70-14,535

HALL, William King, 1943-
A QUEUEING THEORETIC APPROACH TO THE ALLOCATION AND DISTRIBUTION OF AMBULANCES IN AN URBAN AREA.

The University of Michigan, Ph.D., 1969
Business Administration

University Microfilms, Inc., Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

A QUEUEING THEORETIC APPROACH TO THE ALLOCATION AND DISTRIBUTION OF AMBULANCES IN AN URBAN AREA

by

William King Hall

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in
The University of Michigan
1969

Doctoral Committee

Professor W. Allen Spivey, Chairman
Professor Ralph L. Disney
Assistant Professor Marvin J. Karson
Associate Professor Thomas J. Schriber

ABSTRACT

A QUEUEING THEORETIC APPROACH TO THE
ALLOCATION AND DISTRIBUTION OF
AMBULANCES IN AN URBAN AREA

by William King Hall

Chairman: W. Allen Spivey

The problem of ambulance allocation and distribution in an urban area is considered in this research. A mathematical model is developed for the urban emergency medical recovery system, and this model is then utilized to analyze the effects of alternative recovery vehicle distribution and allocation policies on the system response time. Analytical and numerical results are presented which provide guidance in determining the appropriate number and placement of ambulances for an urban area.

Data on emergency occurrence and service processes were collected from four police precincts in the City of Detroit, Michigan. Statistical analyses are conducted on these data to identify probabilistic models for the emergency occurrence and service processes and to estimate the parameters of these models. Both medical

emergencies and police emergencies are analyzed, since many urban ambulance systems respond to both types of emergencies.

These statistical analyses lead to the formulation of the recovery system model as a system of queues in parallel. The resulting model incorporates a Markovian channel selection switch which is conditioned upon the joint state of all queues in the system and upon the type of emergency requiring service. Theoretical results are developed for this queueing system with a semi-Markov emergency occurrence process and a negative exponential emergency service process.

The recovery system model is used to study systems where one, two, three, or four ambulances are assigned to a police precinct. Three distribution policies are examined: placing the recovery vehicles in the regions of maximum demand, placing the recovery vehicles "uniformly" throughout the precinct, and placing all recovery vehicles at a single station within the precinct. Numerical results are developed for both single function ambulance systems and dual-function police ambulance systems.

It is found that the probability that one or more vehicles are available at the instant of a medical emergency varies from .78 for a single vehicle system to

approximately 1.0 for a four vehicle system. The vehicle availability declines as police calls are accepted for service, although a "protective" dispatch policy can be utilized to control this decline.

Placing the recovery vehicles in the regions of maximum demand generally results in a higher probability of serving a nearby medical emergency than the other two distribution policies. However, the improvement over the "uniform" distribution policy is quite small.

The results of this research are directly applicable to ambulance service in the City of Detroit. In addition, the methodology can be applied by other urban areas interested in examining potential improvements in the allocation and distribution of recovery vehicles.

ACKNOWLEDGEMENTS

A portion of this research was supported by The University of Michigan Highway Safety Research Institute as part of a contract on emergency medical care sponsored by the National Highway Safety Bureau and the City of Detroit, Michigan. I would like to thank Mr. James O'Day, Head of the HSRI Systems Analysis Group, for his continuing support over the past several years. Special thanks are due to Mr. Lyle Filkins, HSRI Research Engineer, for his counsel and guidance during our professional association.

The members of my doctoral committee have all offered valuable advice throughout this research. I wish to express my sincere gratitude to Professor W. Allen Spivey for his active supervision in the preparation of this dissertation and for his advice and inspiration throughout my graduate education. Professor Ralph Disney spent many hours discussing the development of the mathematical model underlying this research, and his help has been invaluable. Mrs. Joanne Porter typed this manuscript, and her assistance is deeply appreciated.

Finally, I wish to thank my mother Jeanne, who provided me with the opportunity to start my college education, and my wife Valerie, who provided me with the incentive and encouragement to complete this research.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		ii
LIST OF TABLES		v
LIST OF FIGURES		ix
LIST OF APPENDICES		xii
CHAPTER		
I	INTRODUCTION	1
II	DATA ANALYSIS OF EMERGENCY OCCURRENCE AND SERVICE PROCESSES	25
III	DEVELOPMENT OF THE RECOVERY SYSTEM MODEL	95
IV	ALLOCATION AND DISTRIBUTION POLICIES FOR THE SINGLE FUNCTION PUBLIC AMBULANCE SYSTEM	126
V	ALLOCATION AND DISTRIBUTION POLICIES FOR THE DUAL FUNCTION PUBLIC AMBULANCE SYSTEM	163
VI	COMPARISON OF ALTERNATIVE RECOVERY SYSTEMS	197
VII	SUMMARY AND CONCLUSIONS	216
APPENDICES		223
REFERENCES		260

LIST OF TABLES

Table		Page
2.1	Estimated Autocorrelations of Times between Medical Emergencies	42
2.2	Tests of Renewal Hypothesis Based on Sample Periodogram	48
2.3	Statistics for Marginal Distributions of Interarrival Times between Ambulance Calls	49
2.4	Statistics Computed to Test for a Negative Exponential Distribution of Times between Ambulance Calls	53
2.5	Tests for Negative Exponential Distribution of Times between Ambulance Calls Based upon Partial Sums of Order Statistics	54
2.6	Comparison of Estimated and Sample Coefficients of Variation and Kurtosis--Gamma Distribution of Times between Ambulance Calls	56
2.7	Comparison of Estimated Coefficients of Skewness and Kurtosis for Gamma and Log-Normal Hypotheses--Time between Ambulance Calls	57
2.8	Maximum Likelihood Estimates for Parameters of Gamma Model of Interarrival Times between Ambulance Calls	59
2.9	Significance Test of Gamma Distribution of Times between Ambulance Calls Versus Log-Normal Alternative	61
2.10	Contingency Table to Test for Day-Week Interactions--Police Call Occurrence Process	62
2.11	Tests for Trends in Daily and Within Day Police Call Processes	69
2.12	Tests for Renewal Hypothesis for Times Between Police Calls	71

Table		Page
2.13	Tests for Negative Exponential Distribution of Times between Police Calls	72
2.14	Sample Statistics for Times between Police Calls	73
2.15	Frequency Table for Transition Matrix Between Types of Calls	74
2.16	Tests for Negative Exponential Distribution of Time Between Events in Semi-Markov Occurrence Process	79
2.17	Maximum Likelihood Estimates for Mean Occurrence Times in Semi-Markov Arrival Process	81
2.18	Statistics for Empirical Conditional Distributions of Ambulance Service Times	87
2.19	Tests for Negative Exponential Distribution of Ambulance Service Times	88
2.20	Tests for Negative Exponential Distribution of Police Service Times	90
2.21	Sample Statistics for Police Call Service Times	91
4.1	Single Vehicle Equilibrium Probabilities for Various Vehicle Allocations	139
4.2	Sensitivity of Equilibrium Probabilities to Variations in Emergency Occurrence Parameters	143
4.3	Sensitivity of Equilibrium Probabilities to Variations in Emergency Service Parameters	148
4.4	Equilibrium State Probabilities for Two Ambulance Systems	152
4.5	Equilibrium Probabilistic Results for the Two Ambulance System	152

Table		Page
4.6	Equilibrium Probabilities for the Three Ambulance System	159
4.7	Equilibrium Probabilistic Results for the Three Ambulance System	160
5.1	Transition Matrix for Single Vehicle Dual Function System	169
5.2	Equilibrium Vehicle State Probabilities for Single Vehicle Dual Function System	176
5.3	Switching Probabilities for Two Vehicle Dual Function System	179
5.4	Equilibrium Vehicle State Probabilities At Instant of Medical Emergency for Two Vehicle Dual Function System--Distribution Policy I	181
5.5	Equilibrium Vehicle State Probabilities At Instant of Medical Emergency for Two Vehicle Dual Function System--Distribution Policy II	182
5.6	Equilibrium Results for the Two Vehicle Dual Function System--Distribution Policy I	183
5.7	Equilibrium Results for the Two Vehicle Dual Function System--Distribution Policy II	184
5.8	Equilibrium Results for a Four Vehicle Dual Function System	195
6.1	Probability All Recovery Vehicles Busy as Function of Allocation Policy	199
6.2	Vehicle Availability as Function of Distribution Policy (r = 1)	201
6.3	Vehicle Availability as Function of Distribution Policy (r = 2; r = 3)	202
6.4	Effects of Vehicle Allocation and Distribution on Service Times	207
6.5	Effect of Allocation and Distribution Policy on Conditional Probability that Each Vehicle is on Emergency Requiring Less than One Mile Transit Distance	209

Table		Page
7.1	Summary of Statistical Analyses	218
7.2	Summary of Major Allocation, Distribution, and Dispatch Effects	220

LIST OF FIGURES

Figure		Page
1.1	Some Factors Influencing Victim Morbidity/Mortality	5
1.2	Sample Realization of the Emergency Recovery Process	7
2.1	City of Detroit Police Precincts Sampled for Allocation and Distribution Study	30
2.2	Number of Ambulance Calls per Day (November 19, 1968 - December 15, 1968) . . .	36
2.3	Number of Ambulance Calls per Shift (December 1, 1968 - December 14, 1968) . . .	38
2.4	Empirical Distribution Function - Times between Ambulance Calls - Precinct 7	50
2.5	Empirical Distribution Function - Times between Ambulance Calls - Precinct 15	51
2.6	Number of Police Calls per Day - Precinct 15 (November 19, 1968 - December 14, 1968) .	63
2.7	Number of Police Calls per Shift - Precinct 15 (November 19, 1968 - December 14, 1968) .	66
2.8	Realization of Semi-Markov Emergency Occurrence Process	76
3.1	A System of Queues in Parallel	98
3.2	A Sample Service Area and Ambulance Distribution	116
4.1	Estimated Probabilities that a Medical Emergency will Occur in any Scout Car Territory - Precinct 15	137

Figure		Page
4.2	Equilibrium Idle Probability as Function of Mean Time between Emergency Occurrences--Single Ambulance System (δ = .139)	141
4.3	Equilibrium Probability x_2 as Function of Mean Time between Emergency Occurrences--Single Ambulance System	142
4.4	Equilibrium Idle Probability as Function of Mean Emergency Service Times--Single Ambulance System (δ = .139)	146
4.5	Equilibrium Probability x_2 as Function of Mean Emergency Service Times--Single Ambulance System	147
4.6	Two Ambulance Distribution Policy to Achieve Uniform Geographic Coverage	151
4.7	Three Ambulance Distribution Policy to Achieve Coverage of Area of Maximum Demand	156
4.8	Three Ambulance Distribution Policy to Achieve Uniform Coverage of Service Area	157
5.1	Equilibrium Idle Probability for Single Vehicle/Dual Function System	172
5.2	Equilibrium Police Call Service Probability for Single Vehicle/Dual Function System	173
5.3	Equilibrium "Short" Medical Emergency Service Probability for Single Vehicle/Dual Function System	174
5.4	Equilibrium "Long" Medical Emergency Service Probability for Single Vehicle/Dual Function System	175
5.5	Equilibrium Probability that One or More Vehicles are Idle in Two Vehicle/Dual Function System	185
5.6	Equilibrium Probability that One or More Vehicles are Serving Police Calls in Two Vehicle/Dual Function System	186

Figure		Page
5.7	Equilibrium Probability that One or More Vehicles are Serving "Short" Medical Emergencies in Two Vehicle/Dual Function System	187
5.8	Equilibrium Probability that One or More Vehicles are Serving "Long" Medical Emergencies in Two-Vehicle/Dual Function System	188
5.9	Four Vehicle Distribution Policy	192
6.1	Vehicle Availability as Function of Allocation and Dispatch Policy	204
6.2	Vehicle Utilization as Function of Allocation, Distribution, and Dispatch Policy	212
6.3	Conditional Probability of Serving a "Short" Emergency as Function of Dispatch Policy	214
C.1	Computer Program Logic	239

LIST OF APPENDICES

Appendix		Page
A	STATISTICAL TESTS FOR THE POISSON PROCESS	223
B	LAPLACE-STIETJES TRANSFORMS OF DECOMPOSED EMERGENCY OCCURRENCE PROCESSES	229
C	COMPUTER PROGRAM DOCUMENTATION	234

CHAPTER I

INTRODUCTION

1.1 DESCRIPTION OF THE PROBLEM AND RESEARCH METHODOLOGY

The development of effective methods for transporting the victims of emergencies to a hospital is a problem of much concern to modern society. Each year 100,000 Americans die of accidental injuries and 400,000 more suffer permanent disability (Public Health Service [28]). Thousands more are afflicted by sudden acute illness requiring immediate medical attention.

While the prevention of trauma and other medical emergencies is a desirable national goal, it is reasonable to assume that these emergencies will not be eliminated from our society in the near future. Consequently, the development of rapid, effective recovery and treatment procedures is critical if we are to reduce the morbidity (suffering) and mortality caused by such emergencies.

The improvement of ambulance service offers one such opportunity. Although the reduction in suffering which can be attributed to such an improvement is uncertain, one surgeon estimates that 25% of the 166,000 annual traffic accident victims now suffering permanent disability would have been saved from this fate had appropriate care

been rapidly provided (Ross and Rogers [29]). Other studies have estimated that 20% of the total traffic accident fatalities (10,000/per year) could be avoided with more adequate emergency medical care (Bordner et. al. [6]). In a recent retrospective study of fatal traffic accident victims in Washtenaw County, Michigan, it was concluded that 18% of these cases were "potentially salvagable" with more adequate emergency medical care (Frey [16]).

Unfortunately, ambulance service in this country has suffered from a severe lack of planning and coordination. Funeral homes have traditionally provided such service. However, in recent years many of these agencies have found that their ambulance services are unprofitable, and they have subsequently dropped these services. Thus, in some areas the ambulance service has been absorbed by the police department, the fire department, or some hospital-related agency. In others, private ambulance services have been developed, and in some communities, neighborhood volunteer groups have been formed to provide emergency medical recovery services. This unplanned evolution is so bad that a doctor of the American Academy of Orthopedic Surgeons estimates that two-thirds of the 300,000 ambulance crews are not "truly and totally competent" (Ross and Rogers [29]).

The situation in our urban areas is exceptionally acute. Here the large proportion of persons unable to pay for ambulance service has made private and funeral home services especially unprofitable. The fact that thirteen out of our fifteen largest cities now operate some sort of a public ambulance system attests to this situation (Munro [24]).

The higher population density and the unique patterns of injury in urban areas--traffic and industrial accidents, criminal injuries, etc.--create a strong demand for research into public ambulance system improvements. Indeed, the centralized control over such a public service offers an ideal opportunity for planned, systematic improvement.

In this dissertation one such improvement will be explored--the effect of better distribution and allocation of recovery vehicles. Utilizing terminology developed by McLaren (21), the allocation of ambulances will be defined as the number of vehicles assigned to the recovery system. The distribution of these vehicles will be defined as their geographical placement throughout the service area. Changes in either of these variables should immediately affect system response time--the time between the reporting of a medical emergency and the arrival of an ambulance on the scene. The potential for

placing the victim into contact with appropriate medical care faster is thereby increased, and the degree of victim morbidity/mortality is hopefully reduced. This situation is shown in Figure 1.1.

While the effects of alternative distribution and allocation policies can ideally be evaluated in terms of changed morbidity/mortality indices, in practice this is extremely difficult. This is true because victim morbidity/mortality is affected by many other factors. Figure 1.1 illustrates several of these--ambulance attendant training, the victim's pre-emergency physical condition, the severity of the emergency, and the quality of medical treatment once the patient is delivered to the hospital emergency room. If the effects of these factors cannot be controlled in an experiment or removed from the data retrospectively, changes in the system performance criteria cannot be solely attributed to improved ambulance distribution and/or allocation. Furthermore, the performance criteria themselves--reduced victim morbidity/mortality--suffer from severe definitional and measurement problems. For instance, victim morbidity might be measured by injury severity, cost of the emergency to the patient or society, days spent in the hospital, or some combination of these and other factors. Mortality might be measured by cause of death, time until

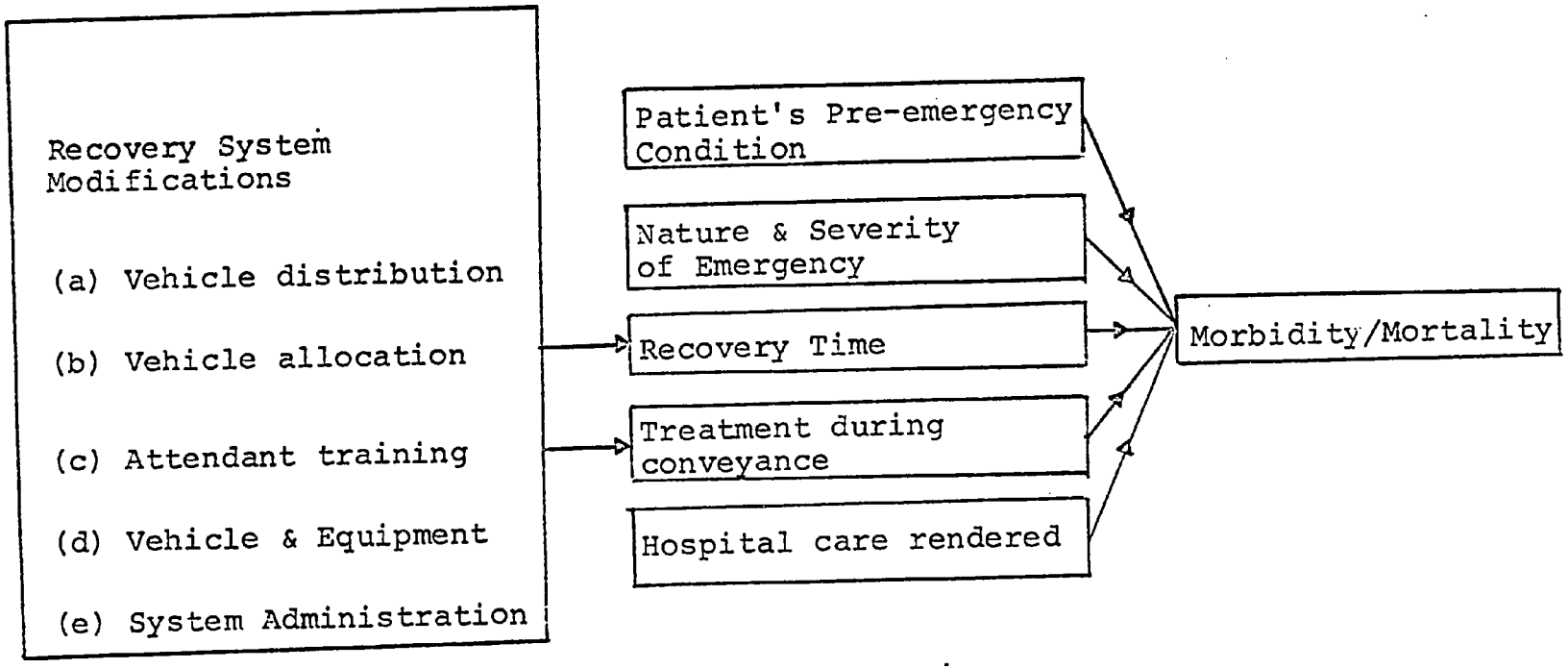


Figure 1.1

Some Factors Influencing Patient Morbidity/Mortality

death, the cost to society due to lost productivity, or some combination of these and other indices.

Unfortunately, neither of these problems has been adequately resolved. Therefore, in this dissertation the effects of alternative distribution and allocation policies will be measured by changes related to the system response time. Relating these changes to quantitative reductions in victim morbidity and mortality is an unresolved research problem which will not be considered here. There is little question, however, that such relationships exist in many cases. Military experience in the rapid evacuation of combat victims attests to this fact.

At the same time, one must remember that reduced response time is not essential in order to improve or maintain the victim's condition for a wide variety of minor injuries and trauma. Unfortunately, inaccurate diagnosis by the layman and overreaction on the part of the person who calls for assistance make the determination of such situations exceptionally difficult until qualified medical personnel examine the victim. Since the recovery system has almost no ability to detect these cases before vehicle dispatch, rapid response to all medical emergencies is a valid intermediate objective.

It is useful to study a sample realization of the emergency recovery process in order to precisely examine the effects of improved vehicle allocation and distribution. Such a sample realization is presented as Figure 1.2.

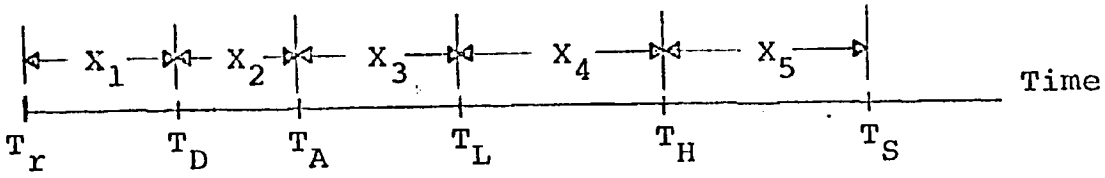


Figure 1.2

A Sample Realization of the Emergency Recovery Process

The various events of interest can be conceptualized as realizations of point stochastic processes occurring over time. The alphabetically subscripted symbol "T" will be utilized to denote random times of occurrence, and the numerically subscripted symbol "X" will be utilized to denote the resulting random time intervals.

The following definitions are introduced:

T_r = Time at which the medical emergency is reported to the recovery system,

T_D = Time at which a recovery vehicle is dispatched to the scene of the emergency,

T_A = Time of ambulance arrival on-the-scene,

T_L = Time of ambulance departure from the scene,

T_H = Time of ambulance arrival at the hospital,

T_S = Time at which ambulance is released from the hospital and returns to service,

X_1 = Time between reporting and dispatch,

X_2 = Transit time to the scene,

X_3 = Time on the scene,

X_4 = Transit time to the hospital,

X_5 = Time out of service at the hospital.

One measure of system performance which can be improved through better distribution and allocation of ambulances is the system response time. This variable, denoted by S, is the sum of the time between reporting and dispatch and the transit time to the scene. That is:

$$S = X_1 + X_2$$

or alternatively, in terms of times of occurrence:

$$S = T_A - T_R$$

It is important to observe that the times of occurrence of events and the time intervals between events should actually be conceptualized as stochastic processes. Consequently the effects of recovery vehicle allocation and distribution policies on the response time S must be measured by their effects on the distribution function of S or by their effects on specified parameters related to this distribution function.

At this time we shall briefly discuss the qualitative effects of vehicle distribution and allocation on the system response time. A quantitative discussion of these effects is presented in Chapter III. We observe that by increasing the number of vehicles in the system (improving the vehicle allocation) the probability that no vehicle is available to respond to an emergency is reduced. In this case the distribution of the random variable X_1-- the time between emergency reporting and vehicle dispatch--is modified so that there is a smaller probability of a longer time until dispatch. Improvements in vehicle distribution have a dual effect. First, improved distribution necessarily means that vehicles are placed closer to potential emergency sites, and it is very likely that the average transit time to the scene is reduced as the travel distance between ambulances and emergencies is decreased. Second, such a reduction in transit time to the scene results in a reduction in the total time the ambulance is out of service. Consequently, it is likely that ambulances will be available a larger portion of the time, thereby reducing the average time between reporting and dispatch.

It is clear that the time interval between reporting and dispatch can be essentially eliminated if the system

is provided with enough ambulances. Furthermore, the transit time to the scene can be reduced to a very small interval if a geographic area is saturated with recovery vehicles. However, economic considerations preclude such policies. In the first place such a system would require many vehicles and would no doubt violate an urban area's budget constraint on public ambulance service. Second, as the number of vehicles is increased, beyond some point the average utilization of these vehicles rapidly declines. It is unreasonable to assume that an urban area would rationally invest in resources which are not being utilized in an efficient manner.

An experimental evaluation of alternative distribution and allocation policies, conducted by making controlled modifications on the operating recovery system, is an appealing method of analysis in this situation. However, the large number of feasible modifications and the disruption in system operation introduced by these modifications makes experimental analysis impossible in this situation.

Consequently, it was decided to model the system and to predict the system response to the above changes by analyzing the resulting model. Two potential types of models can be developed to study the alternatives under consideration in this study. First, one could

develop an optimum-seeking procedure which leads to the selection of an "optimal" allocation and distribution policy for ambulances in an urban area. Alternatively, one could develop a model which "evaluates" the alternative policies under consideration by predicting their effects on the chosen measures of system performance.

The latter procedure has been chosen in this study, since substantial practical and theoretical problems arise in developing an optimum-seeking procedure for this application. For instance, the choice of an appropriate optimization criterion raises several unresolved questions. For example, should one minimize some parameter of the random response time within the cost constraints, or should one minimize cost subject to some requirements on the minimal acceptable response time?

Even when an optimization criterion is available, there are substantial theoretical problems in optimizing the recovery vehicle allocation and distribution in an urban area. In the first place, the system being considered involves several stochastic elements, and the optimum-seeking procedure must incorporate these random elements in a meaningful way. Second, the recovery system is operating in a time-variant fashion; hence the system optimization procedure must also be capable of handling these dynamic properties. As a consequence of

these two problems, it is necessary to integrate concepts of both stochastic and dynamic programming in the development of an optimum-seeking procedure.

Preliminary explorations were made to examine whether Markov-renewal programming (Jewell [18]) could be utilized in this regard. For instance, one might develop a Markov-renewal model where the states consist of the "activity" and location of all recovery vehicles at an instant in time. Random transitions are then made from state to state as emergencies occur and are serviced by the system. In each state one decision from a set of feasible decisions can be made to influence the next state of the system. An optimal operating policy can then be derived by selecting the "best" decision vectors when various costs are imposed upon state transitions.

Two problems arise in applying this model to the present problem. The first is the severe computational problem which arises because of the large number of feasible states. For instance, if four vehicles can participate in three activities (say idle, serving a police call, or serving an ambulance call) in ten feasible locations, the resulting Markov-renewal system takes on 30^4 possible states. This large number results in substantial programming problems. The second problem is more severe, and it arises because of the inability of the Markov-

renewal programming formulation to handle the "waiting" lines of emergencies which potentially arise when no recovery vehicles are available. Any model developed for a study of the recovery system response time in an urban area must consider the problem of emergency occurrences when no vehicles are available in a realistic fashion. Since the Markov-renewal programming formulation cannot easily be modified to resolve this problem, it is of limited usefulness in the present application.

These problems in developing an optimum-seeking procedure led to the development of a model which can be used to evaluate the alternative allocation and distribution policies under consideration. It is felt that the careful exploration of this range of feasible alternatives provides information which the decision makers can combine with economic, political, and administrative considerations to determine the appropriate ambulance system configuration for the region being analyzed.

From the earlier discussion on Markov-renewal programming, it is clear that many useful insights can be gained if we apply the concepts of queueing theory to this problem. The sequence of emergencies occurring over time and space form an arrival process to the queueing system. These are reported, join a "waiting line" until an ambulance is dispatched, and are serviced as the recovery vehicle travels to the scene, recovers

the victim, travels to the hospital, and is eventually released to await another call or to service a call which has been "waiting."

Modifying the recovery vehicle allocation policy is then equivalent to changing the number of service channels in the system. Changes in the vehicle distribution policy potentially alter the distribution of service times and the service channel selection mechanism.

Notice, however, that the operation of such a system differs from that of common queueing models. First, the system consists of multiple channels (vehicles) whose service properties vary from channel to channel and from emergency to emergency. This is true because the transit time to the scene of an emergency is dependent upon both the vehicle distribution and the location of the emergency. Second, the channel selection mechanism does not involve a random selection of the available channels. Instead, this selection is based upon the type of emergency and the availability of the "nearest" service channel. Third, since the arrival process consists of both medical and police emergencies, the detailed interactions and probabilistic properties of these must be considered.

Research has been undertaken in this study to incorporate these features into a queueing theoretic model

which allows alternative recovery vehicle allocation and distribution policies to be examined in detail. In addition, the model which has been developed is based upon the data-analytic base underlying the system operation.

Traditionally, many modelers have approached the problem of model construction by utilizing a tractable model whose mathematical properties are known. The resulting model is then implemented by fitting the parameters of the model to data from the operating system. However, when data is available from this system, it generally contains much more information than that obtained by estimating the parameters in a class of simple distributions. If analysis techniques are correctly applied to this data, models corresponding more closely to the real situation can be derived. Of course, the incorporation of such analysis must be selective, for models which are analytically intractable either must be simulated or simplified in order to obtain meaningful solutions. Nevertheless, this latter approach should result in more realistic models, and it is adopted here whenever possible.

1.2 BACKGROUND INFORMATION

The topics under consideration in this dissertation arose in a study of the public ambulance system in

the City of Detroit. This study was conducted by the University of Michigan Highway Safety Research Institute, and it was funded by the U.S. Department of Transportation National Highway Safety Bureau. Interestingly enough, a study of ambulance distribution and allocation was not a part of the research originally proposed. However initial examination of data from the City's present ambulance system revealed that relatively long response times arise under certain conditions.

Furthermore, a major restructuring of the City of Detroit ambulance system is under consideration. The existing system is managed by the police department. (Police departments operate the ambulance service in seven of the fifteen largest cities in the United States. Hence the results of this study should be applicable to many urban areas in addition to Detroit.) A certain proportion of the police squad cars are station wagons, and these serve as both police cars and ambulances on a first come-first served basis. The City is currently evaluating the merits of abandoning this dual system and adopting an independent public ambulance service. This decision will be primarily based upon medical and economic analyses.[1] However, if such a

---

[1]For instance, there is essentially no "in-ambulance" medical care in the existing system. In addition, city

decision is to be made, it is essential to determine the effects of the distribution and allocation of vehicles in the new system. At the present time these effects are completely unknown. However, it is logical to assume that they are substantially different from those in the existing, dual function system, where allocation is based upon police needs rather than emergency medical needs.

Hence this study will deal directly with distribution and allocation problems in the City of Detroit. Data from the Detroit public ambulance system will be used to make inferences on relevant system parameters, and observed measures of system performance will apply specifically to the City of Detroit. However, the results of this study should apply to other urban areas in two ways. First, the existing and proposed ambulance systems in the City of Detroit do not drastically differ from those operating or under consideration in other urban areas. Hence, the results from this study may be modified to apply elsewhere. Second, the general

---

decision makers feel that a substantial opportunity cost is incurred when the police are diverted from their primary function. This must be added to the real cost of using policemen as ambulance attendants, when police are paid $10,000-$12,000 per year and ambulance attendants are paid at a much lower rate.

methodologies used in this study can be directly applied in other analyses. Since many metropolitan areas are currently reviewing their public ambulance systems, these techniques should be useful in such studies.

Systematic research into ambulance service is a fairly recent activity, and there is little literature pertaining directly to this topic. Savas (30) in a study of New York City ambulance service, utilized GPSS simulation as the basis for a decision to allocate ambulances at two stations within a district rather than one. His model was based upon the assumption of a Poisson arrival process and a log-Normal service process, both with parameters fitted to the data. Dunlap and Associates, in a general study of ambulance service for the Department of Transportation (27), discuss the general problem of ambulance distribution and allocation. They utilize a $M/M/\infty$ queueing model to illustrate the concepts, although the comment is made that the assumption of negative exponential service times conflicts with their sampled data. The assumption of an infinite number of service channels is also a severe compromise of realism in order to achieve greater model tractability.

1.3. DESCRIPTION OF THE EXISTING DETROIT RECOVERY SYSTEM

1.3.1. Dual Function Police System

A short summary of the existing Detroit police ambulance system is necessary to introduce a discussion of the modeling of this system. The City of Detroit is divided into thirteen police precincts. Squad cars and police stations wagons are assigned to these precincts, and these vehicles cross precinct boundaries only in special emergency situations. Since the station wagon ambulances are primarily utilized as squad cars, they, too, operate within precincts. Each squad car is assigned to a scout car territory (a "beat") within the precinct. (There are nine to eighteen of these territories per precinct.) The squad car patrols its assigned territory and responds to emergencies within the territory. Since there are many more territories in each precinct than there are squad cars, each car is likely to respond to calls in several territories during a single shift. Some attempt is made to dispatch cars to nearby territories. This is difficult, however, since real-time vehicle location information in unavailable.

These procedures apply to the station wagon ambulances, except that these vehicles respond to both police

runs and medical emergencies on a first come-first served basis. Hence, there is some possibility that an ambulance will be requested when all station wagons are involved with police matters or when the station wagons are all serving some mixture of police and emergency medical calls.

The additional burden placed on the ambulance system by police calls makes it useful to study the extent to which emergency medical response time can be reduced by eliminating or modifying the dual function of the police ambulances. Any resulting reductions in response time may induce city decision makers to separate the police and ambulance service. More important is the fact that the two services may be separated for other reasons--mainly due to the improved medical capability, reduced cost, and streamlined administration which are possible with such separation. In this situation it is extremely important for the city decision makers to have prior information on the operational response capability of an independent "single-function" ambulance service.

1.3.2. Single Function Public Ambulance System

In a single function public ambulance system there is no a priori need to maintain the precinct distribution and allocation system necessary in the police system.

However, utilization of the precinct service regions might be desirable for command and control purposes if the police were to manage the single function system. On the other hand, if a public ambulance system were to be administered by the city hospitals, locations at these hospitals might be desirable from an administrative viewpoint. Similarly, fire stations are logical ambulance stations if the fire department operates the system.

These decisions are all based upon managerial considerations of the responsible municipal agency. In fact, no agency has been selected to manage a single function public ambulance system if one is developed in the City of Detroit. Hence, this analysis will consider the allocation and distribution problem within the existing precinct boundary constraints on station location.

That is, we shall consider the police precinct as the basic service area. Alternative distribution and allocation policies are then studied with respect to this service region. While placing such constraints on the service area may seem to be unduly restrictive, there is little doubt that some constraints must be placed on the service regions in large metropolitan areas. Otherwise the city-wide vehicle command and control problems would result in reduced system effectiveness.

The majority of the numerical results in this study are obtained for various recovery vehicle allocation and distribution policies as implemented in precinct 15 of the City of Detroit. This service region was selected for two reasons. First, it has emergency occurrence characteristics and emergency service problems which are a blend of those in the congested "inner city" precincts and the less-congested "outer city" precincts. Thus, inferences on the effects of alternative allocation and distribution policies for precincts with both types of characteristics can be drawn from the results of precinct 15. Second, precinct 15 has a larger geographic area than most precincts. (See Figure 2.1) Hence, the results for this precinct also provide some information on the effects of allocation and distribution policies when ambulance service regions larger than the "average" police precinct are utilized.

1.4 ORGANIZATION OF RESULTS

The organization of the presentation is as follows. Chapter II is devoted to an extensive data analysis designed to determine the characteristics of urban emergency occurrence and service processes. This analysis relies heavily upon existing statistical tools for the analysis of point stochastic processes, especially those

documented by Cox and Lewis (10).

In Chapter III the insights obtained from the data analysis are utilized to develop a mathematical abstraction of the emergency medical recovery system which models the effects of alternative vehicle distribution and allocation policies. This model is developed by extending the work of Cinlar (7) on systems of queues in parallel. In addition to providing a model for the recovery system under consideration here, the probabilistic results which are obtained are directly applicable to a large class of unresolved problems in queueing theory where the service channel selection is based upon the simultaneous state of all service channels. (The "shortest service line" selection mechanism is probably the most common of these.)

The resulting recovery system model is utilized in Chapter IV to study various single function ambulance allocation and distribution policies of interest to the recovery system designer. The sensitivity of these policies to the model inputs is also examined through certain parameterizations of the basic model. Policies for dual-function urban ambulance systems are then considered in Chapter V. Results for the two systems are integrated in Chapter VI, where the various recovery vehicle allocation and distribution policies are compared.

A brief summary and suggestions for future research are presented in Chapter VII. Appendix A contains reference information on various statistical tests designed for the analysis of Poisson processes. Appendix B develops the theory which relates the stochastic process consisting of both ambulance and police call occurrences to the process consisting of police calls alone or of ambulance calls alone. Appendix C documents the computer program utilized to develop the numerical results of this study.

CHAPTER II

DATA ANALYSIS OF EMERGENCY OCCURRENCE AND SERVICE PROCESSES

In Chapter I we stressed the importance of integrating the process of model development with the data base which describes the system under examination. This integration should yield a system abstraction from which one can develop better insights into system performance. As a result, the model should ultimately lead to better decisions, and this, of course, is the basic objective in modeling.

The utilization of data analytic techniques in model construction is especially important in this study, since conclusions drawn from the model as to the effects of the alternatives under consideration will be very difficult and expensive to validate experimentally. Therefore, various data analyses will be conducted in this chapter in order to examine the properties of both the emergency occurrence processes and the service processes for these emergencies. The insights gained from these analyses will then be utilized in Chapter III in the model development.

It is important to note that the intent of these data analyses is not to identify or explain the general statistical processes associated with urban emergencies. Such a study would require more extensive data and more detailed analyses than those utilized here. Nevertheless, various interesting hypotheses on urban emergency processes are generated by the results of this chapter, and these could serve as a basis for further research.

As we indicated in Chapter I, the development of models to explain probabilistic phenomena requires giving constant attention to the trade-off between tractability and realism. For instance, consider the problem of selecting a model for the distribution function of a random variable. In certain situations theoretical or asymptotic considerations will provide a class of potential distribution functions, and experimental data can then be utilized to estimate the parameters of these.

In the present analysis no theory has been developed which suggests appropriate models for the emergency occurrence and service processes under consideration. In the absence of theory, two alternatives are available. First, one can use the empirical distribution function as an estimator of the various marginal distribution functions. Although this estimator has

desirable statistical properties when the observations are independent and identically distributed (Cox and Lewis [10]), its utilization in subsequent model analysis will yield results which are analytically intractable. In this case, numerical analysis will be necessary to develop the insights of interest.

Therefore, one may prefer to choose parametric classes of distribution functions which approximate the empirical distribution function. In this case various non-parametric goodness of fit tests are available to examine the adequacy of these classes of functions against general alternatives, and parametric statistical tests can be conducted to examine their "goodness" as compared to well-defined alternatives. Furthermore, parametric statistical inference procedures are available to estimate the parameters in the chosen class of models.

We have chosen the latter approach in this study. As we shall see, the marginal distribution functions of the processes we are considering can be modeled quite adequately with distributions from the "Gamma" class. Furthermore, we shall see in Chapter III that this class of distributions leads to tractable, analytical results which greatly simplify the numerical calculations. Finally, we shall show in Chapter IV that the conclusions

one draws from the model are quite insensitive to parameter variations within this Gamma class.

In addition, it is possible to utilize the statistical inferences of this Chapter to develop a theoretical explanation of the mechanism which generates the emergency processes we are examining. This development is presented in Section 2.2.3 and in Appendix B.

The sampling plan underlying the data collection is discussed in Section 2.1. In Section 2.2 statistical analyses are conducted on both the police emergency and medical emergency occurrence processes and their interactions. The service processes for these two types of emergencies are analyzed in Section 2.3. Results of the data analyses are summarized in Section 2.4.

2.1 SAMPLING PLAN

Since this research was conducted as part of a comprehensive examination of urban emergency medical care, it was impossible to develop an independent sampling plan for the distribution and allocation study. However, a plan was developed which provides the data base necessary for this study as well as information essential to other portions of the research project.

As we have indicated, the City of Detroit and most major cities in the United States are divided into

police precincts for vehicle command and control. Since the police operate the ambulance service in Detroit, the precincts also serve as the basic geographic units for ambulance allocation and distribution.

Four precincts were chosen as sampling units for the entire study. These were selected as a representative sample of the thirteen Detroit precincts in terms of demographic characteristics and accident experience. Two densely populated "inner city" precincts (Precincts 6 and 7) and two more sparsely populated "outer city" precincts (Precincts 15 and 16) were included in order to make comparisons between these two types. Figure 2.1 shows the locations and relevant characteristics of these service regions. Data from the 7th and 15th precincts has been analyzed for the distribution and allocation study.

In these two precincts exhaustive samples of all dispatches for emergency medical runs were obtained and analyzed for the one month period November 18-December 15, 1968. Police emergency runs in precinct 15 were similarly analyzed for the same time period in order to make inferences on the operation of the "dual function" police-ambulance system.

While data from this one month time frame precludes direct inferences on long term characteristics of

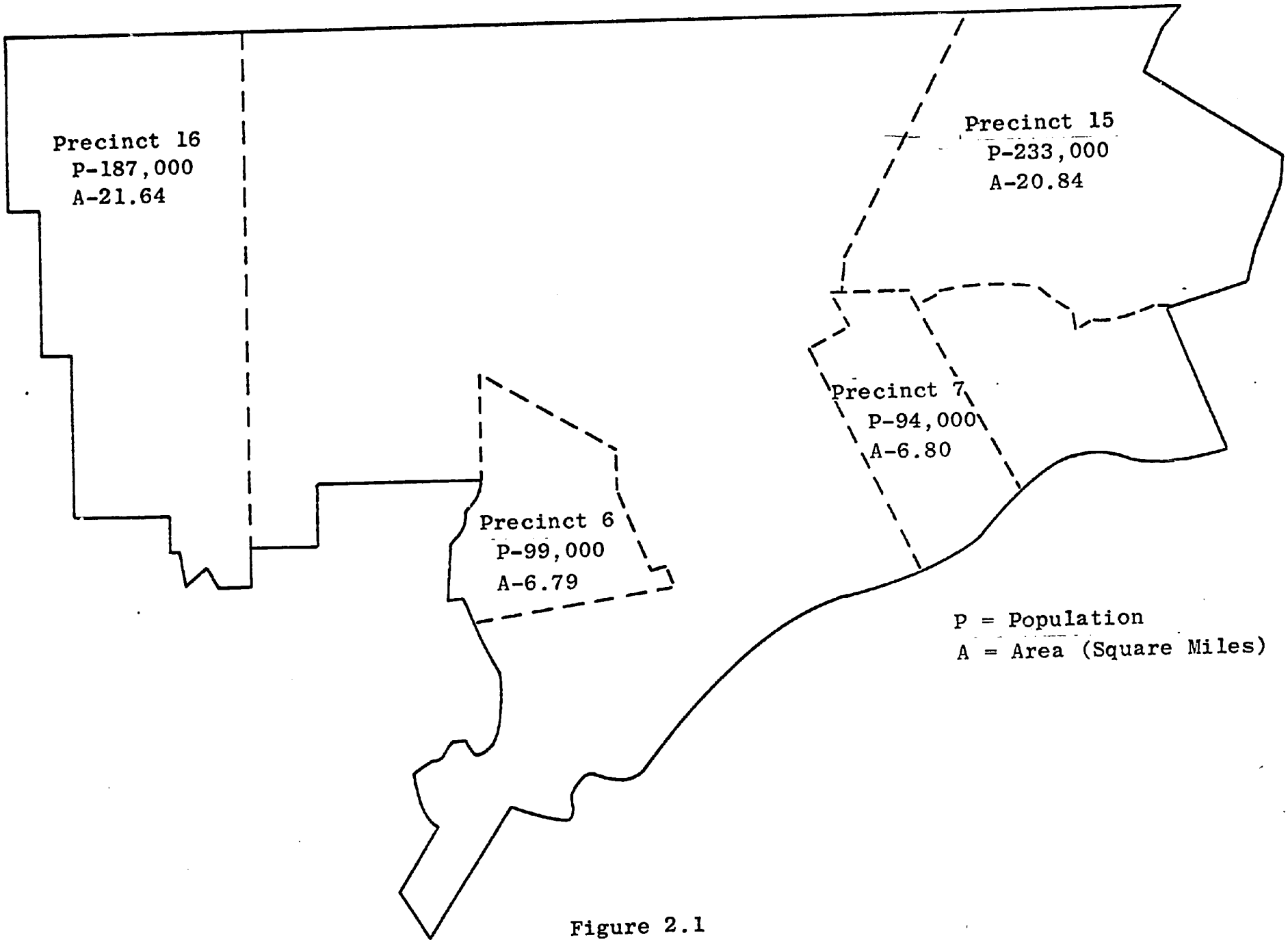


Figure 2.1

City of Detroit Police Precincts
Sampled for Allocation and Distribution Study

the emergency occurrence and service processes, the series are long enough to allow one to examine potential hourly, daily, and weekly trends or cycles. Results from these examinations should provide reasonable indications as to the detailed structure of the long term series, and model parameterization can then be utilized to study system response to changes in such structure.

For instance, suppose we were to infer from the data that the sequence of medical emergency occurrences in November/December, 1968, can be satisfactorily represented as a Poisson process with rate λ. We might then conjecture that the process maintains its Poisson nature throughout the year, but that the rate varies according to seasonal characteristics. The sensitivity of model conclusions to such occurrence rate variations can then be studied by modifying the parameter λ and observing the subsequent system response.

For the November/December time frame and the above geographic regions, data for the allocation and distribution study were taken from two sources. First, the central police communications center routinely fills out a dispatch ticket for each emergency to which a vehicle is dispatched. This ticket contains the following information relevant to this study:

1. Time of emergency report

2. Nature of emergency
3. Location of emergency (street and number)
4. Precinct number and patrol territory of vehicle dispatched
5. Time of dispatch
6. Time mission completed

Second, for the purpose of this study, the operator of each police ambulance filled out a data form for each medical emergency containing the following information:

1. Time and distance to the scene of the emergency
2. Time on the scene
3. Time and distance to the hospital
4. Time at the hospital transferring the victim

The time of emergency reporting on the dispatch ticket is recorded to the nearest minute from an electric clock. The time of dispatch and time of mission completion are automatically imprinted by an electric clock-printer. Hence, these data should be free of biases or other recording errors. The major sources of potential error on the dispatch ticket are in the determination of the nature and location of the emergency. In many cases vehicles are dispatched on medical emergency runs which turn out to be false alarms or minor injuries not requiring an ambulance. Unfortunately, it is nearly impossible to detect these *a priori* at the communications center, since

the caller generally is excited and/or confused, or else he refuses to give enough information to make an accurate diagnosis of the nature of the emergency. Hence, for the purpose of improved ambulance allocation and distribution, information on the dispatch ticket alone was used to determine whether the emergency was a medical emergency or police emergency. For the same reason, emergency location information on the ticket was used to determine the site of the emergency, even though errors sometimes do occur in obtaining and recording this data accurately.

Time data on the ambulance attendant's sheet were recorded from the police officer's watch. Thus, there is a tendency for this data to be biased toward five minute intervals, even though the officers were instructed to make observations to the nearest minute. The data analysis should not be seriously affected by these, since such observations should be within one to two minutes of the true values.

Distance data were recorded from the vehicle odometer. Although this data is generally tabulated to the nearest mile, the fairly large residual errors in the distance time relationships indicate that more accurate distance determinations are unnecessary for this study. We shall discuss this point further in Section 2.3.1.

2.2 ANALYSIS OF THE EMERGENCY OCCURRENCE PROCESSES

The sequence of times between emergency occurrences was partitioned into two sequences--the times between medical emergencies and the times between police emergencies. These sequences were subjected to statistical analyses in order to make reasonable inferences on the stochastic processes underlying the realizations. Analyses to determine the extent and nature of the interactions between the two sequences were also conducted.

2.2.1 Medical Emergency Occurrence Process.

Times between successive medical emergencies occurring within the study precincts during the sampling frames were analyzed as point stochastic processes. These have been studied to examine the following:

(i) The existence and nature of time trends (non-stationarities) in the series.

(ii) The existence and nature of interdependencies in the detrended series (if evidence of trends is found in (i) above).

(iii) The nature of the marginal distribution of inter-emergency occurrence times (and of the joint distributions of such times if significant interdependencies are found in (ii) above).

As we have observed, trends emerging gradually and cycles with periods greater than one month cannot be detected due to limitations in the sampling plan. Instead the effects of these phenomena must be incorporated by a re-parameterization of the system model which has been constructed using the available data. Nevertheless, it is felt that careful analysis of the data does provide much insight into the detailed structure of the series of occurrences and also allows one to make cautious extrapolations to longer time periods.

In order to obtain some preliminary insights into the nature of the emergency occurrence processes in precincts 7 and 15, counts of the daily number of calls to the police communications center for ambulance service were plotted against time (Figure 2.2) for the period November 19 to December 14, 1968. A cursory inspection of this data reveals no strong trends or cyclic tendencies. The major indication is that medical emergencies seem to occur at a slightly higher frequency in precinct 15. No obvious interdependencies between the two series appear in the data.

For the two week period December 1, 1968, to December 14, 1968, the daily counts in precinct 15 were decomposed into counts for each of the three police shifts,

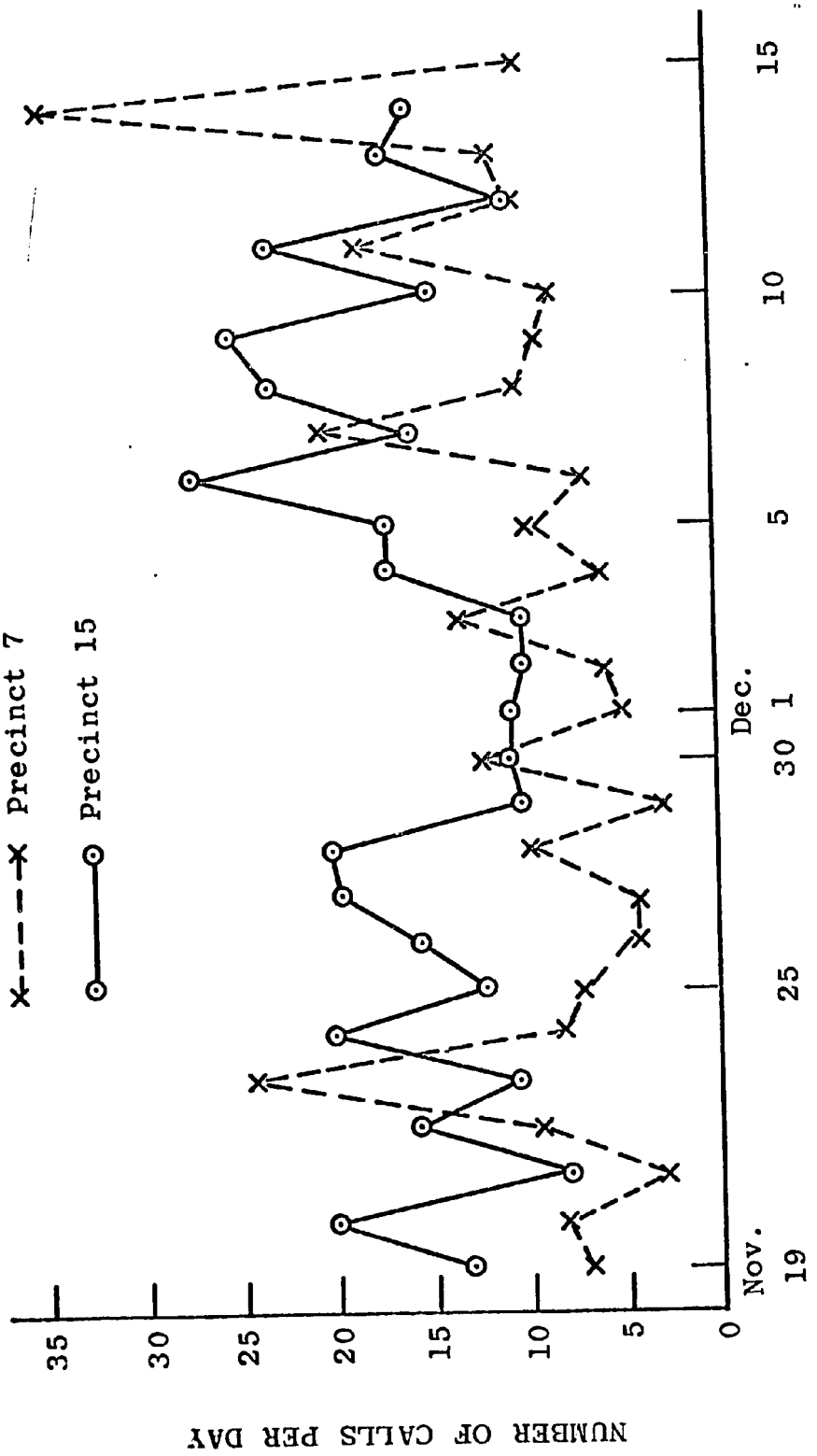


Figure 2.2

Number of Ambulance Calls per Day

November 19; 1968 - December 15, 1968

0000-0800, 0800-1600, 1600-2400. These counts are graphed in Figure 2.3.

No obvious trends or cycles are apparent within each shift. Further, the three shifts appear to be similar in both central tendency and variability.

Since no major daily or within-shift differences were observed in the plotted series, the entire series of emergency occurrences in each precinct over the one month period were subjected to more detailed statistical analysis. Interarrival times between successive calls for ambulance service were computed and analyzed using a program developed by IBM for the statistical analysis of series of events (SASE II). The reader is referred to references (9, 10, 19) for more detailed information on this program. Over the one month time period, 416 calls for emergency ambulance series occurred in precinct 7 and 403 such calls occurred in precinct 15. Hence, these realizations of the series of events should provide sufficient information for accurate analysis of the underlying processes. In addition, the sample records are long enough so that certain asymptotic results can be utilized to test various hypotheses of interest.

