THE UNIVERSITY OF MANITOBA

APPLYING A REGIONAL FLOW MODEL TO HEALTH CARE PLANNING USING UNIVERSAL HEALTH INSURANCE DATA

BY

FRANK FREDERICK ROSS MAYHOOD

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT ....COMPUTER SCIENCE...

WINNIPEG, MANITOBA

October 1977

APPLYING A REGIONAL FLOW MODEL TO HEALTH CARE PLANNING USING UNIVERSAL HEALTH INSURANCE DATA.

by

Frank F. R. Mayhood

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

Master of Science

ABSTRACT

A model of Manitoba's health care delivery system, based on Markov process theory, is discussed in the context of assessing the adequacy of care in various regions. Data processing problems encountered in the course of estimating the model from the universal health insurance data provided by the Manitoba Health Services Commission, are noted. The methods of estimation suggested by S.Zahl [Human Biology 1955 p90] were found to be inadequate for this data and alternatives were developed. The implications of the statistics collected, both for estimation of the model and derived from the model's properties, are explored, and the potential use of this class of model for health care planning is examined.

ACKNOWLEDGEMENTS

I wish to thank the Manitoba Health Services Commission for providing access to their files, and particular thanks to Dennis Roch of the Commission for his helpful suggestions, clarifications, and encouragement throughout the project. I am indebted to Dr. Jana Mossey of the Department of Social and Preventive Medicine and Dr. W. Hoskins of the Department of Computer Science for their perserverance in reading this work and for their helpful comments. Ms. Peg Burh was of great assistance in the preparation of the text. The financial support of the National Research Council through a postgraduate scholarship was much needed and appreciated. For the invaluable advice and council of my advisor, Dr. A. Neil Arnason, I am eternally grateful. Last, but far from least, I wish to acknowledge the efforts of my wife, Linda, whose patience, devotion, and plain hard work made the completion of this thesis possible.

A Regional Flow Model For Health Care Planning

TABLE OF CONTENTS

INDEX OF TABLES AND FIGURES

EQUATION DEFINITIONS

CHAPTER 1. ISSUES

The work presented in this thesis was motivated by three major concerns. The bulk of the theory of Markov processes fills many volumes, and there are oft repeated claims for their abilities as models in the health care field [Zahl 1955, Chiang 1973, Navarro 1970]. However, reports on large scale attempts to apply the theory are almost non-existent. The attempt to apply the Markov process model to a practical situation, solving along the way the many computational problems which arise in such applications, was the primary motivation for this thesis.

The second most important concern was an attempt to provide a workable and useful tool for health care planning in the context of universal health insurance. Direct costs to individuals on a fee for service basis are eliminated in areas where universal health insurance has been accepted. As a result, costs tend to be negotiated between the government regulatory agency and the operational health care facilities, rather than being determined by actual cost considerations. Many of the standard economic models become useless in such a context because they depend upon meaningful and flexible cost structures. In addition, the vast amount of data which is available as a result of the claims procedures generated by the insurance operation should provide the information to substantially improve the quality of modeling tools for health care planning. Many types of model which could not be estimated efficiently before universal health insurance, can now be explored with relative ease.

The third major area of interest was the attempt to provide insight into the functioning of the health care

delivery system in Manitoba, with particular interest in the regional variations in access to care. The data provided by the Manitoba Health Sevices Commission (M.H.S.C.) from the operation of the universal health insurance scheme in 1972, provides a unique opportunity to examine the variations in health care usage by a variety of factors. Regional variations are of particular interest because they may indicate the need for additional facilities or for different kinds of facilities in the various areas.

The remainder of this chapter is devoted to enlarging on some specific themes which are derived from one or more of the above concerns and which recur periodically throughout the rest of the dissertation. Chapter 2 contains a discussion of the relevant theory and introduces the notation commonly used for the class of models which we use in the study. In chapter 3, the computational problems of data reduction for estimation of the model and of applying the theory of chapter 2 are discussed. The strengths and weaknesses of the statistical results and their application to the above issues are discussed in chapter 4. The final chapter contains an examination of some possible uses of the model for planning the evolution of the health care system, or for monitoring its performance.

