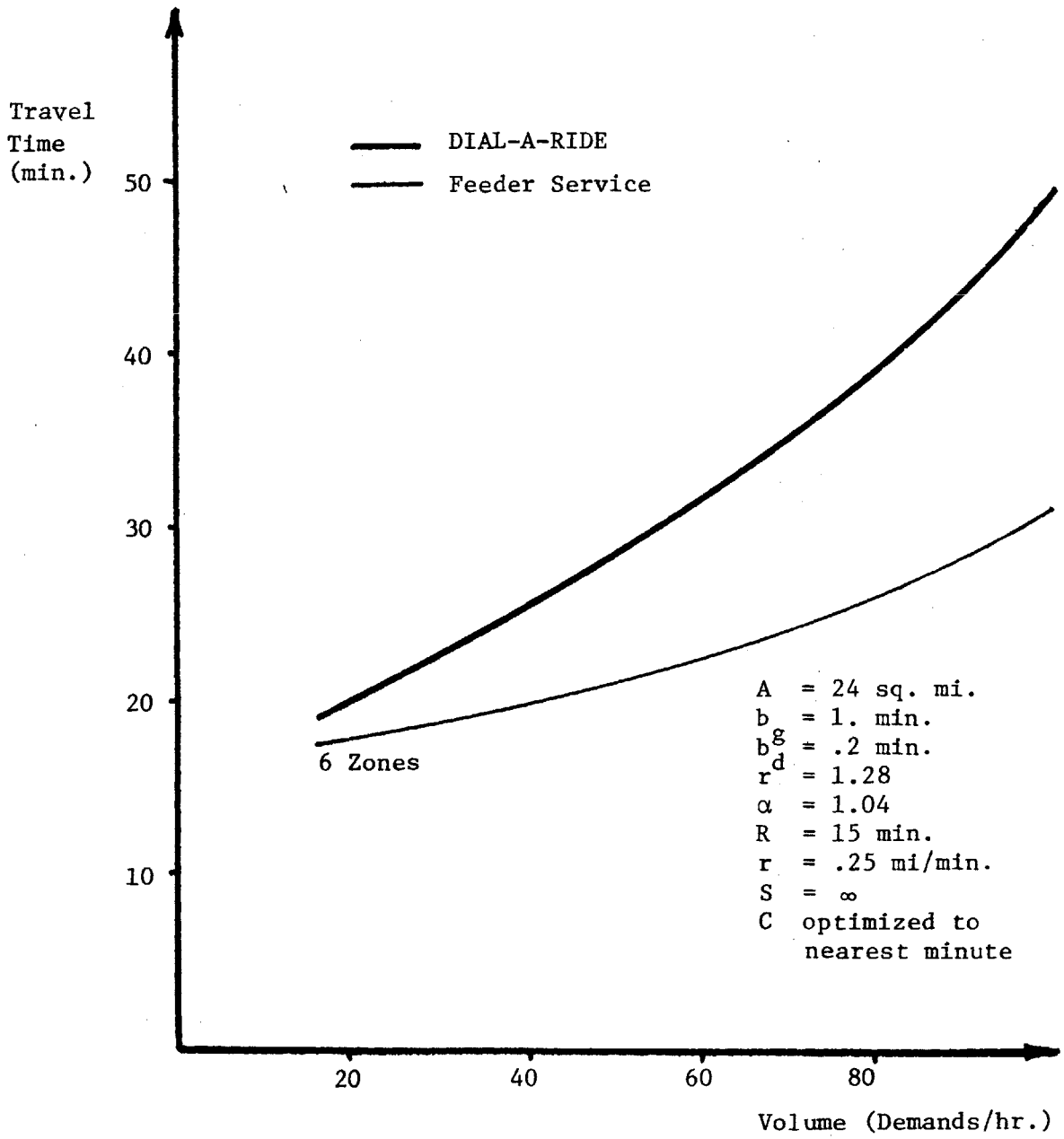20   40   60   80

Volume (Demands/hr.)

Figure 6.6:   Expected   Travel Time with Structured and Unstructured Flexibly Routed Transit Systems (50% Patrons' Origins or Destinations at the CBD).

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

7.1 Introduction

In this thesis, performance models of transit feeder services have been developed and applied. The model development has served to illustrate the use of a particular technique for modelling transit system performance, resulting in models which are approximate, analytic representations of system performance.