Failure to detect any trends in the series by graphical methods leads one to consider the observed interarrival times as realizations of (weakly) stationary

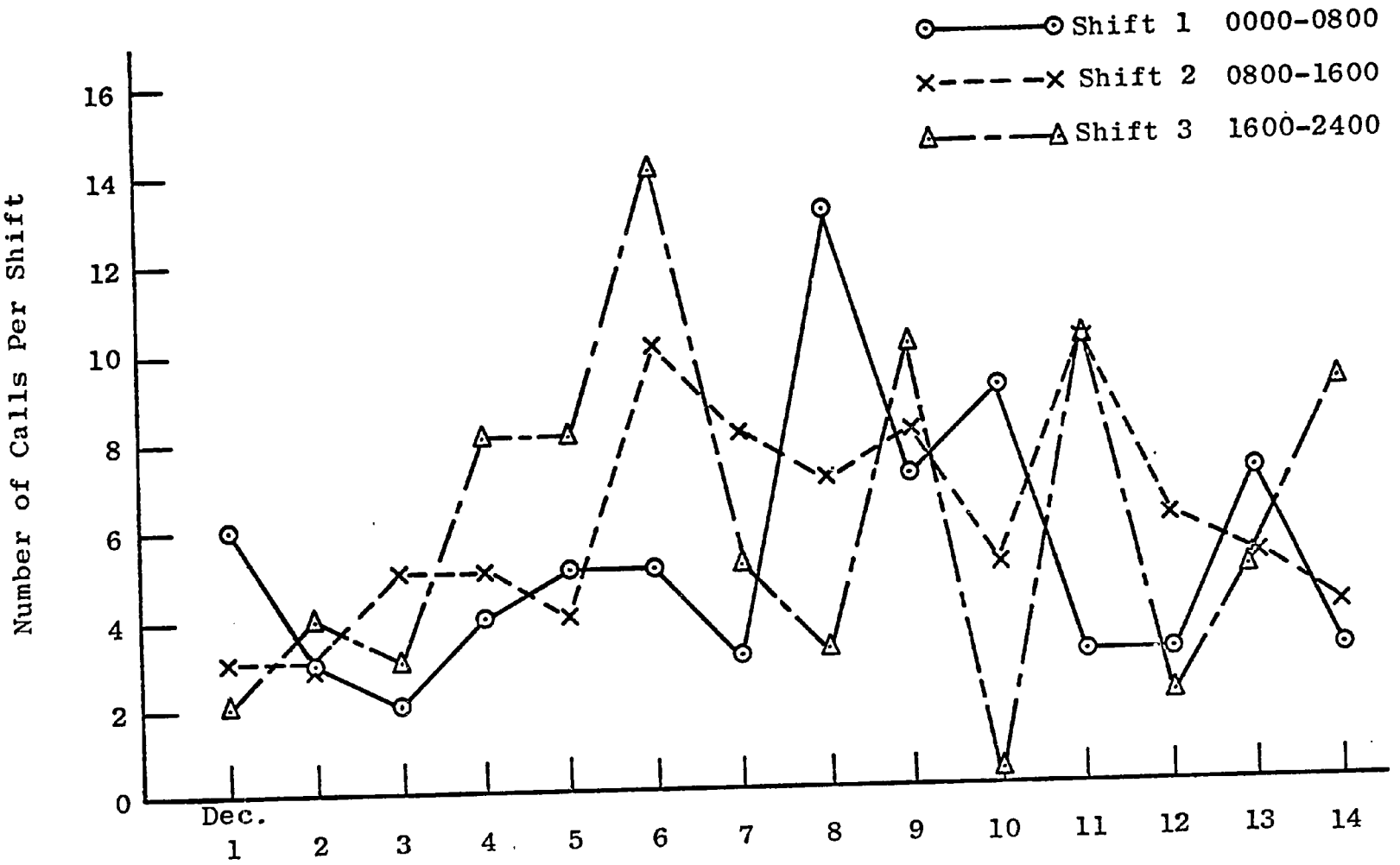


Figure 2.3

Number of Ambulance Calls per Shift

Precinct 15 December 1-December 14, 1968

point processes. That is, we consider the first and second moments of the underlying stochastic processes to be invariant under a time shift. Letting X_i denote the interarrival time between the i^{th} and $(i-1)^{st}$ police call in a precinct, this means that:

$$E(X_i) = E(X_{i+h}) = E(X) \qquad i = 1, 2, \cdots, n$$

$$\text{cov}(X_i, X_{i+h}) = C_h \qquad h = 0, 1, \cdots, n-i$$

where the function C_h is known as the autocovariance function. The autocorrelation ρ_h is then given by C_h/C_0.

It is of immediate interest to study the nature and extent of interdependencies between the interarrival times. Specifically, it is desirable to test the hypothesis that the observations are independent and hence form a renewal process (a sequence of independent identically distributed random variables). Tests of this hypothesis can be based on the sample autocorrelation function or on the sample periodogram. Both are discussed and presented here.

It is well known that for a stationary time series an unbiased estimator of E(X) is given by

$$\bar{X} = \frac{1}{n} \sum_{j=1}^{n} X_j$$

and an asymptotically unbiased estimator of C_h is given by

$$\hat{C}_h = \frac{1}{n-h} \sum_{j=1}^{n-h} (X_j - \bar{X}_h')(X_{j+h} - \bar{X}_h'') \quad h = 1, \cdots, V_{max}$$

where V_{max} is the covariance truncation point and

$$\bar{X}_h' = \frac{1}{n-h} \sum_{j=1}^{n-h} X_j \qquad \bar{X}_h'' = \frac{1}{n-h} \sum_{j=1}^{n-h} X_{j+h}$$

This estimator has been utilized instead of the standard unbiased estimator of C_h since it is robust when small trends are present in the underlying process.

An asymptotic unbiased estimator for the auto-correlation can then be written as

$$\hat{\rho}_h = \frac{\hat{C}_h}{(\hat{C}'_{0,h}\ \hat{C}''_{0,h})^{1/2}}$$

where

$$\hat{C}'_{0,h} = \frac{1}{n-h} \sum_{j=1}^{n-h} (X_j - \bar{X}_h')^2$$

$$\hat{C}''_{0,h} = \frac{1}{n-h} \sum_{j=1}^{n-h} (X_{j+h} - \bar{X}_h'')^2$$

The large sample moments of this estimator are given by Cox and Lewis (10) as:

$$E(\hat{\rho}_h) \doteq 0$$

$$V(\hat{\rho}_h) \doteq \frac{1}{n-h} \sum_{j=-\infty}^{\infty} \rho_j^2 = \frac{1}{n-h} \left(1 + 2 \sum_{j=1}^{\infty} \rho_j^{2)}\right.$$

$$\operatorname{cov}(\hat{\rho}_h, \hat{\rho}_{h+k}) \doteq \frac{1}{n-h} \sum_{j=-\infty}^{\infty} \rho_j \rho_{j+k}$$

The asymptotic normality of estimators of ρ_h has been discussed and demonstrated in Anderson (4). Utilizing the central limit theorem, it is clear that under a renewal hypothesis ($\rho_j = 0$; $j = 1, 2, \cdots, \infty$), the $\hat{\rho}_h$ are asymptotically normal random variables with zero mean and variance equal to $1/(n-h)$. The asymptotic covariance between the sample autocorrelations is of the order $1/n$. Approximate tests of the renewal hypothesis can then be conducted by comparing the statistic $\sqrt{n-h}\ \hat{\rho}_h$ with the upper α^{th} percentile of the normal distribution.

Using these techniques, the sample autocorrelations have been computed for lags up to 100 and tested for the data in precincts 7 and 15 (Table 2.1). In precinct 7 significance at the .05 level was observed at lags of 39, 64, 72, 78, and 94. In precinct 15, significance was observed at lags of 2, 29, 39 and 49. We observe that in both cases the observed empirical probabilities of rejection are less than .05, the significance level chosen for this analysis. Hence this analysis lends support to the renewal hypothesis.

Table 2.1

Estimated Autocorrelations of Times Between Medical Emergencies

	Precinct 15		Precinct 7	
Lag (h)	$\hat{\rho}_h$	$\sqrt{n-h}\,\hat{\rho}_h$	$\hat{\rho}_h$	$\sqrt{n-h}\,\rho_h$
1	-0.011	-0.223	0.085	1.733
2	0.098	1.965	0.025	0.511
3	0.024	0.478	-0.017	-0.350
4	-0.061	-1.214	0.012	0.250
5	-0.054	-1.084	-0.045	-0.912
6	-0.050	-0.991	-0.001	-0.024
7	-0.013	-0.265	-0.038	-0.760
8	0.025	0.499	-0.011	-0.220
9	0.062	1.225	-0.020	-0.396
10	0.079	1.571	0.080	1.603
11	0.005	0.100	0.023	0.458
12	0.015	0.287	0.061	1.228
13	-0.034	-0.667	0.033	0.666
14	0.018	0.359	0.011	0.216
15	-0.021	-0.412	0.038	0.762
16	0.020	0.398	0.013	0.262
17	-0.008	-0.162	-0.062	-1.235
18	-0.009	-0.183	-0.008	-0.166
19	0.018	0.350	0.009	0.176
20	0.001	0.026	0.007	0.145
21	-0.009	-0.178	0.023	0.454
22	-0.008	-0.163	0.041	0.814
23	-0.007	-0.139	-0.031	-0.621
24	0.031	0.611	-0.077	-1.520
25	0.068	1.318	-0.063	-1.253

Table 2.1 (continued)

	Precinct 15		Precinct 7	
Lag (h)	$\hat{\rho}_h$	$\sqrt{n-h}\ \hat{\rho}_h$	$\hat{\rho}_h$	$\sqrt{n-h}\ \hat{\rho}_h$
26	-0.061	-1.181	-0.006	-0.128
27	0.040	0.781	-0.074	-1.463
28	-0.007	-0.132	0.042	0.835
29	0.148	2.862	-0.041	-0.814
30	0.028	0.542	-0.020	-0.384
31	-0.019	-0.362	0.055	1.083
32	0.097	1.863	0.017	0.333
33	0.010	0.185	0.024	0.460
34	-0.036	-0.700	-0.003	-0.060
35	-0.068	-1.304	-0.050	-0.977
36	-0.022	-0.424	0.001	0.021
37	-0.002	-0.036	0.013	0.256
38	0.014	0.275	-0.064	-1.248
39	0.123	2.350	-0.111	-2.149
40	0.016	0.314	-0.021	-0.404
41	0.075	1.422	0.004	0.078
42	-0.016	-0.303	-0.096	-1.860
43	0.006	0.115	-0.073	-1.405
44	-0.048	-0.906	-0.049	-0.949
45	-0.028	-0.535	-0.023	-0.446
46	-0.010	-0.192	-0.072	-1.379
47	0.006	0.119	0.035	0.681
48	0.006	0.112	-0.046	-0.885
49	0.113	2.126	-0.063	-1.205
50	-0.036	-0.670	0.000	0.008
51	0.016	0.304	-0.041	-0.787
52	-0.008	-0.158	0.019	0.370
53	-0.011	-0.201	-0.010	-0.195
54	-0.010	-0.190	0.018	0.336

Table 2.1 (continued)

Lag (h)	Precinct 15 $\hat{\rho}_h$	Precinct 15 $\sqrt{n-h}\ \hat{\rho}_h$	Precinct 7 $\hat{\rho}_h$	Precinct 7 $\sqrt{n-h}\ \hat{\rho}_h$
55	0.004	0.071	-0.041	-0.779
56	0.045	0.844	0.013	0.252
57	0.024	0.455	0.033	0.617
58	0.038	0.714	0.002	0.046
59	-0.044	-0.820	0.016	0.293
60	-0.043	-0.804	0.098	1.857
61	-0.022	-0.415	-0.051	-0.965
62	-0.006	-0.104	0.019	0.363
63	0.034	0.633	-0.019	-0.361
64	-0.054	-0.995	0.183	3.439
65	0.067	1.227	-0.036	-0.680
66	-0.038	-0.690	0.059	1.095
67	0.005	0.096	0.013	0.247
68	-0.050	-0.919	0.027	0.501
69	-0.083	-1.514	-0.008	-0.155
70	-0.008	-0.153	-0.084	-1.556
71	0.060	1.090	-0.027	-0.494
72	-0.035	-0.643	0.129	2.392
73	0.083	1.510	-0.005	-0.097
74	0.056	1.010	0.017	0.313
75	-0.058	-1.049	0.009	0.158
76	-0.012	-0.224	0.029	0.526
77	0.030	0.541	-0.050	-0.918
78	0.000	0.006	0.112	2.067
79	-0.036	-0.656	-0.049	-0.894
80	-0.052	-0.937	-0.003	-0.064
81	0.009	0.153	0.024	0.438
82	-0.030	-0.530	-0.012	-0.226
83	0.050	0.887	-0.006	-0.112

Table 2.1 (continued)

	Precinct 15		Precinct 7	
Lag (h)	$\hat{\rho}_h$	$\sqrt{n-h}\ \hat{\rho}_h$	$\hat{\rho}_h$	$\sqrt{n-h}\ \hat{\rho}_h$
84	-0.029	-0.526	0.103	1.879
85	-0.014	-0.256	-0.001	-0.017
86	0.006	0.107	-0.054	-0.972
87	0.018	0.324	0.004	0.080
88	-0.021	-0.376	0.081	1.473
89	-0.046	-0.807	-0.062	-1.130
90	-0.075	-1.327	-0.014	-0.249
91	-0.065	-1.154	-0.056	-1.015
92	-0.034	-0.598	-0.029	-0.528
93	-0.020	-0.354	0.025	0.441
94	0.030	0.525	0.144	2.582
95	0.024	0.414	0.082	1.474
96	0.029	0.505	0.023	0.408
97	0.008	0.143	0.097	1.724
98	-0.027	-0.468	-0.038	-0.685
99	-0.008	-0.145	-0.013	-0.232
100	-0.054	-0.933	-0.049	-0.873

Since the sample autocorrelations are correlated for all finite record lengths, the statistical test utilized above is inaccurate to the extent that it ignores this dependence. Hence a test based upon the sample spectrum of the intervals is utilized to supplement the analysis in the time domain.

The spectrum of a stochastic process is developed by taking the Fourier transform of the autocovariance function. Unsmoothed spectral estimates can be computed by forming the periodogram $I(\omega)$,

$$I(\omega) = \frac{1}{2\pi} \sum_{\nu=+\mu}^{\mu} e^{i\nu\omega} \kappa(\nu/\mu) \hat{C}_\nu \qquad |\omega| \leq \pi$$

where $\kappa(\nu/\mu)$ is the lag window of the estimate

μ is the truncation point of the estimate

For real-valued processes the above relation becomes

$$I(\omega) = \frac{1}{\pi} \left[1 + 2 \sum_{\nu=1}^{\mu} \kappa(\nu/\mu) \cos(\omega\nu) \hat{C}_\nu\right] \qquad \omega \leq \pi$$

The Bartlett (triangular) lag window and a spectral truncation point of n were utilized in computations of the periodograms, thereby yielding the estimator

$$I(j) = \frac{1}{\pi} \left[1 + 2 \sum_{\nu=1}^{n} \left(\frac{n-\nu}{n}\right) \hat{C}_\nu \cos\left(2\pi\nu\frac{j}{n}\right)\right]$$

$$j = 1, 2, \cdots, n/2$$

If a series is a realization of a sequence of independent, normally distributed random variables, then Cox and Lewis (10) show that $\{I(1), \cdots, I(n/2)\}$ forms a sequence of independent exponentially distributed random variables with parameter $\lambda = \pi/C_0$. Since the intervals between ambulance calls are by definition non-negative, the normal assumption cannot hold. However, under a renewal hypothesis, the above result holds asymptotically as a result of the central limit theorem.

Hence an asymptotic test of the renewal process can be conducted by testing to determine if the sample periodogram follows a Poisson process. There are many tests for Poisson processes, and some of the more common of these are summarized in Appendix A. Results of applying these tests to periodograms calculated from the intervals observed in precincts 7 and 15 are presented in Table 2.2.

All statistical tests indicate that the values assumed by the sample periodograms are consistent with a Poisson process. Hence one is led to the inference that the observed interarrival times follow a renewal process.

A renewal process is completely identified if the marginal distribution of interarrival times is known. Therefore inference procedures to examine this distribution are developed in this section. It is well known that the empirical distribution function is a consistent

Table 2.2

Tests of Renewal Hypothesis Based on Sample Periodogram

Statistic	Value Attained for Precinct 7	Value Attained for Precinct 15	.05 Significance Level
Upper-sided Kolmogorov-Smirnoff	.122	.694	1.224
Lower-sided Kolmogorov-Smirnoff	1.044	.866	1.224
Two-sided Kolmogorov-Smirnoff	1.044	.866	1.358
Anderson-Darling	1.898	.818	2.492

estimator of the underlying marginal distribution function. The empirical distribution functions $F_n(x)$ for interarrival times between ambulance calls in precincts 7 and 15 are presented in Figures 2.4 and 2.5 respectively.

Various statistics for these empirical distributions are recorded in Table 2.3.

Table 2.3

Statistics for Marginal Distributions of Interarrival Times Between Ambulance Calls

Statistic	Observed Value Precinct 7	Observed Value Precinct 15
Mean	98.072 min	101.166 min
Variance	18738.207 min^2	15679.672 min^2
Coefficient of Variation	1.396	1.238
Coefficient of Skewness	3.603	2.449
Kurtosis	20.950	10.462
Range	1-1120 min	1-866 min

Visual comparison of these statistics indicates that times between ambulance calls in the two precincts have approximately the same means, but that times between calls in precinct 15 have slightly less variability and skewness. There is no apparent reason for these differences, and in fact, one might expect more significant

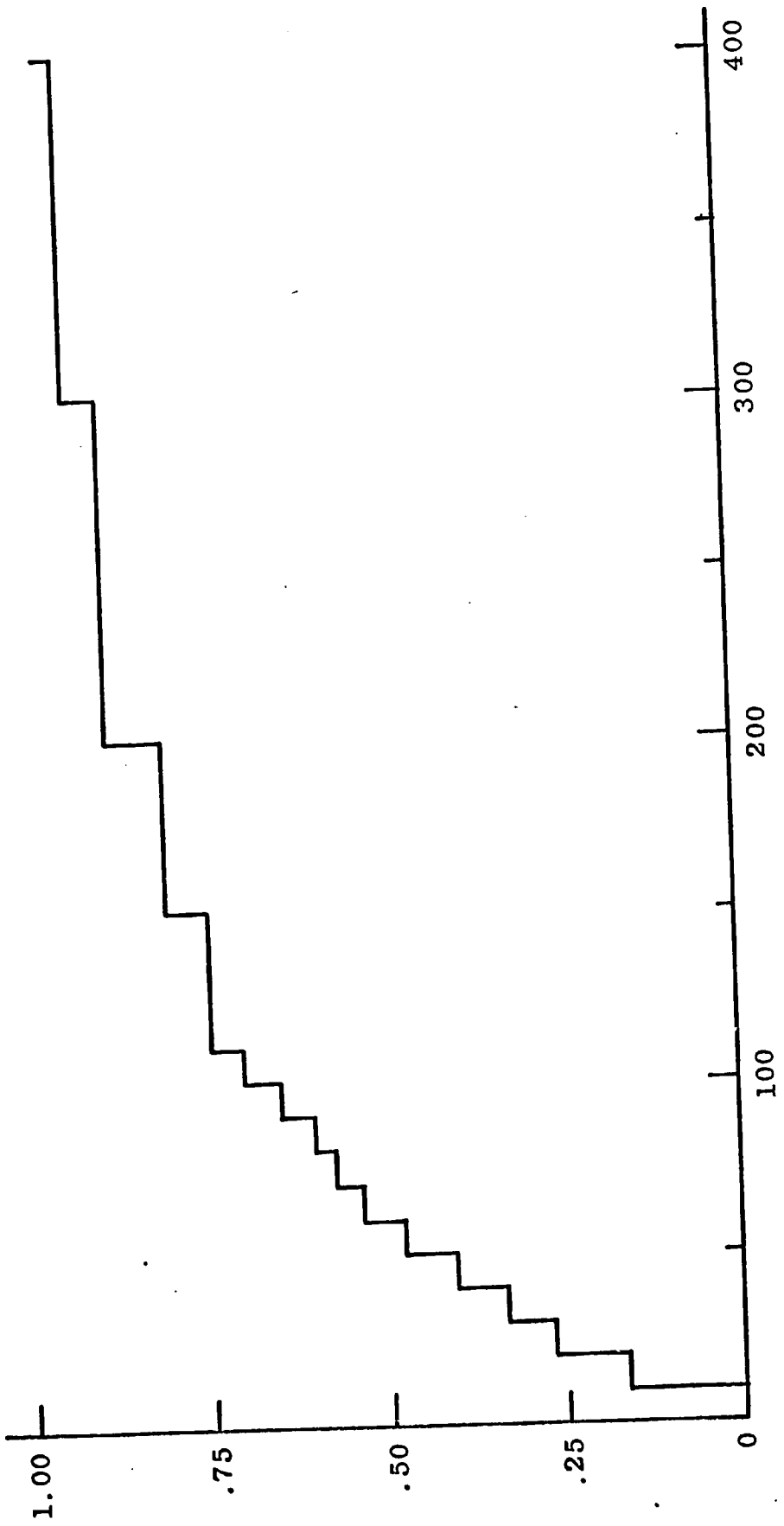


Figure 2.4

Empirical Distribution Function - Times between Ambulance Calls - Pct. 7

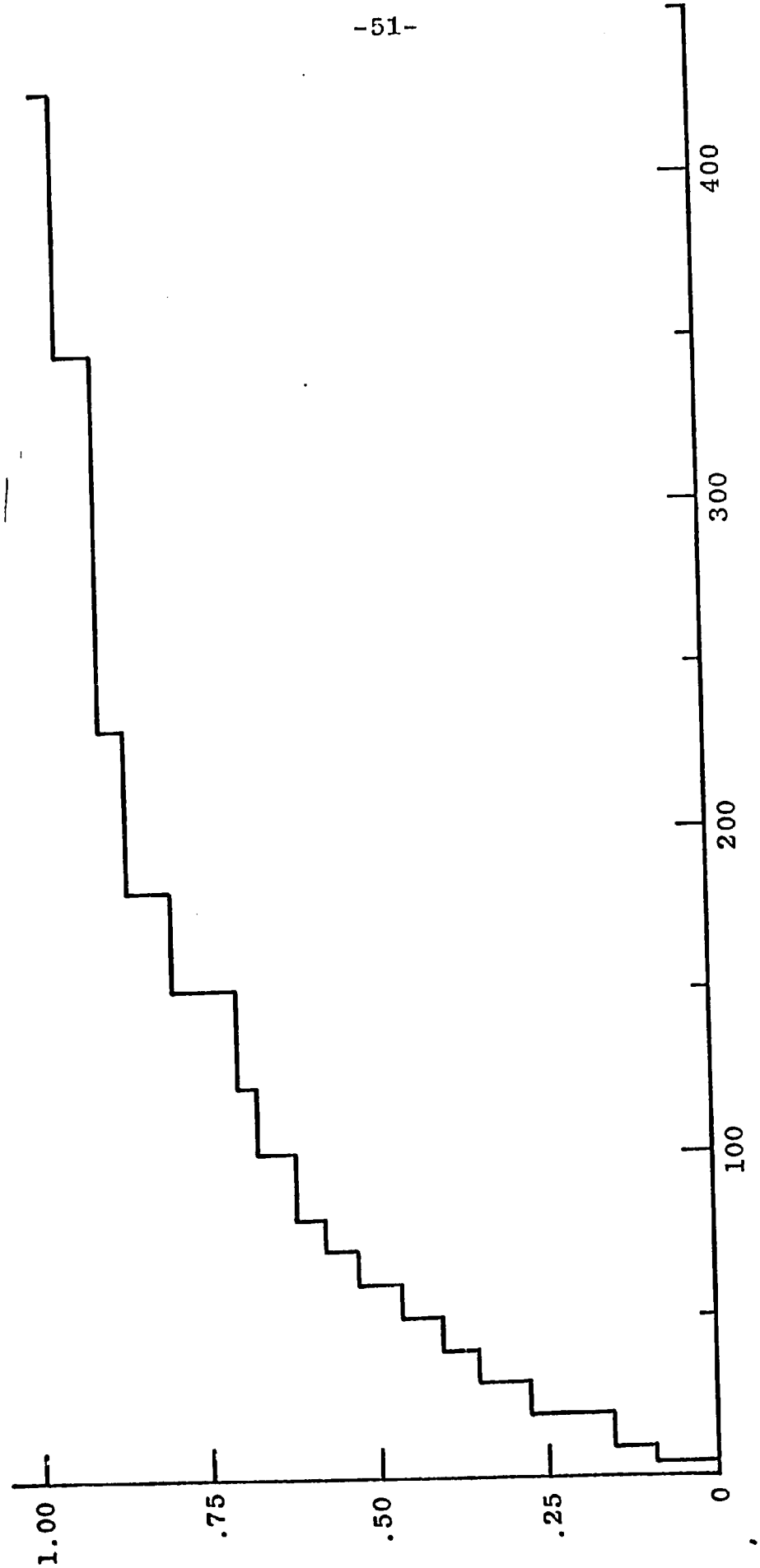


Figure 2.5

Empirical Distribution Function - Times between Ambulance Calls - pct. 15

variations because of the demographic differences between the two precincts.

Although the sample coefficients of variation are larger than one might expect for a negative exponential distribution of interarrival times,[1] statistical tests were conducted to determine if these differences could be attributed to chance variation. Here again, the statistical tests developed in Appendix A for Poisson processes are valid, since a negative exponential distribution of times between events is equivalent to a Poisson distribution of the number of events occurring in any time interval. Statistics computed for these tests are presented in Table 2.4.

The tests were repeated on the partial sums of the order statistics of the interarrival times, since this procedure gives increased power against a wide range of alternatives (see Appendix A). Results are presented in Table 2.5.

We observe that rejection of the Poisson hypothesis of interarrivals in precinct 15 is indicated by all tests except the upper-sided Kolmogorov test. Results in precinct 7 are not as striking, but ample evidence is provided for rejection of the Poisson hypothesis.

[1] The negative exponential distribution with parameter λ has mean $1/\lambda$ and variance $1/\lambda^2$. Hence the coefficient of variation is equal to one.

Table 2.4

Statistics Computed to Test for a Negative Exponential Distribution of Times Between Ambulance Calls

Statistic	Value Attained for Precinct 7	Value Attained for Precinct 15	.05 Significance Level
Upper-sided Kolmogorov-Smirnoff	.555	.361	1.224
Lower-sided Kolmogorov-Smirnoff	1.319*	1.799*	1.224
Two-sided Kolmogorov-Smirnoff	1.319	1.799*	1.358
Anderson-Darling	2.222	3.137*	2.492
Moran	669.000*	624.715*	(465,451)**

* Attained values exceed the .05 significance level.

**Significance level for the Moran test is dependent upon sample size. Hence it differs for each precinct. The significance levels shown are based upon 400 degrees of freedom and thus are slightly higher than the exact significance levels.

Table 2.5

Tests for Negative Exponential Distribution of Times between Ambulance Calls Based upon Partial Sums of Order Statistics

Statistic	Value Attained for Precinct 7	Value Attained for Precinct 15	.05 Significance Level
Upper-sided Kolmogorov-Smirnoff	.208	.198	1.224
Lower-sided Kolmogorov-Smirnoff	2.883*	2.552*	1.224
Two-sided Kolmogorov-Smirnoff	2.883*	2.552*	1.358
Anderson-Darling	20.742*	14.802*	2.492

*Attained value exceeds the .05 significance level.

In Appendix A it is shown that the MORAN test is constructed to provide high power against Gamma alternatives to the Poisson hypothesis. Thus, the large values attained by the Moran statistics (Table 2.4) provide some indication that an adequate model of the marginal distribution of interarrival times between ambulance calls can be developed by using the Gamma distribution.

Further evidence in favor of the Gamma hypothesis is obtained by comparing the sample coefficients of skewness and kurtosis with estimates of their theoretical values in the two precincts. Denoting the coefficient of variation by ξ_0, the coefficient of skewness by ξ_1, and the kurtosis by ξ_2, it is well-known that for the Gamma distribution (Jackson [17]):

$$\xi_1 = 2\xi_0 \qquad \xi_2 = 6\xi_0^2$$

Estimates of these parameters can be found by substituting the empirical coefficients of variation, and these are compared with the corresponding sample estimates in Table 2.6.

Inspection of Table 2.6 indicates close agreement between the coefficients of skewness estimated by the two techniques. The estimates of kurtosis are also similar in precinct 15, although significant variation occurs in precinct 7.

Table 2.6

Comparison of Estimated and Sample Coefficients of Variation and Kurtosis--Gamma Distribution of Times Between Ambulance Calls

	Precinct 7	Precinct 15
ξ_0	1.396	1.238
$\xi_1 = 2\xi_0$	2.79	2.48
Sample Estimate of ξ_1	3.60	2.45
$\xi_2 = 6\xi_0^2$	11.70	9.30
Sample Estimate of ξ_2	20.95	10.46

One logical alternative to the Gamma distribution of times between ambulance calls is the log-normal distribution, which might arise if the events leading to these calls combine in a multiplicative fashion. For a log-normal distribution, the coefficient of skewness and the kurtosis are related to the coefficient of variation by the following relations.

$$\xi_1 = 3\xi_0 + \xi_0^3$$

$$\xi_2 = 16\xi_0^2 + 15\xi_0^4 + 6\xi_0^6 + \xi_0^8$$

Table 2.7 contains a comparison of these values for the Gamma and log-normal distribution.

In all cases much closer agreement between the two estimators is obtained under the Gamma hypothesis.

Table 2.7

Comparison of Estimated Coefficients of Skewness and Kurtosis for Gamma and Log Normal Hypotheses--Time Between Ambulance Calls

	Precinct 7	Precinct 15
Sample Estimate of ξ_0	3.60	2.45
Estimate of ξ_1 based on sample ξ_0		
Gamma	2.79	2.48
Log-Normal	6.92	5.63
Sample Estimate of ξ_2	20.95	10.46
Estimate of ξ_2 based on sample ξ_0		
Gamma	11.70	9.30
Log-Normal	147.12	87.82

While these results lead to a qualitative acceptance of the Gamma model over the Log-Normal alternative, it is useful to apply a more sophisticated statistical test. Thus, tests of a Gamma model versus a Log-Normal alternative were conducted using the log likelihood ratio statistic as developed by Jackson. The test statistic is

$$T_g = (n/2) \ln (\hat{\alpha}_2 / \alpha_{2,\hat{\kappa}}) + n\hat{\kappa}(\hat{\alpha}_1 - \alpha_{1,\hat{\kappa}})$$

where $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are the maximum likelihood estimators of the two parameters of the log-normal distribution. These

are respectively the sample mean and the sample variance of the logarithm of the observations. The parameters κ and μ are the maximum likelihood estimators of the parameters of the Gamma distribution. We now proceed to develop these estimators.

The density function of the Gamma distribution is given by

$$f(x;\ \mu,\ \kappa) = \left(\frac{\kappa}{\mu}\right)^{\kappa} \frac{x^{\kappa-1} e^{-\frac{\kappa x}{\mu}}}{\Gamma(\kappa)} \qquad 0 < x < \infty \tag{2.1}$$

and the likelihood function can be written

$$L(x_1, \cdots, x_n;\ \mu,\ \kappa) = \left(\frac{\kappa}{\mu}\right)^{n\kappa} \prod_{i=1}^{n} x_i^{\kappa-1} e^{-\frac{\kappa}{\mu}\sum_{i=1}^{n} x_i} / \{\Gamma(\kappa)\}^n$$

Hence

$$\ell n\ L(x_1, \cdots, x_n;\ \mu,\ \kappa) = n\kappa\ \ell n\left(\frac{\kappa}{\mu}\right) + (\kappa-1) \sum_{i=1}^{n} \ell n\ x_i \tag{2.2}$$

$$- \kappa/\mu \sum_{i=1}^{n} x_i - n\ \ell n \Gamma(\kappa)$$

From 2.2 it is clear that $\sum_{i=1}^{n} x_i$, n and $\sum_{i=1}^{n} \ell n\ x_i$ form sufficient statistics for μ and κ. Maximum likelihood estimates can be found by taking the partial derivatives of 2.2 with respect to κ and μ, and setting these equal to zero, resulting in the equations

$$-\frac{n\hat{\kappa}}{\hat{\mu}} + \frac{\hat{\kappa}}{\hat{\mu}^2} \sum_{i=1}^{n} x_i = 0 \qquad (2.3)$$

$$n \ln\left(\frac{\hat{\kappa}}{\hat{\mu}}\right) + n + \sum_{i=1}^{n} \ln x_i - \frac{\sum_{i=1}^{n} x_i}{\hat{\mu}} - n \frac{d\ln\Gamma(\hat{\kappa})}{d\hat{\kappa}} = 0 \qquad (2.4)$$

From equation 2.3, $\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n}$ and substituting this in 2.4 one obtains

$$\left[\ln \kappa - \frac{d\ln\Gamma(\hat{\kappa})}{d\hat{\kappa}}\right] = \ln \hat{\mu} - \frac{\sum_{i=1}^{n} \ln x_i}{n} \qquad (2.5)$$

This equation can be solved for $\hat{\kappa}$ using tables of $\frac{d\ln\Gamma(\kappa)}{d\kappa}$ (See, for instance, tables of the Digamma function in 1.) Results for precincts 7 and 15 are presented in Table 2.8.

Table 2.8

Maximum Likelihood Estimates for Parameters of Gamma Model of Interarrival Times Between Ambulance Calls

	Precinct 7	Precinct 15
$\hat{\mu}$	98.072	101.166
$\hat{\kappa}$	.745	.770
Estimated Coefficient of Variation ($\hat{\kappa}^{-1/2}$)	1.16	1.14

The remaining parameters in T_g are $\alpha_{1,\hat{\kappa}}$ and $\alpha_{2,\hat{\kappa}}$, and these are estimated by

$$\alpha_{1,\hat{\kappa}} = \psi(\hat{\kappa}) - \ell n(\hat{\kappa}/\hat{\mu})$$

$$\alpha_{2,\hat{\kappa}} = \psi'(\hat{\kappa})$$

where $\psi(x)$ is the digamma function and $\psi'(x)$ is its first derivative.

The estimated asymptotic variance of T_g is given by Jackson as:

$$V(T_g) = n \left[\frac{\psi'''(\hat{\kappa})}{4\{\psi'(\hat{\kappa})\}^2} - \frac{\hat{\kappa}\{\psi''(\hat{\kappa})\}^2}{4\{\psi'(\hat{\kappa})\}^2\{\hat{\kappa}\psi'(\hat{\kappa}) - 1\}} + 1/2\right]$$

where $\psi''(x)$ and $\psi'''(x)$ are respectively the second and third derivatives of the digamma function.

The test is implemented by utilizing the asymptotic normality of T_g, and by comparing $T_g/\sqrt{V(T_g)}$ with the α^{th} percentile of the standard normal distribution. Negative values are expected under the log-normal alternative.

The various statistics and the test statistics are presented in Table 2.9.

In both precincts the Gamma hypothesis is accepted at a .05 significance level.

In summary, we see that the Moran test, a test based on the higher moments of the distribution, and a

Table 2.9

Significance Test of Gamma Distribution of Times Between Ambulance Calls Versus Log-Normal Alternative

Statistic	Precinct 7	Precinct 15
n	416	403
$\hat{\alpha}_1$	3.783	3.842
$\hat{\alpha}_2$	2.069	2.055
$\hat{\kappa}$	.745	.770
$\hat{\mu}$	98.072	101.166
$\psi(\hat{\kappa})$	-1.100	-1.037
$\psi'(\hat{\kappa})$	2.567	2.443
$\psi''(\hat{\kappa})$	-1.388	-1.456
$\psi'''(\hat{\kappa})$	2.650	2.868
$\alpha_{1,\hat{\kappa}}$	3.73	3.19
$\alpha_{2,\hat{\kappa}}$	2.567	2.443
T_g	-28.0	-19.5
$V(T_g)$	224	218
$T_g/\sqrt{V(T_g)}$	-1.86	-1.32

likelihood ratio test based on a log-normal alternative all lend credibility to the assumption of a Gamma distribution of times between ambulance calls in a police precinct.

It would be useful to supplement these statistical analyses with analyses based upon the physical mechanism generating calls for ambulances in an urban area. We shall delay this theoretical approach until Section 2.2.3.

2.2.2 Police Emergency Occurrence Process

Police emergencies occurring in Precinct 15 during the one month period November 19-December 14, 1968, were analyzed to provide some insight into this process. During the period there were 4725 such calls. A plot of the number of these occurring per day is presented in Figure 2.6.

From this figure one observes that the series is highly variable but that no obvious trends exist in these data. To test for interactions between days and weeks, a contingency table was developed using counts from the three complete weeks of police call data. The table is presented as Table 2.10.

Table 2.10

Contingency Table to Test for Day-Week Interactions
Police Call Occurrence Process

	Sun.	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.	TOT
Week 1	181	215	173	195	155	193	207	1319
Week 2	151	179	189	145	185	159	185	1193
Week 3	160	165	181	171	211	199	172	1259
TOT	492	559	543	511	551	551	564	3771

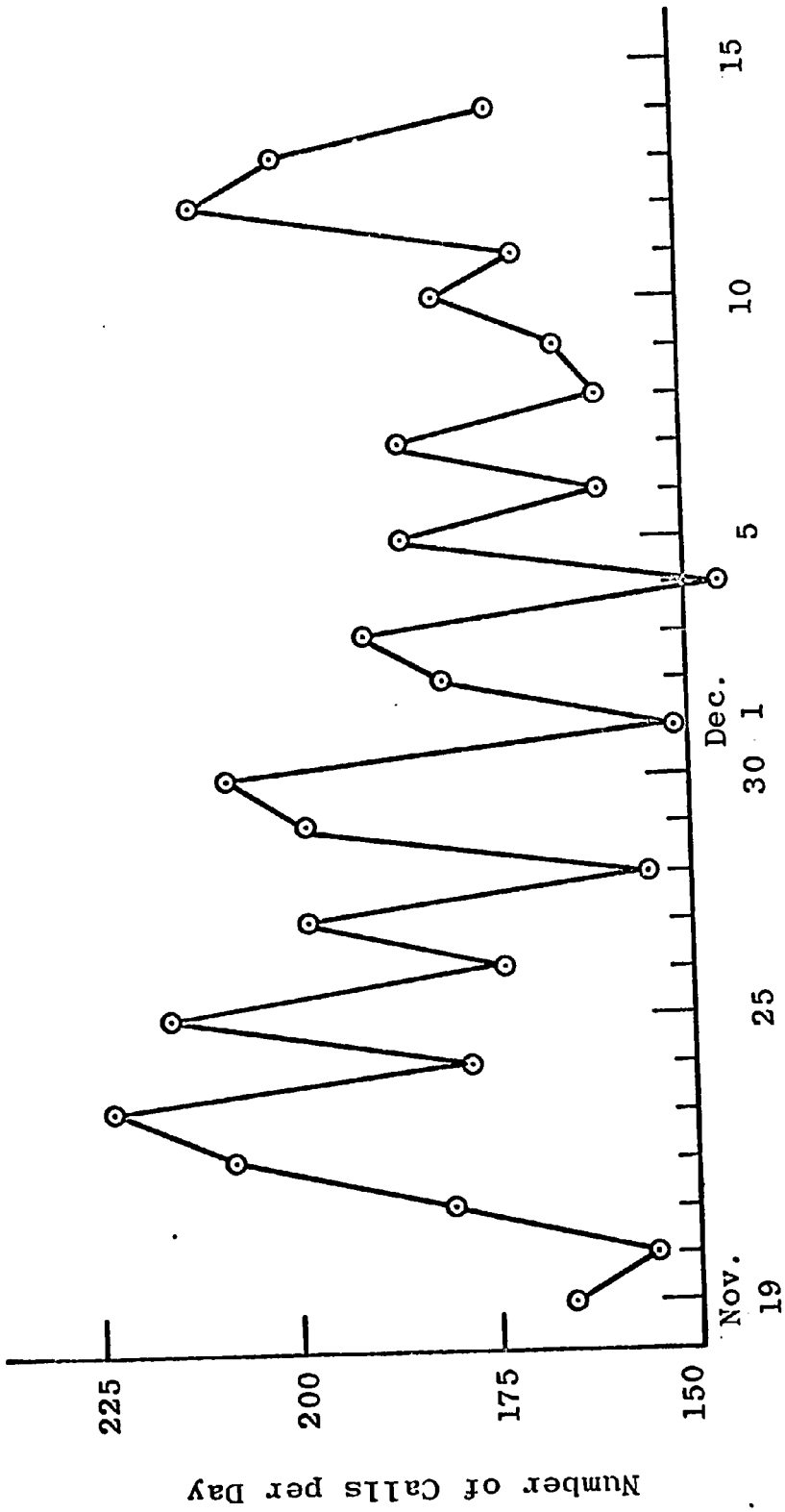


Figure 2.6 Number of Police Calls per Day

Precinct 15 November 19 - December 14, 1969

A Chi-square test was utilized to test for day-week interactions, yielding an attained value of 28.339 with 12 degrees of freedom. The .05 significance level for this test is 21.026 indicating a significant interaction.

However, when one examines the contribution of each cell to the statistic, Thursday of Week 1 is found to contribute 25%--four times the contribution of the next closest cell. This was Thanksgiving Day, and one would expect an abnormally low demand for police services. Hence the cell was corrected to the average of the remaining Thursday counts and the analysis was rerun. The corrected Chi-square statistic was 16.747, indicating no significant interaction. Therefore it was concluded that no significant day-week interactions existed over an average sampling period of this length.

Total daily counts were then tested to ascertain whether these could be distributed uniformly over the seven days of the week. The attained Chi-square statistic was 4.099 with 6 degrees of freedom compared with a .05 significance level of 12.592. This indicates that a uniform weekly distribution of police calls is reasonable for the period and precinct under examination.

The series of counts for the three week period November 19-December 14, 1968, was decomposed into counts for each of the three shifts to study the variation

within each day. The results are shown in Figure 2.7. These indicate that the intensity of police calls in shifts 2 and 3 appears to be greater than that in shift 1. No major differences exist between shifts two and three.

To obtain more insight into the police emergency occurrence process, detailed statistical analyses were conducted on the interarrival times between calls for police assistance on each of the four days November 19-November 22, 1969.[1]

Statistical tests were conducted to examine the observed differences between shifts in more detail. We note that if the intensity of police calls increases during a day, then the interarrival times between such calls must decrease. From this one is led to the conjecture that the expected interarrival time should be a decreasing function of time.

To test this hypothesis,[2] assume that the calls follow a Poisson process with time dependent rate $\lambda(t) = e^{\alpha+\beta t}$. Near $\beta = 0$ this is equivalent to a linear

---

[1] The counts tabulated in Figures 2.6 and 2.7 are based upon these calls as well as all police initiated dispatches. Although information on both types of dispatches is important in determining load on the police system, only calls for assistance are analyzed in detail here.

[2] Discussion follows that of Cox and Lewis (10); pp. 45-47.

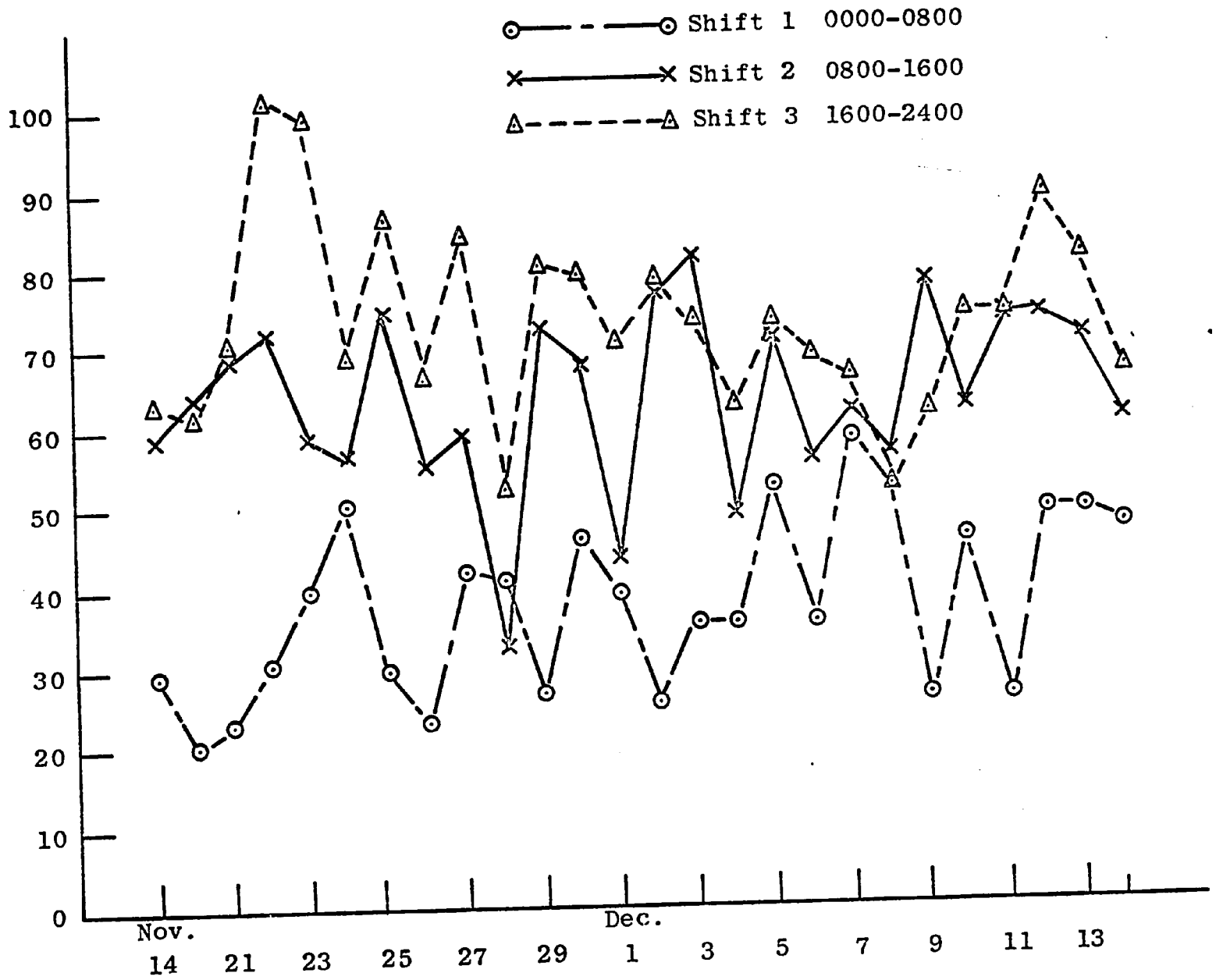


Figure 2.7

Number of Police Calls per Shift

Precinct 15 November 19 - December 14, 1968

trend in the mean of the process. Letting N_t denote the number of events in time t, we may write

$$\Pr\{N_t = n\} = \frac{u_t^n e^{-u_t}}{n!} \tag{2.6}$$

where

$$u_t = \int_0^t \lambda(\nu)d\nu = \int_0^t e^{\alpha+\beta\nu}d\nu = \frac{1}{\beta} e^{\alpha+\beta\nu}\Big]_0^t = \frac{e^{\alpha}}{\beta}(e^{\beta t} - 1)$$

If we observe a time dependent Poisson process for the interval $(0, t_0)$ and if the events occur at times $t_1, t_2, \cdots, t_n$, the likelihood function may be written

$$L(t_1,\cdots, t_n; \lambda) = \lambda(t_1)e^{-u_{t_1}} \lambda(t_2)e^{-(u_{t_2}-u_{t_1})} \cdots$$
$$\lambda(t_n)e^{-(u_{t_n}-u_{t_{n-1}})} e^{-(u_{t_0}-u_{t_n})} = \prod_{i=1}^{n} \lambda(t_i)e^{-u_{t_0}}$$

When $\lambda(t) = e^{\alpha+\beta t}$ we may write

$$L(t_1,\cdots, t_n; \alpha, \beta) = \exp\{n\alpha + \beta \sum_{i=1}^{n} t_i - e^{\alpha}(e^{\beta t_0} - 1)/\beta\}$$

Clearly the two-tuple $(n, \sum_{i=1}^{n} t_i)$ forms a set of sufficient statistics of the parameters α and β. In addition, given β, n is sufficient for α. Hence, when α is treated as a nuisance parameter, the distribution of $z = \sum_{i=1}^{n} t_i$ conditional on n is the appropriate distribution for

inferences on β. This can be found by dividing the joint distribution by the marginal distribution of N_t, given from 2.6 by

$$\Pr\{N_{t_0} = n|\alpha,\beta\} = \frac{1}{n!}\left[\frac{e^{\alpha}}{\beta}(e^{\beta t_0} - 1)\right]^n \exp\left\{\frac{e^{\alpha}}{\beta}(e^{\beta t_0} - 1)\right\}$$

yielding

$$f(z|\alpha,\beta) = \frac{e^{n\alpha}e^{\beta z}\exp\left\{\frac{e^{\alpha}}{\beta}(e^{\beta t_0} - 1)\right\}}{\frac{1}{n!}\left[\frac{e^{\alpha}}{\beta}(e^{\beta t_0} - 1)\right]^n \exp\frac{e^{\alpha}}{\beta}(e^{\beta t_0} - 1)\}} \qquad (2.7)$$

$$= \frac{n!\,\beta^n e^{\beta z}}{(e^{\beta t_0} - 1)^n}$$

We observe that 2.7 is the joint distribution of the order statistics from a truncated exponential distribution with rate parameter β and truncation point t_0. The marginal distribution of this truncated negative exponential random variable is given by

$$f(x;\ \beta,\ t_0) = \frac{\beta e^{\beta x}}{e^{\beta t_0} - 1} \qquad 0 \le x \le t_0,\ \beta \neq 0 \qquad (2.8)$$

Hence, for all $\beta \neq 0$, the conditional distribution of z is the sum of n independent random variables distributed as 2.8 above. For $\beta = 0$ these random variables are uniformly

distributed over the interval $(0, t_0)$ with mean $t_0/2$ and variance $t_0^2/12$. Hence z is asymptotically normal with mean $nt_0/2$ and variance $nt_0^2/12$. A test of the hypothesis $\beta = 0$ can be made by comparing the attained value of $\frac{z - nt_0/2}{t_0\sqrt{n/12}}$ with the upper α^{th} percentile of the standard normal distribution.

This test was conducted on the daily and within-day processes for the four day period November 19-November 22, 1968. Results are presented in Table 2.11.

Table 2.11

Tests for Trends in Daily and Within Day Police Call Processes Precinct 15

Date	Daily	Shifts 2 and 3	Shift 3
11/19/68	2.744*	1.497	N.A.
11/20/68	3.026*	.421	N.A.
11/21/68	1.401	-1.502	N.A.
11/22/68	3.800*	2.041*	-1.343

*Significant at the .05 level.

We observe that evidence of a significant trend is present in the daily series for three of the four days examined. This trend is not present in two of these days when the early morning shift is removed.[1]

[1] The remaining day--November 22--had the largest number of police emergencies of all days during the sampling

These quantitative analyses confirm the earlier observation that significant daily trends do exist in the police emergency occurrence process. However, we note that these generally can be removed by decomposing the series into two components, one consisting of the early morning shift and the other made up of remaining shifts.

Tests for dependencies between interarrival times other than the low frequency trend observed above were based upon the estimated periodogram (refer to Section 2.2.1 for a detailed discussion of these). The relevant test statistics are presented in Table 2.12 below.

For three out of the four days, the null hypothesis of no significant autocorrelations between intervals (the renewal hypothesis) is accepted for all tests. The attained values of the lower-sided Kolmogorov Smirnoff statistic and the Anderson Darling statistic indicate rejection of the hypothesis for the November 22 data. Again, we observe that this day represents an extreme intensity of police calls and thus can be considered as an "outlier."

From these two analyses it was decided to model the decomposed police emergency occurrence process as a renewal process. For the two components, tests were con-

---

period, with a very large number of these occurring during the third shift.

Table 2.12

Tests of Renewal Hypothesis for
Times Between Police Calls
Precinct 15

Statistic	Value for 11/19/68	Value for 11/20/68	Value for 11/21/68	Value for 11/22/68	.05 Significance Level
Upper-Sided Kolmogorov Smirnoff	.459	.599	.286	.099	1.224
Lower-Sided Kolmogorov Smirnoff	1.009	.804	.986	1.341*	1.224
Two-Sided Kolmogorov Smirnoff	1.009	.804	.986	1.341	1.358
Anderson Darling	1.160	1.271	1.624	3.285*	2.492

*Significant at the .05 level.

ducted to determine if the resulting marginal distributions of intervals between calls could be adequately modeled by negative exponential distributions. Data from November 22 were excluded from these analyses because of their deviant nature discussed earlier. Results are presented in Table 2.13.