1.1 Quantity and Quality

The advent of universal health insurance under the control of a government monopoly has shifted the burden of planning from the various institutions and individuals involved in delivering the care, to the government who now pays for it. This shift has provided for a wider outlook, longer term goals, and a political character to decision making which all demand a more complete approach to understanding and to modeling the health care system.

Fortunately, the very nature of the insurance type of implementation has provided a rich source of raw data from which this better understanding could be built. The first step is to isolate those factors which are important ones in regards to the functioning of the system. For health care, there can be little doubt that the quality and quantity of care provided are important outputs of the system. These factors will intimately affect the overall health of the population served, which in turn will affect that population's productivity. These factors will also affect the public's perception of the performance of the health care system, and ultimately of the politicians who control it. Thus the ability to measure the quality and quantity of care delivered, particularly as it relates to the population's perception of its needs, is a desirable characteristic of a model of the health care system.

The overall health status of a population is difficult, if not impossible, to quantify. Various manifestation of ill health are measurable. Such measures as mortality rates, morbidty rates, and counts of services rendered give some indication of the underlying "healthiness" of the population, but these only scratch the surface of deviations from perfect health (whatever that may be). The shortcomings of these statistics or measures of overall "health" have been widely discussed [Goldsmith, 1972]. An alternative index suggested by C.L. Chiang and R. Cohen [1975] is a bi-product of the type of model we discuss in later chapters, and its performance in our situation was examined.

The quantity issue may also be viewed from the point of view of the various facilities which comprise the health care system. These characteristically require a long lead time for either expansion or contraction. Thus accurate forecasts of future requirements, particularly where the population is growing or changing rapidly, are necessary on a medium to long term basis. A model of the health care system should assist in the provision of such forecasts.

1.2 Equality of Access

One of the consequences of the control of health care planning by the political process is the interjection of a new set of goals and priorities. Normally, the health care system would be almost entirely concerned with the provision of adequate care. But because of the political process and our understanding of fair government practice, the issue of equality of access to health care becomes important. Each region in a province such as Manitoba must perceive itself to be receiving fair and adequate health care services, relative to other regions.

Measurement of the degree of access to care is almost impossible. However, manifestations of inequality of access can be anticipated and looked for in the data. Lack of adequate access to a particular type of facility in a particular region is likely to manifest itself in two ways, before it affects the overall health of that region's population. First, those in need of types of care which are inadequately supplied in a particular region can be expected to travel to other regions in search of care. Large numbers of people observed seeking care in other regions would be evidence of a lack of access to that type of care locally. If, in addition to large flows of people seeking care elsewhere, there are no appreciable flows of people into that region to receive care, then there is evidence of a difference in the level of access to health care among

regions. Second, those in need of a particular type of care would be expected to substitute other types of care which are more readily available in the region for those with more limited access in that region.

Both of these measurements depend upon the ability to compare regions on the basis of the health care delivered to those resident in that region, rather than just the health care delivered by the region's various health care facilities. By examining care regionally in this way, instead of the more usual facility "catchment area" approach, and by measuring and interpreting the inter-regional delivery of care, many advantages are realized. Regions may be precisely defined, based on administrative or geographical considerations, without conflict with modeling requirements. This makes the statistics produced more readily interpretable in terms of the physical system modeled, and allows for the treatment of issues such as equality of access as well.

1.3 Data Requirements and Computational Aspects

The types of problems which arise in using data captured for administrative purposes, such as the claim data from a universal health insurance scheme, for research purposes, are well known, if not understood, by those who call themselves Computer Scientists. Any use of data for a purpose other than the one for which it was collected and maintained is fraught with peril. In examining the M.H.S.C. files for the purposes of this study, many of these problems were encountered, some of which were circumvented successfully and some of which were only noted. All have had their impact on the utility of the results. If further attempts to use this type of data in a research context are made, some valuable lessons may be taken from the struggles presented in chapter 3. Most of the examples are particular to the form of the data maintained by the M.H.S.C., but some should be useful whenever the use of universal health

insurance data is contemplated.