To summarize, the technique for constructing models used here has several distinct stages:

. develop an analytic expression for the time and resources required for underlying service processes (i.e. tour lengths or the time to traverse a transit route);

. construct a deterministic queueing system framework of the service, using fluid approximations to the arrival and service processes;

, superimpose correction terms on the deterministic model to account for stochastic phenomena and the integer or indivisible aspects of service; and

. derive measures of level of service and of resources consumed from the resulting model.

Application of this technique results in approximate, analytic performance models which are sufficiently simple to allow manual application.

The results of the modelling efforts are summarized elsewhere (see, in particular, Sec. 1.6). In this chapter, a few comments on the use of

-217-

the models and a short discussion of unresolved research areas are presented.

7.2 The Use of the Models Developed

Application of the performance models developed here may be performed either manually or by computer. Manual application is aided by a programmable electronic calculator in which the tour length expression (Eqs. 3.6, 3.23) and the mean of a truncated normal distribution (Eqs. 3.38, 3.39, 3.42) appear as sub-programs. An approximation to the cumulative normal distribution function is useful in such applications; one such approximation is mentioned in Sec. 3.5. Application of the models by computer is also relatively simple using the formulae summarized in Sec. 1.5. Even with a one dimensional search routine to find optimal cycle lengths, solution of the models is fairly rapid. As a result, the models could be incorportated in larger programs, such as programs to solve the equilibrium traffic assignment problem.

Supplemental models may also be added to the basic performance model to estimate other system impacts. For example, the vehicle miles of travel per cycle in a flexibly-routed, phased feeder service is simply the sum of tour lengths and travel to and from the depot:

$$\text{VMT}_c = v(R-L+D+G'-(n_g b_g + n_d b_d)) \tag{7.1}$$

and total vehicle miles of travel to serve the area per unit of time is then:

$$\text{VMT} = \text{VMT}_c \cdot M/C \tag{7.2}$$

-218-

With such an estimate of vehicle miles and hours of operation, it is possible to estimate fuel consumption, operating costs, air pollution and other impact factors. Similarly, benefit measures may be calculated from the expected level of service measures (see Williams [71]).

In applications, however, it should be emphasized that the models are only intended to be approximations to actual system performance. Derivation of the models involved a number of assumptions which are only expected to be approximately correct. The experiments with validation of the models (Sec. 5.2) and tests of the modelling assumptions (Chap. 4) indicate that the models are fairly accurate and robust. However, neither the model of fixed route transit service (Sec. 6.2) nor the model of unphased, flexibly routed feeder service (Sec. 3.7) have been validated. Moreover, a few comments about the types of applications in which models are not expected to be good approximations may be useful to users:

. Short tours

Simulation experiments indicate that the choice of a starting point is not critical with tours of 3 or more stops (Sec. 3.2). For very short tours, the choice of a starting point, reflected by the length of the rendezvous period, R, may be relatively important. Sec. 5.3 discusses a heuristic correction to improve the accuracy of the estimation of the rendezvous period. In addition, for tours of only one stop, the errors inherent in the fluid approximations become relatively large.

. congested systems with vehicle capacity limitations.

Fluid approximations to queueing phenomena are not as accurate in congested systems as in uncongested or in oversaturated systems

-219-

[56]. The stochastic corrections introduced in the models due to
schedule constraints are expected to be fairly good approximations
since the variability in the system is "damped" by the state de-
pendent nature of the service process (Sec. 5.3). However, the
corrections introduced to account for the constraint imposed by
vehicle capacity limitations (Eqs. 3.39, 2.42, and 6.9) can be
expected to underestimate the actual effects. Fortunately, sys-
tems are only rarely operated in situations in which vehicle ca-
pacity is regularly exceeded . In such cases, travel time increases
rapidly and demand for the transit service tends to fall. Exist-
ing flexibly routed services which operate near capacity typically
require reservations, so that the variance in the number served is
reduced and capacity is rarely exceeded.*

. dynamic control and optimum vehicle routing

Dynamic control of feeder services might be used to insert
patrons into the pool of eligible patrons during collection tours
or to specify a longer idle period, I, in response to the actual
number of patrons waiting for collection. Optimum vehicle routing
might attempt to maximize the value of some objective function [74].
Deneau [16] has shown that the effect of dynamic patron insertion
is relatively minor, and the experiments reported here show that
benefits of better vehicle routing are apt to be small (Sec. 4.3).
Variation in the idle period, I, simply exchanges wait time at
home for riding time in the vehicle (including wait time in the
vehicle at the depot). However, while the effects of these control

*The models developed here may be applied to advanced reservations systems
with appropriate modification of the variance of the arrival process.