Table 2.13

Tests for Negative Exponential Distribution of Times Between Police Calls November 19-21 Precinct 15

Statistic	Attained Value-- Shift 1	Attained Value-- Shift 2 & 3	.05 Significance Level
Upper-Sided Kolomogorov Smirnoff	.905	.734	1.224
Lower-Sided Kolomogorov Smirnoff	.439	.731	1.224
Two-Sided Kolomogorov Smirnoff	.905	.734	1.358
Anderson Darling	1.006	1.257	2.492
Moran	39.746	214.589	(43.9, 216.1)

All tests indicate that the decomposed processes are adequately modeled as Poisson processes. Some sample statistics for these data are presented in Table 2.14.

Table 2.14

Sample Statistics for Times between Police Calls
11/19/68-11/21/68
Precinct 15

Statistic	Shift 1	Shifts 2 and 3
Mean	42.742 min.	17.223 min.
Variance	2675.798 min.	327.420 min.
Coefficient of Variation	1.210	1.051
Coefficient of Skewness	2.267	1.715
Range	(2-219 min.)	(1-94 min.)

These statistics show the large difference that exists between the occurrence rates in the two decomposed series. The mean time between events in shift 1 is almost three times that in shifts 2 and 3. It is well known that the sample mean is the maximum likelihood estimator for the reciprocal of the parameter of the negative exponential distribution, and we shall use this estimator here. Hence, in summary, the data analyses indicate that the police emergency occurrence process is adequately modeled by decomposing the daily series into two sub-series. These series can then be considered as realizations of renewal processes with negative exponentially distributed times between events and with mean times between events of seventeen minutes and forty-three minutes respectively.

2.2.3. Interactions between Police and Medical Emergency Occurrence Processes

Having determined the properties of the two series of events independently, it is natural to examine potential interdependencies between the series. A priori, there is no strong theoretical argument for such dependencies. One would expect that police and ambulance calls evolve independently over time. However, there is some possibility that such calls are interrelated. For instance, in an urban area a large number of ambulance calls may be generated as the result of criminal activity. In this case ambulance calls may follow police calls with a higher frequency than one would expect under a hypothesis of no interaction.

To examine the possibility of interactions between the two types of emergencies, let ambulance calls be indicated by the index 0 and let police calls be indicated by the index 1. Let P_{ij} be the probability that a call is of type j conditional upon the previous call being of type i. For instance P_{10} is the probability that an ambulance call follows a police call. Under the assumption that the two types of call are independent $P_{10} = P_{00} = P_{.0}$, where $P_{.0}$ is the unconditional probability of an ambulance call. Letting n_{ij} equal the number of calls of type j following calls of type i, the one month sample in precinct 15 was

partitioned accordingly. Results are presented in the contingency table below (Table 2.15). For instance, reference to the table indicates that ambulance calls were followed by ambulance calls 42 times during the one month period.

Table 2.15

Frequency Table for Transition Matrix
Between Types of Calls
Precinct 15

	0	1	Total
0	42	373	415
1	373	4,352	4,713
Total	415	4,713	5,128

Estimated Transition Matrix
Between Types of Call

	0	1
0	.101	.899
1	.079	.921

A Chi-square test was used to test for independence in this table. The attained statistic was 2.256 compared with a significance level of 3.84 (1 degree of freedom). The statistical analysis leads us to reject the hypothesis of Markovian transitions between the two types of calls.

The maximum likelihood estimates of the unconditional probabilities of the two types of emergency are then given

by $\hat{P}_{.0} = 415/5128 = .081$ and $\hat{P}_{.1} = 4713/5128 = .919$.

We observe that these probabilities reflect the influence of both calls from the public for police assistance and police-initiated vehicle dispatches. These police-initiated activities include emergencies detected by police vehicles, and dispatches authorizing the release of vehicles for activities such as court appearances, vehicle maintenance, and meals. Since the latter events are not emergency occurrences and can be "scheduled" in advance, it is useful to examine system operation when these are removed from the series of emergency occurrences. In this case, the proportion of ambulance calls increases to approximately .214; no significant interactions are present in the resulting series.

A realization of this series is presented below

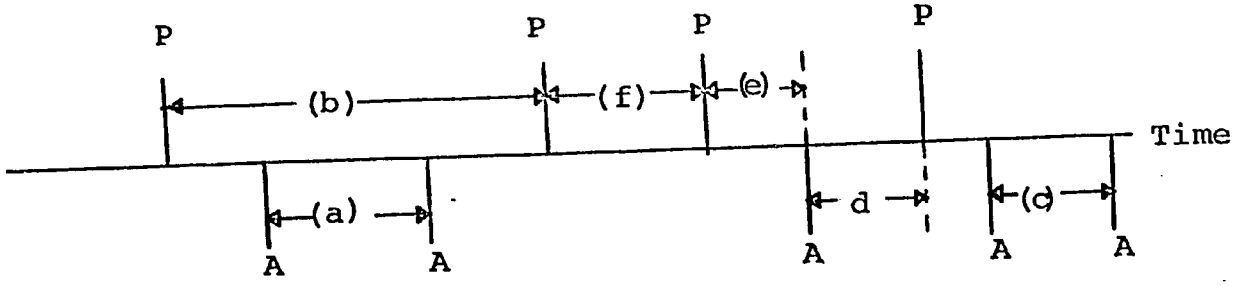


Figure 2.8

Realization of Semi-Markov Emergency Occurrence Process

We have been examining the statistical properties of the times between ambulance calls (the event a) and the times between police calls (the event b). However, a complete

description of the interactive process requires us to consider the stochastic processes which yield the times between the following four types of events.

(i) Ambulance call followed by ambulance call (event c).

(ii) Ambulance call followed by police call (event d).

(iii) Police call followed by ambulance call (event e).

(iv) Police call followed by police call (event f).

In Chapter III we shall see that the resulting stochastic process is called a semi-Markov process. The detailed properties of this type of process are discussed in Section 3.1. In this section we shall develop the statistical inferences necessary to utilize this model in the analysis of a recovery system which handles both ambulances and police calls.

Almost no literature exists on techniques for statistical inferences on semi-Markov processes. Estimators are required for the state transition probabilities (P_{ij}) and for the distribution functions describing the times spent making transitions, denoted $F_{ij}(x)$. Moore and Pyke (22) discuss the problem in some detail, and they have develcped certain non-parametric and maximum likelihood estimators for these processes. Further, they have shown that when the transition probabilities are func-

tionally independent of the distribution function, these two components of the process can be estimated separately. In this case the maximum likelihood estimators of the P_{ij} are equal to n_{ij}/n, and the estimates for these developed earlier in this section apply exactly. Since we have found that the transitions between events can be represented by independent Bernoulli random variables, the transition probabilities are estimated by .214, the unconditional probability of an ambulance call. We denote this parameter by π, and its maximum likelihood estimator by $\hat{\pi}$.

The development of parametric inferences for the distribution functions of the random times between events requires the specification of a class of potential distribution functions and the estimatron of the parameter of this class. When the series of emergency occurrences is studied in this detail, the resulting sample sizes available to estimate the distributions of interarrival times for all four events are quite small. Nevertheless, data for a single week were examined to see what insights could be obtained into the underlying processes. The statistical tests of Appendix A were utilized to determine the usefulness of the negative exponential distribution as a model for the interarrival time. Results are

presented below (Table 2.16). In this table E_{ij} denotes the event where emergency type j follows emergency type i.

Table 2.16

Tests for Negative Exponential Distribution of Time Between Events in Semi-Markov Occurrence Process

Statistic	E_{00}	E_{01}	E_{10}	E_{11}
Upper-Sided Kolomogorov-Smirnoff	.571	.163	1.048	.319
Lower-Sided Kolomogorov-Smirnoff	1.029	.604	.988	1.933*
Two-Sided Kolomogorov-Smirnoff	1.029	.604	1.048	1.933*
Anderson-Darling	1.494	.966	1.879	5.144*
Moran	2.833	40.435	65.881	227.196
Sample Size	18	49	49	202

*Significant at the .05 level.

These results indicate that the negative-exponential distribution provides an adequate model for $F_{00}(x)$, $F_{01}(x)$, and $F_{10}(x)$. However, departures from the exponential process are indicated for $F_{11}(x)$. Closer inspection of the data from this process indicates that these departures are due to the occurrence of a small number of unusually long times. It is probable that the shorter times arises as the result of multiple dispatches to the same police call

and that the longer times represent single vehicle dispatches to an emergency.

In this case the interarrival times in the overall process might be modeled as a compound process, consisting of two negative exponential distributions, one with an arrival rate corresponding to multiple vehicle emergencies, and one with a longer arrival rate corresponding to the less frequent single vehicle dispatches. However, this level of sophistication is unnecessary in the present analysis, for the implementation of such a model yields numerical results in close agreement with a much simpler model. In fact, utilization of the negative exponential distribution as a model for $F_{11}(x)$ yields results which are extremely accurate. (We shall discuss this further in Section 5.1, p.171.) Therefore, a negative exponential approximation for $F_{11}(x)$ is used for parameter estimation in this section.

Maximum likelihood estimates for the mean time between the four types of events, based on a negative exponential distribution, are presented in Table 2.17.

We note that the semi-Markov process is completely defined when these four estimates and the estimate of the proportion of ambulance calls in the series are specified.

The fact that the negative exponential distribution provides an adequate model for the interarrival times in this semi-Markov emergency occurrence process is not

Table 2.17

Maximum Likelihood Estimates for Mean Occurrence Times in Semi-Markov Arrival Process

Event	Estimated Mean Time
E_{00}	14.33 min.
E_{01}	16.20 min.
E_{10}	23.08 min.
E_{11}	14.55 min.

surprising. In fact, it is very reasonable to assume that the occurrences of police and medical emergencies are distributed randomly over time, and this assumption leads to the negative exponential distribution of times between events which has been identified in this section.

The question remains, however, whether the semi-Markov model with negative exponential times between events is consistent with the Gamma distribution of times between ambulance calls developed in Section 2.2.1 and with the negative-exponential distribution of times between police calls developed in Section 2.2.2. In Appendix B we derive the Laplace-Stieltjes transforms of the distribution functions of these two processes from the Laplace-Stieltjes transform the semi-Markov

emergency occurrence process discussed in this section. We then utilize these results to indicate why the earlier empirical results on the decomposed ambulance and police processes provide reasonable models for these phenomena.

2.3. ANALYSIS OF EMERGENCY SERVICE PROCESSES

We turn now to a discussion of the service processes for both medical and police emergencies. The sequences of service times are again analyzed as stochastic processes. However, in this situation non-stationarities and interdependencies should not be present in the series, since successive emergencies are generally served by different vehicles under different environmental conditions. Thus there is strong prior information that the service intervals are realizations of independent random variables.

However, the intervals may not be identically distributed. In fact, if distance to the scene affects service time, then intervals corresponding to different distances should have different distributions. Since vehicle allocation affects recovery system operation through this distance-service time relationship, it is important to study the concept in some detail. Such an analysis is conducted in Section 2.3.1. in the consideration of the ambulance service time process. The police service time

process is discussed in 2.3.2.

2.3.1. Medical Emergency Service Process

As we saw in Chapter I, the ambulance service time can be conceptualized as the sum of four independent random variables--transit time to the scene, time on the scene, transit time to the hospital, and time at the hospital. We also observed that the effect of vehicle redistribution is to reduce the average distance between recovery vehicles and the sites of medical emergencies. Theoretically, such reductions should result in decreased transit times to emergencies and consequently decreased service times for these emergencies.

In this section we shall use data from the City of Detroit to analyze these distance-time relationships in more detail. This analysis will lead to a model for the medical emergency service process which incorporates the effects of alternative vehicle distribution policies. Since there is no reason to expect substantial differences in the distance-time relationships for different precincts, data from all four precincts (6, 7, 15, and 16) which was recorded during the period November 18-December 15, 1968 are included in the analysis.

To obtain some preliminary insight into the relationship between travel distance and travel time to

the scene, the product moment correlation coefficient between these two variables was computed. The sample correlation coefficient for a sample of 558 runs was .492, indicating that approximately 25% of the variation in travel times could be explained by distance. This figure was lower than one might have anticipated, and it indicates that the effect of improved vehicle allocation on transit time to the scene of emergencies may not be great in an urban area. To study this problem in more detail a sub-sample of size 50 from precinct 15 was subjected to further statistical analysis. The use of transformations (log, square root) on the dependent variable (travel time to the scene) did little to improve the quality of the linear regression between travel time and distance.

Examination of a scatter plot between the two variables confirmed the tendency of travel time to be biased toward five minute intervals. It was felt that this bias might be confounding the underlying relation. Thus, recorded five minute times were uniformly "redistributed" about a mean of five minutes. This procedure did nothing to improve the empirical relationship between the variables or transformations of these.

The inability to reduce the unexplained variation between travel time and distance by manipulating the data

led to a more detailed consideration of the underlying relationship. From this one can hypothesize at least two explanations for the lack of a high correlation between these variables in an urban area. First, within a large city there are many uncontrolled factors which contribute to the transit time to the scene of emergencies. These include adverse weather, heavy traffic, vehicular mechanical problems, errors in locating the incident, and unforeseen obstacles such as blocked railroad crossings, etc. The effects of these uncontrolled factors and their interactions can conceivably dominate the effects of travel distance for many situations. In these situations, those changes in vehicle distribution resulting in minor reductions in average transit distance to the scene of emergencies will not significantly reduce response time.

Second, it is conceivable that ambulance response characteristics are different for "nearby" and more distant emergencies. A significant portion of the transit time to nearby emergencies is apt to involve the accomplishment of certain "fixed" tasks--acknowledgement of the call, determination of the route, and initiation of the run. Travel to such emergencies will generally proceed over regular city streets. On the other hand, when longer distances separate the vehicle and the emergency,

expressway travel at higher velocities is likely to be utilized.

These considerations seem to indicate that precise relationships between transit time and the transit distance should not be expected in urban areas when distance is measured on an interval scale. Alternatively, the relationship might be studied using distance measured in a ordinal scale constructed from the original interval scale. To examine this hypothesis, empirical conditional distributions of emergency service times[1] for various ranges of travel distances were constructed. Table 2.18 presents sample statistics for the empirical distributions of service time conditioned on distances to the scene of less than one mile, two to three miles, and four or more miles.

The similarity of service time statistics for distances greater than two miles tends to support the hypothesis

---

[1] We have previously observed that the service time can be conceptualized as the time from dispatch until the vehicle is released from the hospital. In the data analysis of this section this definition has been modified in two ways. First, there are many calls which result in no conveyance. The service times for these are taken as the time from dispatch until the driver of the vehicle reports he is again available. Second, it was found that an abnormally long time is spent by the police making hospital transfers. This time interval was not present in emergency calls answered by a professional ambulance service which was implemented on an experimental basis as part of the project. Hence, the time spent at the hospital has been excluded from these estimated service times.

Table 2.18

Statistics for Empirical Conditional Distributions of Ambulance Service Times

Statistic	Distance to Scene		
	0-1 Mile	2-3 Miles	4+ Miles
Mean (min.)	18.406	20.361	20.769
Variance (min.^{2})	87.893	90.469	120.722
Coefficient of Variation	.509	.467	.529
Coefficient of Skewness	1.504	1.860	1.955
Range (min.)	(4, 60)	(6, 70)	(1, 75)
Sample Size	224	228	108

of differing response characteristics for "nearby" and more distant emergencies. The similarity of the mean service times for all three distance intervals indicates that vehicle redistribution cannot dramatically reduce emergency service times in an urban area. This is true because transit time to the scene comprises only about one-third of the service time. Hence reductions in transit time result in proportionally smaller reductions in total emergency service time.

As a result of these analyses, it was decided that ambulance service times could be adequately modeled by

conditioning on distances to the scene of less than or equal to one mile and on distances greater than one mile. The empirical distribution functions of the resulting conditional distributions were then tested to determine if these followed negative exponential distributions. The results of these statistical tests are presented in Table 2.19.

Table 2.19

Tests for Negative Exponential Distribution of Ambulance Service Times

Statistic	Less Than or Equal to One Mile	Greater Than One Mile	.05 Significance Level
Upper-Sided Kolmogorov Smirnoff	.058	.341	1.224
Lower-Sided Kolmogorov Smirnoff	.742	.570	1.224
Two-Sided Kolmogorov Smirnoff	.742	.570	1.358
Anderson Darling	.915	.341	2.492
Moran	53.897	78.664	(260,442)

The results of these tests indicate that the conditional service time distributions are fitted very well

by negative exponential distributions. The mean service time for distances less than or equal to one mile is 18.406 minutes and the mean time for distances greater than one mile is 20.472 minutes.

Although the difference between these mean times is not large, it will be insightful to study the effect of this difference on alternative allocation and distribution policies. In addition, the effect of reparameterizing the model to increase (and decrease) these mean differences should show the sensitivity of the model to changes in the distance-time relationships studied in this section.

2.3.2. Police Emergency Service Process

Since our objective in conducting an analysis of police emergencies is to examine how these emergencies influence the operation of a "dual function" police ambulance system, it is unnecessary to study the relationship between this process and distance to the scene. In addition, there is no prior reason to expect interdependencies in the series of service times. Therefore the series can be analyzed as a renewal process.

Service times for 973 sequential police calls were obtained from precinct 15 for the period November 18-November 29, 1968. Statistical analyses were then

conducted on these to determine whether the marginal distribution of service times could be modeled by a negative exponential distribution. The results of these analyses are presented in Table 2.20.

Table 2.20

Tests for Negative Exponential Distribution of Police Service Times Precinct 15

Statistic	Attained Value	.05 Significance Level
Upper-Sided Kolmogorov Smirnoff	.010	1.224
Lower-Sided Kolmogorov Smirnoff	.722	1.224
Two-Sided Kolmogorov Smirnoff	.722	1.358
Anderson Darling	1.195	2.492
Moran	416.51	1049.0

All tests indicate that the hypothesis of a negative exponential marginal distribution should be accepted. Hence we shall utilize this distribution as a model for service times for police calls.

Sample statistics for the above data are presented in Table 2.21. The mean time for servicing police calls was thirty-eight minutes compared with mean time of 18-20

Table 2.21

Sample Statistics for
Police Call Service Times
Precinct 15

Statistic	Attained Value
Mean	38.282 min.
Variance	522.979 min.^{2}
Coefficient of Variation	.597
Coefficient of Skewness	1.049
Range	(1,171) min.

minutes for ambulance calls.[1]

2.4 SUMMARY OF DATA ANALYSES

We now briefly summarize the statistical inferences which have been drawn from the various data analyses of this chapter. Data on the interarrival times between medical emergencies in two city precincts indicate that these are renewal processes. The marginal distribution of times between events can be modeled as a Gamma distribution with a mean time between emergencies estimated at 98-101 minutes. The coefficients of variation of the distributions were estimated to be 1.14-1.16.

---

[1] Again, the time for servicing ambulance calls was estimated after excluding time spent at the hospital.

Analysis of data on police emergencies in precinct 15 indicates that a daily trend exists, but that this trend can be removed by decomposing the series into two components--sequences of events in shift one and those in shifts two and three. Both sub-series are adequately modeled by negative exponentially distributed times between events. The estimated mean time between police calls is forty-three minutes in shift one and seventeen minutes in shifts two and three.

Transitions between police and ambulance calls were found to occur at random. The unconditional probability of an ambulance call was estimated at .214 when only police emergencies are considered. The distributions of the random times between state transitions were found to follow a semi-Markov process with negative-exponential distributions of times between events.

Data on the service processes for medical emergencies in four precincts were utilized to make inferences on the properties of these. It was found that distance to the scene (and hence recovery vehicle distribution) influences emergency service times, but not as a continuous interval variable. Analyses of the empirical conditional distributions of emergency service times indicated that transit distances of less than one mile resulted in measurably different average service times

than those greater than one mile. No other significant differences were observed. The observed differences were attributed to the varying procedures for responding to "nearby" and more distant emergencies. Both conditional distributions were found to be adequately modeled as negative exponential random variables.

Data on police emergency service processes were analyzed for a twelve day period in precinct 15. Statistical tests indicated that this service process also obeys a negative exponential distribution. The mean service time for police calls was thirty-eight minutes compared with a corrected mean time of 18-20 minutes for ambulance calls.

From these inferences one concludes that the unconditional emergency service time for a recovery vehicle is a random variable formed by the weighted sum of the three component service times. The density function of this random variable is therefore a mixture of negative exponential random variables. The mixing probabilities are functions of the dispatch policy, the placement of recovery vehicles, and the distribution of emergencies. We shall discuss these in more detail in Chapter III.

We now utilize the insights gained from the data analysis to develop a recovery system model suitable for

examining the effects of recovery vehicle allocation and distribution policies. The parameter estimates obtained in this chapter will be utilized as guidelines for the implementation of this model. In addition, the sensitivity of the model outputs to parameter variation will be explored in some detail.

CHAPTER III

DEVELOPMENT OF THE RECOVERY SYSTEM MODEL

The objective of this chapter is to develop a mathematical model suitable for a study of the recovery vehicle distribution and allocation problem. This model will be used to examine:

(1) The effect of changing the number of recovery vehicles in the system.

(2) The effect of changing the distribution of these vehicles throughout the service region.

(3) The effect of modifying or eliminating the dispatch policy which assigns both police and emergency medical calls to the recovery vehicles.

We have chosen to model the emergency recovery problem as a system of queues in parallel under a multi-queue channel selection mechanism. Cinlar (7) has developed a general theory for parallel queues when the service channel selection mechanism is either (a) Markovian, (b) dependent upon the arrival process, or (c) dependent upon the state of a single queue. In the present section we extend these results to include channel selection switches based upon the joint states of all queues

in the system. We then apply this model to the recovery vehicle allocation and distribution problem. Finally, we show how the effects of the three factors listed above can be studied using the insights obtained from this model.

3.1 DEVELOPMENT OF THE GENERAL THEORY

Let arrivals to a system of parallel queues form a semi-Markov process defined by the bivariate stochastic process $\{X_n, Z_n\}$ where

X_n = Interarrival time between the n^{th} and $(n-1)st$ occurrence. That is, $X_n = T_n - T_{n-1}$ $(n \geq 1)$ where T_j is the instant of arrival of the j^{th} customer.

Z_n = Type of the n^{th} arrival. Assuming m classes of customers, Z_n assumes the states 0, $\cdots$, m - 1.

The theory of semi-Markov processes is well documented, and we shall review only the portions of this theory which are essential to this discussion. The application of semi-Markov processes to queueing systems has been considered by Neuts (25) and Cinlar (8). Such processes constitute a generalization of the use of renewal processes as models of arrival and/or service processes.

We define the two-dimensional stochastic process $\{X_n, Z_n; n \geq 0\}$ as that satisfying the initial conditions

$$P_r\{X_0 = 0\}=1 \quad P_r\{Z_0 = j\} = p_0(j) \qquad j = 0, \cdots, m-1$$

$$p_0(j) \geq 0 \qquad \sum_{j=0}^{m-1} p_0(j) = 1$$

and

$$P_r\{X_n \leq x, Z_n = j | X_j, \cdots, X_{n-1}; Z_0, \cdots, Z_{n-1}\} =$$

$$P_r\{X_n \leq x, Z_n = j | Z_{n-1}\}$$

Such a stochastic process is called a semi-Markov process. One may think of this process as a Markov chain $\{Z_n; n \geq 0\}$, where the time the system spends in any state is a renewal random variable depending on the present and previous states.

Let

$$A_{ij}(x) = P_r\{X_n \leq x; Z_n = j | Z_{n-1} = i\} \qquad n \geq 1$$

and let $A(x)$ be the matrix whose (i,j) element is $A_{ij}(x)$, $(i,j) = 0, \cdots, m-1$. Then the $A_{ij}(x)$ are probability distribution functions satisfying

$$\sum_{j=0}^{m-1} A_{ij}(\infty) = 1 \qquad i = 0, 1, \cdots, m-1$$

The following results follow immediately from these definitions:

$$P_r\{X_n \le x | Z_{n-1} = i\} = \sum_{j=0}^{m-1} A_{ij}(x)$$

$$P_r\{Z_n = j | Z_{n-1} = i\} = A_{ij}(\infty)$$

$$P_r\{X_n \le x | Z_n = j,\ Z_{n-1} = i\} = \frac{A_{ij}(x)}{A_{ij}(\infty)} = A_{ij}^{-1}(\infty)A_{ij}(x) = F_{ij}(x)$$

We now consider the system of parallel queues represented in Figure 3.1 below.

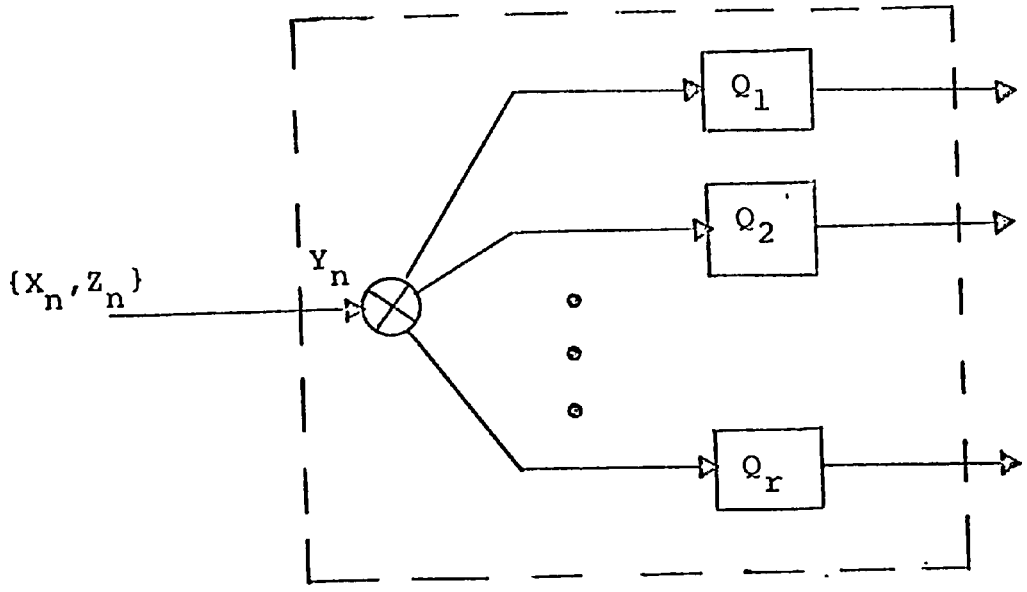


Figure 3.1 System of Queues in Parallel

Arrivals, obeying the semi-Markov process $\{X_n, Z_n\}$ are routed either to one of the r parallel queueing channels $Q_1, \cdots, Q_r$, or they are routed past the system by the switch Y_n. That is, the channel-selection switch is a random variable Y_n taking on the possible values

1, ···, r if the arrival is routed into the system and taking on the value 0 if the arrival is routed past the system. Let S_n^j, j = 1, ···, r denote the state of the j^{th} queue at the instant of the n^{th} arrival to the system. Then the r dimensional process $\{S_n, \cdots, S_n^r\}$ is an imbedded stochastic process whose realizations consist of the vectors $(s_1, \cdots, s_r)$ where $s_j = 0, \cdots, N_j - 1$; j = 1, ···, r. Here N_j is the number of possible states assumed by the j^{th} queueing sub-system. We shall assume that N_j is finite for all j.

For instance, S_n^j might be the number of customers in the j^{th} queueing sub-system at the instant of the n^{th} arrival to the system. In this case $N_j - 1$ is the capacity of this sub-system.

Throughout this development it will be assumed that service times in each channel follow a negative exponential distribution with mean $1/\beta_j$ and that these service processes are independent. (In Section 3.2 we shall extend this formulation to allow the mean service times in each channel to depend upon the type of emergency being serviced.)

We now proceed to develop a probability model for the switch Y_n. Let

$$P_r\{Y_n = k | Y_0, \cdots, Y_{n-1}; Z_0, \cdots, Z_n; S_0^1, \cdots, S_n^1; \cdots;$$

$$S_0^r, \cdots, S_n^r\} = P_r\{Y_n = k | Z_n, S_n^1, \cdots, S_n^r\}$$

$$k = 0, 1, \cdots, r$$

That is, we assume that the probability a customer is assigned to the k^{th} queue or is rejected by the system depends only on the properties of the customer and on the joint state of the r queueing sub-systems at the instant of the customer's arrival. There is no dependence on the "history" of the queueing system or upon the "history" of the arrival stream.

This type of channel selection switch prevails in many applications. For instance, a customer of a certain type may demand a class of service which can be rendered only by a certain sub-set of service channels. Alternatively, a customer seeks service by selecting the channel with the shortest waiting line.

We shall adopt the notation

$$P_r\{Y_n = k | Z_n = j; S_n^1 = s_1, \cdots, S_n^r = s_r\} = b_k(j, s_1, \cdots, s_r)$$

$$= b_K(j, \underline{s}) \quad \text{where } \underline{s} = (s_1, \cdots, s_r) \quad k = 0,1,\cdots,r$$

Also we shall represent the vector-valued stochastic process $\{S_n^1, S_n^2, \cdots, S_n^r\}$ by $\underline{S}_n$.

Under these assumptions it can be shown that the process $\{Z_n, S_n^1, \cdots, S_n^r, X_n; n = 0, 1, \cdots,\}$ is semi-Markov.

Proof.

$$P_r\{Z_n = k, \underline{S}_n = \underline{t}, X_n \leq x | Z_0, \cdots, Z_{n-1}; \underline{S}_0, \cdots, \underline{S}_{n-1};$$

$$X_1, \cdots, X_{n-1}\} = \int_0^x P_r\{\underline{S}_n = \underline{t} | Z_0, \cdots, Z_n; \underline{S}_0, \cdots, \underline{S}_{n-1};$$

$$X_1, \cdots, X_{n-1}; X_n = \nu\}x \qquad (3.1)$$

$$d_\nu P_r\{Z_n = k, X_n \leq \nu | Z_0, \cdots, Z_{n-1}; \underline{S}_0, \cdots, \underline{S}_{n-1};$$

$$X_0, \cdots, X_{n-1}\}$$

Since the service times are negative exponential, $\underline{S}_n$ depends only on $\underline{S}_{n-1}$, X_n, and Y_{n-1}. However, the switch Y_{n-1} depends only on Z_{n-1} and $\underline{S}_{n-1}$. Hence, given X_n, Z_{n-1}, and $\underline{S}_{n-1}$, the probability structure of $\underline{S}_n$ is completely determined.

Similarly, since $\{X_n, Z_n\}$ is semi-Markov, its probability structure depends only on Z_{n-1}. Hence, (3.1) can be written

$$P_r\{Z_n = k, \underline{S}_n = \underline{t}, X_n \leq x | Z_0, \cdots, Z_{n-1}; \underline{S}_0, \cdots, \underline{S}_{n-1};$$

$$X_1, \cdots, X_{n-1}\} = \int_0^x P_r\{\underline{S}_n = \underline{t} | \underline{S}_{n-1}, Z_{n-1}, X_n = \nu\}$$

$$d_\nu P_r\{Z_n = k, X_n \leq \nu | Z_{n-1}\} = P_r\{Z_n = k, \underline{S}_n = \underline{t}, X_n \leq x | Z_{n-1}, \underline{S}_{n-1}\} \quad (3.$$

as required.

An immediate consequence of this result is that the process $\{Z_n, \underline{S}_n; n \geq 0\}$ is a Markov chain, and that the process $\{\underline{S}_n; n \geq 0\}$ is a Markov chain if and only if $m = 1$, that is, if and only if the arrival process is a renewal process (sometimes denoted a GI process in queueing literature).

We shall now derive the transition probabilities for the process $\{Z_n, \underline{S}_n, X_n\}$ and for the underlying Markov chain $\{Z_n, \underline{S}_n\}$. To develop these, let

$$\gamma_k(\underline{t}, \underline{s}, x) = P_r\{\underline{S}_n = \underline{t} | \underline{S}_{n-1} = \underline{s}, Y_{n-1} = k, X_n = x\}$$
$$k = 0, \cdots, r$$

Because of the independence of service times in each of the r queues, this relation can be written

$$\gamma_k = P_r\{S_n^1 = t_1 | S_{n-1}^1 = s_1, Y_{n-1} = k, X_n = x\}x$$
$$P_r\{S_n^2 = t_2 | S_{n-1}^2 = s_2, Y_{n-1} = k, X_n = x\}x$$
$$\vdots$$
$$xP_r\{S_n^r = t_r | S_{n-1}^r = s_r, Y_{n-1} = k, X_n = x\}$$

$$= \prod_{i=1}^{r} \phi_k(t_i, s_i, x) \text{ where}$$

$$\phi_k(t_i, s_i, x) = P_r\{S_n^i = t_i | S_{n-1}^i = s_i, Y_{n-1} = k, X_n = x\}$$

Then

$$P_r\{Z_n = j,\ \underline{S}_n = \underline{t},\ X_n \le x | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s}\}$$

$$= \int_0^x P_r\{\underline{S}_n = \underline{t} | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s},\ X_n = x\} d_x P_r\{Z_n = j,$$

$$X_n \le x | Z_{n-1} = i\}$$

But

$$P_r\{\underline{S}_n = \underline{t} | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s},\ X_n = x\}$$

$$= \sum_{k=0}^{r} P_r\{\underline{S}_n = \underline{t},\ Y_{n-1} = k | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s},\ X_n = x\}$$

$$= \sum_{k=0}^{r} P_r\{Y_{n-1} = k | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s}\} P_r\{\underline{S}_n = \underline{t} | Y_{n-1} = k,$$

$$\underline{S}_{n-1} = \underline{s},\ X_n = x\} = \sum_{k=0}^{r} b_k(i, \underline{s}) \gamma_k(\underline{t}, \underline{s}, x)$$

$$= \sum_{k=0}^{r} b_k(i, \underline{s}) \prod_{\ell=1}^{r} \phi_k(t_\ell, s_\ell, x)$$

Thus

$$P_r\{Z_n = j,\ \underline{S}_n = \underline{t},\ X_n \le x | Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s}\}$$

$$= \int_0^x \sum_{k=0}^{r} b_k(i, \underline{s}) \prod_{\ell=1}^{r} \phi_k(t_\ell, s_\ell, x)\ dA_{ij}(x)$$

$$= \sum_{k=0}^{r} b_k(i, \underline{s}) \int_0^x \prod_{\ell=1}^{r} \phi_k(t_\ell, s_\ell, x) dA_{ij}(x)$$

From this relation, the one-step transition probabilities for the underlying Markov chain $\{X_n, Z_n\}$ can be written

$$P_r\{Z_n = j,\ \underline{S}_n = \underline{t} \mid Z_{n-1} = i,\ \underline{S}_{n-1} = \underline{s}\}$$

$$= \int_0^\infty \sum_{k=0}^{r} b_k(i,\underline{s}) \prod_{\ell=1}^{r} \phi_k(t_\ell,\ s_\ell,\ x)\, dA_{ij}(x) \qquad (3.3)$$

When there is only one type of arrival, m = 1 and $A_{ij}(x) = G(x)$. Then the one-step transition probabilities for the Markov chain $\{\underline{S}_n\}$ can be written

$$P_r\{\underline{S}_n = \underline{t} \mid \underline{S}_{n-1} = \underline{s}\} = \int_0^\infty \sum_{k=0}^{r} b_k(\underline{s}) \prod_{\ell=1}^{r} \phi_k(t_\ell,\ s_\ell,\ x)\, dG(x)$$

These results can be conveniently put in matrix form as follows. Assume for notational convenience that all queues in the system have the same number of states. Then $N_j = p$ and $S_n^j = 0, 1, \cdots, p-1$, for $j = 1, \cdots, r$. Let B_k be a $mp^r \times mp^r$ diagonal matrix whose $(ixp^r + p^{r-1}s_1 + p^{r-2}s_2 + \cdots + s_r + 1)^{st}$ entry is $b_k(i,\underline{s})$. Let $dA(x)$ be a mxm matrix whose (i, j)th entry is $dA_{ij}(x)$ and let $\phi_k(x)$ be a $p^r x p^r$ matrix whose $[p^{r-1}s_1 + p^{r-2}s_2 + \cdots + s_r + 1,\ p^{r-1}t_1 + p^{r-2}t_2 + \cdots + t_r + 1]$ entry is

$$\prod_{\ell=1}^{r} \phi_k(t_\ell,\ s_\ell,\ x) = \gamma_k(\underline{t},\ \underline{s},\ x)$$

Similarly let $P(x)$ be a $mp^r \times mp^r$ matrix whose

$[p^r i + p^{r-1}s_1 + \cdots + s_r + 1,\ p^r j + p^{r-1}t_1 + \cdots + t_r + 1]$ entry is $P_r\{Z_n=j, \underline{S}_n=\underline{t},\ X_n \le x \mid Z_{n-1}=i,\ \underline{S}_{n-1}=\underline{s}\}$

Then $P(x) = \sum_{k=0}^{r} B_k \int_0^x \Phi_k(x) \,\Phi\, dA(x)$ where Φ stands for the Kronecker product of two matrices.

For instance, consider a simple queue with no storage ($r = 1$ and $p = 2$). Let the semi-Markov arrival process consist of one type of event ($m = 1$) with distribution function $G(x)$, and assume customers are always accepted for service when $S_n = 0$. Then

$$B_0 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \Phi_0(x) = [\gamma_0(i,j,x)] \quad \Phi_1(x) = [\gamma_1(i,j,x)]$$

and

$$P(x) = \begin{bmatrix} \int_0^x \gamma_1(0,0,x)\,dG(x) & \int_0^x \gamma_1(0,1,x)\,dG(x) \\ \int_0^x \gamma_0(1,0,x)\,dG(x) & \int_0^x \gamma_0(1,1,x)\,dG(x) \end{bmatrix}$$

From the matrix relation for $P(x)$ one can quickly write

$$P = \sum_{k=0}^{r} B_k \int_0^\infty \Phi_k(x) \Phi \, dA(x)$$

where P is a $mp^r \times mp^r$ matrix whose general element is $\Pr\{Z_n = j, \underline{S}_n = \underline{t} \mid Z_{n-1} = i, \underline{S}_{n-1} = \underline{s}\}$.

The imbedded Markov chain is completely specified from these relationships if the initial distribution is known. Let $p_o(i,\underline{s})$ denote the probability the system is in the state $(i, s_1, \cdots, s_r)$ at the beginning of operation, and let p_o be a row vector of size mp^r whose $(p^r i + p^{r-1} s_1 + \cdots + s_r + 1)^{st}$ element is $P_o(i,\underline{s})$.

Similarly let $p_n(i,\underline{s})$ at the instant before the n^{th} arrival, and let p_n be the corresponding row vector. From the theory of Markov chains it is immediately clear that

$$p_n = p_o P^n$$

From this one can easily calculate the distribution of the state of the system as

$$\Pr\{S_n\} = \sum_{i=0}^{m-1} \Pr\{S_n, Z_n = i\}$$

Steady state behavior of the system can be determined using the limiting probabilities $p = \lim_{n\to\infty} p$. These limiting probabilities exist and are independent of the initial distribution as a consequence of the general theory of Markov chains. They may be found from the relation

$$p = pP$$

$$p\underline{1} = 1$$

where $\underline{1}$ is a column vector of size mp^r whose elements are all ones.

3.2 APPLICATION TO DISTRIBUTION/ALLOCATION PROBLEM

In the previous section a complete set of results were derived for a system of r queues in parallel fed by

a semi-Markov arrival process through a switch which depends on the arrival process and on the state of the entire system formed by the r queues at the instant of arrival. In this section we utilize the features of this model to study the emergency vehicle distribution and allocation problem.

We define our system as that geographic region served by r emergency medical recovery vehicles. In the City of Detroit and in other urban areas, these vehicles are assigned to precincts within the city for command and control purposes. Hence, we shall take the precinct as the basic unit defining the system.

Within the precinct, we assume that emergencies occur according to the semi-Markov process $\{X_n, Z_n\}$, where

X_n = Interrival time between the n^{th} and $(n-1)$st emergency entering the system.

$$Z_n = \begin{cases} 0 & \text{if } n^{th} \text{ emergency is an emergency medical call.} \\ 1 & \text{if } n^{th} \text{ emergency is a police call.} \end{cases}$$

Then the emergency occurrence process can be described by the matrix $A(x)$, where

$$A(x) = \begin{bmatrix} A_{00}(x) & A_{01}(x) \\ A_{10}(x) & A_{11}(x) \end{bmatrix}$$

and $A_{ij}(x)$ = Probability n^{th} emergency is of class j and time until n^{th} emergency is less than x given that $(n-1)^{st}$ emergency is of class i.

In sections 2.2.1-2.2.2, it was shown that the sequence of emergency occurrences in a precinct obeys a Gamma process. Then for a single function ambulance system m = 1 and

$$A_0(x) = \int_0^x dA_0(x), \quad dA_0(x) = \left(\frac{\kappa}{\mu}\right)^{\kappa} \frac{x^{\kappa-1} e^{\frac{-\kappa x}{\mu}}}{\Gamma(\kappa)}$$

In a dual function ambulance system, m=2 and from Section 2.2.3

$$A_{i0}(x) = \Pi \int_0^x dF_{i0}(x) \quad A_{i1}(x) = (1-\Pi) \int_0^x dF_{i1}(x) \quad dF_{ij}(x) = \lambda_{ij} e^{-\lambda_{ij} x} \quad (3.4)$$

The parameters λ_{ij}, κ, μ, and Π are estimated as in Chapters II.

The set of states which each of the r vehicles assume can be described as follows. We require that at most one emergency can be in the queueing sub-system defined by each recovery vehicle at any one time. Therefore, when an emergency is assigned to a vehicle, service on this emergency begins immediately. This assumption is equivalent to saying that at most r emergencies can be in the recovery system at one time. In addition to simplifying the system state space description, this

assumption is likely to correspond to reality. When an emergency occurs and no vehicles are available in the recovery system, the dispatcher will do everything possible to avoid making the emergency "wait" for service. This includes dispatching recovery vehicles from outside the system (vehicles from other precincts, fire rescue squads, or private ambulances), or possibly dispatching vehicles which are not equipped as ambulances (squad cars, detective cars, etc.).

Clearly such dispatch procedures introduce additional costs and time delays, and consequently it is desirable to design the recovery system to avoid "overflows" of the type described above. We shall show how such designs can be implemented and indicate how their effects can be measured in the next section.

Under the assumption that a single recovery vehicle is assigned to at most one emergency at any time, we observe that the vehicle can reside in any of three basic states: idle, serving a police call, or serving a medical emergency. We have shown in Chapter II that police calls and ambulance calls have distinct service time characteristics, and thus must be analyzed separately. Since we are interested in the effects of improved vehicle distribution on medical emergency service improvements, we extend the vehicle state "serving a

medical emergency" to include service requiring transit times to the scene of variable durations.

In Chapter II, an extensive data analysis of the empirical transit distance--transit time relationships revealed that a major source of transit time variation could be removed by conditioning on transit distances of less than one mile and on those distances greater than one mile. The service times for each of these conditional distributions were found to follow negative exponential distributions.

We shall now use this information in developing the recovery system model. Assume, in general, that there are J distinct classes of emergency medical service, each corresponding to a different set of transit distances to the scene. Assume further that ambulance service time for each of these J sets of distances has a negative exponential distribution with mean $1/\beta_j$, $j = 1, \cdots, J$. Our data analysis for the City of Detroit indicates that $J = 2$ for a precinct service area, with distances less than one mile and those greater than one mile defining the classes of service.

In the notation of section 3.1, the state of vehicle i at the instant of the n^{th} emergency is denoted by S_n^i. From the above analysis we see that S_n^i can assume four levels--idle, serving a police emergency,

serving an ambulance call with distance to the scene less than one mile, and serving an ambulance call with distance to the scene greater than one mile. In this case, N_i, the total number of states a vehicle can assume, is equal to four for all vehicles in the recovery system. These states are assigned numerical codes below:

$$s_n^i = \begin{cases} 0 & \text{vehicle i is idle.} \\ 1 & \text{vehicle i is serving a police call.} \\ 2 & \text{vehicle i is serving an emergency with distance to the scene less than one mile.} \\ 3 & \text{vehicle i is serving an emergency with distance to the scene greater than one mile.} \end{cases}$$

A realization of the state space of the r vehicle recovery system is represented by the r-tuple $(s_1, \cdots, s_r)$, where s_j, $j = 1, \cdots, r$ takes on the values 0, 1, 2, 3 according to the rules developed above.

Since the police call service processes and the conditional ambulance service processes were found to follow negative exponential distributions in Chapter II, equations for transitions between system states can be derived using the results of 3.1. We shall systematically proceed to develop these in the remainder of this section.

To begin the discussion, we shall analyze in detail the switch which routes emergencies to the recovery vehicles for service.[1] We observe that this switch must take on

[1] Initially we shall describe the state space of this switch by a two-tuple. However, later in this section we shall map this two-tuple into a vector, thereby making these results comparable with those of Section 3.1.

values indicating the number of the vehicle dispatched as well as the class of service demanded by the emergency. This latter decision depends upon the type of emergency, the location of the emergency, and the placement of the vehicle which is assigned to service the emergency. Since it is desirable to dispatch the closest available vehicle to medical emergencies, the switch must consider this factor in routing an emergency to a vehicle for service.

We shall assume that the channel selection depends only on the type of the emergency Z_n and on the state of the system $S_n^1, \cdots, S_n^r$ at the instant of dispatch. This assumption rules out the lenient dispatcher who looks back over the day's activities of each vehicle and assigns a vehicle which hasn't been heavily utilized. Since such a procedure is not compatible with improved allocation and distribution, we shall not consider it here.

Let the random variable Y_n take on values in the 2-tuple (j,k), where j denotes the vehicle selected to serve the n^{th} emergency, $j = 1, \cdots, r$ and k denotes the class of service demanded by this emergency. $k = 1, 2, 3$.[1] For instance, $y_n = (3, 1)$ means recovery vehicle 3 is assigned to a police call; $y_n = (6, 3)$ means recovery vehicle 6 is assigned to a medical emergency requiring a transit distance to the scene greater than one mile. To

[1] Corresponding, respectively, to police calls, ambulance calls requiring less than one mile transit distance, and ambulance calls requiring more than one mile transit distance.

this set of two-tuples, we add the feasible state (0); this state is assumed when emergencies are rejected by the system or are routed to vehicles outside of the system.

We shall now examine in detail the properties of $\Pr\{Y_n = (j,k) | Z_n, S_n^1, \cdots, S_n^r\}$. We shall consider first these properties when $Z_n = 1$; that is, when the arrival is a police call. Then the set of possible values for Y_n are $[0, (1,1), \cdots, (r,1)]$, since the class of service k is equal to one for all police calls. We note that $Y_n = 0$ when the system is saturated with emergencies, i.e., when $S_n^j = s_j > 0 \; \forall \; j$, $j = 1, \cdots, r$. In addition, there is some probability that a police call will not be assigned to an ambulance, even when this ambulance is in an idle state.

This is true since squad cars are also allocated to the precinct to handle police calls. Hence it is desirable to incorporate this probability structure into the model. We shall do this by assuming that the probability that a police call is routed to the ambulance system is related to the number of idle ambulances--that is, it is proportional to the fraction of the s_j's which are zero. Denote the number of idle ambulances by ϕ, where ϕ takes on the values $0, 1, \cdots, r$. Then P_ϕ is the probability a police call is routed into the system. We further assume that

each of the idle ambulances is equally likely to service a police call. Then

$$\Pr\{Y_n = (j,1) \mid Z_n = 1,\ S_n^1,\ \cdots,\ S_n^r\} = \begin{cases} 1/\phi\, P_\phi & \text{if } s_j = 0 \\ 0 & \text{if } s_j > 0 \end{cases} \qquad j=1,\cdots,r$$

where ϕ is the number of idle ambulances at the instant of the n^{th} arrival.

Notice that by varying the structure of P_ϕ, a wide variety of potential switching devices can be modeled. For instance, if this probability is a constant, then $P_\phi = 1/\phi$. Alternatively, the probability may decline linearly with decreasing ambulance availability. Other models of interest can easily be incorporated.

We turn now to the probabilistic properties of the switch when the call is a medical emergency; i.e., when $Z_n = 0$. Here the emergency is accepted into the recovery system unless all vehicles are utilized at the instant of arrival. That is $Y_n = 0$ if and only if $Z_n = 0$ and $S_n^j = s_j > 0, \forall j,\ j = 1, \cdots, r$. When an emergency enters the system, the selection of a vehicle to service it and the service time distribution for this vehicle depend entirely on the state of the system, the vehicle placement, and the location of the emergency. Hence we must consider these factors in detail to develop this probability structure.

For this purpose we shall divide the service region (precinct) into D distinct sub-regions. The location of the r recovery vehicles can then be described by a vector $\underline{a} = (a_1, \cdots, a_r)$ where a_ℓ is the sub-region distribution of the ℓ^{th} vehicle and takes on the values 1, 2, $\cdots$, D for $1 = 1, \cdots, r$. Using this formulation there are $\binom{D}{r}$ possible distributions when no two vehicles are placed in the same sub-region, and D^r possible distributions when any number up to r may be placed in the sub-region.[1] Let b_d, $d = 1, \cdots, D$ denote the probability that a medical emergency occurs in sub-region d. We assume that the geographical occurrences of emergencies over the region is an independent, time homogeneous random variable obeying the probability law b_d, $\sum_{d=1}^{D} b_d = 1$.

Hence, given $\underline{a}$ and $\underline{b} = (b_1, \cdots, b_d)$, the probabilities associated with the conditional switch are completely determined. For example, we shall work these out for a simple case with D = 4 and r = 2 in this section. Let $\underline{a} = (1,4)$, and let $\underline{b} = (b_1, b_2, b_3, b_4)$. As before, we utilize the "two level" service time distribution for medical emergencies.

Hence S_n^j takes on the possible values 0, 1, 2, 3. When $Z_n = 0$, S_n^j takes on only the values 0, 2, or 3,

---

[1] A situation which prevails when a single, permanent station is utilized to house the vehicles.

corresponding to idle, serving medical emergency requiring transit distance less than one mile, and serving an emergency requiring transit distance greater than one mile.

A schematic diagram of the situation is presented in Figure 3.2.

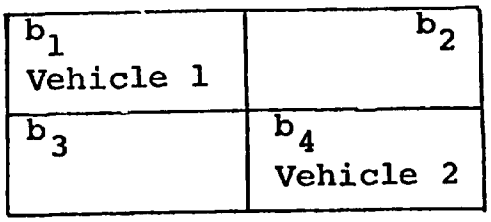


Figure 3.2
Sample Service Area and Ambulance Allocation

We assume that all travel distances are less than one mile unless the vehicle travels between sub-regions one and four. In this latter situation, the transit distance is assumed to be greater than one mile. When both vehicles are idle and equi-distant from an emergency, we assume that a random vehicle selection mechanism is used.

For instance, consider $\Pr\{Y_n = (1,2) \mid Z_n=0, S_n^1=0, S_n^2=0\}$. Here both vehicles are idle. Thus we dispatch vehicle one on a "short" run if the emergency is in region one (with probability b_1). We also dispatch vehicle one on a "short" run one-half of the time when the emergency is in regions two or three. The resulting conditional probability is $b_1 + \frac{b_2+b_3}{2}$.

The following switching matrices can be developed using similar reasoning:

$\Pr\{Y_n = (1,2) \mid Z_n = 0,\ S_n^1,\ S_n^2\}$

S_n^1 \ S_n^2	0	2	3
0	$b_1 + \frac{b_2 + b_3}{2}$	$b_1 + b_2 + b_3$	$b_1 + b_2 + b_3$
2	0	0	0
3	0	0	0

$\Pr\{Y_n = (1,3) \mid Z_n = 0,\ S_n^1,\ S_n^2\}$

S_n^1 \ S_n^2	0	2	3
0	0	b_4	b_4
2	0	0	0
3	0	0	0

$\Pr\{Y_n = (2,2) \mid Z_n = 0,\ S_n^1,\ S_n^2\}$

S_n^1 \ S_n^2	0	2	3
0	$b_4 + \frac{b_2 + b_3}{2}$	0	0
2	$b_2 + b_3 + b_4$	0	0
3	$b_2 + b_3 + b_4$	0	0

$$\Pr\{Y_n = (2,3) \mid Z_n = 0,\ S_n^1,\ S_n^2\}$$

S_n^1 \ S_n^2	0	2	3
0	0	0	0
2	b_1	0	0
3	b_1	0	0

The switching mechanisms for systems involving three or more vehicles become more complicated. These will be derived for realistic model implementation in Chapter IV.