The computational problems associated with applying the theory of Markov processes were found to be substantial. Some of the theoretical formulae were found to be of little computational use in their published forms. The methods described for computing the various estimates of parameters and derived statistics should be of use to anyone who contemplates applying this class of model, whatever the application.

Chapter 2. A Markov Flow Model

2.1 Theoretical Groundwork

The mathematical objects named for A.A. Markov (1856-1922) have received a great deal of attention from mathematicians in the last 50 years. The resulting theory has filled many books (Cox & Miller, Bharucha-Reid, Chiang) some of which are listed in the bibliography. The volume of results makes it impractical to present even the barest outline of them here, but a basic understanding of the concepts is necessary to the discussion which follows, and a brief theoretical discussion is a convenient way to introduce the notational conventions which should facilitate the presention of our conclusions. The attempt is to communicate the essence of the theory which forms a framework for the study, and refer the reader to the appropriate references for a more rigorous approach.

Markov processes are mathematical constructs which arise in the context of the study of objects or systems which change from state to state. As mathematical entities, they may be axiomatized and their properties developed analytically from those axioms. To apply a Markov process to a real world situation is to assume that the axioms are true and use the derived properties to describe, forecast, or analyze the real world. Thus, in an application of the theory of Markov processes to the study of the health care delivery system in Manitoba, one must expend a fair amount of effort in determining the extent to which the mathematical axioms can be said to hold, and the robustness of the derived properties to violations of these axioms. The first step in this process is to identify the set of axioms which are assumed to be true and phrase them in the terms of the real world situation. That task is attempted in the remainder of

this chapter.

2.1.1 Basic Definitions

In a modeling effort, the mathematical axioms of the model are used to simplify the complexity of a real world system into a form which yields to mathematical analysis. The results of the analysis can then be used to draw hypotheses about the performance of the real world system. A Markov process simplifies in two primary ways. Firstly, it assumes that there exists a set of mutually exclusive categories or "states" such that at any instant, the system can be said to be <u>in</u> a particular state. The state of a system may be one dimensional (like the number of people in a queue) or multi-dimensional (like the number of people alive and in a particular age cohort for each of a set of cohorts).

In mathematical terms there must exist a set X of states, a time space (partially ordered) set $\mathcal{T}$ and a function:

$$X : \mathcal{T} \rightarrow \mathcal{X} \qquad \text{(E2.1.1)}$$

The function X(t) gives the state of the system at each point ,t, in the time space, $\mathcal{T}$. This classification function reduces the real world system to movements between states in a well defined set of states, X. But this is still far too general to yield much useful from analysis. The nature of these movements from state to state must be controlled. A Markov process controls the mechanism of movements by assuming that the classification into states is a very special one in that knowledge of the state of the system at a particular time t^0 completely summarizes the

past history of the system at all times $t < t^0$ as far as the future evolution of the system is concerned. More precisely:

$$Pr(X(t_1) = x \mid X(t) \quad t < t_0) = Pr(X(t_1) = x \mid X(t_0)) \qquad \text{(E2.1.2)}$$

The conditional probability that the system is in a particular state x at some later time t^1 given the entire history of the system prior to time t^0 is equal to the conditional probability that the system will be in state x at t^1 given only the state $X(t^0)$ which it was in at time t^0. This "memorylessness" of the transitions between states of a Markov process is the cornerstone of the theory which permits statements to be made about the future evolution of the system.

Markov processes can be classified according to the cardinality of their state and time spaces X and $\mathcal{T}$. If X and $\mathcal{T}$ are both countable, that is they can both be considered as equivalent to subsets of the integers, then the process is called a Markov Chain. These will be useful later [2.1.6] as approximations to other, more complex, processes. The type of process which will be of primary interest in what follows has a finite state space X, and a continuous time space. In this case, one usually replaces by the set of integers {1,2,...NS} where NS is the number of elements in the original state space. The conditional probability of the system being in state j ($1 \leq j \leq NS$) at time t^1 given that it was in state i ($1 \leq i \leq NS$) at time $t^0 \leq t^1$ is written $P_{ij}(t^1, t^0)$:

$$P_{ij}(t_1,t_0) = Pr(X(t_1)=j \mid X(t_0)=i) \qquad \text{(E2.1.3)}$$

Such a process is said to be <u>stationary</u> or <u>time homogeneous</u>

if these transition probabilities are functions of $\Delta t = t^1 - t^0$ alone:

$$p'_{ij}(\Delta t) = p_{ij}(t_1, t_0) \qquad \text{(E2.1.4)}$$

This means that the probability of being in j at time t^1 given that the system was in state i at time t^0 does not depend on the absolute location along the time axis of t^0. Associated with the transition probabilities $p_{ij}(t^1, t^0)$ are the instantaneous rate functions $\gamma_{ij}(t)$ defined by:

$$\lim_{\Delta t \to 0} \frac{p_{ij}(t+\Delta t, t)}{\Delta t} = \gamma_{ij}(t) \qquad i \neq j \qquad \text{(E2.1.5)}$$

which implies:

$$p_{ij}(t+\Delta t, t) = \gamma_{ij}(t)\ \Delta t + o(\Delta t) \qquad i \neq j \qquad \text{(E2.1.6)}$$

where : $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0.$

It is logical to require :

$$\lim_{\Delta t \to 0} p_{ij}(t+\Delta t, t) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \qquad \text{(E2.1.7)}$$

thus (E2.1.5) cannot be used to define $\gamma_{ii}(t)$. Instead using (E2.1.6) as a model, define:

$$-\gamma_{ii} = \gamma_i = \sum_{j=1}^{NS} \gamma_{ij} \qquad \text{(E2.1.8)}$$

then:

$$p_{ii}(t+\Delta t, t) = (1 + \gamma_{ii}(t))\ \Delta t + o(\Delta t) \qquad \text{(E2.1.9)}$$

since the system must be in one of the states at each instant. For convenience, define the transition probability

matrix $P(t^1,t^0)$ and the instantaneous rate matrix $\Gamma(t)$ as:

$$P(t_1,t_0) = \left(P_{ij}(t_1,t_0)\right)$$
$$\Gamma(t) = \left(\gamma_{ij}(t)\right) \qquad \text{(E2.1.10)}$$

If the process is stationary (E2.1.4) then $P(t^1,t^0)=P(t)$ and $\Gamma(t)=\Gamma$ is independent of time.

For the processes which will interest us, it is possible to view the system in two ways, each with it's own insights into the real system. These models involve the movements of individuals or particles about a finite set of states. The model of an individual's movements has a one dimensional state space and X(t) gives the state that that individual or particle occupied at time t. One may also be interested in the system as it describes a set of individuals moving about simultaneously in the same state space. This process is multidimensional, its classification function X(t) gives the number of individuals in each state at time t. It is a vector of population levels of dimension NS, the number of states in the one dimensional model. If the individuals or particles move independently of one another then the $P(t^1, t^0)$ and $\Gamma(t)$ matrices will be related in a very simple way, the $P(t^1,t^0)$ for the multidimensional process being the convolution of $P(t^1,t^0)$ for the one dimensional process, and the rate matrices will be essentially related in the same way. Knowledge of either will determine the properties of the other.

This concludes the definition of a Markov process and the notation particular to the class of processes with which we will be concerned. For further details see the references [eg. 6,7] in the bibliography. The next sections present the useful properties that can be derived from these definitions for this class of finite discrete state

space, continuous time Markov processes.

2.1.2 Properties: The Chapman-Kolmorgorov Equation

A fundamental property of a Markov process is that the state space exhausts the possible configurations of the system. Thus, in order that the system be in state j at some future time t^1, it must have been in some other state at each intermediate time s. The Chapman-Kolmorgorov equation formalizes this property as:

$$p_{ij}(t_1,t_0) = \sum_{k=1}^{NS} p_{ik}(s,t_0)\, p_{kj}(t_1,s) \qquad t_0 < s < t_1 \qquad \text{(E2.1.11)}$$

If the process is stationary this can be written in matrix form as:

$$P(t + t_0) = P(\,t\,)\, P(\,t_0\,) \qquad \text{(E2.1.12)}$$

which relates the theory of Markov processes and the theory of semigroups of operators. More importantly, when compared to $a^{x+y} = a^x a^y$, (E2.1.12) strongly hints at the functional form of P(t) in the time homogeneous case.