-220-

factors may be small, they may be of great interest to operators. Unfortunately, these effects are not captured by the models in their present state of development.

. the order of stops in zonal service tours.

One of the options presented for serving intra-zonal trips in a service area was to both collect and distribute these patrons on a collection or delivery tour (Sec. 3.8). The expression for the tour length (Eq. 3.7) does not reflect the constraint that such patrons must be collected before being delivered. Since the number of intral-zonal trips was assumed to be relatively small, this constraint was not expected to have a major effect. In specific situations, it may be useful to insert heuristic correction terms.*

As a final note, the assumptions and the limitations of the models are viewed here in perspective with the accuracy of input data. It is difficult to make accurate estimates of demand, costs, vehicle speeds, boarding times, and other area characteristics in a planning situation. In the midst of this pervasive uncertainty, the inaccuracy resulting from the approximations and limitations of the performance models is unlikely to be the principal source of prediction uncertainty. Unfortunately, this offers little comfort to a planner who enjoys security.

7.3 Future Research

As in all research efforts, a number of areas for future research work exist at the conclusion of this study. A few of these areas are

---

*For example, one might estimate the tour length as the average of a single tour with both collections and deliveries and of two tours, first with collections and deliveries from the depot, and then with delivery of intra-zonal trips.

noted below:

- General comparison of flexibly or fixed route transit services
  and other operating options.

    Chapter 5 and 6 present some comparisons of alternative service

  options based solely upon differences in expected travel times.

  A more general comparison would consider differences in system

  costs, patron utility (including the disutility of transfers), and

  social benefit. Such a comparison could have important policy im-

  plications.

- Application of the models to particular transit services.

    More extensive application of the models would result in a body

  of engineering knowledge concerning both the models' and services'

  characteristics and would suggest areas of further development

  to the models.

- Use of the models in deterministic simulation.

    Rather than assuming that input parameters and options are con-

  stant and finding the steady state system performance (Sec. 3.3.),

  it is possible to use time dependent input parameters and determine

  the level-of-service and other performance measures of a service

  by means of a deterministic simulation. Such an analysis could be

  represented graphically in the same manner as the steady state si-

  tuation (Figs. 3.2, 3.3 and 3.5). The formulae for such a simula-

  tion model have not been developed here. Moreover, the accuracy

  of such a deterministic simulation is an interesting research ques-

  tion.

. Investigation of alternative transit network structures and service component mixes.

Analysis of alternative designs for integrated transit designs is in its infancy. The models developed in this thesis should be useful tools for such studies.

. Further application of the modelling technique.

The modelling techniques developed here could be useful in the analysis of urban service systems, package delivery services, port operations, and other situations. The technique is most useful in situations with relatively complicated service or arrival process in which either the service or arrival process is state-dependent, so that the variability of the number in the system is reduced. For example, application of the technique to taxi services is a natural extension of the models developed here.

References

1) ARILLAGA, B. and D.M. MEDVILLE (1974). "Demand, Supply and Cost Model-ling Framework for Demand Responsive Transportation Systems," Trans-portation Research Board Special Report 147, 32-47.

2) BARNETT, A. (1974). "On Controlling Randomness in Transit Operations," Transportation Science Vol. 8 pp. 102-111.

3) BATCHELDER, J., L.S. ENGLISHER, B.C. KULLMAN and K.L. SOBEL (1976). "Operational Implications of a Major Modal Diversion to Transit," U.S. Department of Transportation Report DOT-TST-76-72.

4) BEARDWOOD, J., J. HALTON, and J. HAMMERSLEY (1959), "The Shortest Path Through Many Points," Proc. Camb. Phil. Soc., vol. 59 p.299.

5) BECHTEL INC. (1975). "Evaluation of Systems Design for Arterial/Personal Transit Service." Final Report submitted to Santa Clara County Transit District.

6) BILLHEIMER, J.W., R. BULLEMER and M. HOLOSZYC (1976)." Macroanalysis of the implications of major modal shifts in integrated regional trans-portation networks." Final Report, Appendices: Phase I; U.S. Depart-ment of Transportation Report DOT-TST-76-66.