To simplify the rotation we shall map the two dimensional state space of the random variable Y_n into one dimension. To do this, let $\ell = 0$ if $Y_n = 0$, and let $\ell = 3(j-1) + k = 3j + k - 3$, for $j = 1, \cdots, r$ and $k = 1, 2, 3$. Then we may write

$$\Pr\{Y_n = (j,k) \mid Z_n = i,\ S_n^1,\ S_n^2,\ \cdots,\ S_n^r\} = b_\ell(i,\ \underline{s})$$

where ℓ takes on the values $0, 1, 2, \cdots, 3r$.

In order to completely derive the system transition probabilities, it is only necessary to define

$$\phi_\ell(t_i, s_i, x) = \Pr\{S_n^i = t_i \mid S_{n-1}^i = s_i,\ Y_{n-1} = \ell,\ X_n = x\}$$

To do this we note that if $\ell = 3(j-1) + k$ and $i = j$, then

the $n-1^{st}$ emergency was routed to the i^{th} vehicle. In this case $S^i_{n-1} = 0$; that is, the vehicle was idle at the time of the emergency reporting. The state of the i^{th} vehicle at time n, t_i, is then either equal to 0 or k depending upon whether or not service on the $(n-1)^{st}$ emergency is completed in the interval between emergencies. Due to the exponential nature of the vehicle service process, the probabilities of these two events are $1 - e^{-\beta_k x}$ and $e^{-\beta_k x}$ respectively. Hence we may summarize this case as follows:

$$\Pr\{S^i_n = t_i \mid S^i_{n-1} = 0,\ Y_{n-1} = 3(i-1) + k,\ X_n = x\} =$$

$$\begin{cases} e^{-\beta_k x} & t_i = k \\ 1-e^{-\beta_k x} & t_i = 0 \\ 0 & \text{otherwise} \end{cases}$$

The other possible situation occurs when $\ell = 3(j-1) + k$, $i \neq j$; that is, when the $(n-1)^{st}$ emergency is not routed to the i^{th} vehicle for service. If the i^{th} vehicle is idle ($S^i_{n-1} = 0$) then it remains idle ($S^i_n = 0$) with probability 1. On the other hand, if the vehicle is servicing an emergency of class h, h = 1, 2, 3 at this time, then $S^i_{n-1} = h$, and $S^i_n = 0$ with probability $1-e^{-\beta_h x}$ or $S^i_n = h$ with probability $e^{-\beta_h x}$. Hence we may summarize this case as follows:

$$\Pr\{S_n^i = t_i \mid S_{n-1}^i = 0,\ Y_{n-1} = 3(j-1) + k,\ X_n = x\}$$

$$= \begin{cases} 1 \text{ for } t_i = 0 \\ 0 \text{ otherwise} \end{cases}$$

$$\Pr\{S_n^i = t_i \mid S_{n-1}^i = h,\ Y_{n-1} = 3(j-1) + k,\ X_n = x\}$$

$$\begin{cases} 1-e^{-\beta_h x} & \text{for } t_i = 0 \\ e^{-\beta_h x} & \text{for } t_i = h,\ h = 1, 2, 3 \\ 0 & \text{otherwise} \end{cases}$$

These results completely describe the structure of the conditional probabilities $\phi_\ell(t_i\ s_i, x)$. These probabilities can be combined with the $b_\ell(i, \underline{s})$ using relation (3.3) to yield the transition matrices describing system performance.

From result (3.2), it is clear that the process $\{Z_n, S_n^1, \cdots, S_n^r\}$ is Markov with $2(4)^r$ states.

The total number of states for various numbers of recovery vehicles is tabulated below. Since an independent ambulance system does not service police calls, the number of vehicle states and number of emergency states are both decreased by one, yielding $(3)^r$ total system states. These values are also tabulated:

r	$2(4)^r$	3^r
1	8	3
2	32	9
3	128	27
4	512	81

As the number of vehicles in the system increases, computer procedures will be necessary to implement the resulting model. We shall discuss these in more detail in Chapters IV and V.

3.3 MODEL UTILIZATION

We shall now show how the model developed in the previous two sections can be utilized to obtain insights into the three problems posed at the beginning of this chapter.

(1) Determination of the effect of eliminating or modifying the dispatch policy which assigns both police and emergency medical calls to the system of recovery vehicles.

(2) Determination of the effect of alternative recovery vehicle distribution throughout the service region.

(3) Determination of the effect of changing the number of recovery vehicles in the system.

We shall sequentially discuss these three problems in this section.

(1) Elimination or modification of "dual-function" dispatch policy.

This can be easily accomplished by modifying the stochastic switch Y_n so that $Z_n = 1$ with probability one --that is, so that only medical emergencies are considered by the system.

In this case the process $\{S_n^1, S_n^2, \cdots, S_n^r\}$ is Markovian, and steady state probability results can be obtained for various distribution and allocation policies. These can be compared with results obtained by utilizing variants of the "dual" switch to measure system improvements.

Of special interest in these comparisons are the steady state probabilities that the system is saturated, i.e., $\Pr\{\underline{S} = \underline{s}, S_j > 0, \forall j, j = 1, \cdots, r\}$ and the steady state probabilities that k ambulances are idle, $k = 0, 1, \cdots, r$.

The effects of modifying the dispatch policy can be measured by studying the changes in the above probabilities when $P_{\phi/\phi}$ is modified. Uniform and linearly decreasing values of P_ϕ merit special attention.

(2) Changing distribution of vehicles throughout the recovery system.

We have seen that the distribution of the r recovery

vehicles in the system is completely specified by selecting a "distribution" vector $\underline{a} = (a_1, \cdots, a_r)$, where the elements of this vector take on values 1, 2, $\cdots$, D corresponding to the sub-regions recovery vehicles are assigned to.

We have also seen that there are $\binom{D}{r}$ or D^r possible distributions depending upon whether multiple vehicles can be placed into a single sub-region. It is clear that an exhaustive exploration of these possibilities would require a tremendous computational effort.

However, it is possible to select a subset of this number for evaluation. Several obvious elements of this subset are apparent:

(a) Placement of the r vehicles in the r sub-regions with the largest probabilities of medical emergencies, that is, with the largest b_j, $j = 1, \cdots, D$.

(b) Placement of the r vehicles "uniformly" throughout the region on the basis of geographic considerations.

(c) Placement of the r vehicles according to some heuristic combination of (a) and (b) above. For instance, one might locate r/2 vehicles in sub-regions of highest potential demand and place the remaining r/2 vehicles uniformly over the service area.

(d) Placement of all vehicles at one or possibly two sites within the region. This distribution may be necessary if an independent non-police ambulance service is implemented, since in this case a "patrolling" vehicle may be uneconomical or infeasible. For instance, if the fire department assumes responsibility for service, fire stations are obvious candidates for ambulance stations. Similarly, if the hospitals assume responsibility, the hospitals in the service region are potential fixed ambulance sites. Although vehicle distribution at such fixed stations may force the system to operate sub-optimally, it is extremely important to study system performance under such constraints.

For these four basic distribution "strategies" and variants thereof, the system performance can be measured in several ways including the steady state saturation probabilities, the steady state idle probabilities, and the steady state probabilities that j vehicles are serving medical emergencies of classes two or three, $j = 1, \cdots, r$.

This latter measure is of interest since improved distribution policies should increase the probability of serving an emergency which requires shorter transit distance to the scene and hence a shorter service time.

(3) Changing the number of recovery vehicles in the system.

The effects of changing r, the number of recovery vehicles, can be studied by utilizing those measures examined in (1) and (2) above. The one complicating factor is that the availability of additional recovery vehicles provides a much number of feasible distribution policies. Nevertheless, the joint effects or the increased number of vehicles and the basic distribution policies outlined in (2) can be studied. The change in the steady state probabilities that vehicles are idle or that the system is saturated are especially interesting measures of the effectiveness of increasing the number of vehicles.

CHAPTER IV

ALLOCATION AND DISTRIBUTION POLICIES FOR THE SINGLE FUNCTION PUBLIC AMBULANCE SYSTEM

In this chapter we shall utilize the model developed in Chapter III to study the operation of the single function public ambulance system. Analytical and numerical results are presented for systems in which one, two, or three ambulances are assigned to a police precinct. Various parameterizations of the basic model are also studied to determine the sensitivity of the results to variations in the model parameters. Numerical results were calculated on the University of Michigan IBM 360/67 computing system. The Fortran program which was written to perform the calculations is documented in Appendix C.

4.1 RESULTS FOR A SINGLE AMBULANCE SYSTEM

To begin the discussion, calculations for the single ambulance system ($r = 1$) are carried out in some detail to illustrate the techniques of Chapter III and to obtain analytical insights into the system performance.

For the single ambulance system the switching probabilities can be written:

$$\Pr\{Y_n = \ell | Z_n = i;\ S_n^1, \cdots, S_n^r\} = \Pr\{Y_n = \ell |\ Z_n = 0;\ S_n^1 = k\}$$

where Y_n takes on the values 0, 2, or 3, corresponding respectively to situations where the ambulance rejects the n^{th} emergency, serves the emergency and travels less than one mile to the scene, or serves the emergency and travels more than one mile to the scene. Similarly the state of the vehicle S_n^1 takes on the values 0, 2, or 3, corresponding to states when the vehicle is idle, serving an emergency requiring less than one mile transit distance, or serving a emergency requiring more than one mile transit distance. The resulting probabilities $\Pr\{Y_n = \ell | Z_n = 0;\ S_n^1 = k\}$ are given by the following matrix:

Y_n \ S_n^1	0	2	3
0	0	1	1
2	δ	0	0
3	$1-\delta$	0	0

where δ is the probability that the emergency requires a transit distance of less than one mile. As we have seen, this probability is a function of both the vehicle distribution and the geographical distribution of emergencies.

For this system, the probabilities

$$\Pr\{S_n^1 = k \mid Y_{n-1} = \ell;\ S_{n-1}^1 = s_1, \cdots, S_{n-1}^r = s_r;\ X_n = x\}$$

can be written

$\Pr\{S_n^1 = k \mid Y_{n-1} = \ell;\ S_{n-1}^1 = j;\ X_n = x\}$, and these are tabulated below:

$Y_{n-1} = 0$

S_{n-1}^1 \ S_n^1	0	2	3
0	1	0	0
2	$1-e^{-\alpha x}$	$e^{-\alpha x}$	0
3	$1-e^{-\beta x}$	0	$e^{-\beta x}$

$Y_{n-1} = 2$

S_{n-1}^1 \ S_n^1	0	2	3
0	$1-e^{-\alpha x}$	$e^{-\alpha x}$	0
2	0[1]	0	0
3	0	0	0

$Y_{n-1} = 3$

S_{n-1}^1 \ S_n^1	0	2	3
0	$1-e^{-\beta x}$	0	$e^{-\beta x}$
2	0	0	0
3	0	0	0

[1] When various conditioning events are impossible (as in this case), values of zero are assigned to the corresponding probabilities.

In the above conditional probabilities, α is the service rate for medical emergencies requiring less than one mile transit distance, and β is the service rate for medical emergencies requiring more than one mile transit distance.

The probabilities $\Pr\{S_n^1 = k | Y_{n-1} = \ell;\ S_{n-1}^1 = j\}$ are then derived as follows:

$$\Pr\{S_n^1 = k | Y_{n-1} = \ell;\ S_{n-1}^1 = j\}$$

$$= \int_0^\infty \Pr\{S_n^1 = k | Y_{n-1} = \ell;\ S_{n-1}^1 = j;\ X_n = x\} dA_0(x)$$

where from Chapter III, we recall that

$$dA_0(x) = \left(\frac{\kappa}{\mu}\right)^\kappa \frac{x^{\kappa-1} e^{-\frac{\kappa x}{\mu}}}{\Gamma(\kappa)}$$

These are tabulated below as a function of P_y where:

$$P_y = \int_0^\infty e^{-yx} \left(\frac{\kappa}{\mu}\right)^\kappa \frac{x^{\kappa-1} e^{-\frac{\kappa x}{\mu}}}{\Gamma(\kappa)}\, dx$$

$$= \left(\frac{\kappa}{\mu}\right)^\kappa / \Gamma(\kappa) \int_0^\infty x^{\kappa-1} e^{-\left(\frac{\kappa}{\mu} + y\right)x}\, dx$$

$$= \left(\frac{\kappa}{\mu}\right)^\kappa / \Gamma(\kappa) \cdot \Gamma(\kappa) / \left(\frac{\kappa}{\mu} + y\right)^\kappa = \left[\frac{\kappa/\mu}{\kappa/\mu + y}\right]^\kappa$$

$Y_{n-1} = 0$

S^1_{n-1} \ S^1_n	0	2	3
0	1	0	0
2	$1-P_\alpha$	P_α	
3	$1-P_\beta$	0	P_β

$Y_{n-1} = 2$

S^1_{n-1} \ S^1_n	0	2	3
0	$1-P_\alpha$	P_α	0
2	0	0	0
3	0	0	0

$Y_{n-1} = 3$

S^1_{n-1} \ S^1_n	0	2	3
0	$1-P_\beta$	0	P_β
2	0	0	0
3	0	0	0

Using the notation of Chapter III, we let the matrices in the above table be denoted Φ_k. The corresponding switching matrices B_k can then be written:

$$B_1 = \begin{bmatrix} C & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} \delta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 1-\delta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The matrix of system transition probabilities $\Pr\{S_n^1 = k | S_{n-1}^1 = j\}$ is given by

$$P = \sum_{k=1}^{3} B_k \Phi_k$$

$$= \begin{bmatrix} \delta[1 - P_\alpha] + [1 - \delta][1 - P_\beta] & \delta P_\alpha & [1 - \delta]P_\beta \\ 1 - P_\alpha & P_\alpha & 0 \\ 1 - P_\beta & 0 & P_\beta \end{bmatrix}$$

Inspection of this matrix indicates that the elements could be derived by methods less sophisticated than those utilized here for this simple system. For instance, the element p_{11} is the probability that the system is idle at time of the n^{th} emergency given that it was idle at the time of the $(n-1)^{st}$ emergency. This is equal to the probability that the $(n-1)^{st}$ emergency required less than one mile transit distance and that subsequent service was completed plus the probability

that the $(n-1)^{st}$ emergency required more than one mile transit distance and that service was completed. In the notation of this chapter this quantity is simply $\delta[1 - P_\alpha] + [1 - \delta][1 - P_\beta]$. The other elements of the matrix can be derived by a similar argument, although this informal procedure becomes increasingly difficult for more complex systems.

We recall from Chapter III that in an ergodic system the equilibrium probabilities for each state can be derived from the following system of linear relations:

$$\underline{x}' P = \underline{x}'$$

$$\underline{x}'\underline{1} = 1$$

$$\underline{x}' \geq 0$$

where the nxl column vector $\underline{x}$ is the vector of equilibrium probabilities and $\underline{1}$ is an nxl column vector, all of whose elements are equal to one.

The associated system of n + 1 equations in n unknowns can be rewritten as:

$$\underline{x}'(P-I) = \underline{0}$$

$$\underline{x}'\underline{1} - 1 = 0$$

or equivalently

$$
\begin{aligned}
(P_{11} - 1)x_1 + \quad & P_{12}x_2 + \cdots + P_{1n}x_n = 0 \\
P_{21}x_1 + (P_{22} - 1)x_2 + \cdots & + P_{2n}x_n = 0 \\
& \vdots \\
P_{n1}x_1 + \cdots & + (P_{nn}-1)x_n = 0 \\
x_1 + \cdots & + x_n = 1
\end{aligned}
\qquad (4.1)
$$

where $\sum_{i=1}^{n} P_{ij} = 1$; $j = 1, \cdots, n$. The equilibrium probabilities x_i are the non-negative solutions to this system.

Inspection of (4.1) reveals that the system is clearly redundant, since the first equation can be written as a linear combination of equations 2 through n. We shall eliminate the redundancy by forming an equivalent system as suggested in by Dantzig (12). This is carried out by finding $\lambda_i \geq 0$ such that $P_{ij} - \lambda_i \geq 0$; $i = 1, \cdots, n$. Then we form a system of full rank with n equations in n unknowns by subtracting each of the equations 1 through n above from the $(n + 1)^{st}$ equation multiplied by λ_i.

The resulting system can be written

$$
\begin{aligned}
(\lambda_i + 1 - P_{11})x_1 + (\lambda_1 - P_{12})x_2 + \cdots + (\lambda_1 - P_{1n})x_n &= \lambda_1 \\
& \vdots \\
(\lambda_n - P_{n1})x_1 + \cdots + (\lambda_n + 1 - P_{nn})x_n &= \lambda_n
\end{aligned}
$$

or in matrix form

$$(I - G)\ \underline{x} = \underline{\lambda}$$

where G is a non-negative square matrix whose $(i,j)^{th}$ element is $P_{ij} - \lambda_i$, and where $\underline{\lambda}$ is an (nx1) column vector whose i^{th} element is $\lambda_i \geq 0$.

The solution of this system can be written

$$\underline{x} = (I - G)^{-1} \underline{\lambda}$$

or equivalently

$$\underline{x} = (I + G + G^2 + \cdots)\underline{\lambda} = I\underline{\lambda} + G\underline{\lambda} + G^2\underline{\lambda} + \cdots$$

Since G is non-negative, all powers of G are non-negative. Furthermore, since $\underline{\lambda}$ is non-negative, each vector in the infinite series above is non-negative. Hence the solution $\underline{x}$ is non-negative as required.

We now apply these results to the transition matrix P for the single ambulance system. In this case the reduced system of full rank can be written as three linear equations in three unknowns:

$$[\lambda_1 + 1 - \delta(1 - P_\alpha) - (1 - \delta)(1 - P_\beta)]x_1 + [\lambda_1 + P_\alpha - 1]x_2 + [\lambda_1 + P_\beta - 1]x_3 = \lambda_1$$

$$[\lambda_2 - \delta P_\alpha]x_1 + [\lambda_2 + 1 - P_\alpha]x_2 + [\lambda_2]x_3 = \lambda_2$$

$$[\lambda_3 - (1 - \delta)P_\beta]x_1 + [\lambda_3]x_2 + [\lambda_3 + 1 - P_\beta]x_3 = \lambda_3$$

Letting $\lambda_1 = 1 - P_\beta$, $\lambda_2 = \lambda_3 = 0$, we obtain:

$$[2-P_\beta-\delta(1-P_\alpha)-(1-\delta)(1-P_\beta)]x_1 + [P_\alpha-P_\beta]x_2 = 1-P_\beta$$

$$[-\delta P_\alpha]x_1 + [1-P_\alpha]x_2 = 0$$

$$[-(1-\delta)P_\beta]x_1 + [1-P_\beta]x_3 = 0$$

which can be solved for $\underline{x}$ yielding:

$$x_1 = \frac{[1-P_\alpha][1-P_\beta]}{[1-P_\alpha] + \delta[P_\alpha-P_\beta]}$$

$$x_2 = \frac{\delta P_\alpha[1-P_\beta]}{[1-P_\alpha] + \delta[P_\alpha-P_\beta]} \qquad (4.2)$$

$$x_3 = \frac{[1-\delta][1-P_\alpha]P_\beta}{[1-P_\alpha] + \delta[P_\alpha-P_\beta]}$$

where x_1 is the equilibrium probability of being idle, x_2 is the equilibrium probability of serving a "short" emergency (transit distance less than one mile), and x_3 is the equilibrium probability of serving a "long" emergency (transit distance greater than one mile).

In Chapter II the maximum likelihood estimates of the parameters κ, μ, α, and β were found to be:

$$\hat{\kappa} = .77$$

$$\hat{\mu} = 101.166$$

$$\hat{\alpha} = 1/18.406 = .054$$

$$\hat{\beta} = 1/20.472 = .049$$

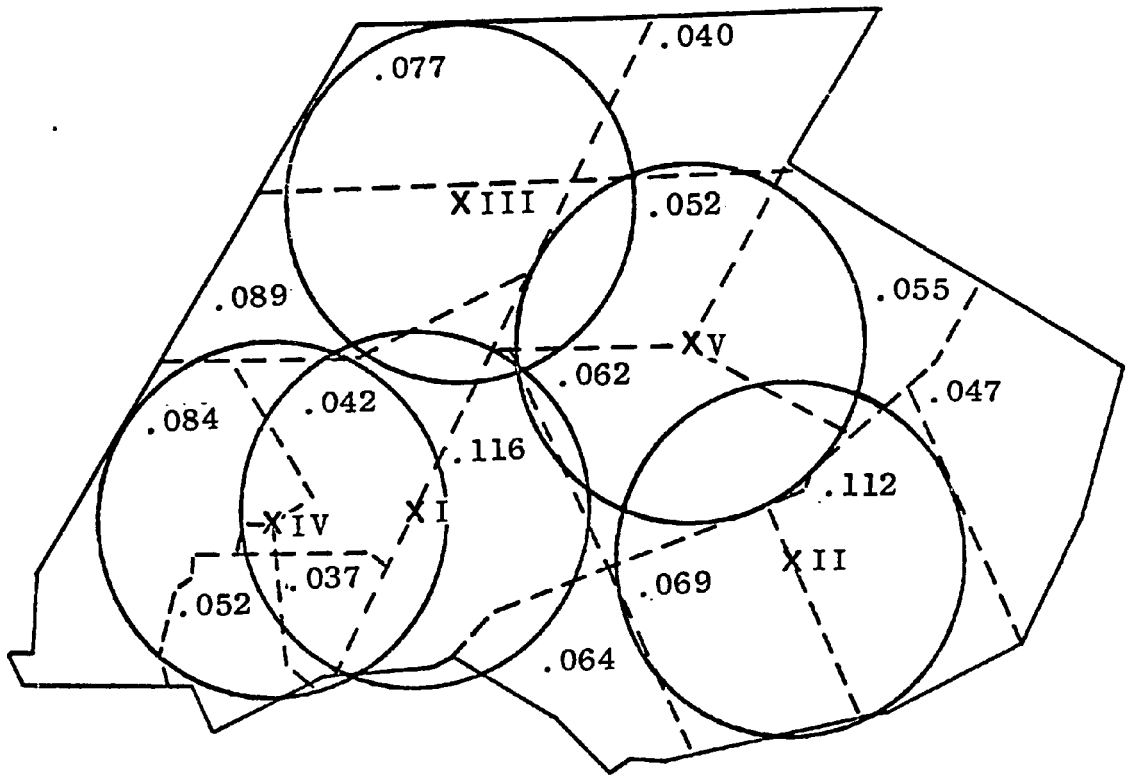
and substituting these in 4.2 we find the maximum likelihood estimates of P_α and P_β to be

$$\hat{P}_\alpha = .200$$
$$\hat{P}_\beta = .214$$

The remaining parameter in the model is δ, the probability that a medical emergency will occur within one mile of the single recovery vehicle. This probability depends upon the geographic distribution of emergencies and upon the distribution of the recovery vehicles within the service region. We shall now consider the effects of these factors in more detail.

The geographical distribution of emergencies in a police precinct was estimated from the observed distribution over the one-month sampling period. Counts of the numbers of emergencies occurring in each patrol car territory were used to develop the estimates. The resulting estimated probabilities that a medical emergency will occur in any patrol car territory are shown for precinct 15 in Figure 4.1.

We observe that the estimated probabilities in Figure 4.1 show considerable variation over the precinct. The maximum (.116) and the minimum (.037) occur in adjacent territories, although these territories are of



Scale - 1 mile

Sample Size = 403

Figure 4.1
Estimated Probabilities that a Medical Emergency
Will occur in any Scout Car Territory
Precinct 15

substantially different size. The probabilities of emergencies occurring in various regions of interest can then be derived by interpolating from these estimates. For instance, placing the single recovery vehicle in the beat of maximum demand (circle I) results in an estimated δ of .196. Estimates of δ for other reasonable ambulance locations are tabulated below.

Location	$\hat{\delta}$
I	.195
II	.196
III	.166
IV	.173
V	.139

These results indicate that δ varies from .139 to .196 for the five distribution policies studied above. The effect of a shift in this parameter (when other parameters are maintained at their maximum likelihood estimates) is tabulated below (Table 4.1).

From this tabulation we observe that the equilibrium idle probability (x_1) is only slightly affected by the change in δ. This result is apparent from equation which shows that x_1 is approximately invariant for small values of $P_\alpha - P_\beta$.

δ	x_1	x_2	x_3
.100	.7878	.0197	.1925
.139	.7883	.0274	.1843
.196	.7891	.0387	.1722
.250	.7898	.0494	.1608

Table 4.1

Single Vehicle Equilibrium Probabilities for Various Vehicle Allocations

On the other hand, the equilibrium probabilities x_2 and x_3 vary approximately as linear functions of δ. These results are also anticipated for small $(P_\alpha - P_\beta)$, and in fact, such linear relations hold exactly for $P_\alpha = P_\beta$.

For the range of δ corresponding to the vehicle distributions in Figure 4.1, the equilibrium probability of being on a short run varies from .0274 to .0387. This result again indicates that vehicle distribution does not drastically affect the system service characteristics. Furthermore, such a change in δ increases the equilibrium idle probability by only .0008. Hence, changes in the location of a single vehicle do not significantly modify the probability that the vehicle will be available at the instant of an emergency.

It is useful to study the sensitivity of these results to changes in the emergency occurrence parameters (κ and μ) and to changes in the emergency service parameters (α and β). Although the joint effect of changes of all four parameters is of interest, the emergency occurrence and service processes will be parameterized separately in this section for computational convenience.

Results for a range of variations in κ and μ are presented below (Table 4.2). The service parameters are maintained at their maximum likelihood estimates for these calculations. Values of x_1, x_2, and x_3 are tabulated for each (μ, κ, δ) combination.

Specific results of interest for these parameterizations are presented in Figures 4.2 and 4.3. We note in Figure 4.2 that the equilibrium idle probability increases monotonically as a function of the mean time between emergencies μ and κ (the reciprocal of the square of the coefficient of variation). Variations in δ do not affect the equilibrium idle probability appreciably.

In the neighborhood of the maximum likelihood estimates ($\hat{\mu} = 101$, $\hat{\kappa} = .77$), a 20% error in μ results in a maximum error in x_1 of approximately .04, and a 20% error in κ results in a maximum error of approximately .05. Since the gradients of the function are as large

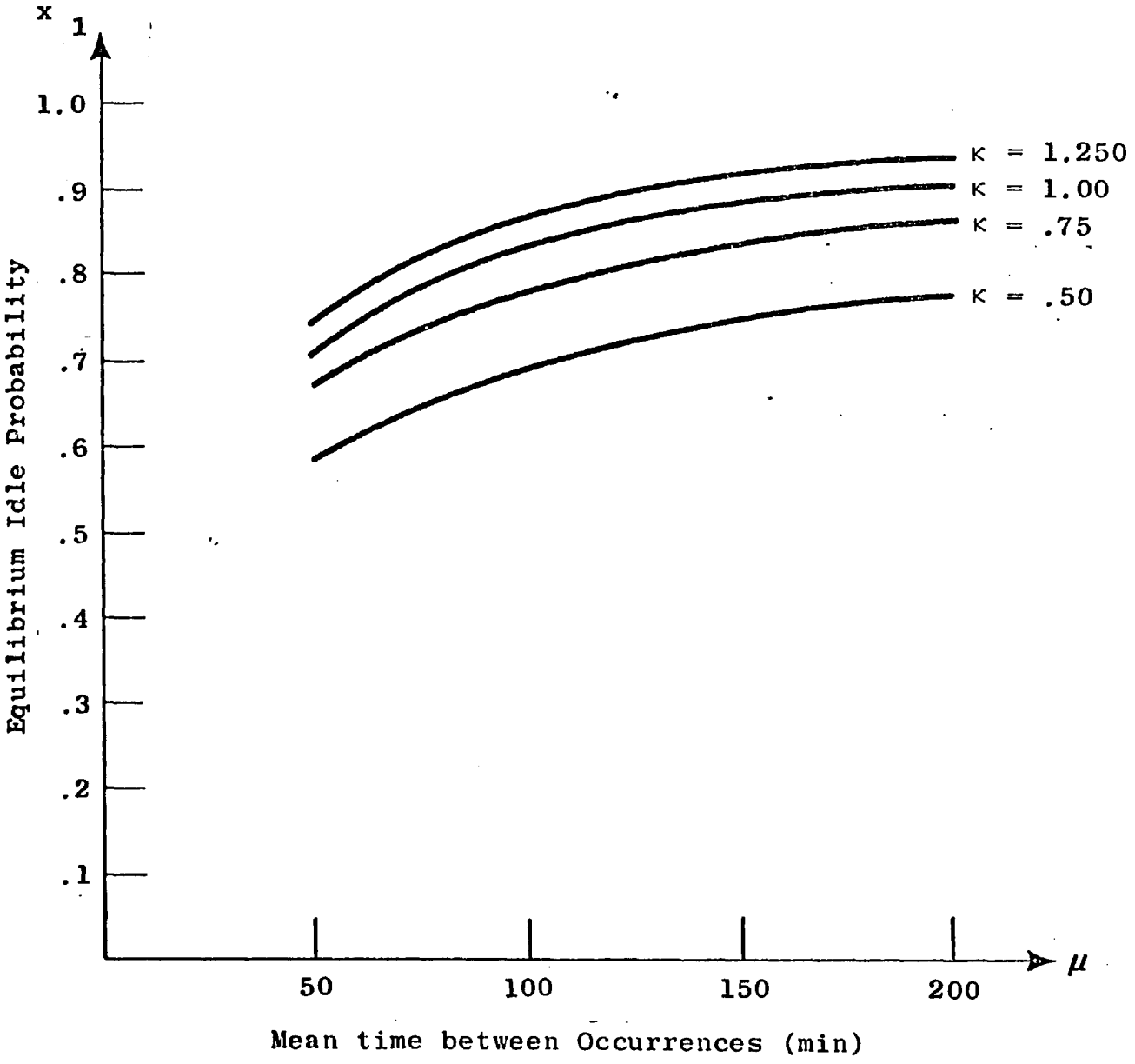


Figure 4.2
Equilibrium Idle Probability as Function
of Mean Time between Emergency Occurrences
Single Ambulance System $\delta = .139$

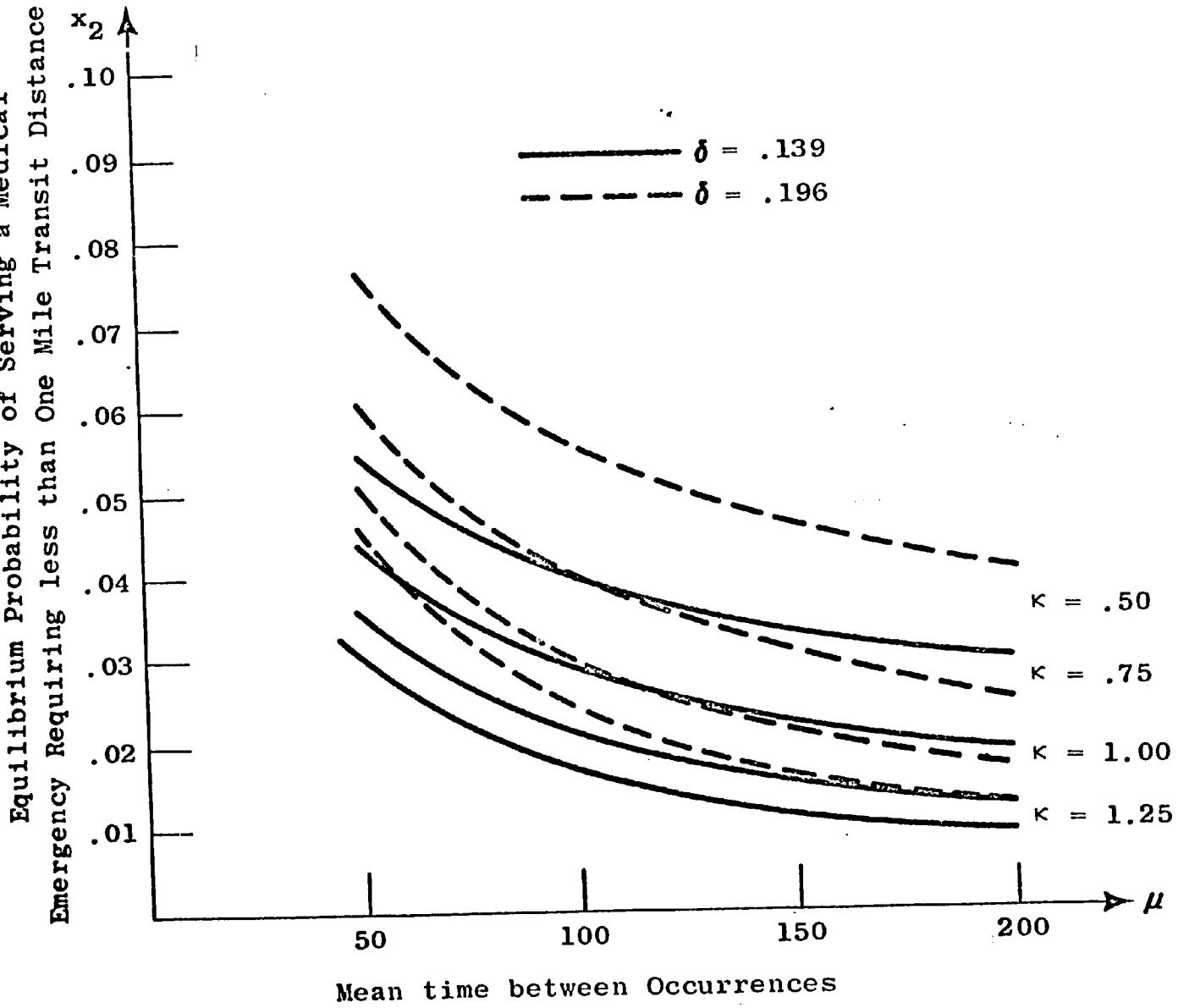


Figure 4.3
Equilibrium Probability x_2 as Function
of Mean Time between Emergency Occurrences
Single Ambulance System

δ= .139

κ \ μ		50	100	150	200
.50	x_1	.591	.698	.749	.781
	x_2	.054	.040	.033	.029
	x_3	.355	.262	.218	.190
.75		.666	.782	.834	.864
		.043	.028	.022	.018
		.291	.190	.144	.118
1.00		.713	.832	.882	.908
		.037	.021	.015	.012
		.250	.147	.103	.080
1.250		.745	.865	.912	.935
		.032	.017	.011	.008
		.223	.118	.077	.057

δ= .196

r \ μ		50	100	150	200
.500	x_1	.592	.698	.750	.782
	x_2	.076	.056	.047	.041
	x_3	.332	.246	.203	.177
.750		.667	.783	.834	.864
		.061	.040	.031	.025
		.272	.177	.135	.111
1.000		.714	.833	.882	.909
		.052	.030	.021	.017
		.234	.137	.097	.074
1.250		.746	.866	.912	.936
		.046	.024	.016	.012
		.208	.110	.072	.052

Table 4.2

Sensitivity of Equilibrium Probabilities
to Variations in Emergency
Occurrence Parameters

in this neighborhood as anywhere over the region of Figure 4.2, parameter specification errors should result in smaller effects elsewhere.

Inspection of Figure 4.3 reveals that the equilibrium probability a vehicle is serving an emergency requiring less than one mile transit distance decreases monotonically as a function of δ, κ, and μ. Here a 20% error in μ (again in the neighborhood of the maximum likelihood estimates) results in a maximum error in x_2 of about .005 for $\delta = .196$ and about .007 for $\delta = .139$. A similar error in κ results in a maximum error in x_2 of approximately .007 for $\delta = .196$ and approximately .0095 for $\delta = .139$.

We observe that for both equilibrium probabilities x_1 and x_2, specification errors in κ seem to affect the estimates more than errors in μ. We also note that the a priori assumption of a Poisson emergency occurrence process ($\kappa = 1$) results in appreciable errors in the equilibrium probabilities. The maximum percentage error in x_1 is 6.4% (.05/.78); the maximum percentage error in x_2 is 23% (.009/.039). Hence, we are able to conclude that the equilibrium idle probability is less affected (relative to its maximum likelihood value) by parameter specification errors.

We shall now study the effects of variations in the conditional service rates α and β. Since $P_\alpha - P_\beta$ appears in the earlier expressions for equilibrium probabilities, we shall conduct the parameterization in terms of α and the ratio β/α. Results are tabulated for a range of these parameters in Table 4.3. The range of α is equivalent to mean service times from 15 to 40 minutes. The maximum mean service time for emergencies requiring transit distances greater than one mile is 80 minutes (corresponding to the pair $\alpha = .025$, $\beta/\alpha = .50$).

A plot of the equilibrium idle probability x_1 as a function of α and β/α is presented in Figure 4.4. We observe that this probability is barely affected by variations in δ. The sensitivity to minor parameter misspecifications in β/α and α is also small.

A plot of the equilibrium probability of serving an emergency requiring less than one mile transit distance (x_2) as a function of α and β/α is presented in Figure 4.5. Here we observe a minor sensitivity to changes in the ratio β/α, but a somewhat larger sensitivity to specification errors in α. We also observe that the probability is much more sensitive to variations in δ.

These sensitivity analyses of the emergency occurrence parameters and emergency service parameters indicate

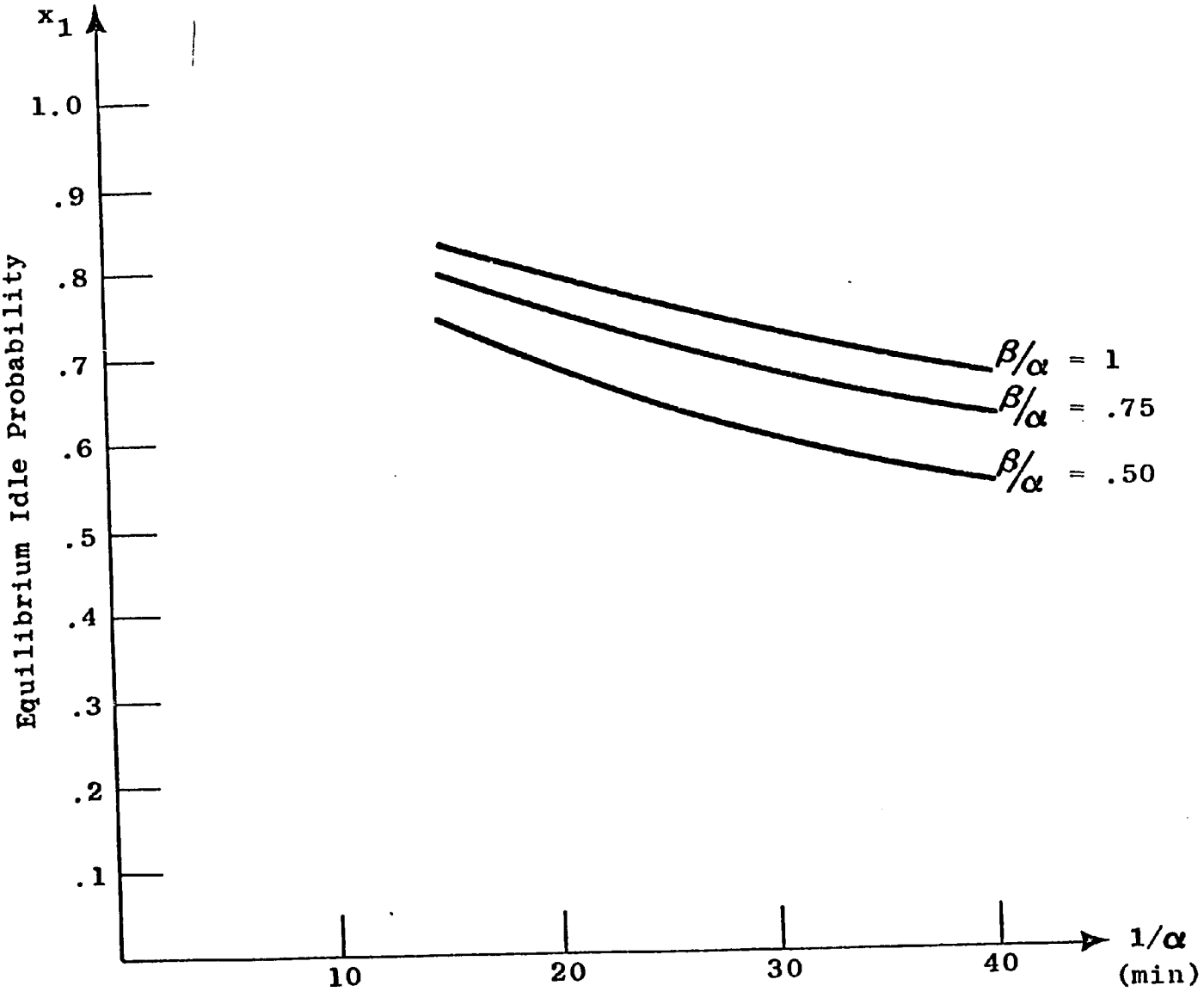


Figure 4.4
Equilibrium Idle Probability as Function
of Mean Emergency Service Times
Single Ambulance System δ = .139

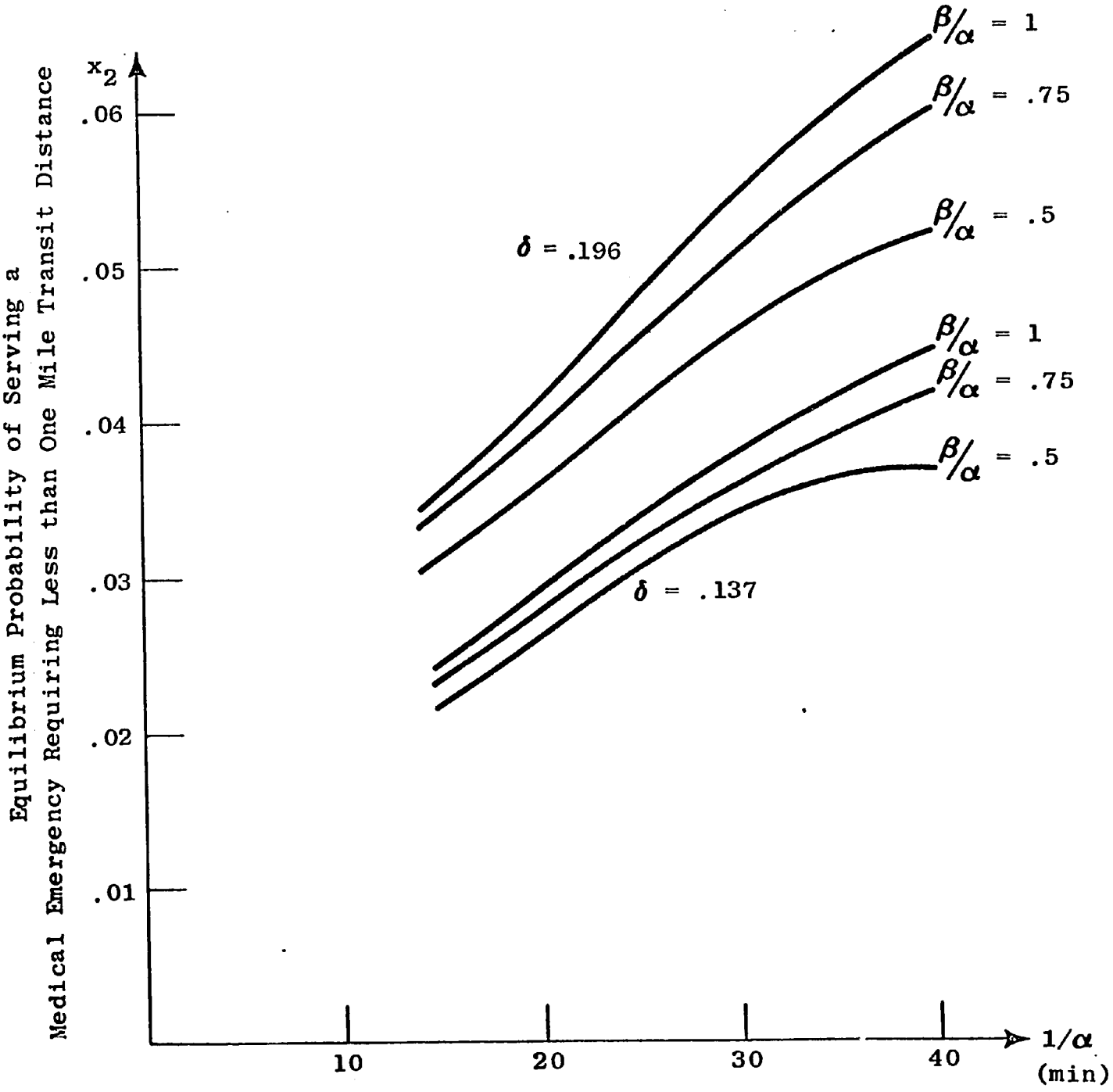


Figure 4.5
Equilibrium Probability x_2 as Function
of Mean Emergency Service Times
Single Ambulance System

$\delta = .139$

β/α \ α		.067 (15 min.)	.05 (20 min.)	.033 (30 min.)	.025 (40 min.)
1	x_1	.827	.789	.724	.674
	x_2	.024	.029	.038	.045
	x_3	.149	.182	.238	.281
.75		.795	.751	.678	.623
		.023	.028	.036	.042
		.182	.221	.286	.335
.50		.739	.687	.604	.543
		.022	.026	.032	.037
		.239	.287	.364	.420

$\delta = .196$

β/α \ α	.067	.05	.033	.025
1	.827	.789	.724	.673
	.034	.041	.054	.064
	.139	.170	.222	.262
.75	.797	.754	.681	.626
	.033	.039	.051	.060
	.170	.207	.268	.314
.50	.745	.694	.610	.550
	.030	.036	.046	.052
	.225	.270	.344	.398

Table 4.3

Sensitivity of Equilibrium Probabilities to Variations in Emergency Service Parameters

that the various system equilibrium probabilities do not drastically change with small variations in the parameters. Hence we are led to conclude that our model is reasonably "robust" with respect to such changes. The equilibrium idle probabilities are insensitive to changes in δ and are only moderately sensitive to changes in the other four parameters. On the other hand, the equilibrium "busy" probabilities x_2 and x_3 are more sensitive to changes in δ, while variations in the other four parameters introduce only minor variations into these values.

4.2 RESULTS FOR A TWO AMBULANCE SYSTEM

The matrix of switching probabilities for a two ambulance system is presented below:

S_n^2, S_n^1 / Y_n	(0,0)	(0,2)	(0,3)	(2,0)	(2,2)	(2,3)	(3,0)	(3,2)	(3,3)
0	0	0	0	0	1	1	0	1	1
(1,2)	P_1	0	0	ρ	0	0	ρ	0	0
(1,3)	P_2	0	0	$1-\rho$	0	0	$1-\rho$	0	0
(2,2)	P_3	Ψ	Ψ	0	0	0	0	0	0
(2,3	$1-P_1-P_2-P_3$	$1-\Psi$	$1-\Psi$	0	0	0	0	0	0

Here P_i represents the probability of a specified assignment when both ambulances are idle at the time of the emergency occurrence. The parameters Ψ and ρ represent the probabilities ambulances two and one respectively are

dispatched to emergencies within one mile when the other vehicle is busy.

From Figure 4.1 we observe that the single vehicle distribution policies I and II result in non-intersecting one-mile service regions which include the areas of heaviest demand. For the two vehicle distribution involving these vehicles, the estimated switching parameters are $\hat{P}_1 = \hat{\rho} = .195$, $\hat{P}_3 = \hat{\Psi} = .196$, and $\hat{P}_2 = .334$.

A second distribution of interest places the vehicles to obtain more uniform geographical coverage as in Figure 4.6. Parameter estimates for this distribution are $\hat{P}_1 = \hat{\rho} = .176$, $\hat{P}_3 = \hat{\Psi} = .163$, and $\hat{P}_2 = .347$.

A third distribution of interest is one where the two vehicles are constrained to operate from a single station. The 15^{th} precinct police station was chosen to represent this distribution. Service region I in Figure 4.1 illustrates the resulting system. Parameter estimates for this distribution are $\hat{P}_1 = \hat{P}_3 = .097$, $\hat{\rho} = \hat{\Psi} = .195$, and $\hat{P}_2 = .403$.

The equilibrium state probabilities for these three two-ambulance systems are presented in Table 4.4.

From these results several quantities of interest can be calculated. Certain of these are presented in Table 4.5 below.

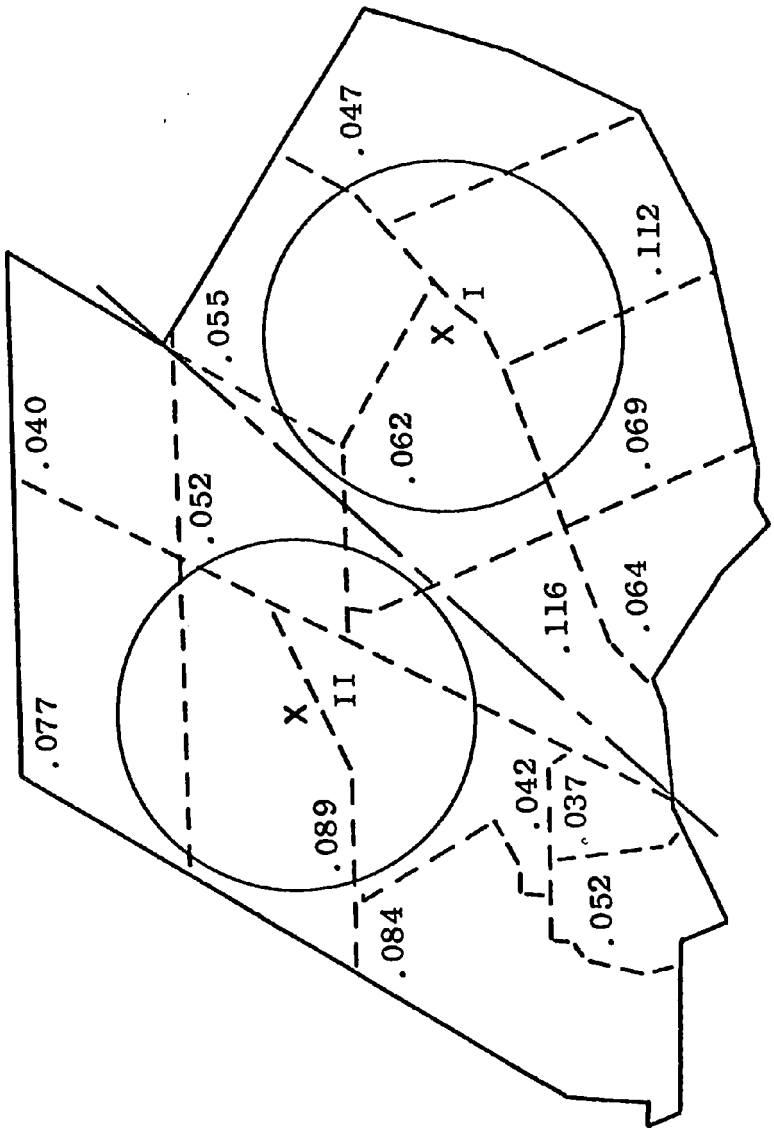


Figure 4.6
Two Ambulance Distribution Policy to Achieve Uniform Geographic Coverage

State (S_n^2, S_n^1)	Distribution Policy I	II	III
(0,0)	.7657	.7656	.7649
(0,2)	.0378	.0343	.0207
(0,3)	.0724	.0745	.0841
(2,0)	.0383	.0319	.0207
(2,2)	.0021	.0016	.0011
(2,3)	.0025	.0022	.0019
(3,0)	.0623	.0695	.0841
(3,2)	.0024	.0023	.0019
(3,3)	.0165	.0182	.0206

Table 4.4

Equilibrium State Probabilities for
Two Ambulance System

Equilibrium Result	Distribution Policy I	II	III
Probability both vehicles idle	.7657	.7656	.7649
Probability one vehicle idle	.2108	.2102	.2096
Probability one or more vehicles idle	.9765	.9758	.9745
Probability both vehicles on "short" run (less than one mile)	.0021	.0016	.0011
Probability one vehicle on "short" run	.0810	.0707	.0452
Probability one or more vehicles on "short" run	.0831	.0723	.0463
Probability both vehicles on "long" run (greater than one mile)	.0165	.0182	.0206
Probability one vehicle on "long" run	.1396	.1485	.1720
Probability one or more vehicles on "long" run	.1561	.1667	.1926

Table 4.5

Equilibrium Probabilistic Results for
the Two Ambulance System

Inspection of Table 4.5 indicates that both recovery vehicles will be idle at the instant of a medical emergency about 76% of the time; one or more ambulances will be idle at the instant of a medical emergency about 98% of the time. Both of these equilibrium probabilities are essentially independent of the distribution policy.

The equilibrium probabilities that one or more vehicles will be serving an emergency requiring less than one mile transit distance to the scene at the instant of a medical emergency vary from .0831 to .0463 depending upon the chosen distribution policy. Distribution policies I and II have nearly equal values and both result in measurable improvements over policy III. A similar result holds for the equilibrium probability that one or more vehicles are serving an emergency requiring more than one mile transit distance to the scene at the instant of a medical emergency.