2.1.3 Properties: The Kolmorgorov Differential Equations

Using the definitions of the instantaneous rate functions [E2.1.6, E2.1.9] we may rewrite E2.1.11 as:

$$\begin{aligned} p_{ij}(t + \Delta t,t_0) &= \sum_{k=1}^{NS} p_{ik}(t,t_0)p_{kj}(t + \Delta t,t) \\ &= p_{ij}(t,t_0)\,[\,(1 + \gamma_{jj}(t))\,\Delta t + o(\Delta t)\,] \\ &\quad + \sum_{k\neq i}^{NS} p_{ik}(t,t_0)\,[\,\gamma_{kj}(t)\Delta t + o(\Delta t)\,] \end{aligned} \qquad \text{(E2.1.13)}$$

The first term is the probability of being in state i at t^o and getting to state j by time t, then remaining in state j for at least t time units. The remaining terms are the respective probabilities of transferring into state j from some other state, given that at time t^o the system was in

state i, in the interval $[t, t+\Delta t]$. Subtracting $P_{ij}(t+\Delta t, t^0)$ from both sides, dividing by t and taking the limit as $\Delta t \to 0$ yields the differential equations:

$$\frac{\partial p_{ij}(t,t_0)}{\partial t} = \sum_{k=1}^{NS} \gamma_{kj}(t)\, p_{ik}(t,t_0) \qquad \text{(E2.1.14a)}$$

In matrix form for fixed t^0:

$$\frac{\partial P(t)}{\partial t} = P(t)\, \Gamma(t) \qquad \text{with } P(t_0,t_0) = I \qquad \text{(E2.1.14b)}$$

This system of homogeneous first order differential equations are known as the Kolmorgorov forward differential equations of the Markov process. Unfortunately for the case of general time dependent rates ($\gamma_{ij}(t)$), the solution to these equations may not exist. However, for the stationary case, (ie with constant, time independent rates γ_{ij}) the solutions to E2.1.14 can be shown to exist and to be unique under quite general circumstances [Feller, 1940]. Under these conditions the solutions may be written in the form:

$$P(t) = e^{\Gamma t} = \sum_{n=0}^{\infty} \frac{(\Gamma t)^n}{n!} \qquad \text{(E2.1.15)}$$

This equation will be of little direct use as a method for computing the solution [see 3.3.3] because of the alternating nature of the infinite series and the roundoff problems this creates. It does represent a complete solution to the mathematical problem presented by the model.

2.1.4 Properties: Length of Stay

Each time a system enters a particular state, it remains in that state for some interval of time. This "length of stay" is a random variable whose expected value will represent the average amount time spent in a state i before leaving it for some other state. It is natural to expect

that this expected value will be closely related to the transition rates from state i. For a stationary Markov process this relationship is particularly simple. Letting L_i be the length of stay in state i we write, following Arnason[1975]:

$$G_i(T) = Pr(L_i \leq T \mid \text{currently in state i})$$

then:

$$G_i(T + \Delta t) = G_i(T)\ Pr(\text{ no transitions in } [T, T + \Delta t])$$

$$= G_i(T)\ [\ (1 + \gamma_{ii})\ \Delta t + o(\Delta t)]$$

therefore:

$$\frac{\partial G_i(t)}{\partial t} = G_i(t)\ \gamma_{ii}$$

implying:

$$G_i(t) = e^{\gamma_{ii} t} \qquad \text{(E2.1.16)}$$

Since 1-G(t) is the cumulative distribution function of L_i, this implies that L_i is exponentially distributed with parameter $-\gamma_{ii}$. The mean of the exponential distribution is $\mu_i = 1/\gamma_i$ which simply relates the average length of stay to the instantaneous transition rates.

The variance of the length of stay in state i is $1/\gamma_{ii}^2$ by the properties of the exponential distribution. This implies that the standard deviation ($\sqrt{\text{variance}}$) is equal to the mean, a property which will be of use later when we examine the model assumptions.