7) BOREL, E. (1924). Elements de la Theorie des Probabilities, Paris: Librairie Scientific (3rd Edition).

8) BREUER, M.W.K.A., W.E. HENDRIKSE and J.L. deKROES (1974). "The bus taxi, a study of different algorithms for alloting passengers to vehicles." Proceedings of 2nd IFAC/IFIP/IFORS Symposium on Traffic Control and Transportation Systems, 109-119.

9) CARLSON, R.C. (1976). "Anatomy of a Systems Failure: Dial-a-ride in Santa Clara, California." Transportation (5).

10) CHRISTOFIDES, N. (1976). "Worst Case Analysis of a New Heuristic for the Travelling Salesman Problem," Presented at the Symposium on New Directions and Recent Results in Algorithms and Complexity, Pittsburgh, Pa.

11) CLARENS, G.C. and V.F. HURDLE, (1975). "An Operating Strategy for a Commuter Bus System," Trans. Sci. 9(1).

12) CLAYTON, A.S.H. (1941) "Road Traffic Calculations" Journal of the In-stitute of Civil Engineers 16(7).

13) DAGANZO, C.F. (1976), "Unpublished Lecture Notes on Geometrical Pro-pability and Networks."

14) DACANZO, C.F., C.T. HENDRICKSON and N.H.M. WILSON (1977). "An Approximate Analytic Model of Many-to-one Demand Responsive Transportation Systems." Proceedings of the Seventh International Symposium on Transportation and Traffic Theory, Kyoto, Japan pp. 743-772.

15) DARROCH J.N. (1964) "On the Traffic Light Queue," Ann. Mathematical Statistics (35) pp. 380-388.

16) DENEAU, T.M. (1976) "An Analytic Model for a Many-to-One Dial-a-Ride System" unpublished BS thesis M.I.T.

17) DOMENCICH, T.A. and D. McFADDEN, (1975) Urban Travel Demand New York: American Elsevier Publishing Co.

18) EILON, S., C.D.T. WATSON-GANDY and N. CHRISTOFIDES (1971) "Distribution Management," New York: Hafner.

19) ENGLISHER, L.S. and SOBEL, K.L. (1978). "A Methodology for the Analysis of Local Paratransit Options," paper presented at the 56th Annual Transportation Research Board Meeting .

20) FAIRTHORNE, D. (1963) "The Distances Between Pairs of Points in Towns of Simple Goemetrical Shapes," Proc of the 2nd Intl. Sym. on the Theory of Traffic Flow, London, pp. 391-406.

21) FELLER, W. (1968) An Introduction to Probability Theory and Its Applications, Vols. I-II, New York: John Wiley & Sons, Inc.

22) FLORIAN, M. (1977) "Analysis of Transportation Supply in a Spatial Context," paper presented at Intl. Sym. on Travel Supply Models Montreal, Canada.

23) FLUSBERG, M., and N.H.M. WILSON (1976) "A Descriptive Supply Model for Demand Responsive Transportation System Planning," Proceedings of the Seventh Annual Meeting of the Transportation Forum 425-431.

24) Gakenheimer, R. (1976), Transportation Planning as Response to Controversy Cambridge: M.I.T. Press.

25) GERRARD, M.H. (1974) "Comparison of Taxi and Dial-A-Bus Services," Transportation Science 8(2).

26) GOLDEN, B. (1976) Large Scale Vehicle Routing and Related Combinatorial Problems unpublished Ph.D. thesis, M.I.T.

27) GORDON, S. and R. deNEUFVILLE (1973) "Design of Air Transportation Networks," Transportation Research 7(3).

28) GUENTHER, K.W., (1976), "Demand-Responsive Transportation in Ann Arbor," Trans. Res. Rec. #608.

29) HAIGHT, F.A. (1964), "Some Probability Distributions Associated with Commuter Travel in a Homogeneous Circular City," Oper. Res. 12, pp. 964-975.

30) HAUER, E. (1971) "Optimal Form of a Class of Collection-Distribution Networks" Fifth Intl. Sym. on Trans. and Traffic Flow.

31) HEATHINGTON, K.W., F.W. DAVIS JR., D.P. MIDDENFORF and J.D. BROGAN (1974). "Demand-Responsive Transportation Systems in the Private Sector," Transportation Research Record 522, 47-55.