4.3 RESULTS FOR A THREE AMBULANCE SYSTEM

We conclude this chapter by presenting results equivalent to those of the previous sections for a three vehicle independent ambulance system. The matrix of switching probabilities for such a system is presented below.

S_n^3, S_n^2, S_n^1 \ Y_n	0	(1,2)	(1,3)	(2,2)	(2,3)	(3,2)	(3,3)
000	0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
002	0	0	0	a_1	a_2	a_3	a_4
003	0	0	0	a_1	a_2	a_3	a_4
020	0	b_1	b_2	0	0	b_3	b_4
022	0	0	0	0	0	ρ	$1-\rho$
023	0	0	0	0	0	ρ	$1-\rho$
030	0	b_1	b_2	0	0	b_3	b_4
032	0	0	0	0	0	ρ	$1-\rho$
033	0	0	0	0	0	ρ	$1-\rho$
200	0	c_1	c_2	c_3	c_4	0	0
202	0	0	0	ψ	$1-\psi$	0	0
203	0	0	0	ψ	$1-\psi$	0	0
220	0	ν	$1-\nu$	0	0	0	0
222	1	0	0	0	0	0	0
223	1	0	0	0	0	0	0
230	0	ν	$1-\nu$	0	0	0	0
232	1	0	0	0	0	0	0
233	1	0	0	0	0	0	0
300	0	c_1	c_2	c_3	c_4	0	0
302	0	0	0	ψ	$1-\psi$	0	0
303	0	0	0	ψ	$1-\psi$	0	0
320	0	ν	$1-\nu$	0	0	0	0
322	1	0	0	0	0	0	0
323	1	0	0	0	0	0	0
330	0	ν	$1-\nu$	0	0	0	0
332	1	0	0	0	0	0	0
333	1	0	0	0	0	0	0

Hence δ_i is the probability of a specified assignment when all vehicles are idle at the instant of an emergency occurrence. The parameters a_i, b_i, and c_i are the probabilities of specified assignments when two of the three vehicles are idle; ρ, ψ, and ν are the probabilities of specified assignments when only one vehicle is available.

We first consider the distribution where the three vehicles are placed to maximize the coverage of the region of heaviest demand as in Figure 4.7. For this vehicle placement on the following parameter estimates were derived:

$\hat{\delta}_1 = .146$ $\hat{a}_1 = .146$ $\hat{b}_1 = .146$ $\hat{c}_1 = .196$

$\hat{\delta}_2 = .107$ $\hat{a}_2 = 348$ $\hat{b}_2 = .105$ $\hat{c}_2 = .275$

$\hat{\delta}_3 = .146$ $\hat{a}_3 = .174$ $\hat{b}_3 = .174$ $\hat{c}_3 = .195$

$\hat{\delta}_4 = .225$ $\hat{a}_4 = .332$ $\hat{b}_4 = .575$ $\hat{c}_4 = .334$

$\hat{\delta}_5 = 151$

$\hat{\delta}_6 = .225$

$\hat{\rho} = .231$ $\hat{\psi} = .195$ $\hat{\nu} = .196$

A second distribution of interest places the vehicles in non-intersecting one-mile service regions to achieve more uniform coverage of the precinct. The resulting distribution is shown in Figure 4.8. The

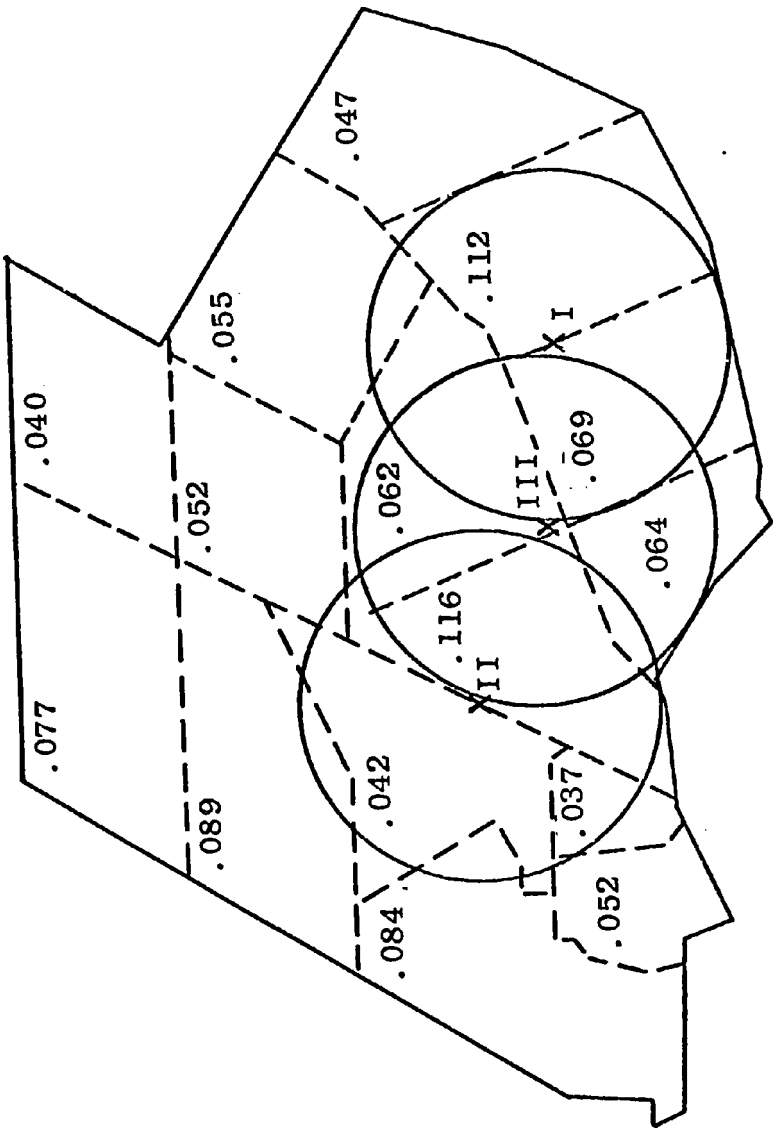


Figure 4.7
Three Ambulance Distribution Policy to Achieve Coverage of Area of Maximum Demand

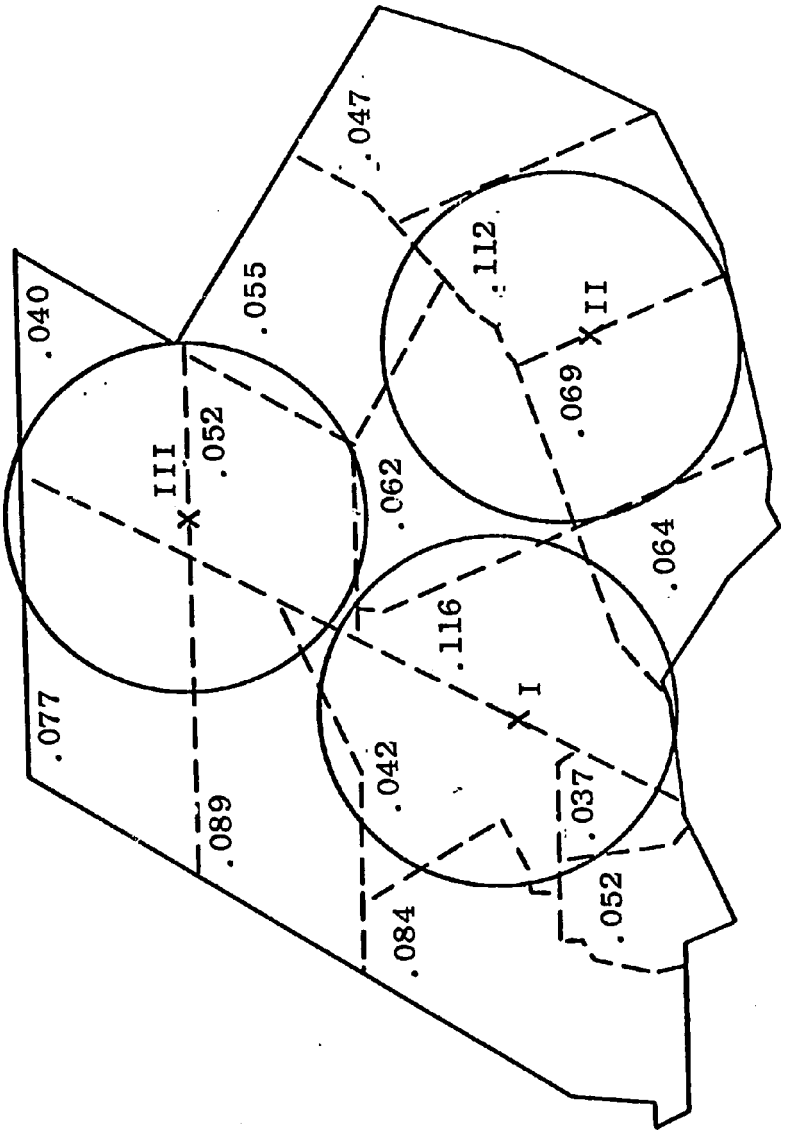


Figure 4.8
Three Ambulance Distribution Policy to Achieve Uniform Coverage of Service Area

following parameter estimates were derived for this model.

$$\hat{\delta}_1 = .195 \quad \hat{a}_1 = .196 \quad \hat{b}_1 = .195 \quad \hat{c}_1 = .195$$
$$\hat{\delta}_2 = .136 \quad \hat{a}_2 = .298 \quad \hat{b}_2 = .466 \quad \hat{c}_2 = .334$$
$$\hat{\delta}_3 = .196 \quad \hat{a}_3 = .143 \quad \hat{b}_3 = .143 \quad \hat{c}_3 = .196$$
$$\hat{\delta}_4 = .138 \quad \hat{a}_4 = .365 \quad \hat{b}_4 = .266 \quad \hat{c}_4 = .275$$
$$\hat{\delta}_5 = .143$$
$$\hat{\delta}_6 = .192$$
$$\hat{\rho} = .143 \quad \hat{\psi} = .196 \quad \hat{\nu} = .195$$

The third distribution studied here places all three vehicles at a single point, again at the 15th precinct police station. Parameter estimates are as follows:

$$\hat{\delta}_1 = \hat{\delta}_3 = \hat{\delta}_5 = .065 \quad \hat{a}_1 = \hat{a}_3 = \hat{b}_1 = \hat{b}_3 = \hat{c}_1 = \hat{c}_3 = .097$$
$$\hat{\delta}_2 = \hat{\delta}_4 = \hat{\delta}_6 = .268 \quad \hat{a}_2 = \hat{a}_4 = \hat{b}_2 = \hat{b}_4 = \hat{c}_2 = \hat{c}_4 = .403$$
$$\hat{\rho} = \hat{\psi} = \hat{\nu} = .195$$

The equilibrium state probabilities for these three alternative distributions are given in Table 4.6. Results derived from these probabilities are tabulated in Table 4.7.

(s_n^3, s_n^2, s_n^1)	DISTRIBUTION POLICY I	II	III
000	.7683	.7710	.7552
002	.0285	.0360	.0165
003	.0237	.0303	.0561
020	.0281	.0363	.0165
022	.0011	.0018	.0004
023	.0010	.0031	.0017
030	.0453	.0290	.0533
032	.0024	.0023	.0017
033	.0011	.0022	.0035
200	.0289	.0266	.0167
202	.0014	.0013	.0004
203	.0017	.0019	.0017
220	.0013	.0013	.0004
222	.0001	.0001	.0000
223	.0001	.0002	.0001
230	.0024	.0016	.0017
232	.0002	.0002	.0001
233	.0004	.0004	.0004
300	.0474	.0398	.0550
302	.0026	.0029	.0017
303	.0033	.0037	.0066
320	.0035	.0024	.0017
322	.0002	.0002	.0001
323	.0004	.0005	.0003
330	.0061	.0030	.0065
332	.0006	.0006	.0004
333	.0010	.0009	.0016

Table 4.6

Equilibrium Probabilities for the Three Ambulance System

Equilibrium Result	Distribution Policy		
	I	II	III
Probability Three Vehicles Idle	.7683	.7710	.7552
Probability Two Vehicles Idle	.2019	.1980	.2141
Probability One Vehicle Idle	.0279	.0275	.0280
Probability One or More Vehicle Idle	.9981	.9965	.9973
Probability Three Vehicles on "Short" Run (less than one mile)	.0001	.0001	.0000
Probability Two Vehicles on "short" run	.0043	.0050	.0015
Probability One Vehicle on "short" run	.1005	.1146	.0610
Probability One or More Vehicles on "short" run	.1049	.1197	.0625
Probability Three Vehicles on "Long" run (greater than one mile)	.0010	.0009	.0016
Probability Two Vehicles on "long" run	.0119	.0104	.0177
Probability One Vehicle on "long" run	.1305	.1139	.1749
Probability One or More Vehicles on "long" run	.1434	.1252	.1942

Table 4.7

Equilibrium Probabilistic Results
For the Three Ambulance System

Inspection of Table 4.7 indicates that all three recovery vehicles will be available at the instant of a medical emergency 75% to 77% of the time. One or more vehicles will be available about 99% of the time. Results for the distribution policies I and II are nearly identical; the equilibrium idle probabilities for distribution policy III (placing all vehicles at a single station) do not significantly differ from those of policies I and II.

The equilibrium probabilities that one or more vehicles will be serving an emergency requiring less than one mile transit distance to the scene at the instant of a medical emergency vary from .1197 to .0625 depending upon the chosen distribution policy. Distribution policies I and II have nearly equal values, and both result in measurable improvements over policy III.

The equilibrium probabilities that one or more vehicles are on runs requiring transit distances greater than one mile vary from .1252 to .1942. Again, results for policies I and II are substantially better than those for policy III.

We shall delay the comparison of the one, two and three ambulance systems until Chapter VI. In Chapter VI we shall compare these three systems with each other,

and with results for the dual function ambulance system which is studied in the next chapter.

CHAPTER V

ALLOCATION AND DISTRIBUTION POLICIES FOR THE DUAL FUNCTION PUBLIC AMBULANCE SYSTEM

As we have observed in Chapters I and II, many urban areas operate an ambulance system in conjunction with the police department. In such systems, "dual function" police-ambulance vehicles respond to police calls and medical emergencies on a first come-first served basis. Since the number of police calls being served by such systems is much greater than the number of ambulance calls, one might conjecture that a dual function system will not serve medical emergencies as effectively as a single function ambulance system. (As we have observed, such a loss of effectiveness will result in decreased vehicle availability and/or increased medical service time.)

We shall now examine this conjecture by applying the results of Chapter III to dual function recovery systems incorporating one, two, and more than two ambulances. We shall examine these three systems in Sections 5.1, 5.2, and 5.3 respectively. All numerical results

are developed from a general purpose computer program designed for the analysis of these alternative systems. This program is documented in Appendix C.

5.1 RESULTS FOR A SINGLE VEHICLE DUAL FUNCTION AMBULANCE SYSTEM

For the single vehicle, dual function ambulance system the switching probabilities can be written

$$\Pr\{Y_n = \ell \mid Z_n = i;\ S_n^1, \cdots, S_n^r\} = \Pr\{Y_n = \ell \mid Z_n = i;\ S_n^1\}$$

where Y_n takes on the values 0, 1, 2, or 3, corresponding respectively to situations where the system rejects the emergency, serves the emergency as a police call, serves the emergency as an ambulance call requiring less than one mile transit distance, or serves the emergency as an ambulance call requiring more than one mile transit distance to the scene. The random variable Z_n takes on values 0 or 1, corresponding to situation where the n^{th} emergency is an ambulance call or a police call. The vehicle state S_n^1 takes on values 0, 1, 2, 3, corresponding to states where the vehicle is idle, serving a police call, serving an ambulance call requiring less than one mile transit distance, or serving an ambulance call requiring more than one mile transit distance.

The matrix of switching probabilities can then be written as follows:

Y_n \ (S_n^1, Z_n)	(0,0)	(0,1)	(1,0)	(1,1)	(2,0)	(2,1)	(3,0)	(3,1)
0	0	$1-\phi_1$	1	1	1	1	1	1
1	0	ϕ_1	0	0	0	0	0	0
2	δ	0	0	0	0	0	0	0
3	$1-\delta$	0	0	0	0	0	0	0

where ϕ_1 is the probability a police call is accepted by the system when the recovery vehicle is idle, and δ is the probability that a medical emergency requires a transit distance to the scene of less than one mile.

The conditional vehicle transition probabilities $\Pr\{S_n^1 | Y_{n-1}, S_{n-1}^1, X_n = x\}$ can be written:

$Y_{n-1} = 0$

S_{n-1}^1 \ S_n^1	0	1	2	3
0	1	0	0	0
1	$1-e^{-\gamma x}$	$e^{-\gamma x}$	0	0
2	$1-e^{-\alpha x}$	0	$e^{-\alpha x}$	0
3	$1-e^{-\beta x}$	0	0	$0^{-\beta x}$

$Y_{n-1} = 1$

S^1_{n-1} \ S^1_n	0	1	2	3
0	$0-e^{-\gamma x}$	$e^{-\gamma x}$	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

$Y_{n-1} = 2$

S^1_{n-1} \ S^1_n	0	1	2	3
0	$1-e^{-\alpha x}$	0	$e^{-\alpha x}$	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

$Y_{n-1} = 3$

S^1_{n-1} \ S^1_n	0	1	2	3
0	$1-e^{-\beta x}$	0	0	$e^{-\beta x}$
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

Here γ is the police call service rate, and α and β are the ambulance call service rates as defined in Chapter IV.

Similarly, the conditional state transition probabilities $\Pr\{S_n^1, Z_n | S_{n-1}^1, Z_{n-1}, Y_{n-1}\}$ are derived as in equation 3.2:

$$\Pr\{S_n^1, Z_n | S_{n-1}^1, Z_{n-1}, Y_{n-1}\} = \int_0^\infty \Pr\{S_n^1 | S_{n-1}^1, Y_{n-1}, X_n\} \times$$

$$d_x \Pr\{X_n, Z_n | Z_{n-1}\}$$

We have seen (3.4) that $d_x \Pr\{X_n, Z_n | Z_{n-1}\}$ can be represented by the matrix

$$\begin{bmatrix} \Pi dF_{00}(x) & (1-\Pi) dF_{01}(x) \\ \Pi dF_{10}(x) & (1-\Pi) dF_{11}(x) \end{bmatrix}$$

The resulting one step system transition probabilities $\Pr\{S_n^1, Z_n | S_{n-1}^1, Z_{n-1}\}$ are then found from the following relation:

$$\Pr\{S_n^1, Z_n | S_{n-1}^1, Z_{n-1}\}$$

$$= \sum_{\ell=0}^{3} \Pr\{S_n^1, Z_n | S_{n-1}^1, Z_{n-1}, Y_{n-1} = \ell\} \Pr\{Y_{n-1}$$

$$= \ell | Z_{n-1}, S_{n-1}\}$$

These are presented for this system in Table 5.1.

Where

$$p_y^{ij} = \int_0^\infty e^{-yx} \lambda_{ij} e^{-\lambda_{ij} x} dx = \left(\frac{\lambda_{ij}}{\lambda_{ij} + y}\right) \quad i,j=0,1$$

Equilibrium state probabilities can be derived from the one step transition probabilities by techniques identical to those of Section 4.1.

From Chapter II we recall that the maximum likelihood estimates for the semi-Markov emergency occurrence parameters are:

$$\hat{\lambda}_{00} = 1/14.3$$
$$\hat{\lambda}_{01} = 1/16.2$$
$$\hat{\lambda}_{10} = 1/23.1$$
$$\hat{\lambda}_{11} = 1/14.6$$

The police call service rate is estimated as $\hat{\gamma} = 1/38.3 = .026$.

The model parameters for ambulance operation in the dual system are estimated as in Chapters II and IV. The maximum likelihood estimate for the proportion of total calls which are ambulance calls was found to be $\Pi = .214$.

We now consider the parameter ϕ_1 in more detail. We have previously observed that the dual function

(s_{n-1}^1, z_{n-1}) \ (s_n^1, z_n)	(0,0)	(0,1)	(1,0)	(1,1)	(2,0)	(2,1)	(3,0)	(3,1)
(0,0)	$\delta\pi(1-P_\alpha^{00})+(1-\delta)\pi(1-P_\beta^{00})$	$\delta(1-\pi)(1-P_o^{01})+(1-\delta)(1-\pi)(1-P_\beta^{01})$	0	0	$\delta\pi P_\alpha^{00}$	$\delta(1-\pi)P_\alpha^{01}$	$(1-\delta)\pi P_\beta^{00}$	$(1-\delta)(1-\pi)P_\beta^{01}$
(0,1)	$\pi(1-\phi_1)+\pi\phi_1(1-P_\gamma^{10})$	$(1-\pi)(1-\phi_1)+\phi_1(1-\pi)(1-P_\gamma^{11})$	$\pi\phi_1 P_\gamma^{10}$	$(1-\pi)\phi_1 P_\gamma^{11}$	0	0	0	0
(1,0)	$\pi(1-P_\gamma^{00})$	$(1-\pi)(1-P_\gamma^{01})$	πP_γ^{00}	$(1-\pi)P_\gamma^{01}$	0	0	0	0
(1,1)	$\pi(1-P_\gamma^{10})$	$(1-\pi)(1-P_\gamma^{11})$	πP_γ^{10}	$(1-\pi)P_\gamma^{11}$	0	0	0	0
(2,0)	$\pi(1-P_\alpha^{00})$	$(1-\pi)(1-P_\alpha^{01})$	0	0	πP_α^{00}	$(1-\pi)P_\alpha^{01}$	0	0
(2,1)	$\pi(1-P_\alpha^{10})$	$(1-\pi)(1-P_\alpha^{11})$	0	0	πP_α^{10}	$(1-\pi)P_\alpha^{11}$	0	0
(3,0)	$\pi(1-P_\beta^{00})$	$(1-\pi)(1-P_\beta^{01})$	0	0	0	0	πP_β^{00}	$(1-\pi)P_\beta^{01}$
(3,1)	$\pi(1-P_\beta^{10})$	$(1-\pi)(1-P_\beta^{11})$	0	0	0	0	πP_β^{10}	$(1-\pi)P_\beta^{11}$

Table 5.1

One Step Transition Probabilities for Single Vehicle, Dual Function System

recovery vehicle is one of several vehicles assigned to service police calls in a precinct. During actual operation the total number of vehicles and the number of dual function vehicles both vary from shift to shift and from day to day because vehicles become unavailable due to maintenance, court appearances, and other special missions.

One would expect that each available vehicle should service approximately the same proportion of all police calls. Inspection of vehicle availability statistics in Precinct 15 for the November-December data collection period revealed that an average of eleven vehicles were available during each shift. If only one of these were a station wagon (the dual function recovery vehicle), then on a random selection basis we would expect this to be dispatched to about one of every eleven police calls when it is available. As vehicles are assigned to other functions, this proportion should increase. Hence we shall parameterize the model to indicate the sensitivity to variations in ϕ_1 from .10 to .50.

Since the emergency service rates are greater than the emergency occurrence rates, the system is ergodic, and the equilibrium state probabilities exist. These can be written $\Pr\{S_n^1, Z_n\}$, and of course they are independent

of n. The equilibrium vehicle state probabilities at the instant of a medical emergency $Pr(S_n^1 | Z_n = 0)$ can be written:

$$Pr(S_n^1 | Z_n = 0) = Pr(S_n^1, Z_n = 0)/Pr(Z_n = 0) = \frac{1}{\Pi} Pr(S_n^1, Z_n = 0)$$

These have been computed for the two basic single vehicle distribution policies of Chapter 4 ($\hat{\delta} = .139$ and $\hat{\delta} = .196$), and the results are presented in Table 5.1. Plots of the various probabilities as functions of ϕ_1 are presented in Figures 5.1 to 5.4. As before, a "short" medical emergency is one requiring less than one mile transit distance to the scene and a "long" medical emergency is one requiring more than one mile transit distance to the scene.

Notice that when $\phi_1 = 0$, the equilibrium results for this model should be identical to those for the single function, single vehicle system described in Section 4.1. Comparison of Tables 4.1 and 5.1 reveals that the corresponding values are approximately equal. The small differences can be attributed to numerical errors and to differences in the semi-Markov and Gamma arrival processes which were specified as inputs to the two systems. From this we can conclude that a semi-Markov model with negative-exponential "inter-event" times is an adequate repre-

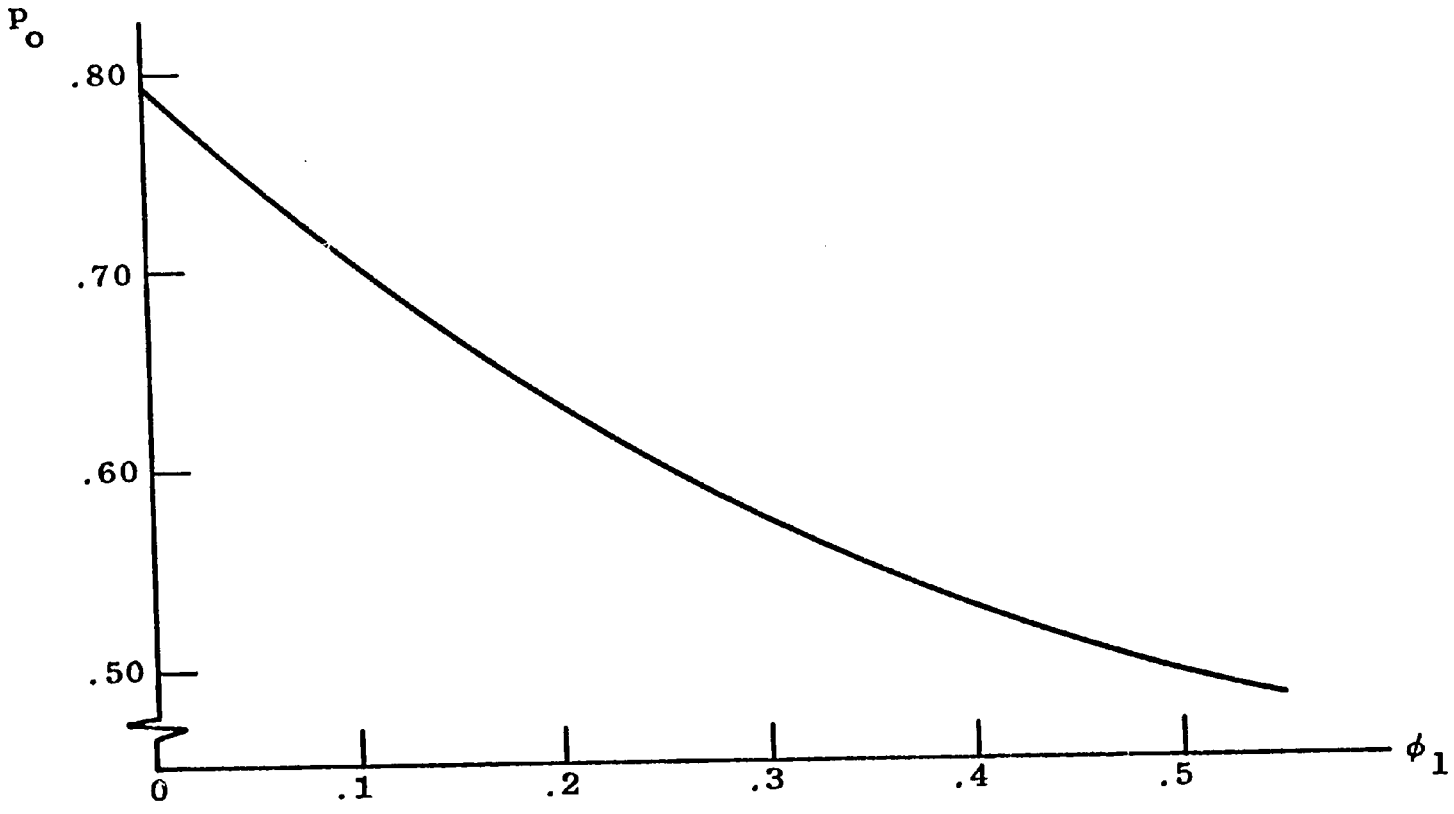


Figure 5.1

Equilibrium Idle Probability For

Single Vehicle/Dual Function System

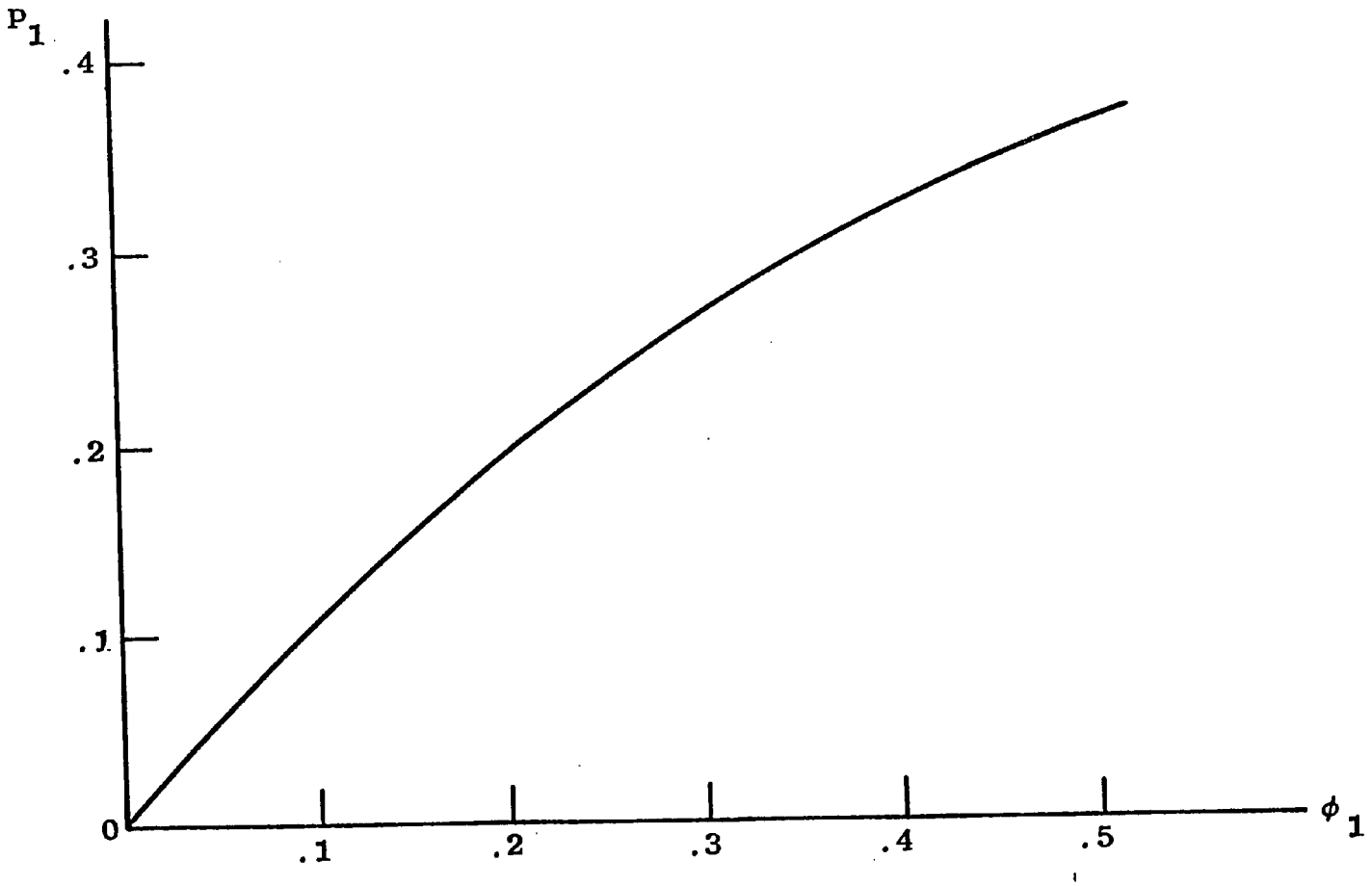


Figure 5.2

Equilibrium Police Call Service Probability for Single Vehicle/Dual Function System

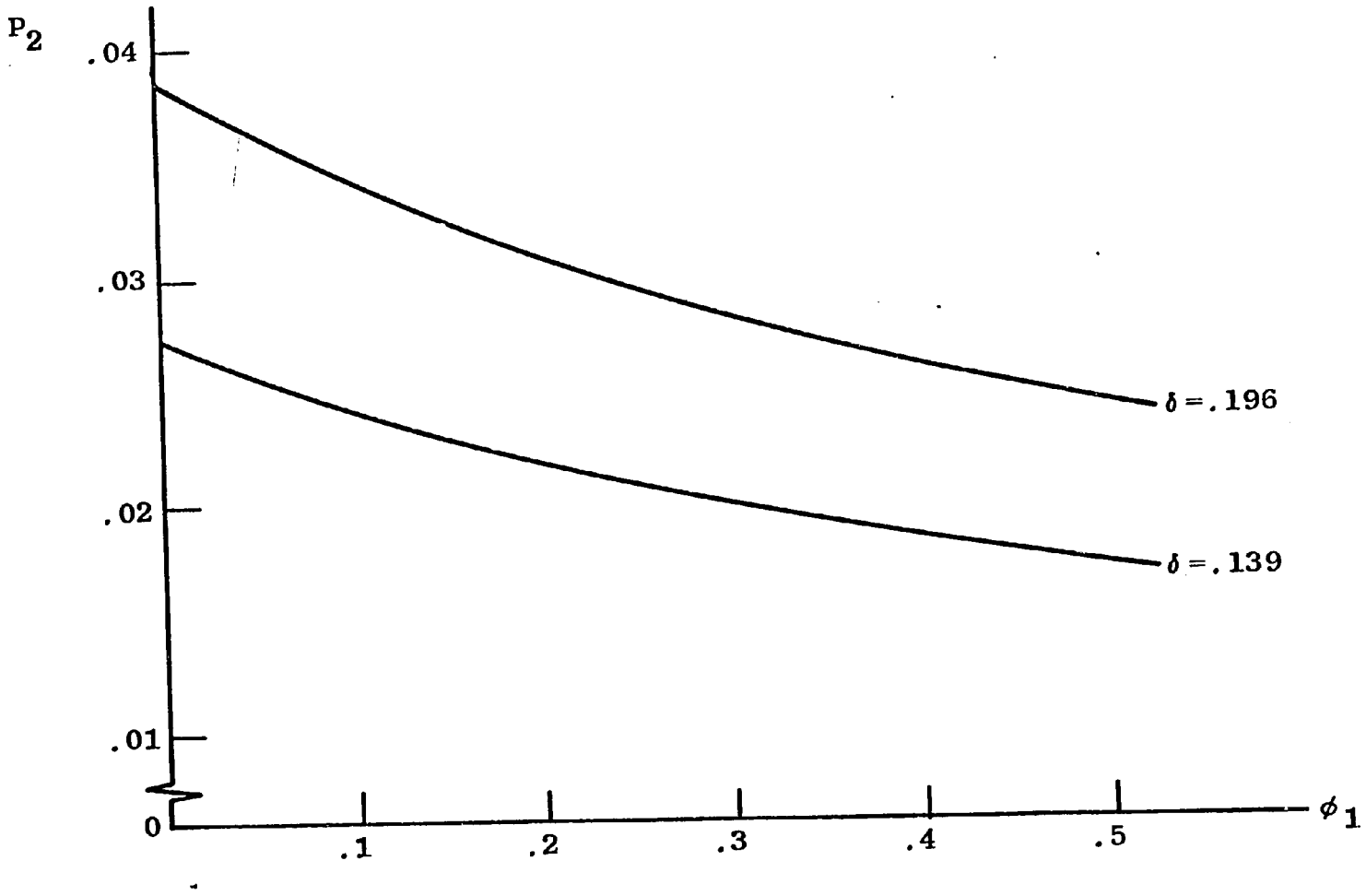


Figure 5.3

Equilibrium "Short" Medical Emergency

Service Probability for Single Vehicle/Dual Function System

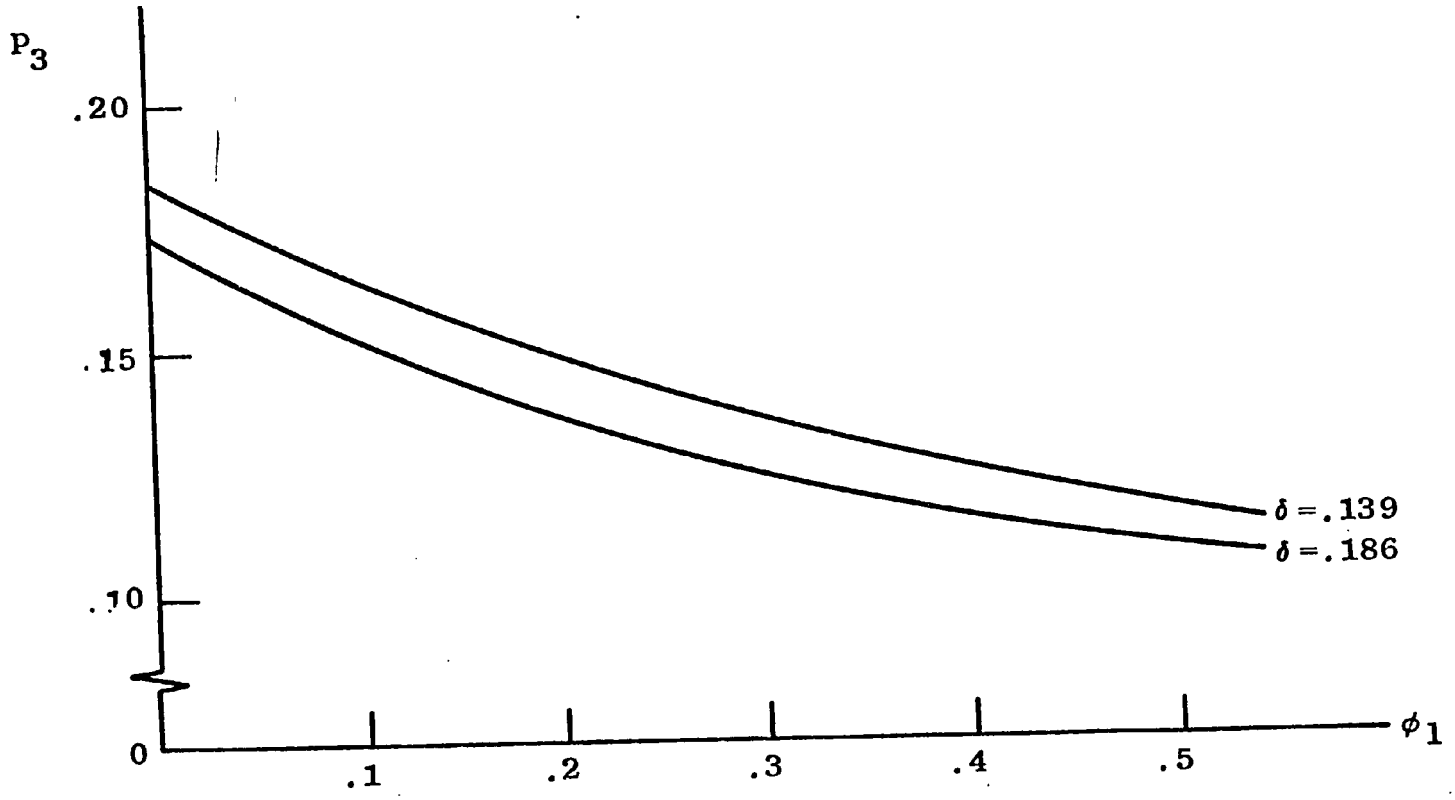


Figure 5.4

Equilibrium "Long" Medical Emergency Service Probability for Single Vehicle/Dual Function System

Table 5.2

Equilibrium Vehicle State Probabilities For Single Vehicle Dual Function System

$\delta = .139$

	Equilibrium Vehicle State Probability			
ϕ_1	P_0	P_1	P_2	P_3
0	.7874	0	.0272	.1854
.1	.6987	.1126	.0242	.1645
.2	.6307	.1990	.0218	.1485
.3	.5767	.2675	.0200	.1358
.4	.5330	.3231	.0184	.1255
.5	.4967	.3691	.0172	.1169

$\delta = .196$

ϕ_1	P_0	P_1	P_2	P_3
0	.7883	0	.0384	.1733
.1	.6994	.1127	.0341	.1538
.2	.6312	.1992	.0308	.1388
.3	.5772	.2678	.0282	.1269
.4	.5334	.3234	.0260	.1172
.5	.4971	.3694	.0242	.1093

sentation of the dual-function emergency occurrence process. This conclusion strengthens the inferences drawn from the statistical analyses which were conducted in Section 2.2.3 to partially identify the underlying process. In addition, these results further indicate that the Gamma distribution is an adequate model for the distribution of times between ambulance calls.

We observe that variations in δ do not affect the vehicle idle probability or the police call service probability. However, variations in ϕ do introduce substantial changes in these two quantities. The equilibrium idle probability decreases from .79 when $\phi_1 = 0$ to .49 when $\phi_1 = .5$.

The equilibrium probabilities of serving either of the two types of medical emergencies are quite insensitive to variations in ϕ_1. As we saw in Chapter IV, these quantities are moderately sensitive to changes in δ. From Figures 5.3 and 5.4 we observe that this sensitivity is essentially uniform for all ϕ_1.

We shall delay a more detailed discussion of these results until the system comparisons of Chapter VI.

5.2 RESULTS FOR A TWO-VEHICLE, DUAL FUNCTION AMBULANCE SYSTEM

The matrix of switching probabilities for the two vehicle, dual function ambulance system are presented in Table 5.3. These are derived by applying the decision rules of Section 3.2.

Here the p_i represent the probabilities of various vehicle assignments when both vehicles are idle at the time of an ambulance call, δ and γ represent the probabilities of the assignments when one of the vehicles is available at the time of an ambulance call, ϕ_1 represents the probability the system will accept a police call when one vehicle is available, and ϕ_2 represents the probability of such an acceptance when two vehicles are available.

From previous results (Sections 4.2 and 5.1), estimates are available for all parameters in this system except ϕ_1 and ϕ_2. We shall now consider two reasonable dispatch "strategies" giving rise to values for these parameters. We first consider a situation where the dispatcher assigns idle vehicles to police calls without initially considering the existing system state. As we indicated in the previous section, the resulting system is likely to be operating with both parameters set at a value determined by the total number of vehicles in the

Table 5.3

Switching Probabilities for Two Vehicle Dual Function System

(s_n^2, s_n^1, z_n) \ Y_n	0	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
000	0	0	p_1	p_2	0	p_3	p_4
001	$1-\phi_2$	$\phi_{2/2}$	0	0	$\phi_{2/2}$	0	0
010	0	0	0	0	0	δ	$1-\delta$
011	$1-\phi_1$	0	0	0	ϕ_1	0	0
020	0	0	0	0	0	δ	$1-\delta$
021	$1-\phi_1$	0	0	0	ϕ_1	0	0
030	0	0	0	0	0	δ	$1-\delta$
031	$1-\phi_1$	0	0	0	ϕ_1	0	0
100	0	0	γ	$1-\gamma$	0	0	0
101	$1-\phi_1$	ϕ_1	0	0	0	0	0
110	1	0	0	0	0	0	0
111	1	0	0	0	0	0	0
120	1	0	0	0	0	0	0
121	1	0	0	0	0	0	0
130	1	0	0	0	0	0	0
131	1	0	0	0	0	0	0
200	0	0	γ	$1-\gamma$	0	0	0
201	$1-\phi_1$	ϕ_1	0	0	0	0	0
210	1	0	0	0	0	0	0
211	1	0	0	0	0	0	0
220	1	0	0	0	0	0	0
221	1	0	0	0	0	0	0
230	1	0	0	0	0	0	0
231	1	0	0	0	0	0	0
300	0	0	γ	$1-\gamma$	0	0	0
301	$1-\phi_1$	ϕ_1	0	0	0	0	0
310	1	0	0	0	0	0	0
311	1	0	0	0	0	0	0
320	1	0	0	0	0	0	0
321	1	0	0	0	0	0	0
330	1	0	0	0	0	0	0
331	1	0	0	0	0	0	0

precinct. For this case we shall consider subsequent system performance with values of $\phi_1 = \phi_2$ ranging from 0.0 to 0.5.

A second potential dispatch strategy attempts to conserve recovery vehicles so that these are more likely to be idle at the instant of a medical emergency. The specific strategy we shall consider here allows ϕ_2 to vary from 0.0 to 0.5, but maintains ϕ_1 at a value of 0.1. Under this strategy, when only one recovery vehicle is idle, this vehicle accepts a smaller percentage of police calls for service. In this way, we increase the probability that at least one vehicle is available to service a medical emergency.

Equilibrium vehicle state probabilities are presented for both of these dispatch strategies for distribution policies I and II (defined in Section 4.2) in Table 5.3 and 5.4. Distribution policy III, which places all recovery vehicles at a fixed location, is no longer feasible in a dual function system, since all vehicles are on patrol when they are not serving emergencies.