Related to length of stay are the probabilities that the transition which occurs is to state j given that a transition has occurred and that the system was in state i. Again following Arnason[1975], we have:

$$\Pr(\text{transition is from } i \text{ to } j \mid \text{transition occurs from } i \text{ at } t)$$

$$= \lim_{\Delta t \to 0} \frac{\Pr(i \to j \text{ and a transition occurs from } i \text{ in } [t, t+\Delta t])}{\Pr(\text{transition from state } i \text{ occurs in } [t, t+\Delta t])}$$

$$= \lim_{\Delta t \to 0} \frac{\gamma_{ij}\,\Delta t + o(\Delta t)}{-\gamma_{ii}\,\Delta t + (\Delta t)}$$

$$= \gamma_{ij} / \gamma_i \qquad \text{(E2.1.17)}$$

Thus the state to which the transition occurs, given the time the system leaves state i, is multinomial with parameter vector $P_j = \gamma_{ij} / \gamma_i$ $(i \neq j)$. These two properties permit the simulation of Markov processes in a very simple and efficient manner. One need only know the state of the system at present to generate the time of the next transition, and the type of transition may be determined afterwards.

2.1.5 Properties: Residency Times

Related to the length of stay is another set of interesting measures of a system's behavior: the residency times. The amount of time that the system spends in a state j in an interval of observation [0,T], given that at time 0 it was in state i, is a random variable called the residency time, $r_{ij}(T)$. The expected residency $e_{ij}(T) = E(r_{ij}(T))$ is the average amount of time the system would be expected to be in state j, given that it was initially in state i, in an interval of length T (for the time homogeneous case). Expected residency is related to the transition probabilities P(t) by:

$$e_{ij}(T) = \int_0^T p_{ij}(t)\, dt \qquad \text{(E2.1.18)}$$

Chiang [1973] demonstrates the validity of (E2.1.18) by defining indicator functions:

$$I_{ij}(t) = \begin{cases} 1 \text{ if is in state j at time t and initially in state i} \\ 0 \text{ otherwise} \end{cases}$$

The expected values of these indicator functions are simply the $p^{*}_{ij}(t)$, the probability of the system being in state j, given it was initially in state i. Once it is in state j it will remain for a time instant $[t,t+\Delta t]$. The expectation of the system being in state j for $[t,t+\Delta t]$ is:

$$E(I_{ij}(t))\,\Delta t = p_{ij}(t)\,\Delta t \qquad \text{(E2.1.19)}$$

Summing over all time instants gives E2.1.18.

If we substitute E2.1.15 into E2.1.18 and perform the integration we obtain an infinite series for the expected residency:

$$E(T) = \int_0^T P(t)\,dt = IT + \frac{\Gamma T^2}{2!} + \frac{\Gamma^2 T^3}{3!} + \cdots \qquad \text{(E2.1.20)}$$

Again, this formula will be of little direct use in computing expected residency because of the major computational problems involved.

2.1.6 The Imbedded Chain

Figure 3.2.1 illustrates a possible trajectory (sequence of state changes) for a Markov process in continuous time with discrete states . Notice that because of assumption one, transitions are instantaneous. The system must be in one and only one state for each instant of time. Suppose now that instead of looking at the system as it evolves in continuous time we observe its state only at evenly spaced times $t^{i}=t^{0}+i\ t: i=0,1,\ldots,n$. Then the one step transition

probabilities are given by:

$$p_{ij}(1) = p_{ij}(t_k - t_{k-1}) = p_{ij}(t) = p^*_{ij} \quad \text{independent of k.}$$

Moreover the Markov property is inherited from the continuous time process. Thus these observations correspond to a Markov chain defined on the same state space. This chain is said to be <u>imbedded</u> in the continuous time process. The existence of the imbedded chain permits us to borrow results for Markov chains and use them to describe the more complex continuous process. In particular, we may classify the state in the customary way for Markov chains. A state may be <u>absorbing</u> if the probability of ever leaving the state is zero ($p^*_{ii}=1$), a member of a set of <u>ergodic</u> states if the limit as $T\to\infty$ of the probability of being in that state is nonzero, or <u>transient</u> if this limiting probability is zero.