32) HENDERSON, J.M. and R.E. QUANDT (1971) Microeconomic Theory New York: McGraw Hill Book Co.

33) HURDLE, V.F. (1973) "Minimum Cost Schedules for a Public Transportation Route," Transportation Science 7(2).

34) JOLLIFFE, J.K. and T.P. HUTCHINSON (1975). "A Behavioral Explanation of the Association between Bus and Passenger Arrivals at a Stop," Transportation Science 9(3) pp. 248-282.

35) KARP, R.M. (1977), "Probabilistic Analysis of Partitioning Algorithms for the Travelling Salesman Problem in a Plane," Math. of Oper. Res. 2(5) pp. 209.

36) KENDALL, D.G. and P.S.P. MORAN (1963) Geometrical Probability London: Griffin and Co.

37) KIRBY, R., K.U. BHATT, M.A KEMP, R.G. McGILLIRRAY, and M. WOHL (1974) Para-Transit Washington D.C.: The Urban Institute.

38) LAMPKIN W. and P.D. SAALMANS (1967) "The Design of Routes, Service Frequencies and Schedules for a Municipal Bus Undertaking: a Case Study," Operational Research Quarterly 18(4).

39) LARSON, R.C. (1972), Urban Police Patrol Analysis, Cambridge: M.I.T. Press.

40) LARSON, R.C. and A.R. ODONI (1979). Analysis of Urban Service Systems. Boston: M.I.T. Press. (Forthcoming).

41) LERMAN, S.R., M. FLUSBERG, W. PECKNOLD and N.H.M. WILSON (1977). "A Method for Estimating Patronage of Demand Responsive Transportation Systems." U.S. Department of Transportation Report DOT-TSC-977.

42) LERMAN, S.R. and N.H.M. WILSON (1974). "Analytic Model for Predicting Dial-A-Ride System Performance." Transportation Research Board Special Report 147, 48-53.

43) LITTLE, J.D.C. (1961) "A Proof of the Queueing Formula L=λW," Operations Research, 9(3) pp. 383-387.

44) MANHEIM, M., (1976), "The Emerging Planning Process: Neither Long-Range Nor Short-Range but Adaptive and (Hopefully) Decisive," paper presented at 55th TRB meeting.

45) MANHEIM, M., P. FURTH. and I. SALOMON (1978). "Responsive Transportation Analysis: Evaluation of TSM and Other Strategies Using Pocket Calculators," presented at 56th TRB Meeting.

46) MANSKI, C.J. and J.D. WRIGHT (1976) "Nature of Equilibrium in the Market for Taxi Services," Transportation Research Record 619.

47) MASON, F.J. and J.R. MUMFORD (1972). "Computer Models for Designing Dial-A-Ride Systems." Society of Automotive Engineers, Automotive Engineering Congress, 720216.

48) MASSACHUSETTS BAY TRANSPORTATION AUTHORITY (1976), "Twelth Annual Report," Boston, Ma.

49) McLEOD, M.G. (1972). "The Operation and Performance of a Taxi Fleet." Civil Engineer Thesis, Department of Civil Engineering, Massachusetts Institute of Technology.

50) MENHARD, H.R., M. FLUSBERG, L.S. ENGLISHER and S.R. LERMAN (1977), "Modelling Demand-Responsive Feeder Systems in the UTPS Framework - Users Manual," U.S. Dept. of Trans.

51) MEREWITZ, L. (1977). "On Measuring the Efficiency of Public Enterprises: Bus Operating Companies in the SF Bay Area," Transp. 6(1)

52) MEYER, M. and N.H.M. WILSON (1976). "The Use of Simulation in the Design of a Dial-A-Ride System." Proceedings of 1976 Summer Computer Simulation Conference.

53) MITCHELL, D.J. (1972). "Some Aspects of Demand Actuated Public Transport Services." Doctoral Dissertation, Faculty of Science and Engineering, University of Birmingham (U.K.), Publication 43.

54) MORLOK, E.K. (1977), "Supply Functions for Public Transport," paper presented at the Intl. Sym. on Trans. Supply Models, Montreal Canada.

55) NATH, S. (1969) A Reappraisal of Welfare Economics. London. Routledge and Regan Paul.