Various quantities of interest derived from these equilibrium probabilities are tabulated in Tables 5.5 and 5.6, and these results are plotted in Figures 5.5-5.8. In these results $P_j(k)$ is the equilibrium probability that

Table 5.4

Equilibrium Vehicle State Probabilities
At Instant of Medical Emergency For
Two Vehicle/Dual Function System

Distribution Policy I

	Dispatch Strategy 1						Dispatch Strategy 2			
(s_n^2, s_n^1) \ (ϕ_1, ϕ_2)	(0,0)	(.1,.1)	(.2,.2)	(.3,.3)	(.4,.4)	(.5,.5)	(.1,.2)	(.1,3)	(.1,4)	(.1,.5)
00	.7703	.6575	.5667	.4935	.4343	.3859	.5878	.5337	.4905	.4553
01	0.0	.0536	.0911	.1170	.1349	.1472	.0866	.1123	.1327	.1494
02	.0349	.0292	.0247	.0211	.0182	.0159	.0266	.0247	.0231	.0218
03	.0709	.0610	.0529	.0463	.0409	.0364	.0572	.0543	.0519	.0500
10	0.0	.0543	.0921	.1182	.1361	.1483	.0873	.1128	.1333	.1499
11	0.0	.0096	.0320	.0609	.0924	.1243	.0155	.0201	.0237	.0267
12	0.0	.0033	.0056	.0072	.0083	.0090	.0043	.0051	.0057	.0061
13	0.0	.0114	.0194	.0250	.0289	.0316	.0160	.0196	.0224	.0248
20	.0353	.0295	.0249	.0213	.0184	.0161	.0269	.0249	.0233	.0220
21	0.0	.0033	.0056	.0072	.0083	.0070	.0043	.0051	.0057	.0062
22	.0018	.0015	.0012	.0011	.0009	.0008	.0013	.0012	.0012	.0011
23	.0057	.0048	.0041	.0035	.0031	.0027	.0044	.0041	.0039	.0037
30	.0609	.0528	.0461	.0406	.0360	.0322	.0499	.0476	.0458	.0443
31	0.0	.0108	.0185	.0239	.0277	.0303	.0155	.0191	.0220	.0243
32	.0054	.0045	.0039	.0033	.0029	.0026	.0042	.0039	.0037	.0035
33	.0149	.0129	.0112	.0099	.0087	.0078	.0121	.0114	.0111	.0107

Table 5.5

Equilibrium Vehicle State Probabilities
At Instant of Medical Emergency For
Two Vehicle/Dual Function System

Distribution Policy II

	Dispatch Strategy 1						Dispatch Strategy 2			
(s_n^2, s_n^1) \ (ϕ_1, ϕ_2)	(0.0)	(.1,.1)	(.2,.2)	(.3,.3)	(.4,.4)	(.5,.5)	(.1,.2)	(.1,.3)	(.1,.4)	(.1,5)
00	.7694	.6568	.5661	.4931	.4339	.3856	.5872	.5332	.4901	.4549
01	0.0	.0536	.0911	.1171	.1350	.1472	.0866	.1122	.1326	.1493
02	.0315	.0263	.0222	.0190	.0164	.0144	.0240	.0223	.0208	.0197
03	.0735	.0632	.0548	.0479	.0423	.0376	.0593	.0562	.0537	.0517
10	0.0	.0542	.0919	.1180	.1359	.1481	.0871	.1127	.1331	.1497
11	0.0	.0096	.0320	.0609	.0923	.1242	.0155	.0200	.0237	.0267
12	0.0	.0030	.0051	.0065	.0074	.0081	.0039	.0046	.0051	.0055
13	0.0	.0117	.0199	.0257	.0297	.0325	.0164	.0201	.0230	.0254
20	.0293	.0245	.0207	.0177	.0153	.0133	.0223	.0207	.0194	.0183
21	0.0	.0028	.0047	.0060	.0069	.0075	.0036	.0042	.0047	.0051
22	.0013	.0011	.0009	.0008	.0007	.0006	.0010	.0009	.0009	.0008
23	.0049	.0041	.0035	.0030	.0026	.0023	.0038	.0035	.0033	.0032
30	.0683	.0590	.0513	.0450	.0399	.0356	.0555	.0528	.0507	.0489
31	0.0	.0115	.0196	.0253	.0293	.0321	.0163	.0201	.0230	.0255
32	.0052	.0044	.0037	.0032	.0028	.0025	.0040	.0038	.0035	.0034
33	.0166	.0143	.0124	.0109	.0096	.0086	.0134	.0128	.0122	.0118

Table 5.6

Equilibrium Results for the Two Vehicle/Dual Function System

Distribution Policy I

Dispatch Strategy I

Dispatch Strategy 2

Probability \ ϕ_1, ϕ_2	(0,0)	(.1,.1)	(.2,.2)	(.3,.3)	(.4,.4)	(.5,.5)	(.1,.2)	(.1,.3)	(.1,.4)	(.1,.5)
P_0 (2)	.7703	.6575	.5667	.4935	.4343	.3859	.5878	.5337	.4905	.4553
P_0 (1)	.2020	.2804	.3318	.3645	.3845	.3961	.3345	.3766	.4101	.4374
P_0 (1+)	.9723	.9399	.8985	.8580	.8188	.7820	.9223	.9103	.9006	.8927
P_1 (2)	0	.0096	.0320	.0609	.0924	.1243	.0155	.0201	.0237	.0267
P_1 (1)	0	.1367	.2323	.2985	.3442	.3754	.2140	.2740	.3218	.3607
P_1 (1+)	0	.1463	.2643	.3594	.4366	.4997	.2295	.2941	.3255	.3874
P_2 (2)	.0018	.0015	.0012	.0011	.0009	.0008	.0013	.0012	.0012	.0011
P_2 (1)	.0813	.0746	.0688	.0636	.0592	.0553	.0707	.0678	.0654	.0633
P_2 (1+)	.0831	.0761	.0700	.0647	.0601	.0561	.0720	.0690	.0666	.0644
P_3 (2)	.0149	.0129	.0112	.0099	.0087	.0078	.0121	.0115	.0111	.0107
P_3 (1)	.1429	.1453	.1449	.1426	.1395	.1358	.1472	.1486	.1497	.1506
P_3 (1+)	.1571	.1582	.1561	.1525	.1482	.1436	.1593	.1601	.1608	.1613

Table 5.7

Equilibrium Results for the Two Vehicle/Dual Function System

Distribution Policy II

Probability \ (ϕ_1,ϕ_2)	(.0,.0)	(.1,.1)	(.2,.2)	(.3,.3)	(.4,.4)	(.5,.5)	(.1,.2)	(.1,.3)	(.1,.4)	(.1,.5)
P_0 (2)	.7694	.6568	.5661	.4931	.4339	.3856	.5872	.5332	.4901	.4549
P_0 (1)	.2026	.2808	.3320	.3647	.3848	.3962	.3348	.3769	.4103	.4376
P_0 (1+)	.9720	.9376	.8981	.8578	.8187	.7818	.9220	.9101	.9004	.8925
P_1 (2)	0	.0096	.0320	.0609	.0923	.1242	.0155	.0200	.0237	.0267
P_1 (1)	0	.1368	.2323	.2986	.3442	.3755	.2139	.2739	.3215	.3605
P_1 (1+)	0	.1464	.2643	.3595	.4365	.4997	.2294	.2939	.3452	.3872
P_2 (2)	.0013	.0011	.0009	.0008	.0007	.0006	.0010	.0009	.0009	.0008
P_2 (1)	.0709	.0651	.0599	.0554	.0514	.0481	.0616	.0591	.0568	.0552
P_2 (1+)	.0722	.0662	.0608	.0562	.0521	.0487	.0626	.0600	.0577	.0560
P_3 (2)	.0166	.0143	.0124	.0109	.0096	.0086	.0134	.0128	.0122	.0118
P_3 (1)	.1519	.1539	.1528	.1501	.1466	.1426	.1553	.1565	.1572	.1581
P_3 (1+)	.1685	.1682	.1652	.1620	.1562	.1512	.1687	.1693	.1694	.1699

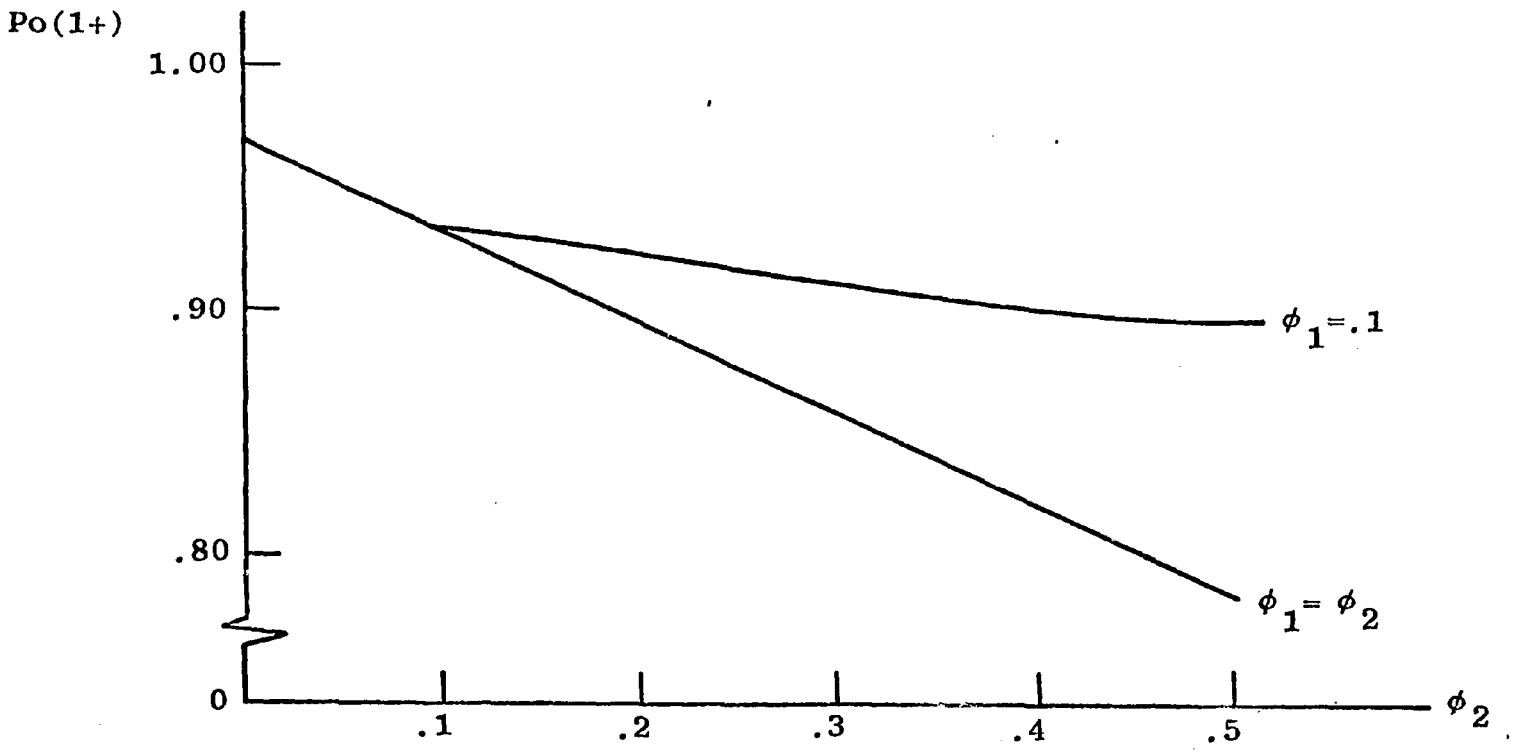


Figure 5.5

Equilibrium Probability That One or More Vehicles are Idle in Two Vehicle/Dual Function System

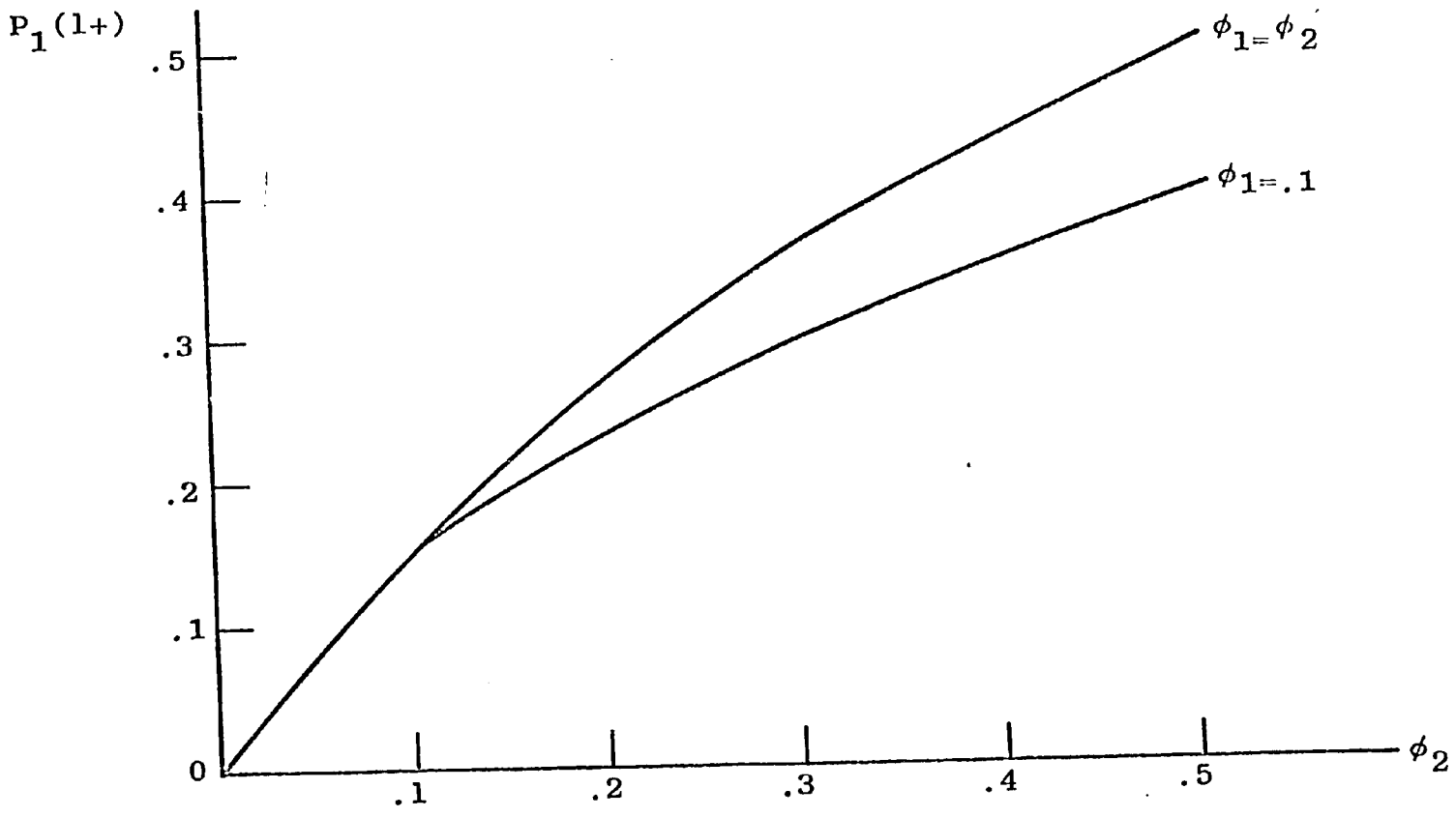


Figure 5.6

Equilibrium Probability That One or More Vehicles are Serving Police Calls in Two Vehicle/Dual Function System

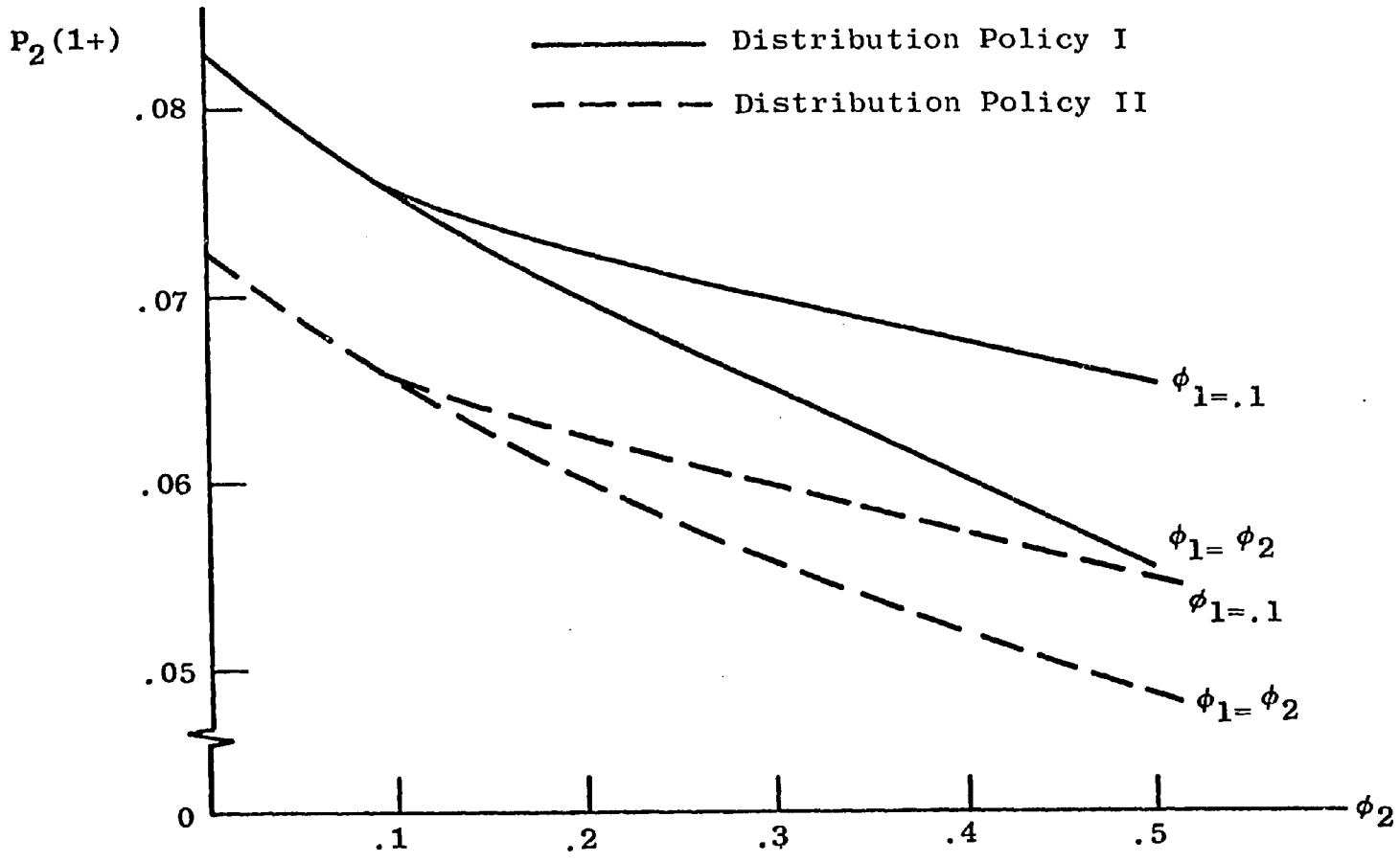


Figure 5.7

Equilibrium Probability that One or More Vehicles Are Serving "Short" Medical Emergencies in Two Vehicle/Dual Function System

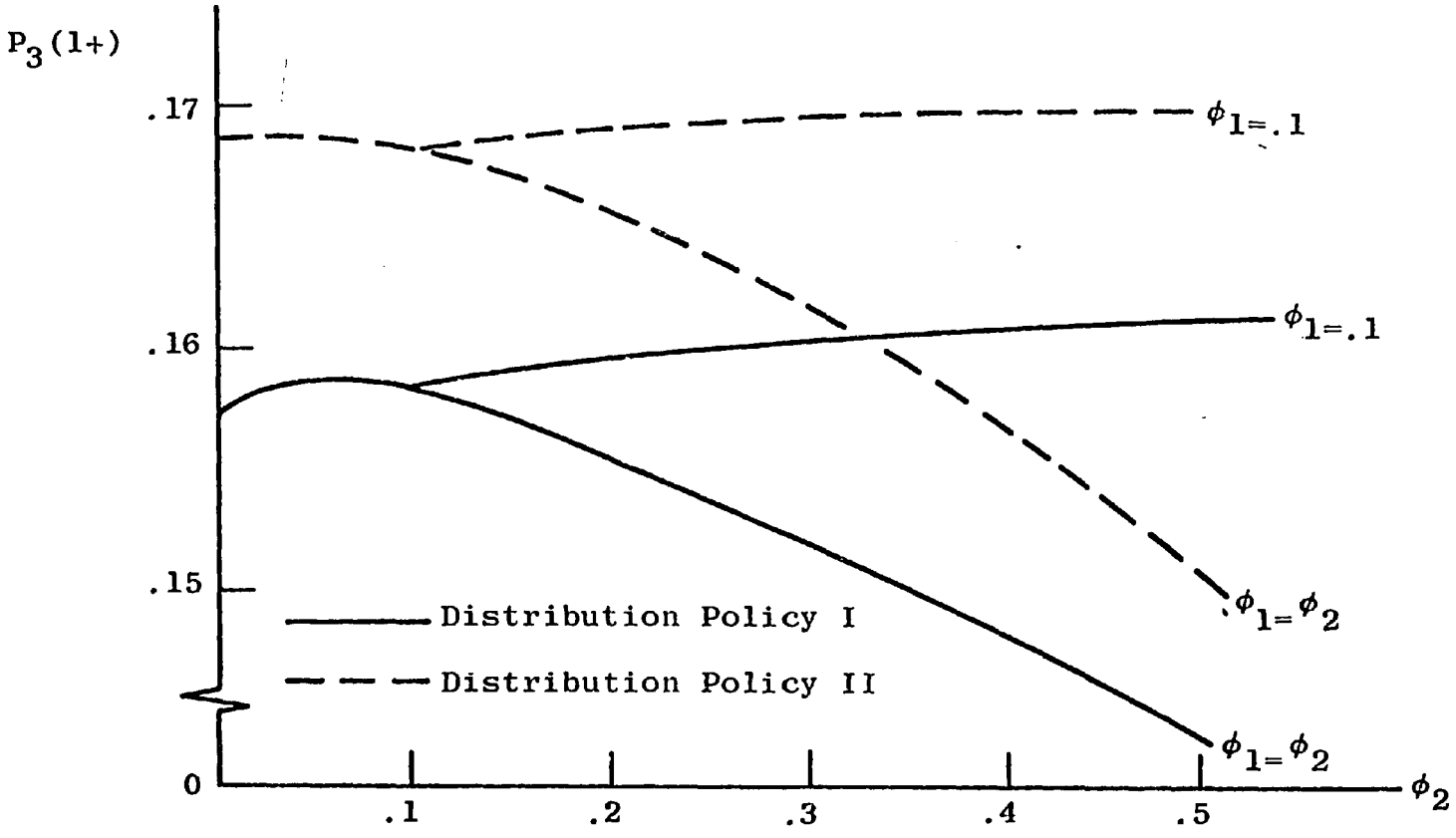


Figure 5.8

Equilibrium Probability that One or More Vehicles Are Serving "Long" Medical Emergencies in Two Vehicle/Dual Function System

k vehicles are in state j, and $p_j(k+)$ is the equilibrium probability that k or more vehicles are in state j. For instance, $P_0(1+)$ is the equilibrium probability that one or more vehicles are idle.

We observe that the equilibrium probabilities that one or more vehicles are idle and that one or more vehicles are serving medical emergencies are essentially unaffected by the distribution policy. On the other hand, the equilibrium probability that one or more vehicles are serving "short" medical emergencies and the equilibrium probability that one or more vehicles are serving "long" medical emergencies are much more sensitive to the chosen distribution policy.

The effects of the two alternative dispatch strategies on the equilibrium results are interesting. The equilibrium idle probability declines almost linearly for strategy I ($\phi_1 = \phi_2$). However, when ϕ_1 is maintained at .1, this decline is much less severe. Similarly, the equilibrium probabilities that one or more vehicles are serving "short" and "long" emergencies are both higher for the second dispatch strategy.

As before, we shall postpone a more detailed discussion of these results until the system comparisons of Chapter VI.

5.3 RESULTS FOR A DUAL FUNCTION SYSTEM WITH MORE THAN TWO VEHICLES

We observed in Chapter III that the number of system states in our model of the dual function ambulance system is $2(4)^r$. When $r = 3$, the size of the state space reaches 128, and when $r = 4$, it becomes 512. Although the model developed in Chapter III is completely general, substantial computational problems arise when the number of potential states is as large as those values described above. For instance, a sample run of the computer program documented in Appendix C with $r = 3$ required eight minutes of execution time and more than 20 pages of virtual memory on an IBM 360/67 computer. The cost for this run was approximately $80.00 under current University billing policies.

Consequently it was decided to approximate the performance of a dual-function system with more than two vehicles to obtain some insights into the effects of such an allocation policy. The approximation was implemented by partitioning the single precinct service region into non-overlapping sub-regions. The techniques of the previous section were then applied to these sub-regions, and the results were aggregated to approximate results for the entire precinct.

These results should be quite accurate, since the likelihood that any vehicle is dispatched on cross-region runs is decreased when a larger number of vehicles is assigned to a service region. As a result, the vehicles operate in smaller sub-regions a larger portion of the time. The approximating procedure merely forces the probability of a vehicle crossing a sub-region boundary to be zero.

In this section we present numerical results for this approximation when r = 4. A value of four was chosen since this is approximately the average number of dual-function vehicles assigned to a police precinct, and since results for r = 3 can be approximated by interpolating between these results and those of Section 5.2.

The fifteenth precinct was divided into the two sub-regions as shown in Figure 5.9, and the system performance for each of these sub-regions was modeled with r = 2. The properties of the emergency occurrence processes in the two sub-regions were developed from those of the regional processes in the following way. It was assumed that the distributions of times between events in the sub-region processes were negative-exponential, as were those in the aggregate process. The mean times between occurrences in the sub-process were then developed by weighting

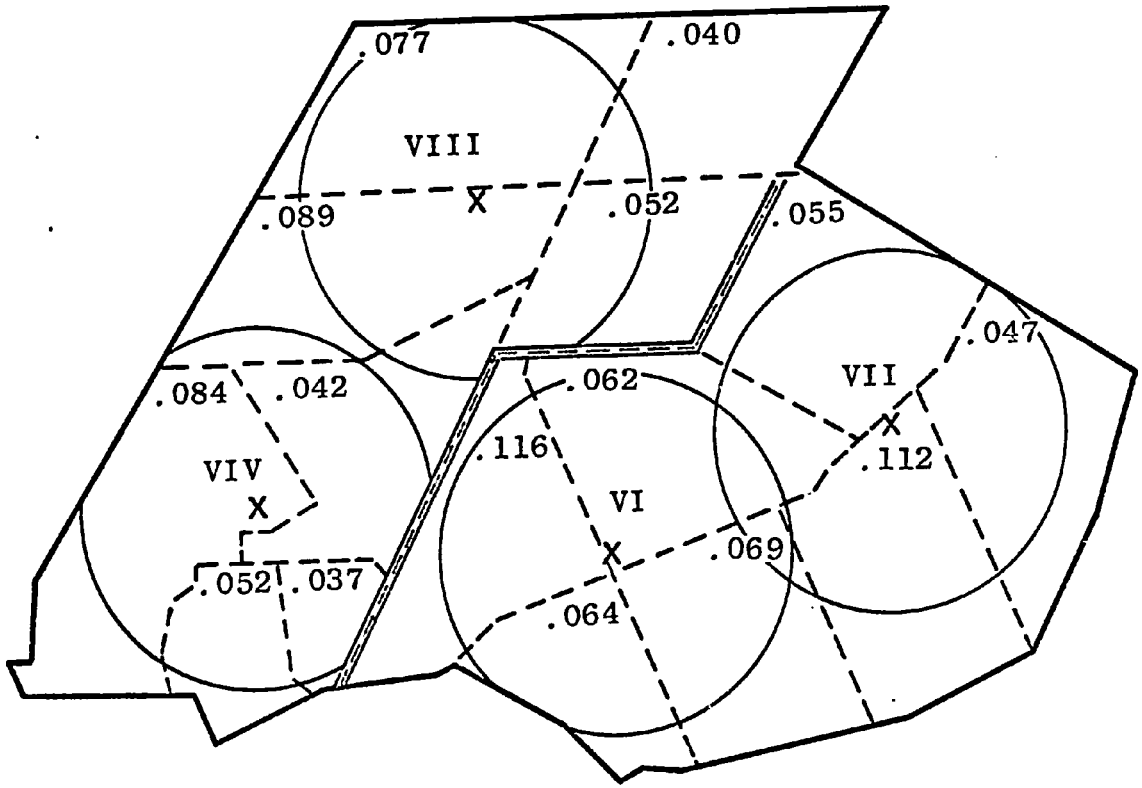


Figure 5.9

Four Vehicle Distribution Policy

the mean times in the aggregate process by the reciprocal of the probabilities of an emergency occurrence in each of the sub-regions.

The resulting mean times between emergency occurrences in the sub-regions are tabulated below:

	Sub-Region 1	Sub-Region 2
Mean Time between Ambulance Call Followed by Police Call (min.)	27.2	30.1
Mean Time between Ambulance Call Followed by Police Call (min.)	30.9	34.1
Mean Time between Police Call Followed by Ambulance Call (min.)	44.0	48.6
Mean Time between Police Call Followed by Police Call (min.)	27.8	30.8
Probability of Ambulance Call	.525	.475

Equilibrium vehicle state probabilities were calculated for the two sub-regions, each with two recovery vehicles distributed as in Figure 5.9. Results were calculated for a single function ambulance system ($\phi_1 = \phi_2 = 0$) and for dual function ambulance systems with $\phi_1 = \phi_2 = .1$ and $\phi_1 = \phi_2 = .5$. Since the equilibrium probabilities in the two sub-regions are independent, the equilibrium probabilities for the precinct can be easily

derived. For instance, the equilibrium probability that all four recovery vehicles are idle at the instant of a medical emergency is equal to the probability that two vehicles are simultaneously available in each sub-region at the instant of a medical emergency. Therefore, the probability is simply the product of the probabilities that two recovery vehicles are idle in each of the sub-regions.

The equilibrium results for the four vehicle approximation are presented in Table 5.7.

We observe that with four recovery vehicles in the system the equilibrium probability that one or more vehicles are idle is approximately one regardless of the dispatch strategy. Similarly, the probabilities that one or more vehicles are on "long" and "short" medical emergencies are insensitive to the alternative dispatch strategies. We shall discuss these results in more detail in the system comparisons of the next chapter.

From this discussion it is clear that the assignments of four vehicles to a police precinct results in a system where at least one vehicle is available at the instant of an ambulance call essentially "all" of the time. This is true even when the vehicle responds to both police and medical emergencies. Hence we shall not consider

Table 5.8

Equilibrium Results for a Four Vehicle Dual Function System

		$\phi_1=\phi_2=0$	$\phi_1=\phi_2=.1$	$\phi_1=\phi_2=.5$
Probability k vehicles are idle at instant of medical emergency				
k	4	.765	.655	.366
	3	.202	.276	.353
	2	.027	.063	.222
	1	.006	.006	.052
	0	0.0	0.0	.007
Probability k vehicles on police call at instant of medical emergency				
k	4	0.0	0.0	.002
	3	0.0	0.0	.025
	2	0.0	.012	.134
	1	0.0	.133	.356
	0	1.0	.855	.483
Probability k vehicles on "short" ambulance call at instant of medical emergency				
k	4	0.0	0.0	0.0
	3	.001	0.0	0.0
	2	.009	.006	.015
	1	.145	.139	.111
	0	.845	.855	.874
Probability k vehicles on "long" ambulance call at instant of medical emergency				
k	4	0.0	0.0	0.0
	3	0.0	0.0	0.0
	2	.005	.005	.004
	1	.089	.074	.098
	0	.906	.921	.898

distribution and allocation policies for systems with more than four vehicles per precinct. The results for the various ambulances systems with one, two, three, and four recovery vehicles are compared in the next chapter.

CHAPTER VI

COMPARISONS OF ALTERNATIVE RECOVERY SYSTEMS

In this chapter we shall utilize the numerical results of Chapters IV and V to compare the effects of alternate ambulance allocation and distribution policies on the recovery system performance. To review briefly, the specific factors we are interested in examining are:

(1) The effects of changing the number of recovery vehicles assigned to the system (the vehicle allocation policy).

(2) The effects of changing the placement of the recovery vehicles assigned to the system (the vehicle distribution policy).

(3) The effects of modifying the dispatch policy which assigns both police and ambulance calls to the recovery vehicles (the vehicle dispatch policy).

Measures of system performance were discussed in the introduction and in Section 3.3. In these sections we noted that improvements in recovery vehicle availability and reductions in recovery vehicle service times are measures directly influenced by improved ambulance

allocation and distribution. The numerical results of the previous two chapters provide equilibrium vehicle state probabilities which are closely related to these measures, and we shall now use these results to develop the appropriate comparisons.

6.1 RECOVERY VEHICLE AVAILABILITY

We have observed that medical emergencies occurring when no recovery vehicles are available either must "wait" for a vehicle to finish its current service, or they must be served by a special vehicle from outside the system. As we have seen, a system which forces a large number of emergencies to wait for recovery vehicles to become available is undesirable, since the resulting waiting times introduce potential, uncontrollable degradations in the victim's medical condition. On the other hand, a system which requires frequent external assistance is likely to incur substantial real and opportunity costs, as well as to cause inefficiencies in the system from which the extra vehicle is "borrowed."

Therefore, in this section we shall examine the probability that all recovery vehicles are busy, or equivalently, the probability that one or more recovery vehicles are available at the time of an ambulance call.

The effects of alternative vehicle allocation, distribution, and dispatch policies are sequentially considered.

6.1.1. Allocation Effects

Qualitatively, one expects that an increase in the number of recovery vehicles will decrease the probability that all vehicles are simultaneously busy. In fact, if enough vehicles are supplied to the recovery system we would expect that this probability should approach zero. Of course, the fixed and operating costs are both greater for a recovery system equipped with a larger number of vehicles, and consequently it is necessary to carefully explore the trade-off between vehicle availability and the allocation policy.

Quantitative results for the single function ambulance system are presented in Table 6.1 below. (Results for the dual-function system are presented in Section

Table 6.1

Probability All Recovery Vehicles Busy as Function of Allocation Policy

Number of Recovery Vehicles	Probability All Recovery Vehicles Busy
1	.2109
2	.0235
3	.0019
4	.0000

6.1.3.) In all cases these results are for that vehicle distribution which maximizes the equilibrium idle probability. The model parameters are set at the maximum likelihood estimates.

Inspection of Table 6.1 reveals that increasing the number of recovery vehicles from one to two decreases the probability that all vehicles are busy by .1874. An increase to three vehicles further decreases this probability by only .0216. When four recovery vehicles are assigned to the police precinct, the probability is essentially reduced to zero.

When two recovery vehicles are assigned to a single function ambulance system, at least one will be available approximately ninety-eight percent of the time. The substantial degradation in vehicle availability when only one vehicle is assigned and the moderate improvement in availability when three are assigned provide further evidence in favor of a two vehicle system.

6.1.2. Distribution Effects

As we have indicated, the major effect of modifying the distribution of recovery vehicles should be to decrease vehicle service times by decreasing the transit time to the scene. This effect will be discussed in Section 6.2. In addition, by decreasing the vehicle service

time, the improved distribution of vehicles may result in some increases in vehicle availability, and we shall discuss these effects at this time. Results are presented here for the single function recovery system only; results for the dual function system are discussed in Section 6.1.3.

Two distribution policies were examined for the single vehicle recovery system. Policy I placed the vehicle in the sub-region of minimum demand, and Policy II placed the vehicle in the sub-region of maximum demand. The equilibrium probability that the vehicle is busy is tabulated as a function of these distribution policies in Table 6.2:

Table 6.2

Vehicle Availability as Function of Distribution Policy

r = 1

Policy	Probability All Recovery Vehicles Busy
I	.2117
II	.2109

Three distribution policies were examined for the two and three vehicle recovery systems. Policy I placed the vehicles in sub-regions to maximize coverage in the areas of greatest vehicle demand. Policy II placed the

vehicles in sub-regions to obtain uniform geographical coverage. Policy III placed all recovery vehicles at a single station to stimulate system operation when a single, fixed ambulance station is utilized. (The chosen station coincides with the one-mile service region of the single vehicle distribution policy II discussed earlier.) The equilibrium probabilities that all vehicles are busy are tabulated as a function of the three distribution policies in Table 6.3.

Table 6.3

Vehicle Availability as Function of Distribution Policy

r = 2; r = 3

	Probability All Recovery Vehicles Busy	
Policy	r = 2	r = 3
I	.0235	.0019
II	.0242	.0035
III	.0255	.0027

Inspection of Tables 6.2 and 6.3 indicates that the effect of vehicle distribution on ambulance availability is essentially negligable for all allocation policies. Placement of the vehicle(s) in the regions of maximum demand (Policy I) results in the lowest probability that all recovery vehicles are busy. However, the improvement

over other distribution policies is very small. Small differences also exist between the other distribution policies. Hence we can conclude the placement of recovery vehicles does not influence ambulance availability in any substantial way for the system under study.

6.1.3. Dispatch Effects

When an ambulance system accepts both police and ambulance calls for service it is reasonable to expect that any given system configuration will service ambulance calls less efficiently than a comparable single function system. We shall now examine the degradation in vehicle availability as a function of the percentage of police calls accepted for service by the recovery system. Comparisons are made of this degradation in recovery systems with one, two, and four vehicles. Results for the three systems are presented in Figure 6.1.

In Figure 6.1 we observe a monotonic increase in the probability all recovery vehicles are busy at the instant of a medical emergency of the proportion of police calls serviced. This increase is present for all three allocation policies examined, but the rate of increase is lower for the systems with a larger number of recovery vehicles. In fact, almost no degradation in this measure

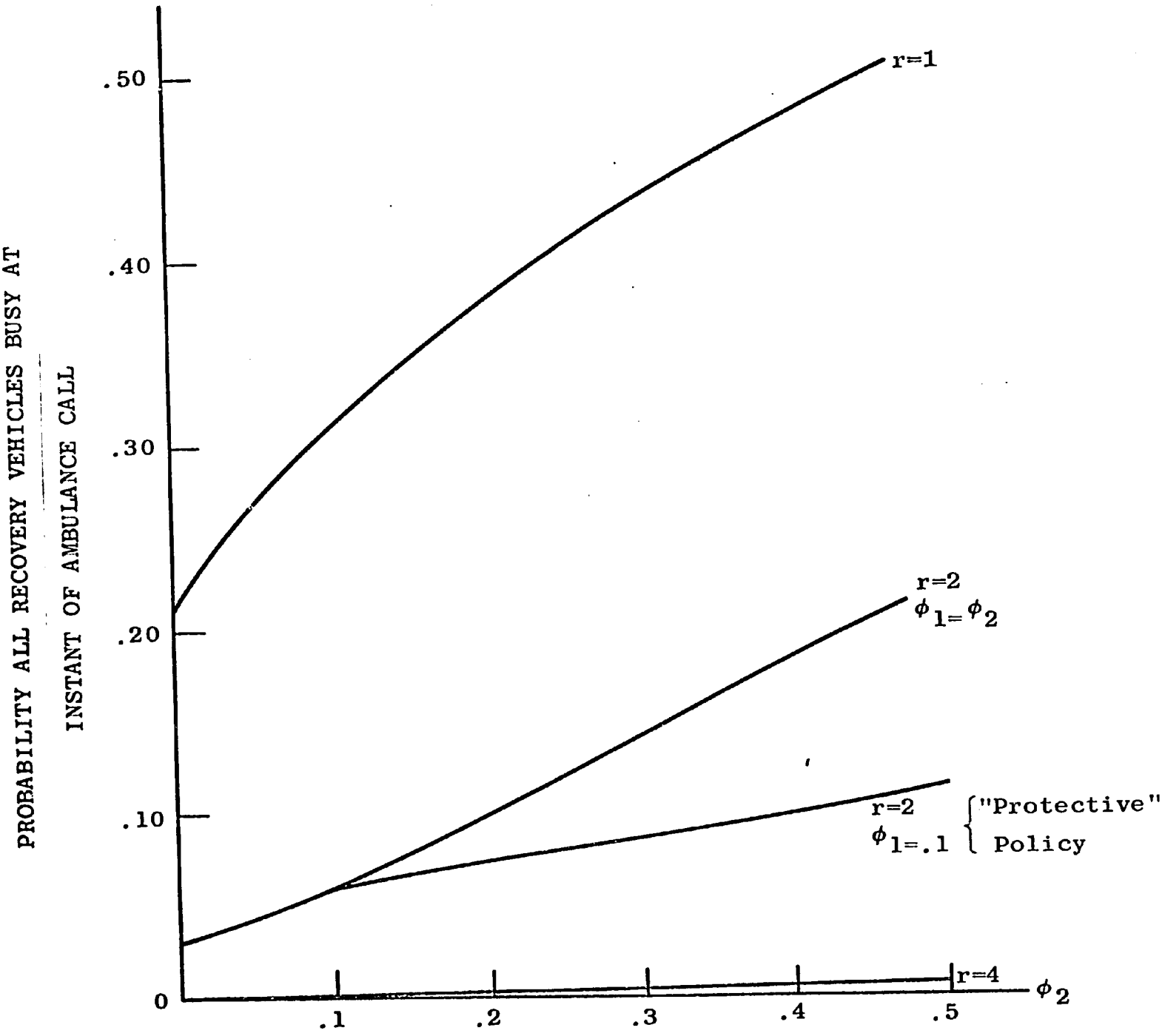


Figure 6.1

VEHICLE AVAILABILITY AS FUNCTION OF ALLOCATION AND DISPATCH POLICY

of vehicle availability is present when four vehicles are assigned to the recovery system.

We have seen that system performance can be substantially improved if one modifies the dispatch policy to "protect" the last available ambulance. Results for such a policy with r = 2 are also presented in Figure 6.1. We observe that this protective policy results in at least a .89 probability that at least one vehicle is available, even when the system is accepting fifty percent of all police calls for service when both vehicles are idle. The equivalent probability is only .78 when the protective policy is not utilized. Hence, we can conclude that a protective dispatch policy does assist in maintaining recovery vehicle availability when a dual function system is utilized. However, some degradation in vehicle availability still is present whenever any police calls are accepted for service.

We also observe that a protective dispatch policy may have a detrimental influence on the availability of vehicles to handle police calls. Since in general there are many more police squad cars than dual function police ambulances, this is not believed to be a serious problem in the system we are examining.

6.2 RECOVERY VEHICLE SERVICE TIMES

As we have seen, the recovery vehicle response time is the sum of two random variables: the waiting time for an ambulance to become available and the transit time to the scene of an emergency. We have also observed that the transit time to the scene is one component of the recovery vehicle service time. Empirical analysis of ambulance service times in the City of Detroit has revealed that these times are exponential, identically distributed, random variables when one conditions on transit distances of less than one mile and on transit distances greater than one mile. The mean service time for calls requiring less than one mile to the scene has been found to be approximately two minutes less than that for calls requiring transit distances greater than one mile to the scene.

Thus, transit times and subsequent system response times are lower in a system where a larger proportion of ambulance calls require less than one mile transit distance to the scene. Consequently we shall base our system comparisons on the equilibrium probabilities that medical emergencies are being serviced as "short" (less than one mile transit distance) runs. We first consider the effects of alternative allocation and distribution

policies. We then consider the effects of alternative dispatch policies.

6.2.1. Allocation and Distribution Effects

The effects of allocation and distribution policies for the single function ambulance system are considered simultaneously in this section. The basic numerical results are presented in Table 6.4.

Table 6.4

Effects of Vehicle Allocation and Distribution on Service Times

Allocation Policy	Distribution Policy	Probability one or more vehicles on run requiring less than one mile transit distance
r = 1	I	.0274
	II	.0387
r = 2	I	.0831
	II	.0723
	III	.0463
r = 3	I	.1049
	II	.1197
	III	.0625
r = 4	I	.15

Inspection of Table 6.4 indicates that the probability one or more vehicles are serving "short" emergencies increases substantially with the number of vehicles, regardless of the distribution policy. Given a distribution

policy with one recovery vehicle, a difference of .0113 in this probability is introduced by shifting from a distribution policy which places the ambulance in the region of minimum demand (I) to a policy which places the ambulance in the region of maximum demand (II). In the two and three vehicle systems, similar differences exist between those distribution policies designed to obtain maximum coverage in the region of maximum demand (I) and those policies designed to obtain uniform coverage of the entire service region (II). For these allocation policies, each of the two distribution plans results in a considerable improvement over a single-station policy (III).

It is possible that the effect of alternative allocation policies on the probability one or more vehicles are serving "short" emergencies is due to the substantially different vehicle availabilities associated with these policies. Consequently, the conditional probabilities that a vehicle is serving a "short" emergency, given that the vehicle is "busy," were calculated for each vehicle for each allocation and distribution policy. Results are presented in Table 6.5.

By conditioning on the event that the vehicle is busy, the effects of differing allocation policies are essentially removed. Therefore, Table 6.5 can be utilized

Table 6.5

Effect of Allocation and Distribution Policy on Conditional Probability that each Vehicle is on Emergency Requiring Less Than One Mile Transit Distance

Allocation Policy	Distribution Policy	Vehicle Number	Probability on Short Emergency Given that Vehicle is Busy
r = 1	I	1	.130
	II	1	.183
r = 2	I	1	.317
		2	.346
	II	1	.272
		2	.284
	III	1	.182
		2	.182
r = 3	I	1	.540
		2	.380
		3	.365
	II	1	.510
		2	.530
		3	.382
	III	1	.229
		2	.229
		3	.229

to study the effects of alternative distribution policies alone.

For r = 1, a .05 increase in the conditional probability of serving a short emergency occurs when the vehicle

is assigned to the region of maximum demand. For r = 2, this probability is .182 when both vehicles are assigned to the same ambulance station. An average increase of .096 occurs when a uniform coverage distribution policy is utilized. Implementation of a distribution policy to increase coverage in the area of maximum demand results in an average increase of .064 over the uniform coverage policy.

For r = 3, the conditional probability is .229 when all three vehicles are assigned to the same station. An average increase of .145 can be obtained by implementing a uniform coverage policy. In this case, however, the maximum coverage policy results in a .046 decrease in the average probability.

From these results we can conclude that the distribution of recovery vehicles does affect the probability that one or more vehicles are on a short run and the probability that any recovery vehicle is on a short run, given that it is busy. Placing all recovery vehicles at a single station results in values for these probabilities which are substantially lower than those for alternative distributions. Placing the vehicle(s) in the regions of maximum demand results in the highest probabilities when r = 1 and r = 2. When r = 3, the uniform coverage distri-

bution results in higher values for the relevant probabilities. However, no substantial difference exists between these latter two distributions for either multiple vehicle allocation policy.

6.2.2. Dispatch Effects

We shall now examine the probabilities that one or more recovery vehicles are on short runs in a dual function system. These are plotted in Figure 6.2 as a function of the proportion of police calls accepted for service when all vehicles are idle. We observe a gradual reduction in this probability as more police calls are accepted by the system. This result holds for all allocation and distribution policies. When no police calls are accepted ($\phi_2 = 0$), results are identical to those of Table 6.4. As we have seen, distribution policy III (placing all vehicles at a single station) is inappropriate in a dual-function system, and results are not presented for this distribution. In addition, the degradation in performance has not been calculated for the three vehicle allocation.

If one adopts the protective dispatch policy discussed in Section 5.2, the reduction in the probability that one or more vehicles are serving "short" emergencies is not as rapid. Results for such a protective policy

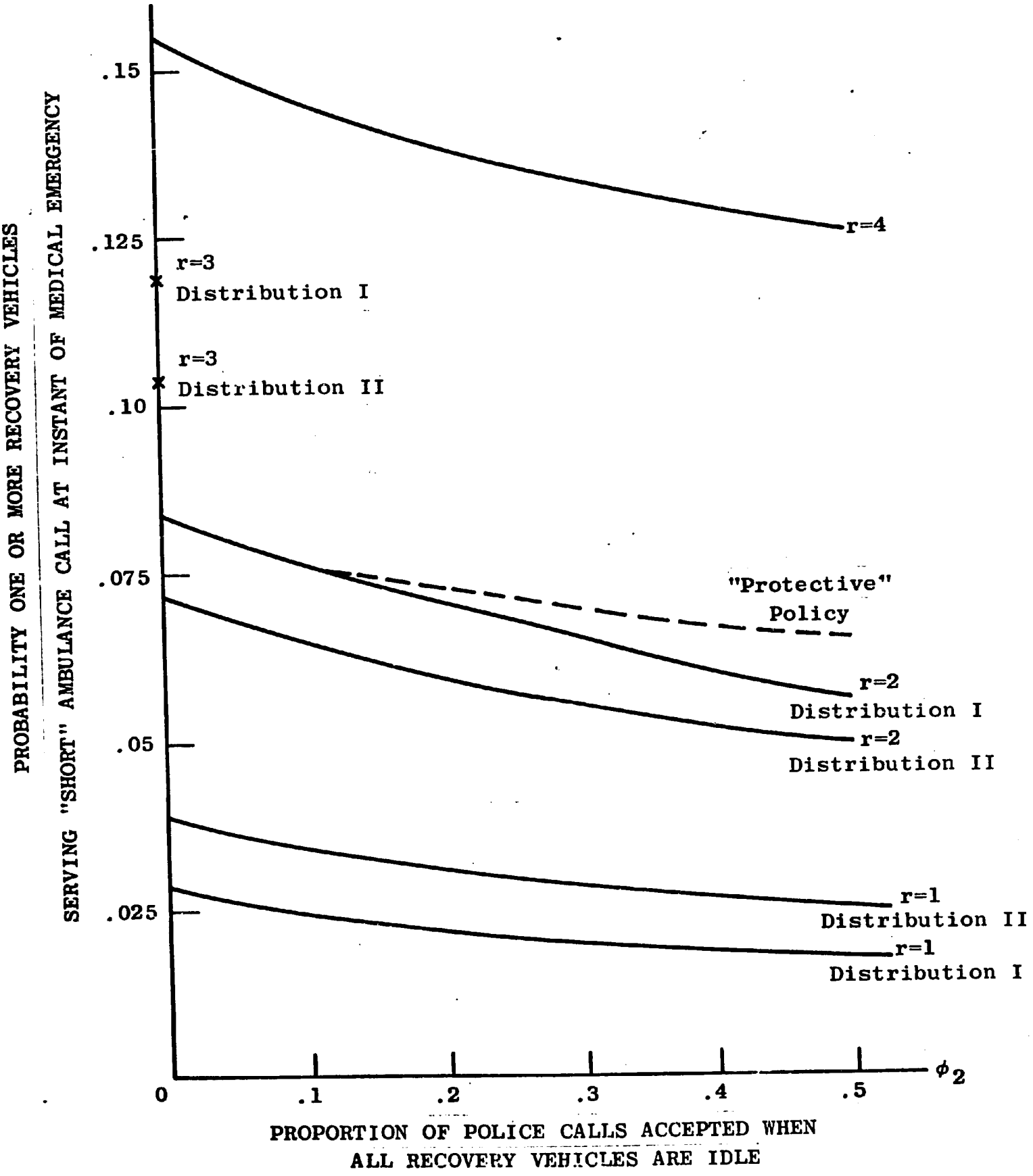


Figure 6.2

VEHICLE UTILIZATION AS FUNCTION OF ALLOCATION, DISTRIBUTION, AND DISPATCH POLICY

are shown by the dotted line in Figure 6.2. Nevertheless, some degradation in this probability is present whenever any police calls are accepted for service.

To study this problem in more detail, the conditional probability that each vehicle is serving a "short" emergency, given that it is busy, was calculated for the various allocation and distribution policies with r = 1 and r = 2. Results are presented in Figure 6.3. (Observe that semi-log plotting has been utilized in this figure.)

When the conditional probabilities are examined in this way, one observes that the degradation occurs at the same rate for each vehicle. This was not true for the degradation in vehicle availability studied in Section 6.1.3.

However, when one conditions on the event that the vehicles are serving medical emergencies, it was found that the conditional probability of serving a "short" emergency is constant regardless of the dispatch policy. From this we can conclude that the dispatch policy does not affect the relative proportions of time spent serving medical emergencies requiring less than one mile transit distance and serving medical emergencies requiring more than one mile transit distance. Instead, the decline in

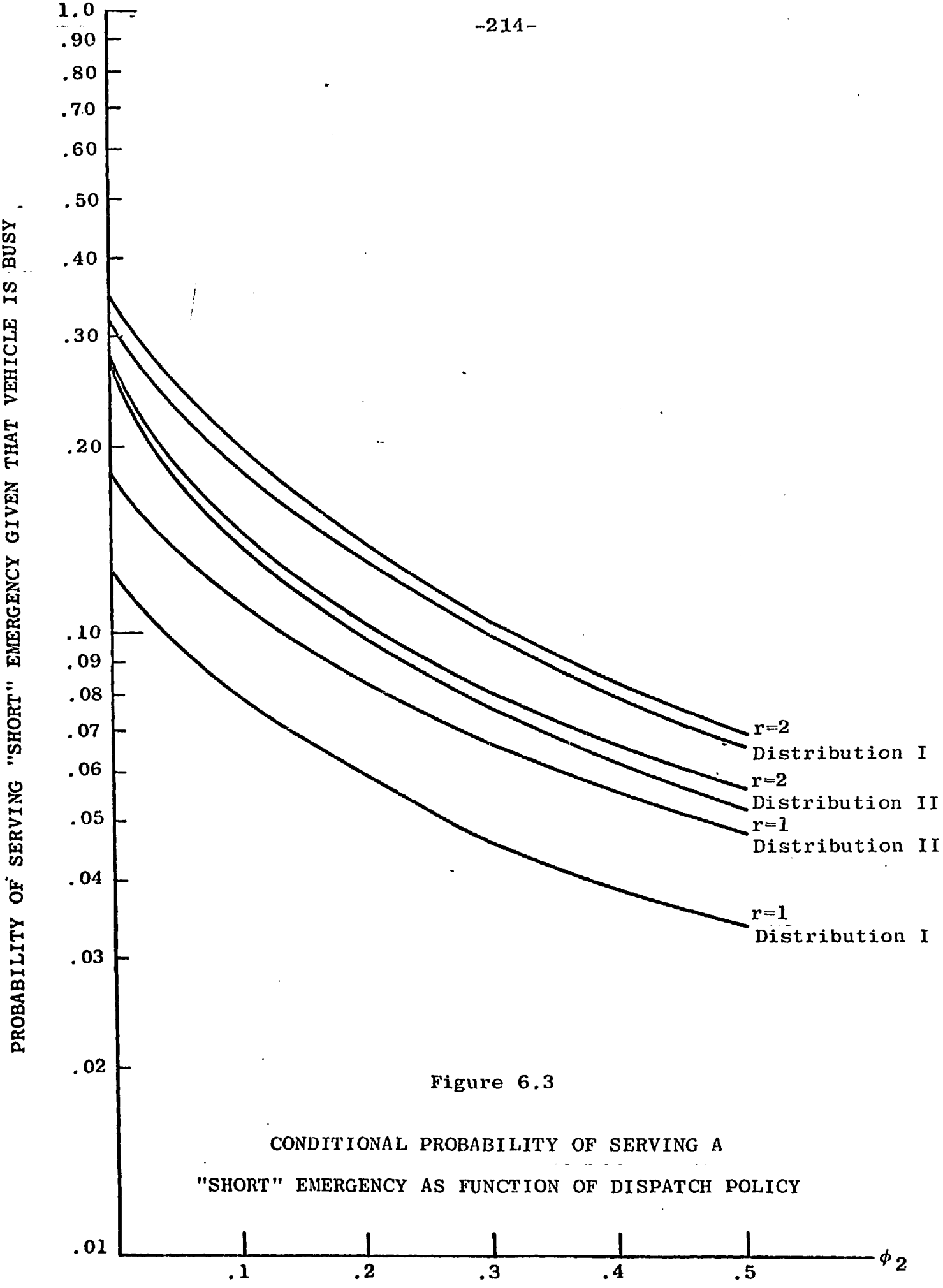


Figure 6.3

CONDITIONAL PROBABILITY OF SERVING A

"SHORT" EMERGENCY AS FUNCTION OF DISPATCH POLICY

those probabilities of Figures 6.2 and 6.3 is introduced by the larger proportion of the vehicles' time which is required to service police calls as more of these are accepted by the system.

CHAPTER VII

SUMMARY AND CONCLUSIONS

The purpose of this study has been to examine improvements in the quality of urban ambulance service which can be attributed to alternative ambulance allocation and distribution policies. While these improvements should contribute to reduced victim mortality or morbidity, we have observed that it is difficult to use such reductions as practical measures of system performance. This is true in part because indices of morbidity and mortality are difficult to define in a quantitative manner. Furthermore, it is extremely difficult to relate changes in these indices to the recovery system modifications we are examining. Consequently, intermediate measures of system performance, which are directly related to the system response time, have been chosen for this analysis. These measures are the probability that a vehicle is available to respond to an emergency and the probability that the recovery vehicle(s) must travel less than one mile to the scene of a medical emergency.

Statistical analyses of emergency occurrence and service processes have been conducted on data obtained

from the City of Detroit to aid in the development of a suitable model for the recovery process and to estimate the parameters of the resulting model. The results of these analyses are briefly summarized in Table 7.1.

Three major problems were resolved in using these data analyses in the model development. The first of these involved the incorporation of transit distance effects into a model of vehicle service times. The fact that no strong relationship was found between expected travel time and expected travel distance led to an abstraction where ambulance service times were considered as independent, identically distributed exponential random variables when one conditions on ranges of transit distances as in Table 7.1.

Second, it was necessary to develop a model for the joint occurrences of police and ambulance calls, since these both are inputs to the dual-function ambulance system. It was found that a semi-Markov process provided a reasonable model for these occurrences. Statistical analyses indicated that the state transitions in this semi-Markov process could be adequately modeled as Bernoulli random variables with independent negative-exponential times between transitions.

The third problem was in the development of a recovery system model which allowed emergencies to be

	Medical Emergency Occurrence Process	Police Emergency Occurrence Process	Medical Emergency Service Process	Police Emergency Service Process
Observed Trends in Time between Events	None	Daily Trend Removed by analyzing time periods 000-0800 and 0800-2400 separately	Not Applicable	Not Applicable
Observed Autocorrelations in Time between Events	None	None when Daily Trend Removed	Not Applicable	Not Applicable
Special Distributional Properties of Time between Events	None	None	Conditioned on Transit Distances to the scene of less than one mile and greater than one mile*	None
Marginal Distribution of Time between Events	Gamma with parameters k, μ	Negative Exponential with parameter λ_1 (0800-2400 shift) and with parameter λ_2 (0000-0800) shift	Negative exponential with parameter α for distance less than one mile and parameter β otherwise	Negative exponential with parameter γ
Parameter Estimates (Maximum Likelihood)	$\hat{K} = .77$ $\hat{\mu} = 101$ min	$\hat{\lambda}_1 = 1/42.7$ $\hat{\lambda}_2 = 1/17.2$	$\hat{\alpha} = 1/18.4$ $\hat{\beta} = 1/20.5$	$\hat{\gamma} = 1/38.3$

*The appropriate distance partitions were determined from the data analysis.

Table 7.1

Summary of Statistical Analyses

serviced in a differential fashion by the recovery system, depending both upon the state of the system and the type of emergency. This was resolved by extending certain results in the theory of systems of queues in parallel to include Markovian channel selection switches which are conditioned on the joint state of all queues in the system. In addition to solving the problem involved in this application, these theoretical results should also be useful in the study of queueing systems with non-random channel assignment rules. These include systems with shortest-service line channel selection rules, preferred server selection rules, and other related channel selection mechanisms.

The queueing theoretic model of the recovery process has been utilized to study the effects of the various allocation, distribution, and dispatch options available to the system designer. Analytical and numerical results have been developed to determine the effects of these alternatives on system performance. These results are briefly summarized in Table 7.2.

Although the purpose of this study has been to model the effects of a range of alternative policies, these results do provide much guidance in selecting a preferred system. We observe that a two vehicle, single function ambulance system operating in a police precinct

	Measure of System Performance	
	Probability one or more vehicles are busy	Probability a vehicle is serving a medical emergency requiring less than one mile transit distance, given that vehicle is busy
Effect of Allocation Policy	One Vehicle .7891 Two Vehicles .9765 Three Vehicles .9981 Four Vehicles 1.000	General increase as function of number of vehicles in system (depending upon distribution policy)
Effect of Distribution Policy	No appreciable Effect	Increase as vehicles are moved from single station to multiple stations
Effect of Dispatch Policy	Monotonic increase as function of proportion of police calls accepted for service; rate of increase less for larger number of recovery vehicles	No effect on relative proportion of time spent serving "short" or "long" medical emergencies

Table 7.2

Summary of Major Allocation, Distribution, and Dispatch Effects

results in one or more vehicles available for service approximately 98% of the time. A rather substantial decrease in availability occurs when a single vehicle system is implemented, and a very small increase occurs when a three vehicle system is utilized. Furthermore, in the

two vehicle system, over 30% of each vehicle's activities will involve servicing "short" runs when the vehicles are distributed at separate stations located in the sub-regions of maximum demand.

If the recovery system must service both police and ambulance calls, availability declines monotonically with the proportion of police calls accepted for service. In this case, it is possible that three or four vehicles will be necessary to maintain adequate availability. However, if one adopts the "protective" dispatch policy discussed in Chapter V, 90% vehicle availability can be maintained in the two-vehicle system.

Several problems which are not considered in this paper could serve as the basis for future research. First, the development of queueing theoretic models incorporating more sophisticated distance--service time relationships is essential for studying distribution effects in larger, more complex service regions. Second, one might examine system performance when the idle recovery vehicles are "re-deployed" to maximize coverage as the other recovery vehicles are assigned to emergencies. Finally, one might consider system performance when the number of vehicles assigned varies either stochastically or deterministically as a consequence of

vehicle maintenance, differential shift assignments, and other special activities. In this case transient effects become more important, and the model analysis must include a detailed consideration of the time-variant system performance.

APPENDIX A

STATISTICAL TESTS FOR THE POISSON PROCESS

In this appendix we briefly document the various statistical methods which were used in Chapter II to test for the "Poissonness" of the emergency occurrence and service processes. This documentation is included for the following reasons:

(1) to provide the readers of this paper with readily accessable reference material;

(2) to bring together reference information as an aid to future researchers who are interested in the statistical analysis of point processes.

It should be pointed out initially that most of the procedures discussed here provide tests of the Poisson hypothesis against very general alternatives. Unfortunately, the power of these tests is difficult to assess, since this is highly dependent upon the specific alternative under consideration. Consequently a variety of tests with differing characteristics are discussed in this appendix, and these are all utilized in Chapter II to accumulate evidence for (or against) the Poisson null hypothesis.

The majority of the tests discussed here are based upon the uniformity of waiting times in the conditional Poisson process. Let t_i be the time of the i^{th} event in a Poisson process. Then when the process is observed until the k^{th} event (occurring at time t_o), the k random variables $U_{(i)} = t_i/t_o$ are independent and uniformly distributed over the interval (0, 1).

To show this consider the joint density of the times $t_1, \cdots, t_k$ conditional on k events occurring in time t_o

$$f(t_1,\cdots,t_K|N_{t_o}=k) = \frac{\lambda e^{-\lambda t_1}\lambda e^{-\lambda(t_2-t_1)}\cdots\lambda e^{-\lambda(t_K-t_{K-1})}e^{-\lambda(t_o-t_K)}}{\frac{(\lambda t_o)^K e^{-\lambda t_o}}{k!}}$$

$$= \frac{k!}{t_o^K} \quad 0 \le t_1 \le t_2 \cdots \le t_K \le t_o$$

But this is exactly the joint distribution of the order statistics from a uniform distribution over the interval $(0, t_o)$. The distribution of the $U_{(i)}$ follow from a simple transformation. Hence tests for a Poisson process can be implemented by testing whether the independent random variables t_i/t_o follow a uniform distribution.

The Kolmogorov-Smirnoff statistic provides one such test by computing the absolute deviations between the

empirical distribution function $[F_n(x)]$ and the cumulative distribution function as defined under the null hypothesis $[F_o(x)]$. Three test statistics can be calculated:

Upper-Sided $\quad D_n^+ = \sqrt{n} \sup_{-\infty \le x \le \infty} \{F_n(x) - F_o(x)\}$

Lower-Sided $\quad D_n^- = \sqrt{n} \sup_{-\infty \le x \le \infty} \{F_o(x) - F_n(x)\}$

Two-Sided $\quad D_n = \sqrt{n} \sup_{-\infty \le x \le \infty} |F_o(x) - F_n(x)|$

The empirical distribution function is given by $F_n(x) = i/n$, while under the Poisson hypothesis, the cumulative distribution function of the $U_{(i)}$ is given by $F_o(x) = t_i/t_o$. The null hypothesis is rejected at level α if $D_n \ge d_{n,\alpha}$ where $d_{n,\alpha}$ is the α^{th} percentile of the distribution of D_n under the null hypothesis. Certain results on the minimum power of the test are known (Massey [20]). References on this test include Cox and Lewis (10), Darling (13), and Parzen (26).

The Anderson-Darling test is based upon the mean square deviation between $F_n(x)$ and $F_o(x)$. This test is derived from the Cramer-Von Mises statistic w_n^2, where

$$w_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F_o(x)]^2 \, dF_o(x)$$

The Anderson-Darling test computes the weighted mean square error using the weight function $g(x) = \frac{1}{x(1-x)}$. When the

data are assumed to follow a Poisson distribution, the resulting statistic is

$$\tilde{w}_n^2 = n\int_0^1 \frac{[i/n - u]^2}{u(1 - u)}\, du$$

$$= -n - 1/n \sum_{i=1}^{n} \{(2_i - 1)\ell n u_{(i)} + [2(n-i) + 1]\ell n(1-u_{(i)})\}$$

The null hypothesis is rejected when $\tilde{w}_n^2 > \tilde{w}_{n,\alpha}$ where $\tilde{w}_{n,\alpha}$ is the α^{th} percentile of the distribution of $\tilde{w}_n^2$. It is known that this test is sensitive to deviations from "Poissonness" in the tails of the distribution (Cox and Lewis [10]), but nothing is known of the power of the test. References on the Anderson-Darling test include Anderson and Darling (2), (3).

The Kolmogorov Smirnoff and Anderson-Darling tests can be extended by applying them to transformations of the ordered times to events. One modification, proposed by Durbin (14), bases the tests upon the partial sums of the ordered interarrival times. It is known that tests based upon these observations give greater power against processes where small interarrival times between events become smaller and large interarrival times become longer. References include Durbin (14), and Cox and Lewis (10).

The final test referenced in Chapter II is the Moran Test. As we pointed out in Section 2.2.1, this test is designed to provide high power against Gamma alternatives to the Poisson hypothesis. Specifically we are testing the hypothesis that $\kappa = 1$ in equation (2.1), since in this case the Gamma distribution simplifies to the negative exponential distribution. Recall from equation (2.5) that the maximum likelihood estimator of κ is given by the solution to:

$$-\ell n\hat{\kappa} + \frac{d\ell n\Gamma(\hat{\kappa})}{d\hat{\kappa}} = \ell n \frac{\prod_{i=1}^{n} x_i^{1/n}}{\bar{x}}$$

It can be shown that the left-hand side of this equation is monotonic increasing near $\hat{\kappa} = 1$, and hence $\hat{\kappa}$ can be written as a monotonic increasing function of the right-hand side (Moran [23]). Therefore a test can be based on the statistic

$$M_n = \ell n \frac{\prod_{i=1}^{n} x_i^{1/n}}{\bar{x}}$$

where we reject the null hypothesis for M_n large. In practice the statistic $\tilde{M}_n$ is utilized, and this given by Bartlett (5) as:

$$\tilde{M}_n = \frac{2n\,[\ell n\bar{x} - 1/n\textstyle\sum \ell n\ x_i]}{1 + \frac{(n+1)}{6n}}$$

Bartlett has shown (5) that under the Poisson hypothesis M_n has approximately a Chi-square distribution with n-1 degrees of freedom. Hence, approximate significance levels can be found by using tables of the Chi-square distribution. References include Cox and Lewis (10).

APPENDIX B

LAPLACE-STIELTJES TRANSFORMS OF DECOMPOSED EMERGENCY OCCURRENCE PROCESSES

In this appendix we shall develop the Laplace-Stieltjes Transforms of the distribution of the times between ambulance calls and of the times between police calls when it is assumed that a semi-Markov process generates these processes. We shall then briefly indicate why the models of Sections 2.2.1 and 2.2.2 provide reasonable representation of these two processes.

As we saw in Section 2.2.3, the semi-Markov process generating the sequence of emergency occurrences is completely identified when the marginal distribution of the times between events and the proportion of ambulance calls in the sequence are known. In Chapter III we indicated that these can be summarized in the matrix A(x), where

$$A(x) = \begin{bmatrix} \Pi F_{00}(x) & (1-\Pi) F_{01}(x) \\ \Pi F_{10}(x) & (1-\Pi) F_{11}(x) \end{bmatrix}$$

The Laplace-Stieltjes transform of this matrix, denoted A*(s), is given by

$$A^*(s) = \begin{bmatrix} \Pi F^*_{00}(s) & (1-\Pi) F^*_{01}(s) \\ \Pi F^*_{10}(s) & (1-\Pi) F^*_{11}(s) \end{bmatrix}$$

where $$F^*_{ij}(s) = \int_0^\infty e^{-sx}\, dF_{ij}(x)$$

To generate the stochastic process consisting of medical emergencies alone, we define the random variable Y_n and let

$Y_n = 0$ if event n is a medical emergency.

$Y_n = 0$ if event n is a police call.

Further define

$$f_0 = Pr\{Y_n = 0 | Z_n = 0\} = 1$$

$$f_1 = Pr\{Y_n = 0 | Z_n = 1\} = 0$$

Let D be the matrix

$$D = \begin{bmatrix} f_0 & 0 \\ 0 & f_1 \end{bmatrix}$$

and let $Q^*(s)$ be the matrix $(I - D)A^*(s)$.

This situation is equivalent to a semi-Markov process being decomposed by a state dependent switch. Cinlar (7) has shown that the LaPlace-Stieltjes transform of the resulting semi-Markov process is given by $B^*(s)$, where

$$B^*(s) = A^*(s)[I - Q^*(s)]^{-1} D$$

In the present case this reduces to the single term given below:

$$B^*(s) = \Pi\left[F_{00}^*(s) + \frac{(1-\Pi)F_{01}^*(s)F_{10}^*(s)}{1 - (1-\Pi)F_{11}^*(s)}\right]$$

$$= \Pi\left[\frac{F_{00}^*(s) + (1-\Pi)[F_{01}^*(s)F_{10}^*(s) - F_{00}^*(s)F_{11}^*(s)]}{1 - (1-\Pi)F_{11}^*(s)}\right]$$

To generate the stochastic process consisting of police emergencies alone, define

$$\tilde{f}_0 = Pr\{Y_n = 1 | Z_n = 0\} = 0$$

$$\tilde{f}_1 = Pr\{Y_n = 1 | Z_n = 1\} = 1$$

Then let $\tilde{D}$ be the matrix $\tilde{D} = \begin{bmatrix} \tilde{f}_0 & 0 \\ 0 & \tilde{f}_1 \end{bmatrix}$ and let $\tilde{Q}^*(s) = (I - \tilde{D})A^*(s)$. Then $\tilde{B}^*(s) = A^*(s)[1 - \tilde{Q}^*(s)]^{-1}\tilde{D}$ is the Laplace-Stieltjes Transform of the distribution of the time between police calls. This is given by

$$\tilde{B}^*(s) = (1-\Pi)\left[F_{11}^*(s) + \frac{\Pi F_{10}^*(s)F_{01}^*(s)}{1 - \Pi F_{00}^*(s)}\right]$$

$$= (1-\Pi)\left[\frac{F_{11}^*(s) + [F_{10}^*(s)F_{01}^*(s) - F_{00}^*(s)F_{11}^*(s)]}{1 - \Pi F_{00}^*(s)}\right]$$

We have seen in Chapter II that $F_{ij}(x)$ can be modeled by an exponential distribution with parameter λ_{ij}. In this case $F^*_{ij}(s) = \frac{\lambda_{ij}}{\lambda_{ij}+s}$, and both $B^*(s)$ and $\tilde{B}^*(s)$ can be represented by ratios of polynomials. It is known that a non-negative random variable whose Laplace transform is the ratio of two polynomials belongs to the "general Erlangian" class of distributions (Cox and Smith [11]). Since the Gamma and negative exponential distributions are members of this class, their empirical acceptance in Chapter II is not surprising.

In fact, we conjecture that the transforms $B^*(s)$ and $\tilde{B}^*(s)$ can be adequately approximated by the Laplace transforms of the Gamma and negative-exponential distribution which were identified in Sections 2.2.1 and 2.2.2 respectively. For instance, the application under consideration in this research resulted in a small value for Π and $\lambda_{00} \doteq \lambda_{11} = \lambda$. Then we might approximate $\tilde{B}^*(s)$ as:

$$\tilde{B}^*(s) = \frac{(1-\Pi)F^*_{11}(s)}{1 - \Pi F^*_{00}(s)} = \frac{(1-\Pi)\left(\frac{\lambda}{\lambda+s}\right)}{1-\Pi\left(\frac{\lambda}{\lambda+s}\right)} = \frac{\lambda(1-\Pi)}{s+\lambda(1-\Pi)}$$

In this case the approximate distribution of the times between police calls is negative-exponential with parameter $\lambda(1-\Pi)$. When maximum likelihood estimators of

λ and Π are substituted into this relation, we obtain a mean time between police calls of 18.4 minutes, a result which closely agrees with that of Section 2.2.2.

APPENDIX C

COMPUTER PROGRAM DOCUMENTATION

The numerical results in Chapters IV, V, and VI of this dissertation were calculated by programming the algorithm of Chapter III on the University of Michigan IBM 360/67 MTS Computing System. The resulting program, which was written in the Fortran G compiler language, is documented in this appendix. The purpose of this documentation is to illustrate the relevant computational techniques and to provide a reference for other researchers wishing to adapt these results to their particular application.

Imput data to the program consists of (1) the number of recovery vehicles in the system, (2) the estimated parameters of the emergency occurrence and service processes, and (3) the parameters of the various "switching" matrices. The output consists of the equilibrium state probabilities, although the Markovian state transition probabilities are optionally available.

The "switching" matrices $\Pr\{Y_n | \underline{S}_n, Z_n\}$ are developed for each vehicle distribution policy through a direct implementation of the decision rules of Section 3.2. The documented program considers only the cases

r=1, r=2; larger values of r can be implemented by a straight-forward generalization.

The conditional state transition probabilities $Pr\{\underline{S}_n, Z_n|\underline{S}_{n-1}, Z_{n-1}, Y_{n-1}\}$ are developed from the formula

$$Pr\{\underline{S}_n, Z_n|\underline{S}_{n-1}, Z_{n-1}, Y_{n-1}\} = \int_0^\infty Pr\{\underline{S}_n|\underline{S}_{n-1}, Y_{n-1}, X_n=x\}\, x \qquad (C.1)$$

$$d_x Pr\{X_n \le x, Z_n|Z_{n-1}\}$$

In Section 3.1 we observed that for the model under consideration, non-zero values of $Pr\{\underline{S}_n|\underline{S}_{n-1}, Y_{n-1}, X_n = x\}$ can be written as the product of terms of the form $1-e^{-\beta_i x}$ and $e^{-\beta_j x}$, where β_i and β_j denote the various emergency service rates. In general this product can be written:

$$\prod_{i=1}^{r_1} (1-e^{-\beta_i x}) \prod_{j=1}^{r_2} e^{-\beta_j x} \qquad r_1 + r_2 = r \qquad (C.2)$$

For instance, when $r_1 = r_2 = 1$, the above product simplifies to

$$e^{-\beta_j x} - e^{-(\beta_i+\beta_j)x}$$

For general density functions $d_x Pr\{X_n \le x, Z_n|Z_{n-1}\}$, the integration of equation (C.1) can be performed numerically. However, since in our model the distribution of times between ambulance calls in a single function

system is Gamma and the distributions of times between various events in a dual function system are negative exponential, the integrations can be performed analytically. We note that each term in the product (C.2) is of the form $e^{-\xi x}$, where ξ is a linear combination of certain emergency service rates. Hence the conditional state transition probabilities $Pr\{\underline{S}_n, Z_n | \underline{S}_{n-1}, Z_{n-1}, Y_{n-1}\}$ consist of linear combinations of terms of the form

$$\int_0^\infty e^{-\xi x} \left(\frac{\kappa}{\mu}\right)^\kappa \frac{x^{\kappa-1} e^{-\frac{\kappa x}{\mu}}}{\Gamma(\kappa)} dx = \left[\frac{\kappa/\mu}{\xi+\kappa/\mu}\right]^\kappa \qquad (C.3)$$

in the pure ambulance system and of terms of the form

$$\int_0^\infty e^{-\xi x} \lambda_{ij} e^{-\lambda_{ij} x} dx = \frac{\lambda_{ij}}{\lambda_{ij}+\xi} \qquad (C.4)$$

in the dual function ambulance system.

To perform the actual computation, each term of the form $1-e^{-\beta_i x}$ and $e^{-\beta_j x}$ was assigned a unique prime number. The product of these terms (C.2) can then be uniquely represented by the product of the associated prime numbers. Finally, the probabilities $Pr\{\underline{S}_n, Z_n | \underline{S}_{n-1}, Z_{n-1}, Y_{n-1}\}$ can be found by performing a table look-up which associates a unique linear combination of elements of the form (C.3) or (C.4) with this resulting product.

The table look-up was performed by calling the subroutine GETA2. The arguments to this subroutine consisted of the prime product calculated above and the two parameters associated with the emergency occurrence process. (The dummy parameter K=1 was used as the second parameter when the time between the events of interest was negative exponential.)

The matrix of products of primes was generated by observing that the matrix of probabilities $Pr\{\underline{S}_n | \underline{S}_{n-1}, Y_{n-1}, X_n = x\}$ can be computed by forming the Kronecker products of identical matrices whose elements are $Pr\{S_n^i | S_{n-1}^i, Y_{n-1}, X_n = x\}$ $i = 1, 2, \cdots, r$. In the Fortran program, these component matrices are denoted by NS, and the resulting matrix of products of primes is denoted by NSS. The matrix of conditional state transition probabilities is denoted by SS.

The Markovian transition probabilities $Pr\{\underline{S}_n, Z_n | \underline{S}_{n-1}, Z_{n-1}\}$ are computed according to the formula

$$Pr\{\underline{S}_n, Z_n | \underline{S}_{n-1}, Z_{n-1}\} = \sum_{j=0}^{\ell} Pr\{\underline{S}_n, Z_n | \underline{S}_{n-1}, Z_{n-1}, Y_{n-1} = j\}$$

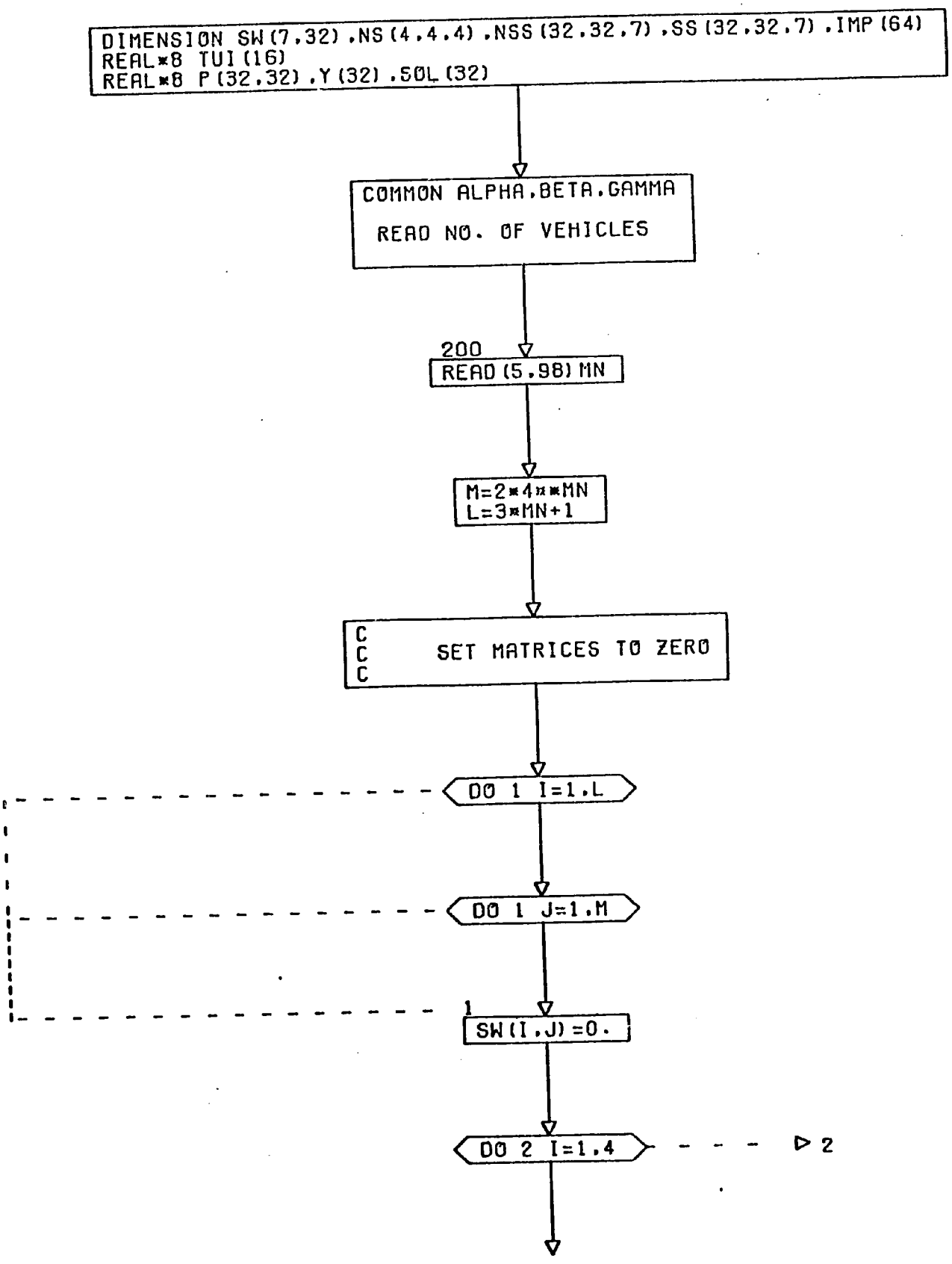
$$Pr\{Y_n = j | \underline{S}_{n-1}, Z_{n-1}\}$$

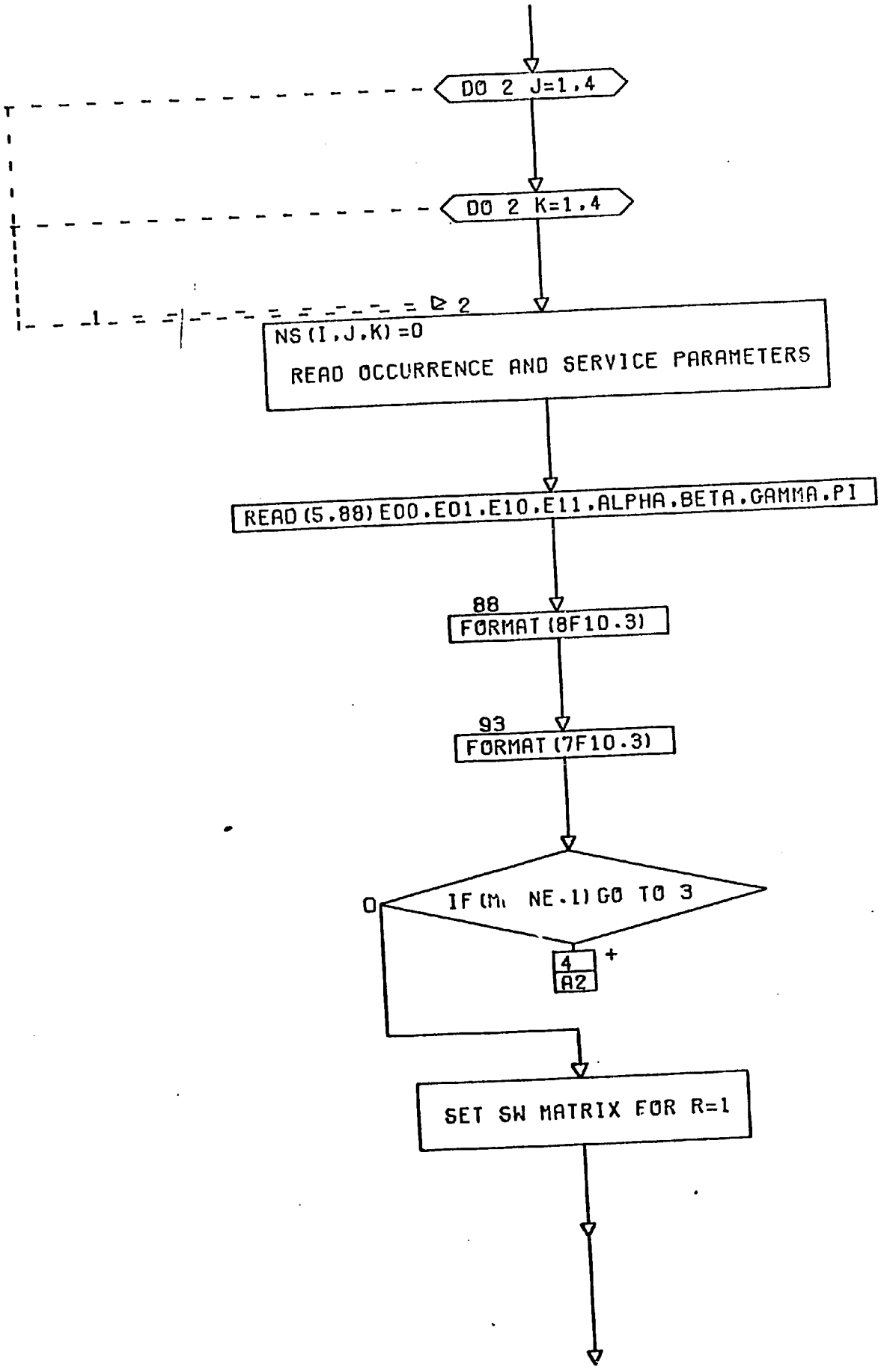
The transition probabilities are stored in the matrix denoted P.

The equilibrium state probabilities were calculated by forming the system of linear equations previously discussed in Section 4.1. The solution to the resulting system was found using the University of Michigan Computing Center Library Subroutine SLE1. This routine uses the Gaussian elimination algorithm documented by Forsythe and Moler in (15), and it performs all computation in double precision arithmetic.

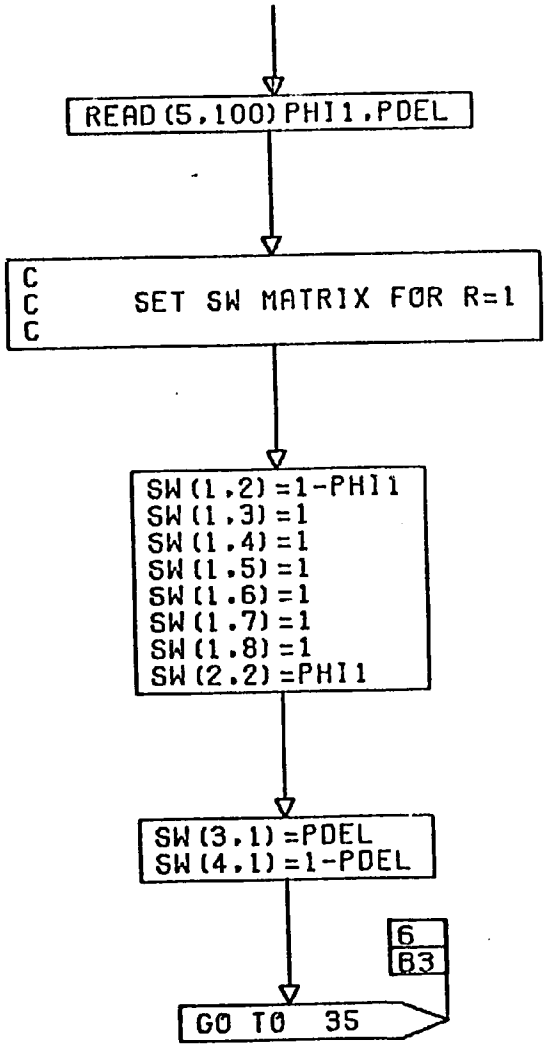
A flow chart showing the program logic is presented as Figure C-1. This flowchart was prepared using the Calcomp Flowgen program on the University of Michigan Highway Safety Research Institute IBM 1130 Computing System.

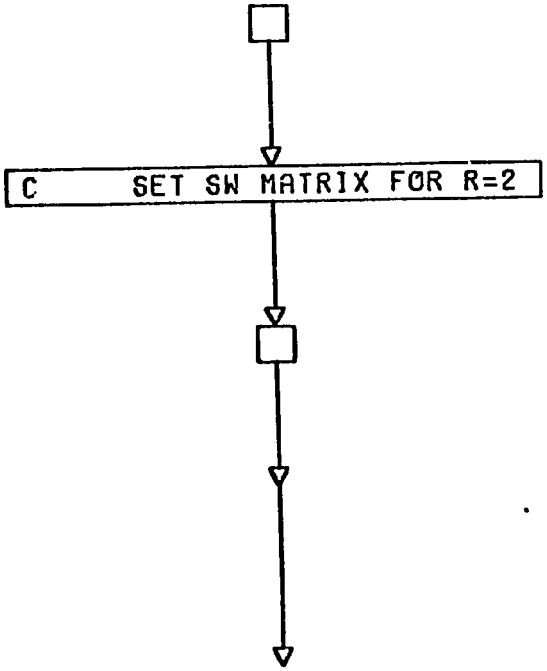
FIGURE C.1
COMPUTER PROGRAM LOGIC



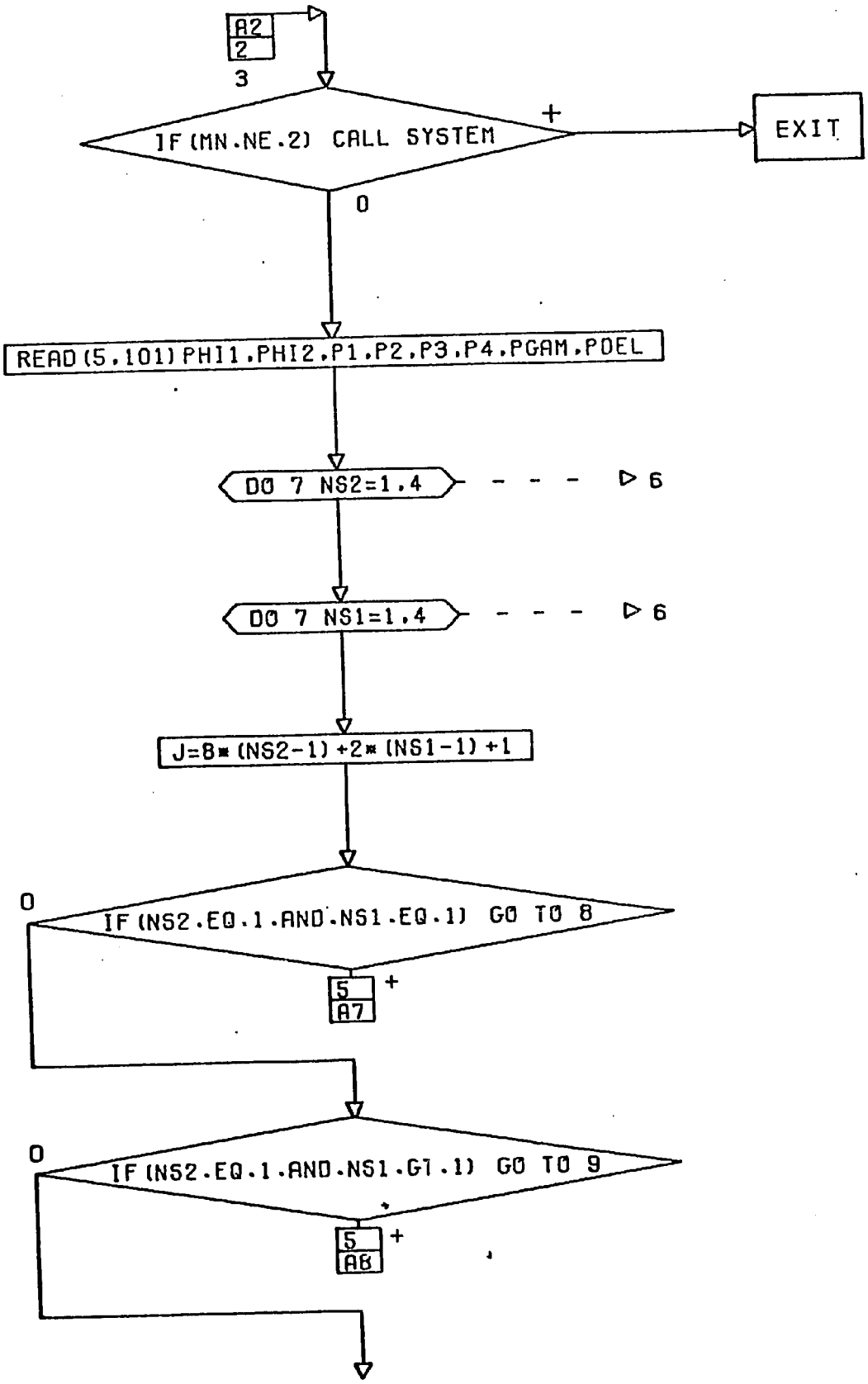

DO 2 J=1,4
DO 2 K=1,4
2
NS(I,J,K)=0
READ OCCURRENCE AND SERVICE PARAMETERS
READ(5,88)E00,E01,E10,E11,ALPHA,BETA,GAMMA,PI
88
FORMAT(8F10.3)
93
FORMAT(7F10.3)
IF(M NE.1)GO TO 3
0
4
A2
+
SET SW MATRIX FOR R=1

CONT. ON PG 3

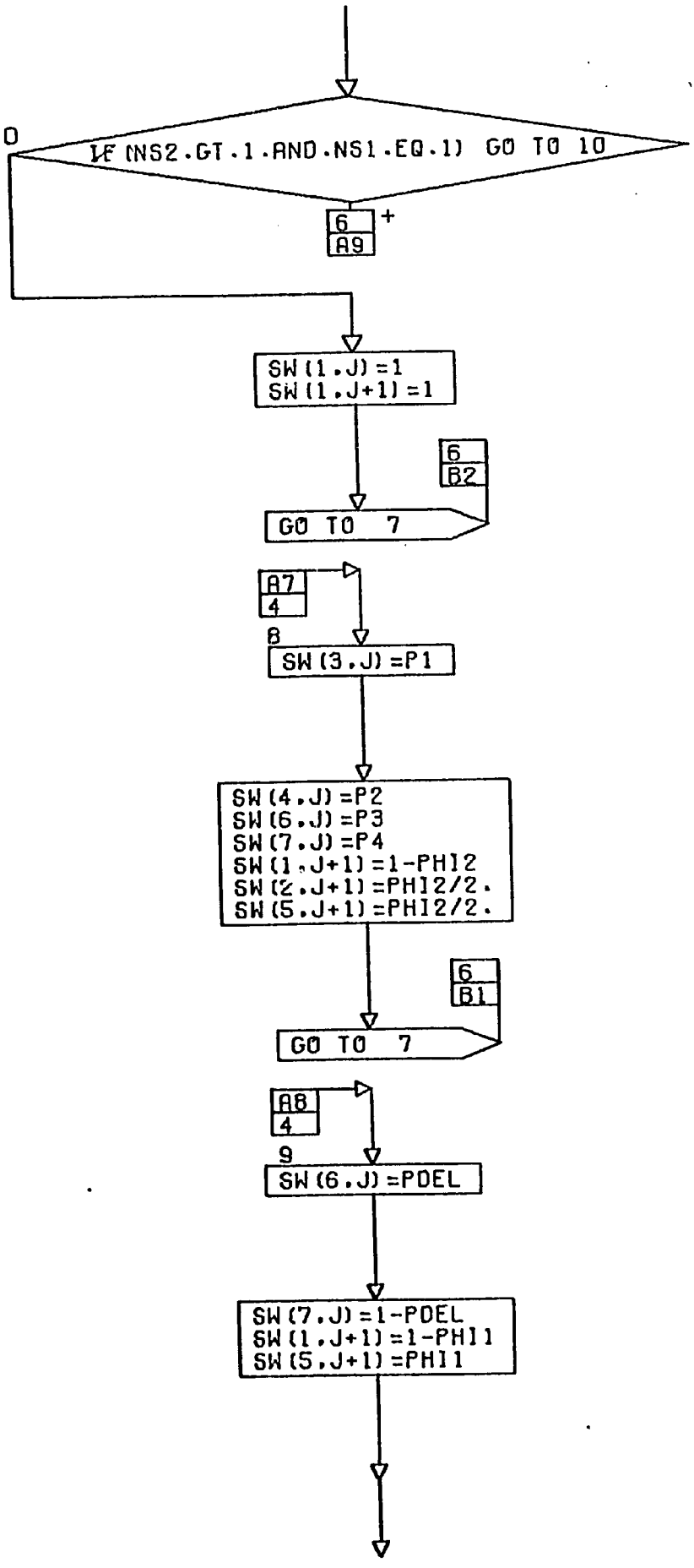

READ(5,100)PHI1,PDEL
C
C SET SW MATRIX FOR R=1
C
SW(1,2)=1-PHI1
SW(1,3)=1
SW(1,4)=1
SW(1,5)=1
SW(1,6)=1
SW(1,7)=1
SW(1,8)=1
SW(2,2)=PHI1
SW(3,1)=PDEL
SW(4,1)=1-PDEL
6
83
GO TO 35


C SET SW MATRIX FOR R=2

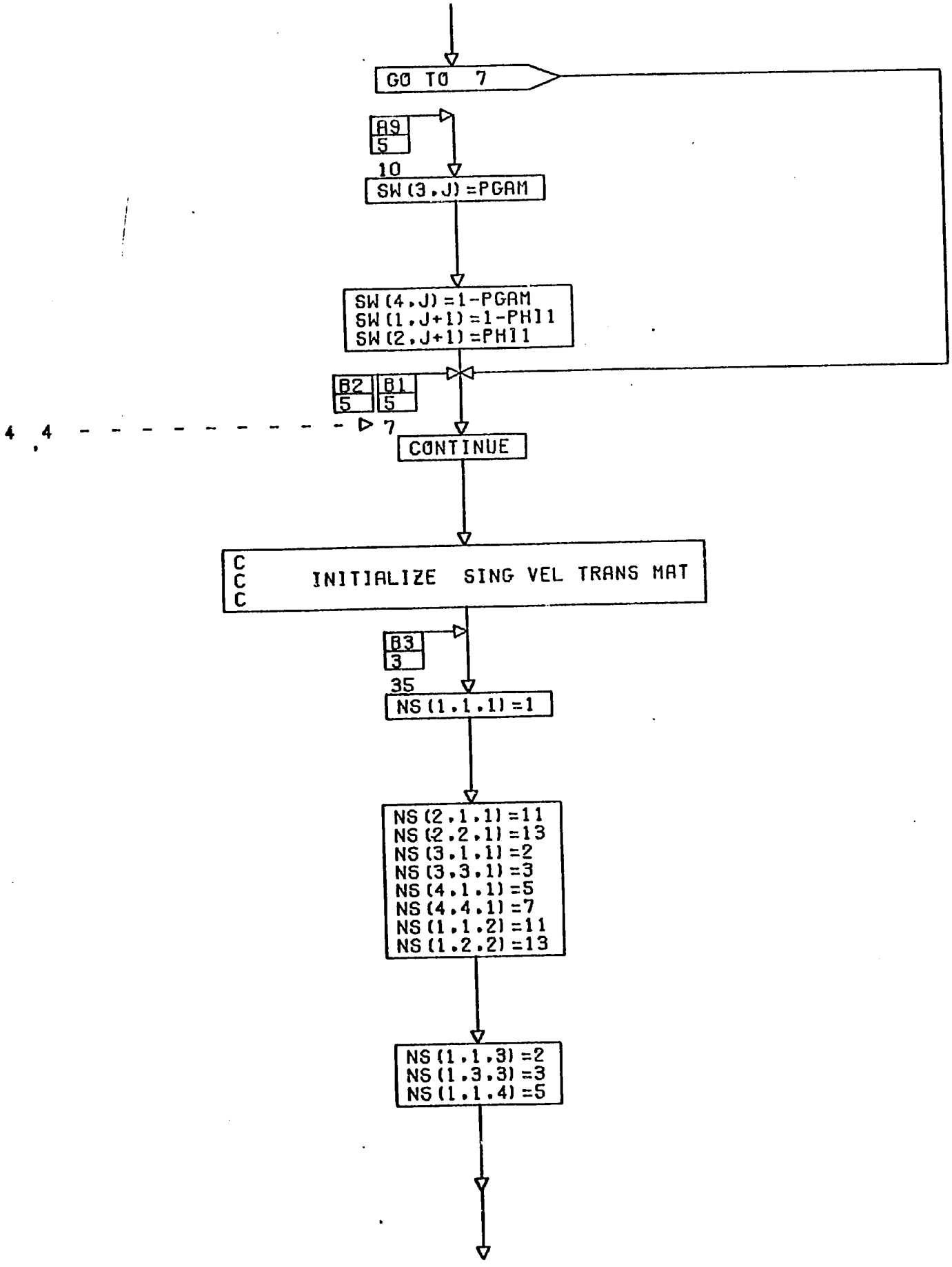
CONT. ON PG 4


A2
2
3
IF (MN.NE.2) CALL SYSTEM
+
EXIT
0
READ(5,101)PHI1,PHI2,P1,P2,P3,P4,PGAM,PDEL
DO 7 NS2=1,4
6
DO 7 NS1=1,4
6
J=8*(NS2-1)+2*(NS1-1)+1
0
IF (NS2.EQ.1.AND.NS1.EQ.1) GO TO 8
5
A7
+
0
IF (NS2.EQ.1.AND.NS1.GT.1) GO TO 9
5
A8
+

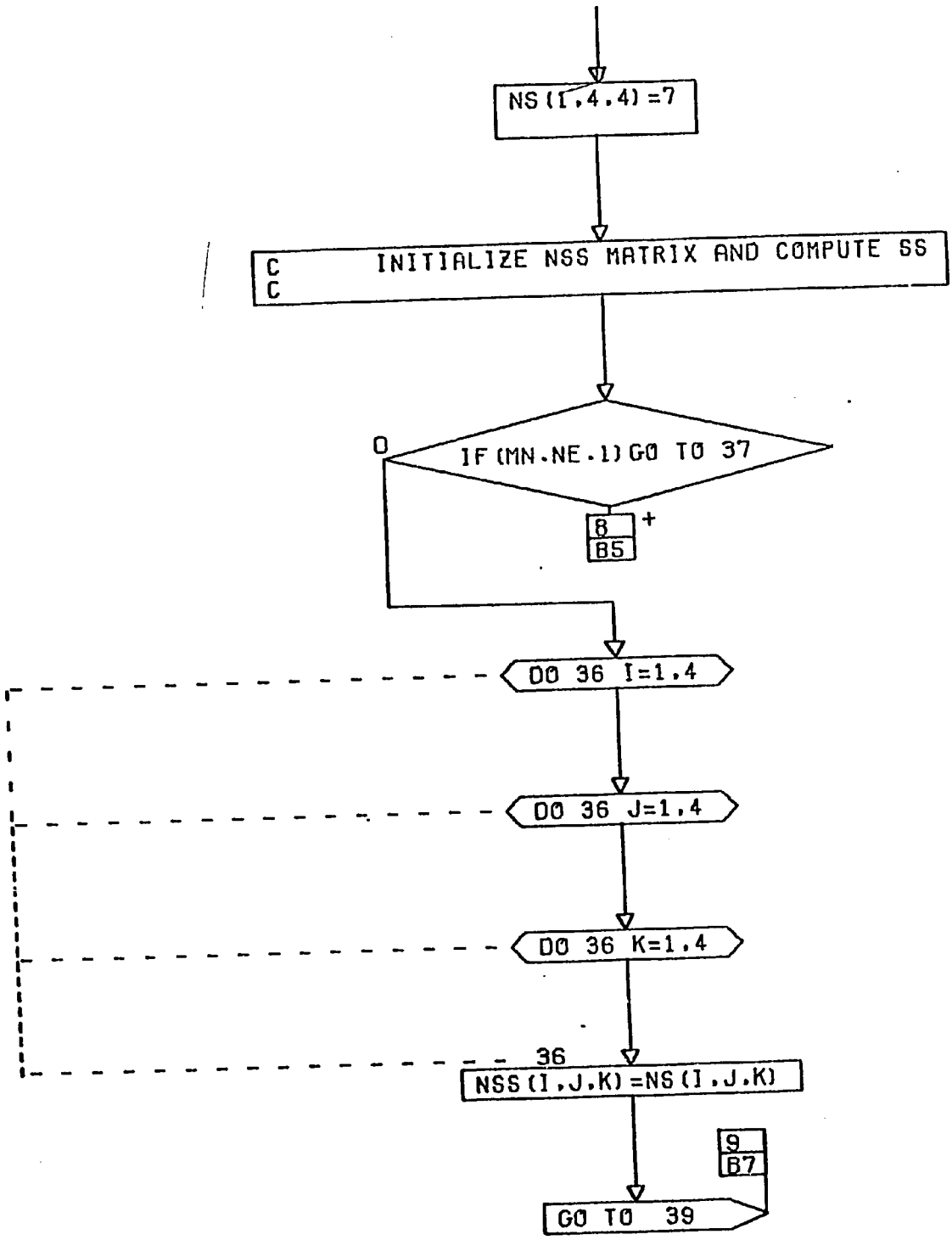
CONT. ON PG 5


0
IF (NS2.GT.1.AND.NS1.EQ.1) GO TO 10
6
A9
+
SW(1,J)=1
SW(1,J+1)=1
6
B2
GO TO 7
A7
4
8
SW(3,J)=P1
SW(4,J)=P2
SW(6,J)=P3
SW(7,J)=P4
SW(1,J+1)=1-PHI2
SW(2,J+1)=PHI2/2.
SW(5,J+1)=PHI2/2.
6
B1
GO TO 7
A8
4
9
SW(6,J)=PDEL
SW(7,J)=1-PDEL
SW(1,J+1)=1-PHI1
SW(5,J+1)=PHI1

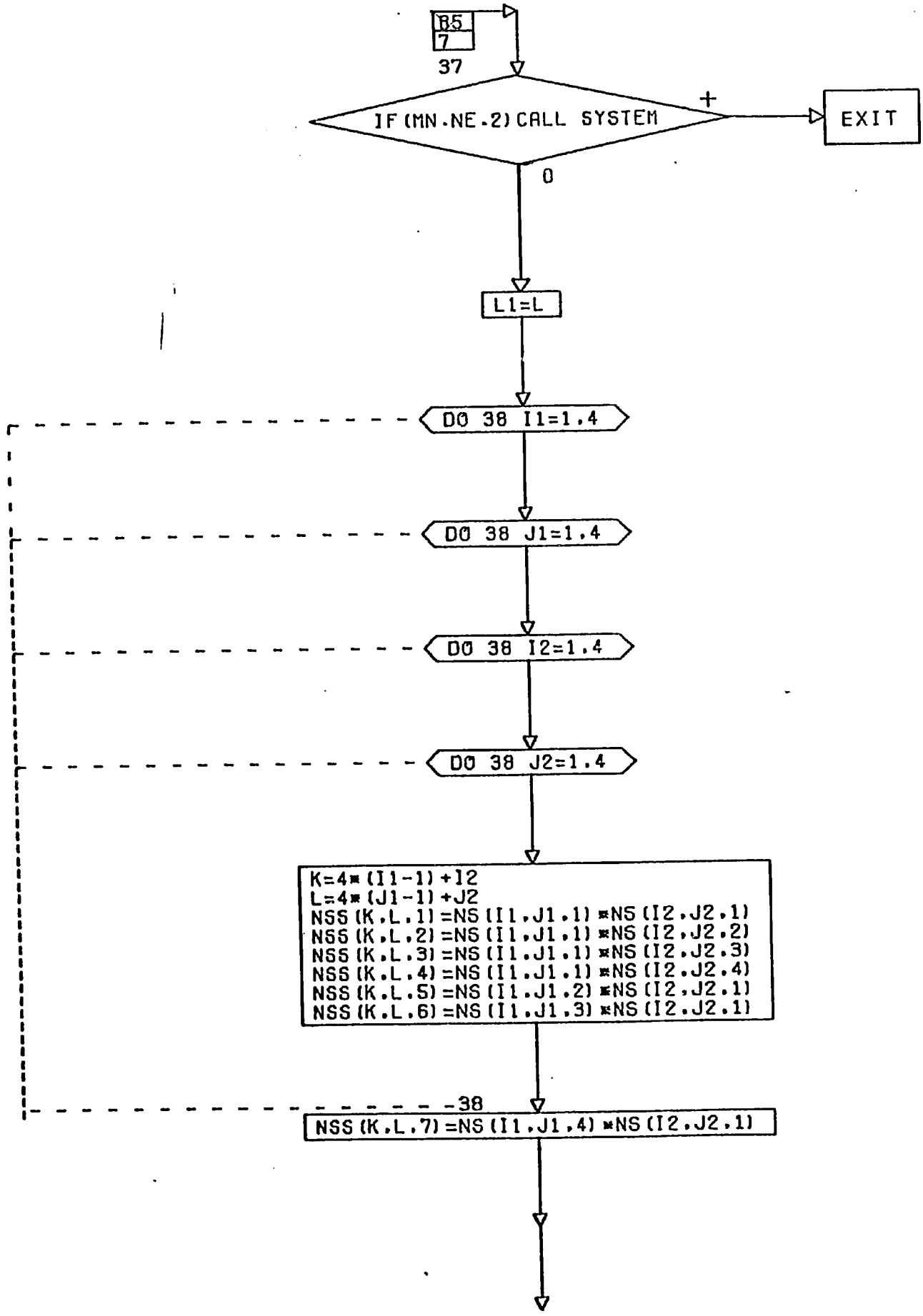
CONT. ON PG 6


GO TO 7
A9
5
10
SW(3,J)=PGAM
SW(4,J)=1-PGAM
SW(1,J+1)=1-PHI1
SW(2,J+1)=PHI1
B2
5
B1
5
4 4
7
CONTINUE
C
C INITIALIZE SING VEL TRANS MAT
C
B3
3
35
NS(1,1,1)=1
NS(2,1,1)=11
NS(2,2,1)=13
NS(3,1,1)=2
NS(3,3,1)=3
NS(4,1,1)=5
NS(4,4,1)=7
NS(1,1,2)=11
NS(1,2,2)=13
NS(1,1,3)=2
NS(1,3,3)=3
NS(1,1,4)=5