56) NEWELL, G.F. (1971), Applications of Queuing Theory. London: Chapman and Hall Ltd.

57) NEWELL, G.F. and POTTS, R.B. (1964). "Maintaining a Bus Schedule" Proc. Second Conf. of the Australian Road Research Board, Melbourne, pp.388-393.

58) NEWELL, G.F. "Dispatching Policies for a Transportation Route," _Trans._ _Science_ 5(1)

59) NEWELL, G.F. (1974) "Unstable Bownian Motion of a Bus Trip," _Proc. of_ _a Symp. on Statistical Mechanics and Statistical Methods in Theory_ _and Application_.

60) NEWMAN, L.A., J.A. WOJNO, and R.D. JUSTER (1977), "Integrated Dial-A-Ride and Fixed Route Transit in Ann Arbor Michigan," U.S. Dept. of Transportation Report #UMTA-MA-06-1083-77-1

61) OSUNA, E.E. and G.F. NEWELL (1972), "Control Strategies for an Idealized Public Transportation System," _Trans. Sci._ (6) pp.52-72.

62) POTTS, R.B. and E.A. TAMLIN (1964) "Pairing of Buses," _Australian Road_ _Research Board_.

63) ROSENKRANTZ, D., R. STEARNS, and P. LEWIS (1974), "Approximate Algorithms for the Travelling Salesperson Problem," _Proc. of the 15th IEEE Symp._ _on Switching and Automata Theory_, pp 33-42.

64) SCHEELE, S. (1977), "A Supply Model for Public Transit Services," paper presented at the _International Symposium on Traffic Supply Models_, University of Montreal.

65) SEN, ASHISH and C. JOHNSON (1977), "On the Form of a Bus Service Frequency Supply Function," _Trans. Res._ 11(1).

66) SIRBU, M.A. (1968). "Waiting Times in a Class of Bulk Queues," Masters Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology.

67) STEIN, D.M. (1977). "Scheduling Dial-A-Ride Transportation Systems," Doctoral Dissertation, Harvard University.

68. U.S. DEPARTMENT OF TRANSPORTATION, URBAN MASS TRANSPORTATION ADMINISTRATION (1972), _UMTA Transportation Planning system (UTPS) Reference_ _Manual_.

69) WARD, D.E. (1975), "A Theoretical Comparison of Fixed Route Bus and Flexibly Routed Subscription Bus Feeder Service in Low Density Areas" NTIS #PB-240-808 U.S. Dept. of Trans, Trans. Sys. Ctr.

70) WEBSTER, F.V. (1958) _Traffic Signal Settings_, Road Research Technical Paper #39, London H.M. Stationary Office.

71) WILLIAMS, H.C.W.L. (1977), "On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit," _Environment and_ _Planning A_, Vol. 9, pp. 285-344.

72) WILSON, N.H.M. and D. ROOS (1967). "An Overview of CARS." Papers of Eighth Annual Meeting Transportation Forum, 129-140.

73) WILSON, N.H.M. J.M. SUSSMAN, B.T. HIGONNET and L.A. GOODMAN (1969). "Simulation of a Computer Aided Routing System (CARS)," Highway Research Record 318, 66-76.

74) WILSON, N.H.M., J.M. SUSSMAN, H. WONG and B.T. HIGONNET (1971), "Scheduling Algorithms for a Dial-A-Ride System," Massachusetts Institute of Technology, Urban Systems Laboratory Report USL TR-70-13.

75) WILSON, N.H.M. (1975), "Research and Development in Demand-Responsive Transportation," Transportation Research Board Report 154, 104-109.

76) WILSON, N.H.M. and C. HENDRICKSON, "Performance Models of Flexibly Routed Transportation Services," paper presented at the Intl. Sym. on Trans. Supply, Montreal, Canada.

77) WILSON, N.H.M. (1977), "Short Range Transit Planning: Current Procedures and a New Approach," working paper.

78) WIRASINGHE, S.C. (1977), "Assignment of Buses in a Coordinated Rail and Bus Transit System," Proc. of 7th Intl. Sym. on Trans. and Traffic Flow Theory, Kyoto, Japan.

79) WIRASINGHE, S.C. (1976), "A Model of a Coordinated Rail and Bus Transit System," Ph.D. Dissertation, Institute of Trans. and Traffic Engr., Univ. of Ca., Berkeley.