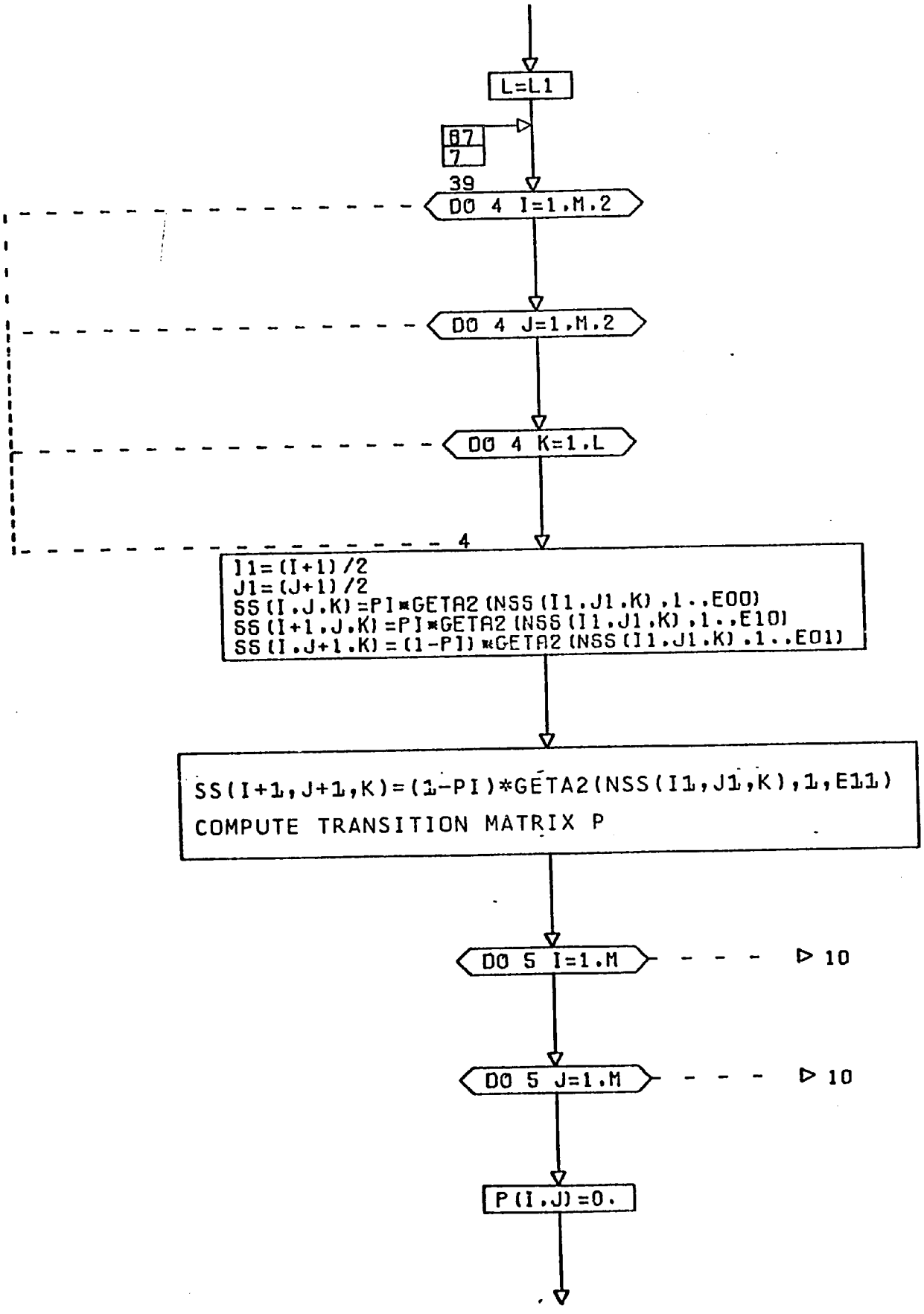
CONT. ON PG 7


NS(I,4,4)=7
C INITIALIZE NSS MATRIX AND COMPUTE SS
C
0
IF(MN.NE.1)GO TO 37
8
B5
+
DO 36 I=1,4
DO 36 J=1,4
DO 36 K=1,4
36
NSS(I,J,K)=NS(I,J,K)
9
B7
GO TO 39

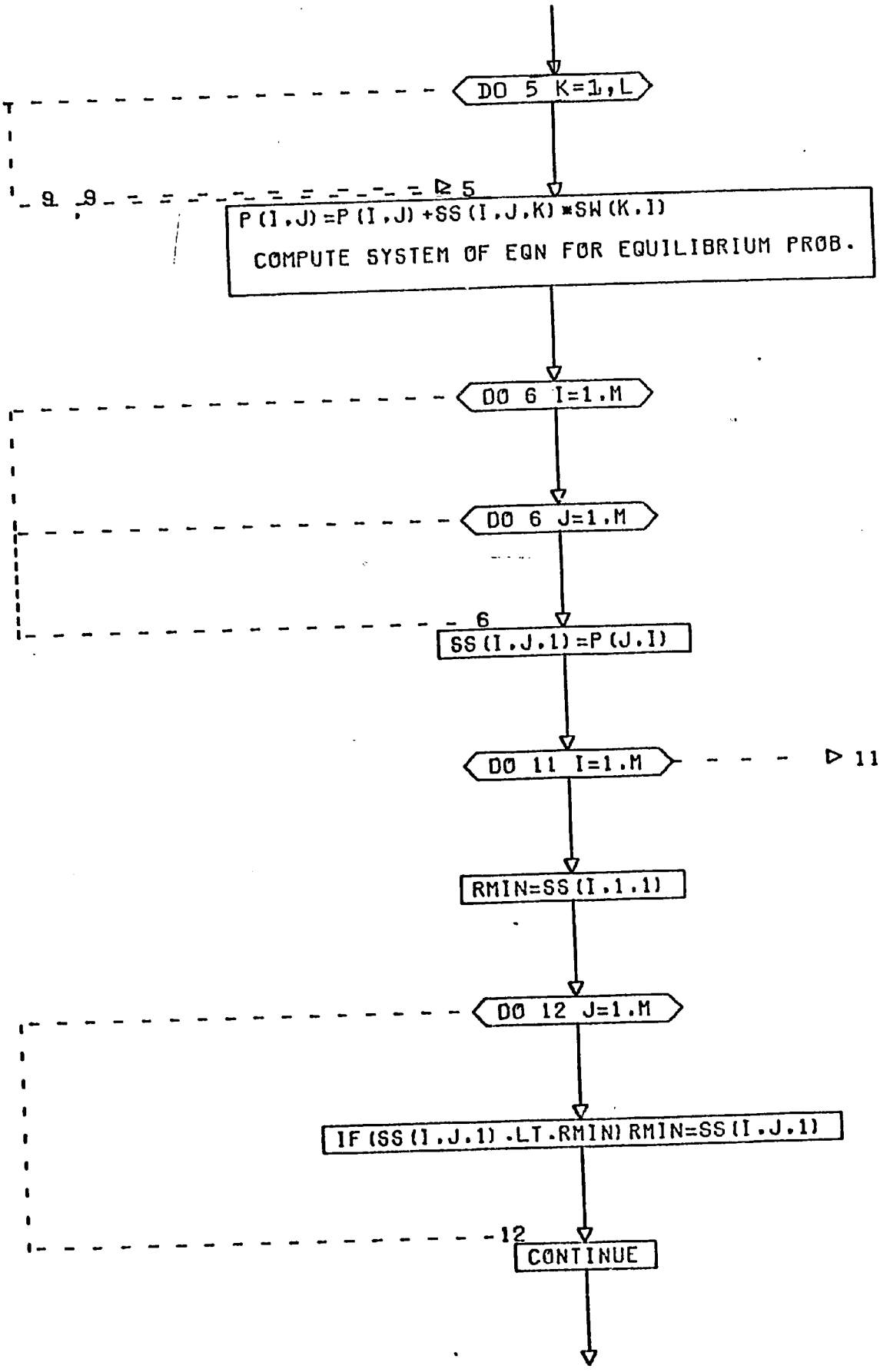
CONT. ON PG 8


B5
7
37
IF(MN.NE.2)CALL SYSTEM
+
EXIT
0
L1=L
DO 38 I1=1,4
DO 38 J1=1,4
DO 38 I2=1,4
DO 38 J2=1,4
K=4*(I1-1)+I2
L=4*(J1-1)+J2
NSS(K,L,1)=NS(I1,J1,1)*NS(I2,J2,1)
NSS(K,L,2)=NS(I1,J1,1)*NS(I2,J2,2)
NSS(K,L,3)=NS(I1,J1,1)*NS(I2,J2,3)
NSS(K,L,4)=NS(I1,J1,1)*NS(I2,J2,4)
NSS(K,L,5)=NS(I1,J1,2)*NS(I2,J2,1)
NSS(K,L,6)=NS(I1,J1,3)*NS(I2,J2,1)
38
NSS(K,L,7)=NS(I1,J1,4)*NS(I2,J2,1)

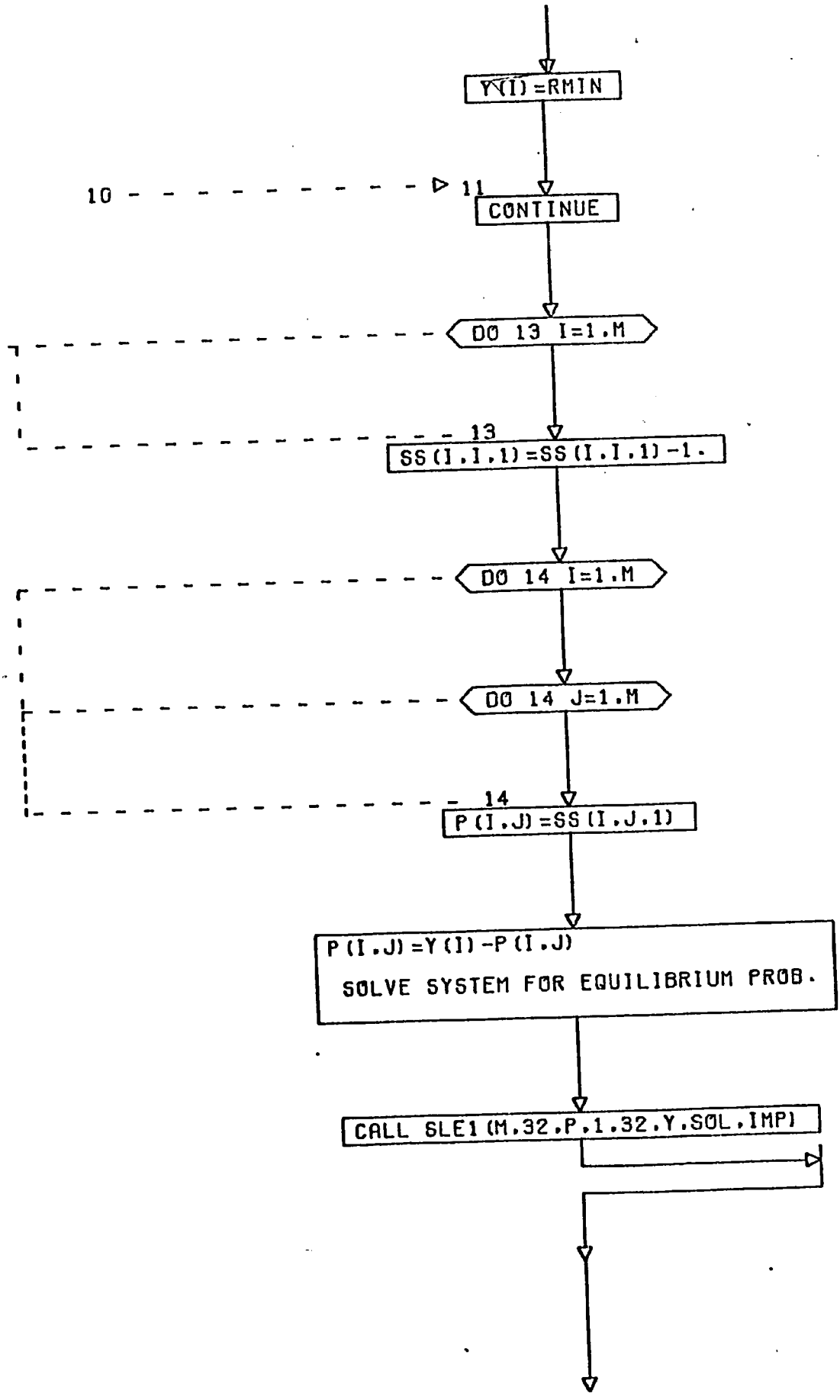
CONT. ON PG 9


L=L1
87
7
39
DO 4 I=1,M,2
DO 4 J=1,M,2
DO 4 K=1,L
4
I1=(I+1)/2
J1=(J+1)/2
SS(I,J,K)=PI*GETA2(NSS(I1,J1,K),1,,E00)
SS(I+1,J,K)=PI*GETA2(NSS(I1,J1,K),1,,E10)
SS(I,J+1,K)=(1-PI)*GETA2(NSS(I1,J1,K),1,,E01)
SS(I+1,J+1,K)=(1-PI)*GETA2(NSS(I1,J1,K),1,E11)
COMPUTE TRANSITION MATRIX P
DO 5 I=1,M
10
DO 5 J=1,M
10
P(I,J)=0.

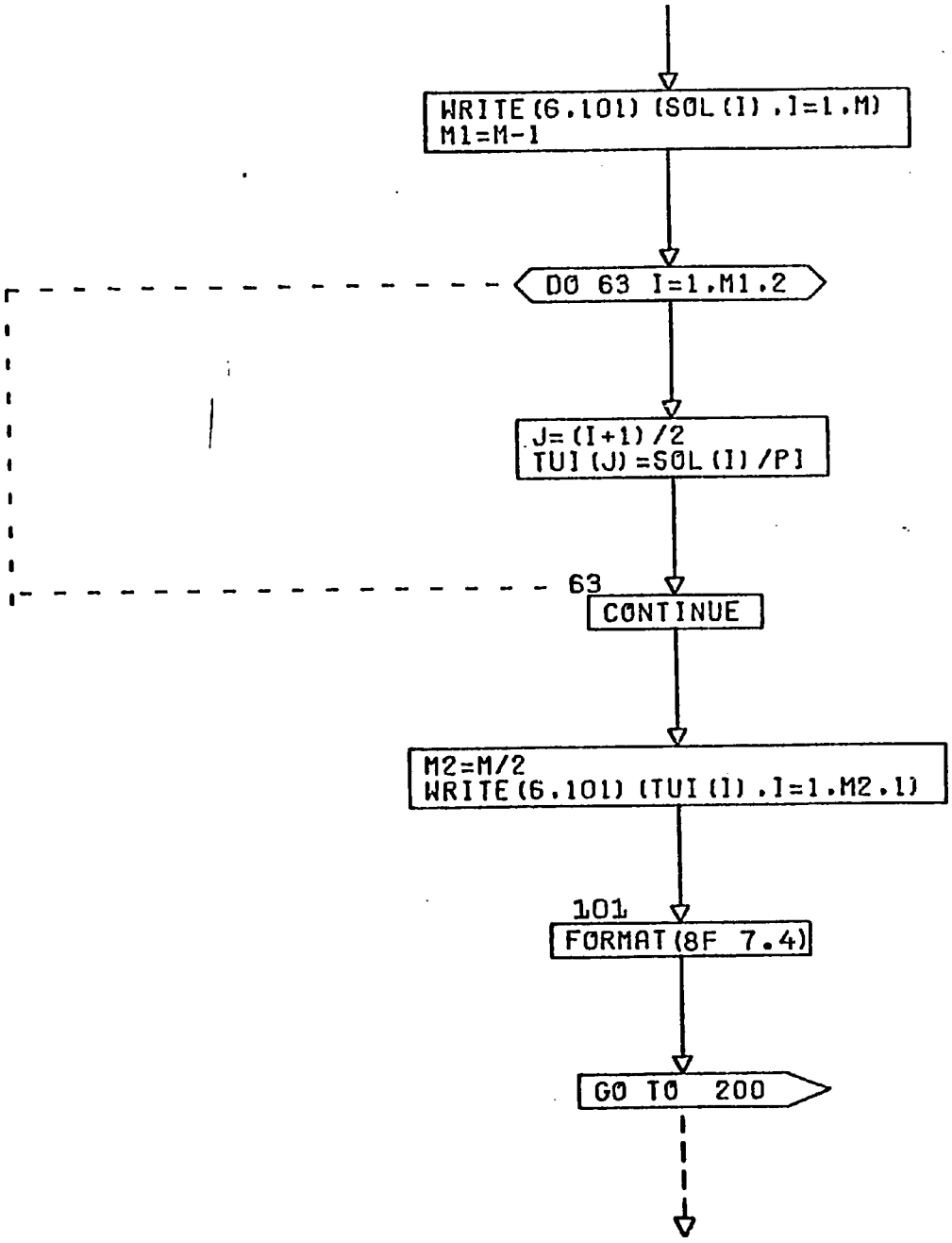
CONT. ON PG 10

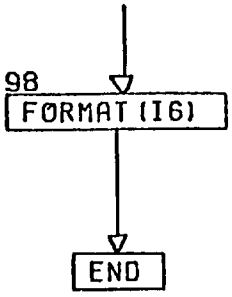

DO 5 K=1,L
9 9 ▷ 5
P(I,J)=P(I,J)+SS(I,J,K)*SW(K,1)
COMPUTE SYSTEM OF EQN FOR EQUILIBRIUM PROB.
DO 6 I=1,M
DO 6 J=1,M
6
SS(I,J,1)=P(J,I)
DO 11 I=1,M
▷ 11
RMIN=SS(I,1,1)
DO 12 J=1,M
IF(SS(I,J,1).LT.RMIN)RMIN=SS(I,J,1)
12
CONTINUE

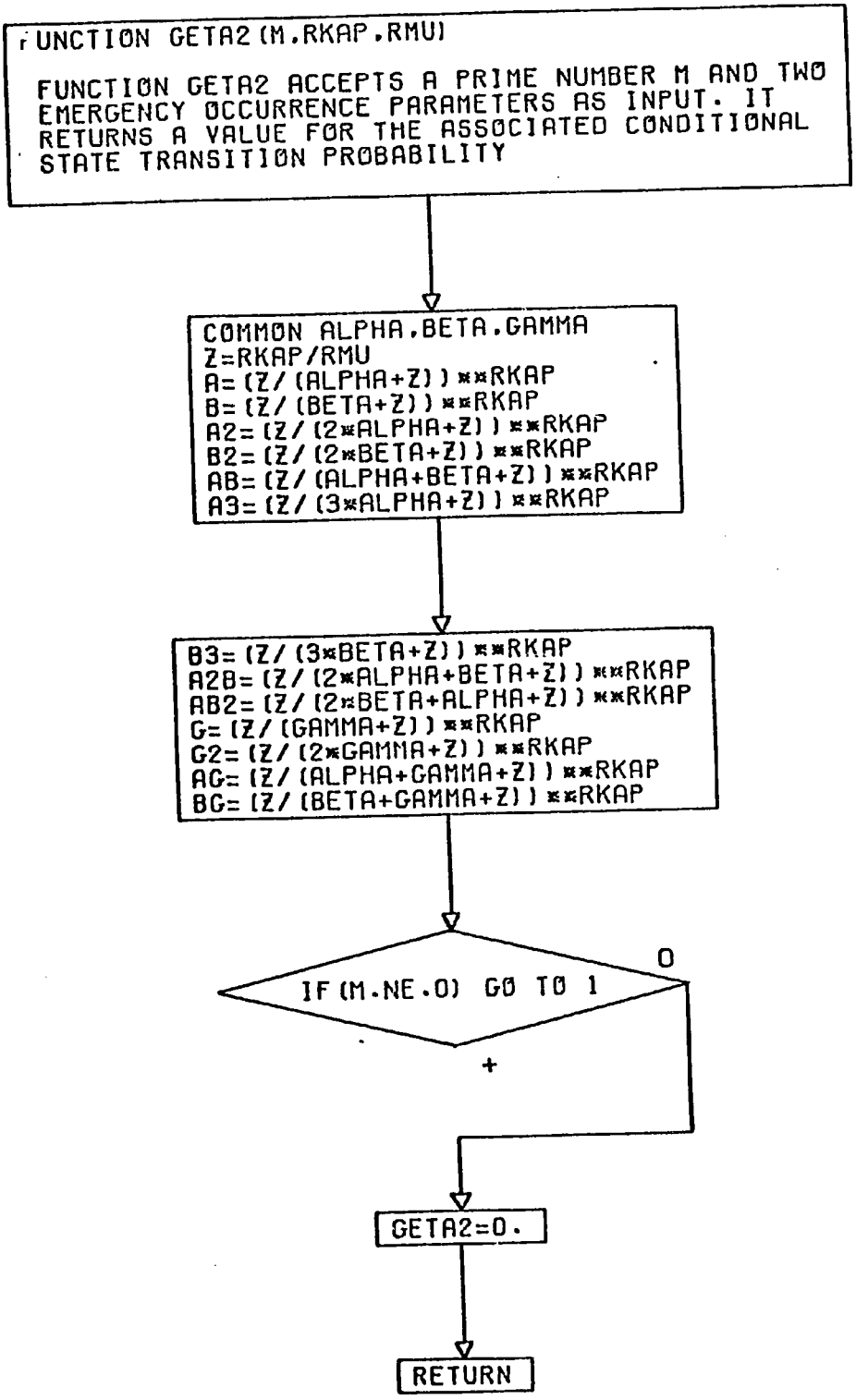
CONT. ON PG 11

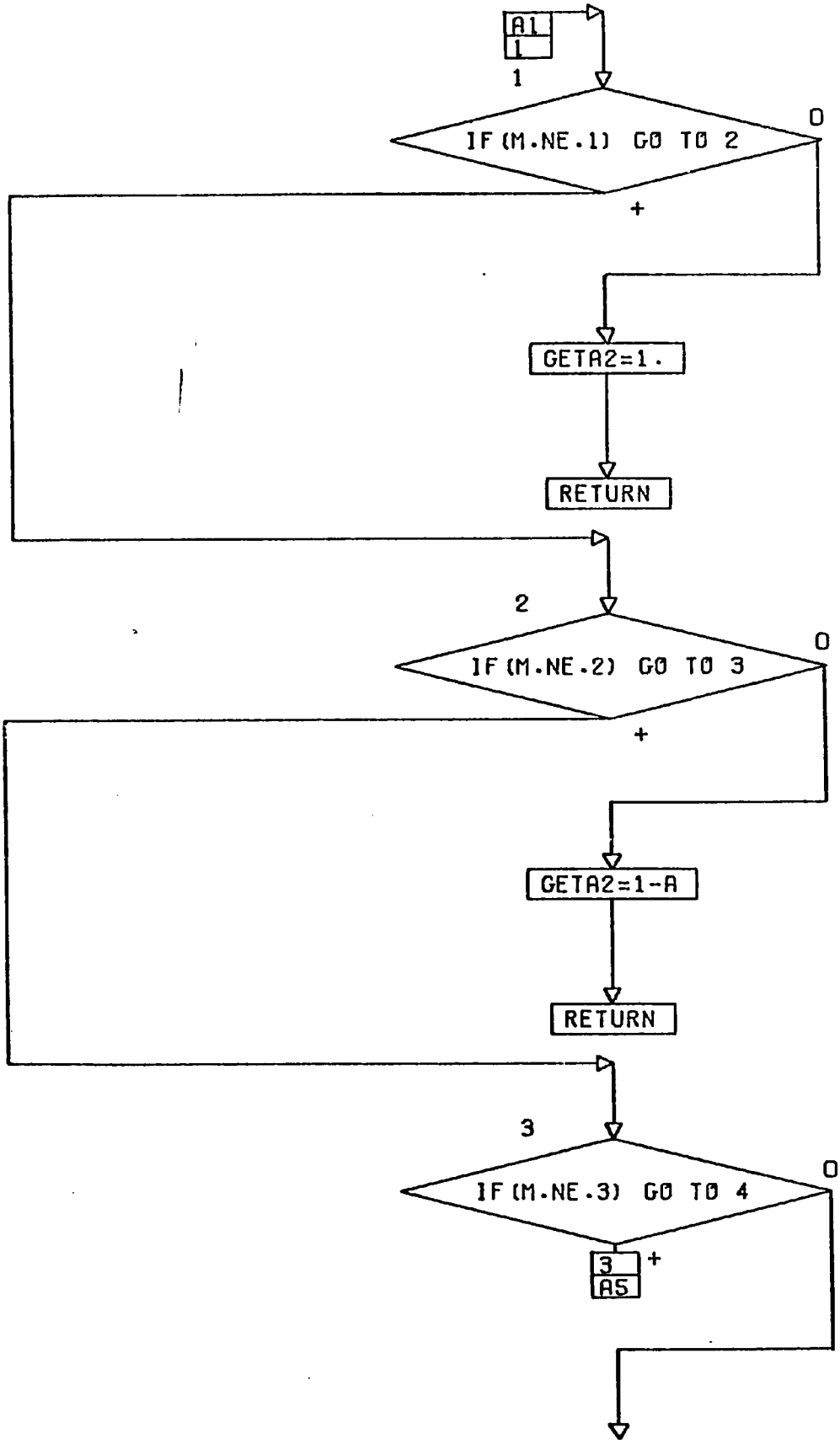

Y(I)=RMIN
10
11
CONTINUE
DO 13 I=1,M
13
SS(I,I,1)=SS(I,I,1)-1.
DO 14 I=1,M
DO 14 J=1,M
14
P(I,J)=SS(I,J,1)
P(I,J)=Y(I)-P(I,J)
SOLVE SYSTEM FOR EQUILIBRIUM PROB.
CALL SLE1(M,32,P,1,32,Y,SOL,IMP)

CONT. ON PG 12

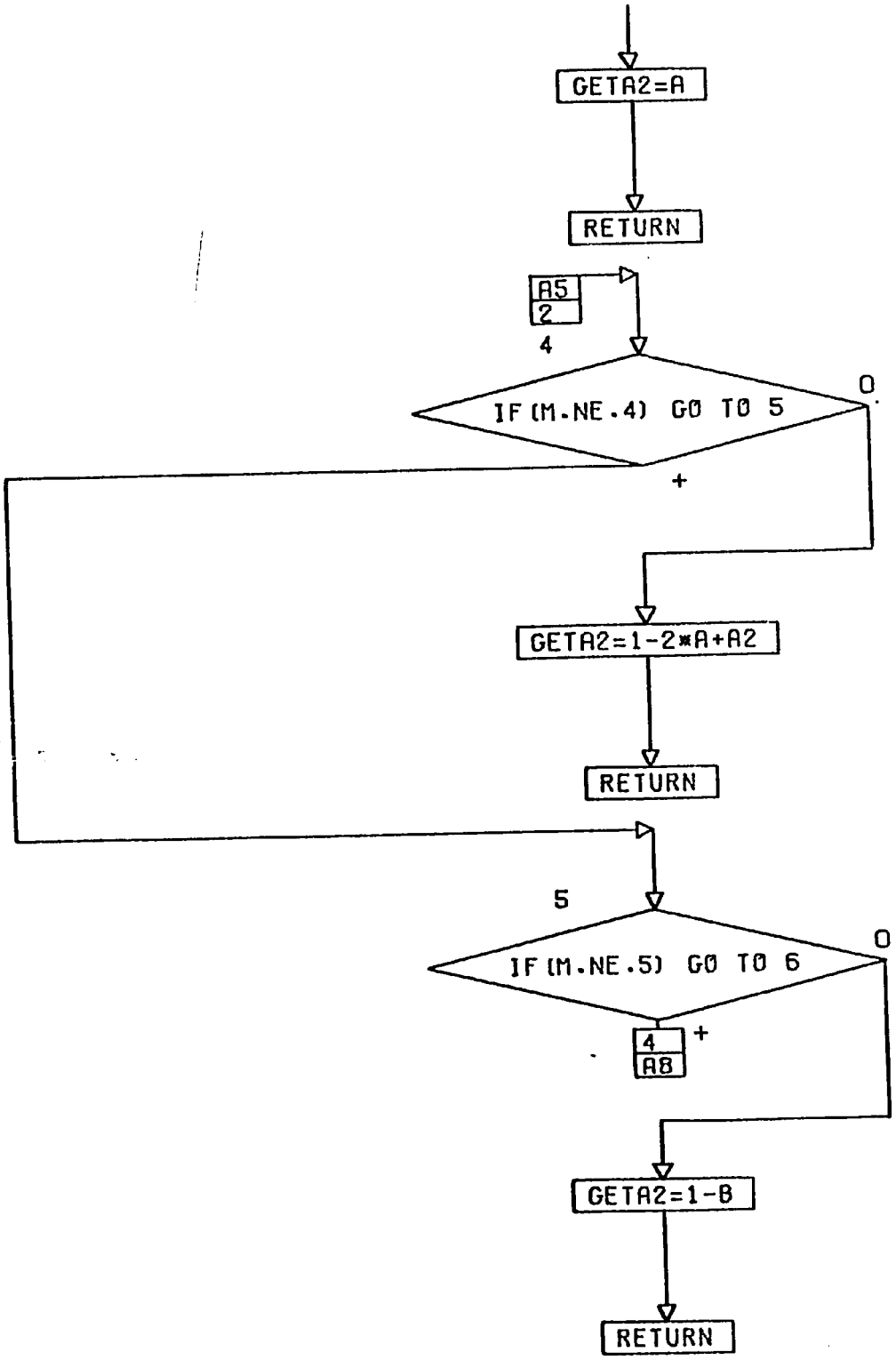

WRITE(6,101) (SOL(I),I=1,M)
M1=M-1
DO 63 I=1,M1,2
J=(I+1)/2
TUI(J)=SOL(I)/P1
63
CONTINUE
M2=M/2
WRITE(6,101) (TUI(I),I=1,M2,1)
101
FORMAT(8F 7.4)
GO TO 200


98
FORMAT (I6)
END

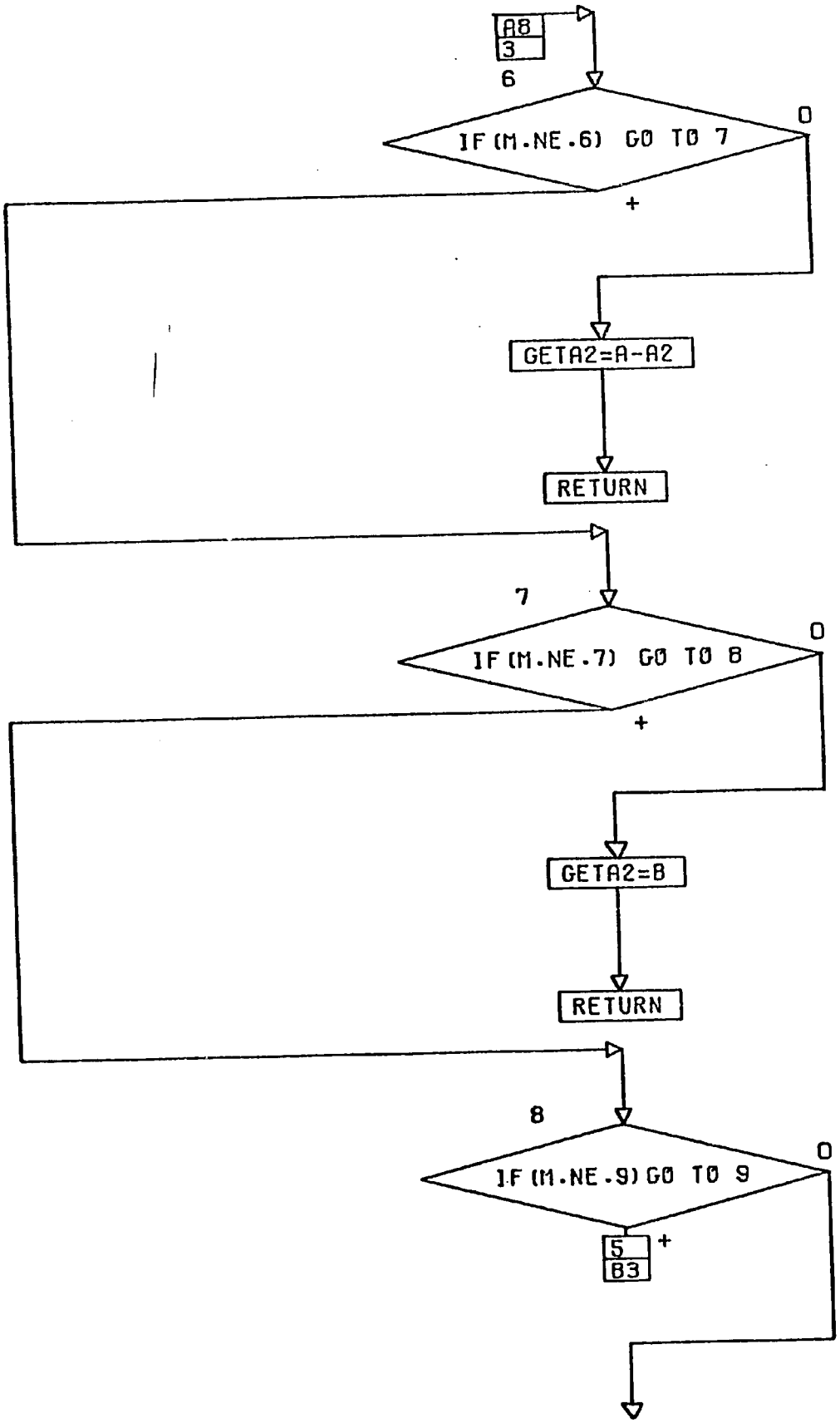

FUNCTION GETA2(M,RKAP,RMU)
FUNCTION GETA2 ACCEPTS A PRIME NUMBER M AND TWO EMERGENCY OCCURRENCE PARAMETERS AS INPUT. IT RETURNS A VALUE FOR THE ASSOCIATED CONDITIONAL STATE TRANSITION PROBABILITY
COMMON ALPHA,BETA,GAMMA
Z=RKAP/RMU
A=(Z/(ALPHA+Z))**RKAP
B=(Z/(BETA+Z))**RKAP
A2=(Z/(2*ALPHA+Z))**RKAP
B2=(Z/(2*BETA+Z))**RKAP
AB=(Z/(ALPHA+BETA+Z))**RKAP
A3=(Z/(3*ALPHA+Z))**RKAP
B3=(Z/(3*BETA+Z))**RKAP
A2B=(Z/(2*ALPHA+BETA+Z))**RKAP
AB2=(Z/(2*BETA+ALPHA+Z))**RKAP
G=(Z/(GAMMA+Z))**RKAP
G2=(Z/(2*GAMMA+Z))**RKAP
AG=(Z/(ALPHA+GAMMA+Z))**RKAP
BG=(Z/(BETA+GAMMA+Z))**RKAP
IF(M.NE.0) GO TO 1
0
+
GETA2=0.
RETURN


A1
1
1
IF (M.NE.1) GO TO 2
0
+
GETA2=1.
RETURN
2
IF (M.NE.2) GO TO 3
0
+
GETA2=1-A
RETURN
3
IF (M.NE.3) GO TO 4
0
3
A5
+

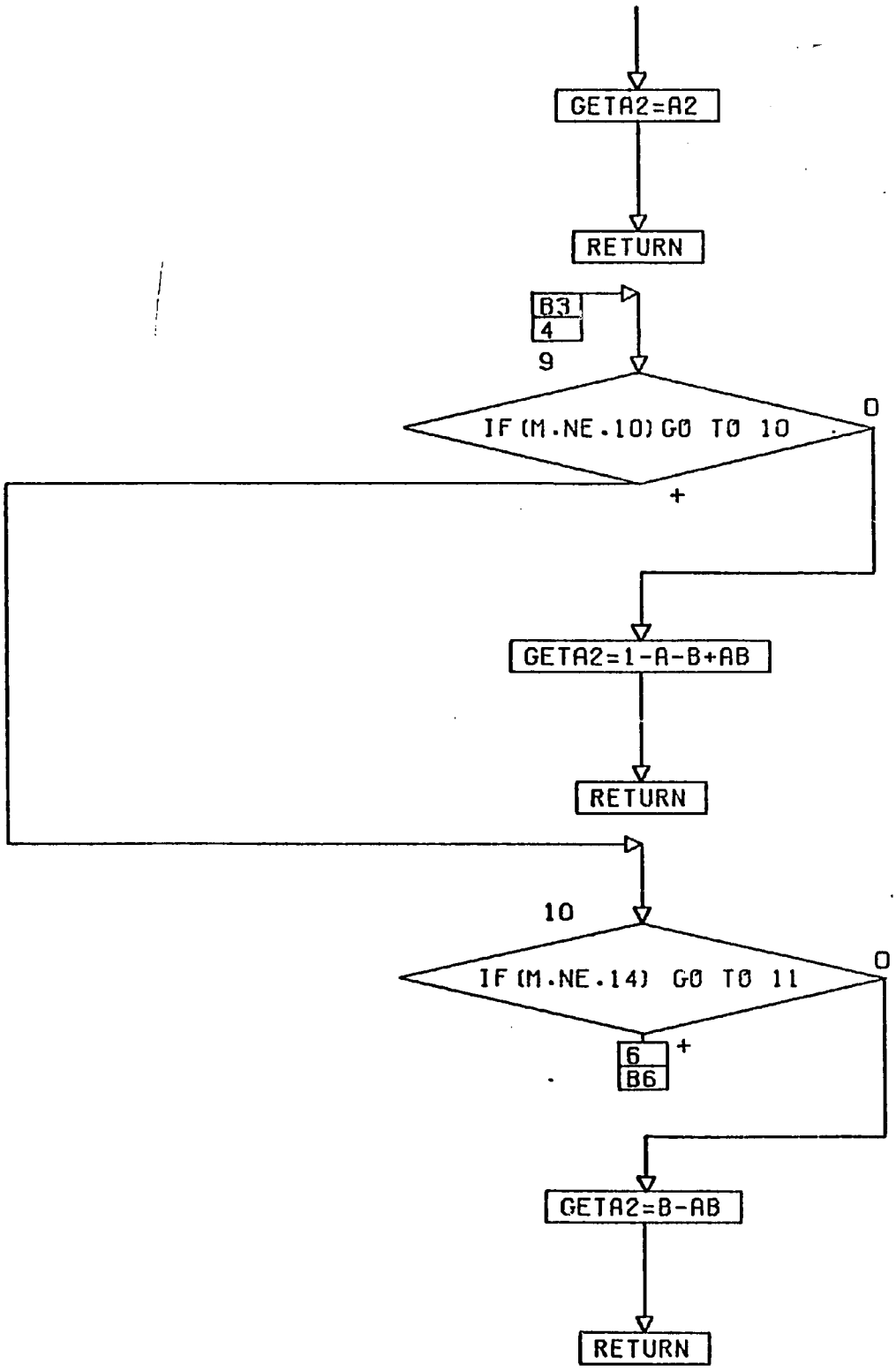
CONT. ON PG 3


GETA2=A
RETURN
A5
2
4
IF (M.NE.4) GO TO 5
0
+
GETA2=1-2*A+A2
RETURN
5
IF (M.NE.5) GO TO 6
0
4
A8
+
GETA2=1-B
RETURN

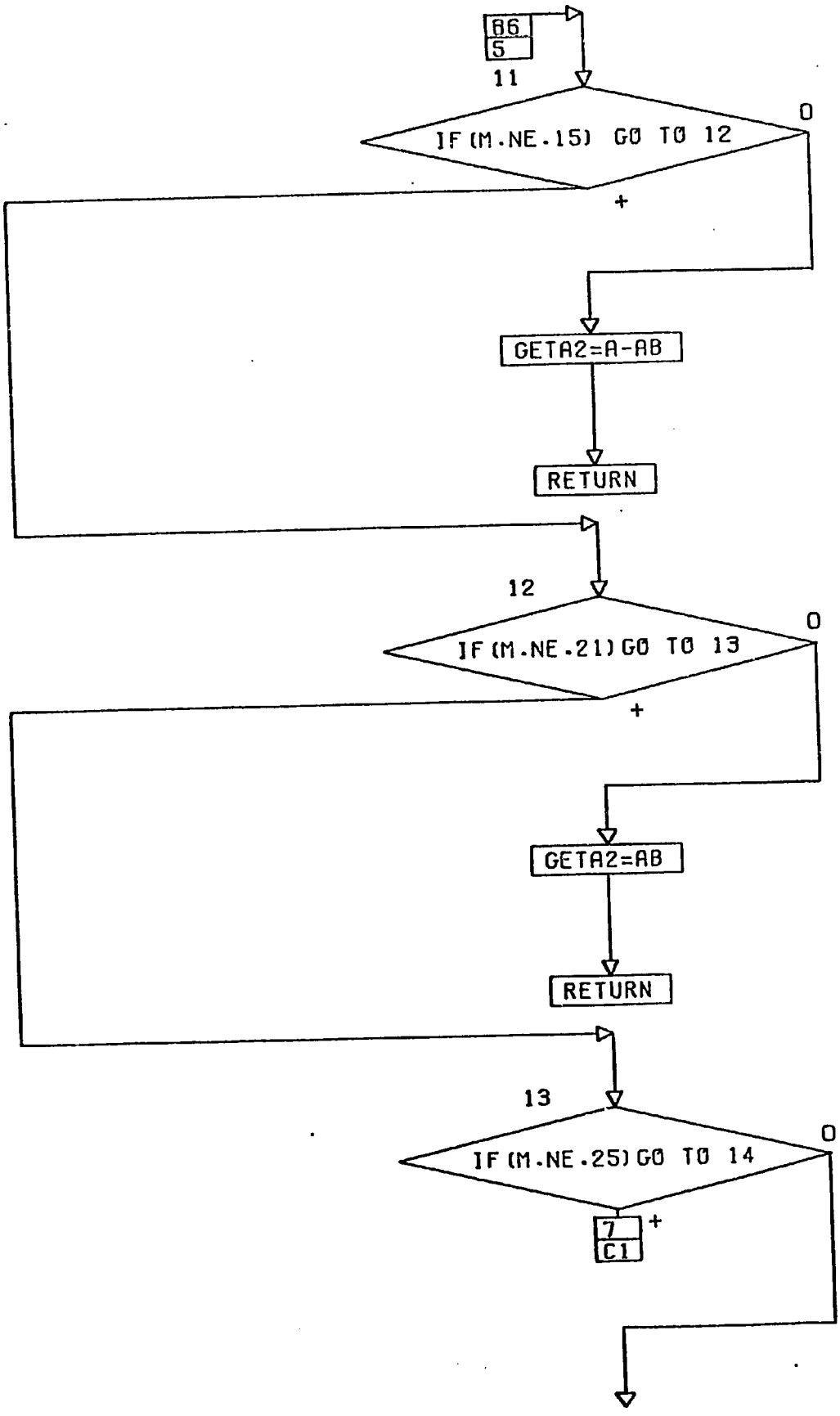
CONT. ON PG 4


A8
3
6
IF (M.NE.6) GO TO 7
0
+
GETA2=A-A2
RETURN
7
IF (M.NE.7) GO TO 8
0
+
GETA2=B
RETURN
8
IF (M.NE.9) GO TO 9
0
5
B3
+

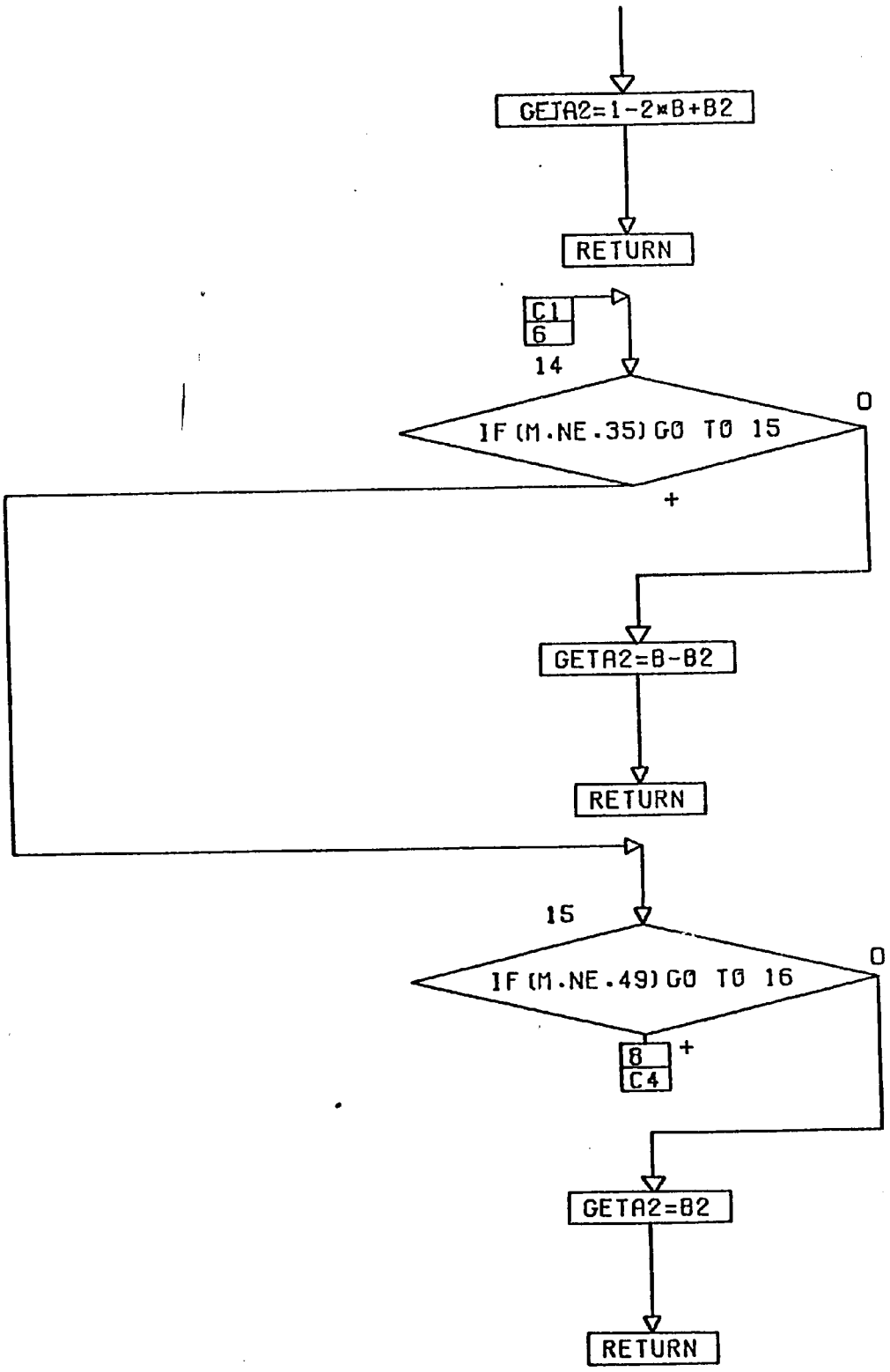
CONT. ON PG 5


GETA2=A2
RETURN
B3
4
9
IF (M.NE.10) GO TO 10
0
+
GETA2=1-A-B+AB
RETURN
10
IF (M.NE.14) GO TO 11
0
6
B6
+
GETA2=B-AB
RETURN

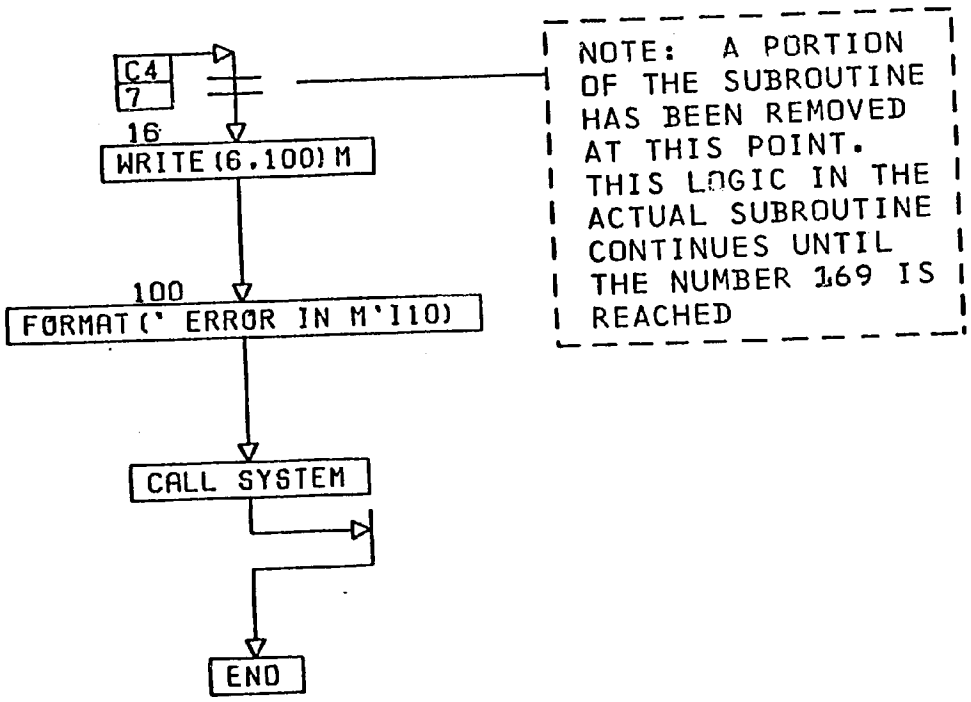
CONT. ON PG 6


A6
5
11
IF (M.NE.15) GO TO 12
0
+
GETA2=A-AB
RETURN
12
IF (M.NE.21) GO TO 13
0
+
GETA2=AB
RETURN
13
IF (M.NE.25) GO TO 14
0
7
C1
+

CONT. ON PG 7


GETA2=1-2*B+B2
RETURN
C1
6
14
IF (M.NE.35) GO TO 15
0
+
GETA2=B-B2
RETURN
15
IF (M.NE.49) GO TO 16
0
8
C4
+
GETA2=B2
RETURN

CONT. ON PG 8


C4
7
16
WRITE(6,100)M
100
FORMAT(' ERROR IN M'I10)
CALL SYSTEM
END
NOTE: A PORTION OF THE SUBROUTINE HAS BEEN REMOVED AT THIS POINT. THIS LOGIC IN THE ACTUAL SUBROUTINE CONTINUES UNTIL THE NUMBER 169 IS REACHED

REFERENCES

1. Abramowitz, M., and Stegun, L.A., ed., Handbook of Mathematical Functions, National Bureau of Standards, 1965.

2. Anderson, T.W., and Darling, D.A., "Asymptotic Theory of Certain Goodness of Fit Criteria Based on Stochastic Processes," Annals Math. Stat., v. 23 (1952), 193-212.

3. Anderson, T.W., and Darling, D.A., "A Test of Goodness of Fit," J. Amer. Stat. Assn., v. 49 (1954), 765-769.

4. Anderson, T.W., and Walker, A.M., "On the Asymptotic Distribution of the Autocorrelations of a Sample from a Linear Stochastic Process, Annals Math. Stat., v. 35. (1964), 1296-1303.

5. Bartlett, "Properties of Sufficiency and Statistical Tests," Proc. Roy. Soc. A, v. 160 (1937), 268-282.

6. Bordner, et. al., Emergency Care Systems Demonstration Projects, Franklin Institute Research Laboratories, Philadelphia, June, 1968.

7. Cinlar, Erhan, Analysis of Systems of Queues in Parallel, Doctoral Dissertation, University of Michigan, 1965.

8. Cinlar, Erhan, "Queues with Semi-Markov Arrivals," J. Applied Prob., v. 4 (1967), 365-379.

9. Cox, D.R., and Lewis, P.A.W., "Computer Program for Statistical Analysis of Series of Events, IBM Research Journal, no. 362, November 12, 1965.

10. Cox, D.R., and Lewis, P.A.W., The Statistical Analysis of Series of Events, London: Metheun, 1966.

11. Cox, D.R., and Smith, W.L., Queues, London: Metheun, 1961.

12. Dantzig, George B., Linear Programming and Extensions, Princeton: Princeton University Press, 1963.

13. Darling, D.A., "The Kolmogorov-Smirnoff, Cramer-von Mises Tests," Annals Math. Stat., v. 28 (1957), 823-838.

14. Durbin, J., "Some Methods of Constructing Exact Tests," Biometrika, v. 48 (1961), 41-55.

15. Forsythe, G. and Moler, C.B., Computer Solution of Linear Algebraic Systems, New York: Prentice-Hall, 1967.

16. Frey, Charles, A Retrospective Study of Traffic Fatalities in Washtenaw County, Michigan, University of Michigan Hospital, 1968, unpublished.

17. Jackson, O.A.Y., "Fitting a Gamma or Log-Normal Distribution to Fibre-diameter Measurements on Wool Tips," Applied Statistics, v. 18 (1969), 70-75.

18. Jewell, William, "Markov-Renewal Programming, Parts I and II," Opns. Res., v. 11 (1963), 938-971.

19. Lewis, P.A.W., "A Computer Program for the Statistical Analysis of Series of Events," IBM Systems Journal, vol. 5, no. 4, 1965, pp. 202-225.

20. Massey, F.J., "A Note on the Power of a Non-Parametric Test," Annals Math. Stat., v. 20 (1950), 440-442.

21. McLaren, Roy C., "Allocation and Distribution of Police Patrol Manpower," in Law Enforcement Science and Technology, ed. by S.A. Yefsky, London: Thompson, 1967.

22. Moore, E.H., and Pyke, R., Estimation of the Transition Distributions of a Markov Renewal Process, Mathematical Note No. 364, Boeing Scientific Research Laboratories, September, 1964.

23. Moran, P.A.P., "The Random Division of an Interval. Part II," J.R. Stat. Soc. B, v. 13 (1951), 147-150.

24. Munro, Donald, Survey of Urban Ambulance Administration, Highway Safety Research Institute, University of Michigan, 1969, unpublished.

25. Neuts, M.R., "The Single Server Queue with Poisson Input and Semi-Markov Service Times," J. Appl. Prob., v. 3, (1966), 202-230.

26. Parzen, Emanuel, Stochastic Processes, San Francisco: Holden-Day, 1962.

27. Pepler, R.D., et. al., Economics of Highway Emergency Ambulance Services, Vol. I and II, Dunlop and Associates, Inc., Darien, Connecticut, July, 1968.

28. Public Health Service, Checklists for Emergency Health Services Workshops, Department of Health, Education, and Welfare, Washington, D.C., undated.

29. Ross, Sid, and Rogers, J.G., "Our Ambulance Crisis, Parade, March 2, 1969, p. 10.

30. Savas, E.S., and Spiegal, M.L., Emergency Medical Service, City of New York, Office of Administration, Office of the Mayor, March 8, 1968.