2.2 The Health Care Flow Model

2.2.1 The State Space

Attempting to define a state space for a Markov model of the health care delivery system in Manitoba is complicated by conflicting goals. If the model is to be useful from the point of view of management of the health care services, the state space must accurately reflect the nature of the demands placed on the real system. The individual or the social scientist is more interestd in the way the system interacts with the individual, and would prefer a state space which was easily interpretable in individual terms. But the Markov assumptions require that the state space have the property of summarizing the characteristics of the system's past.

Also involved in the selection of an appropriate state space are the computational and data requirements involved in estimating and using the model. The number of parameters

to be estimated is of the order of the square of the number of states. Thus the data and computational requirements suggest that the number of states be kept small. Interpretation of the model will also be simpler if the number of parameters is kept reasonably small.

The compromise selected for the purpose of this study does not meet any of the above criteria very well as we shall see, but such is the nature of compromise. The purpose of the study was to explore the potential usefulness of the Markov model as a measure of the health care delivery system's performance. To assess this potential, some connection with existing techniques would be desirable . Thus, the selection of health care states was very much determined by the existing ways of measuring their utilization. Health care delivery was divided on the basis of point of delivery or facility used, rather than by disease treated. For reasons of data availability and compatibility, not all health "facilities" available could be treated, and others had to be considered as one facility even though quite separate. The final list of facilities selected for the study appears in table 2.2.1.

In order to attack the problem of measuring equality of access to health care services the facilities were disaggregated by region. Three regions were chosen as representative of the potential of the model as a tool for assessing equality of access. These were: Brandon, defined as the municipality of Cornwallis and the town of Brandon; Winnipeg, defined as metropolitan Winnipeg (municipality codes 301-314); and the rest of Manitoba defined as everything not in Winnipeg or Brandon. A fourth region, Out of Province, was added in order to satisfy the Markov requirement that the state space be exhaustive. The cross product of these regions with the facilities of table 2.2.1 gives a state

Table 2.2.1 Facility Definitions

No.	LABEL	MEANING
1	Dr.'s Care	at least one visit to a doctor that day
2	Acute Hospital	in an acute care hospital bed
3	Extended Hosp.	in an extended care hospital that day
4	Out-Patient	attends an outpatient clinic at least once
5	Healthy	did not recieve health care and was alive
6	Dead	had died on or prior to this day

Table 2.2.2 Final State Definitions

State	Area	Facility
1	Winnipeg	Dr.'s Care
2	Winnipeg	Acute Care Hospital
3	Winnipeg	Extended Hospital
4	Winnipeg	Out-Patient Care
5	Winnipeg	Healthy
6	Winnipeg	Dead
7	Brandon	Dr.'s Care
8	Brandon	Acute Care Hospital
9	Brandon	Extended Hospital
10	Brandon	Out-Patient Care
11	Brandon	Healthy
12	Brandon	Dead
13	Rest Man.	Dr.'s Care
14	Rest Man.	Acute Care Hospital
15	Rest Man.	Extended Hospital
16	Rest Man.	Out-Patient Care
17	Rest Man.	Healthy
18	Rest Man.	Dead
19	Out Prov.	Dr.'s Care
20	Out Prov.	Acute Care Hospital
21	Out Prov.	Extended Hospital
22	Out Prov.	Out-Patient Care
23	Out Prov.	Healthy
24	Out Prov.	Dead

space with 24 possible states. These states are listed with these corresponding numbers and the labels used in the computation process in table 2.2.2. In reality only 22 of these states will be meaningful. Deaths which occur after people have moved out of province are not reported, nor is any distinction made between acute and extended hospital care for those services rendered out of the province. The states : (24) (Out of Province, Dead) and (22) (Out of Province, Extended Care Hospital) are therefore redundant.

2.2.2 Time Measurement and Observed Population

Transactions are reported to the M.H.S.C. with a date of service. Thus the unit of time chosen for our study was one day, the maximum time resolution of the data. This lack of resolution will be shown to be the cause of some of the computational difficulties encountered [3.4].

The population of interest in the study would ideally have been all those who received health care in Manitoba in 1972. Data problems and the homogeneity restrictions of the model prevent this, however. The actual population used were all those alive, living in Manitoba, and registered with the M.H.S.C on January 1,1972, with the single exception of treaty Indians. Indians were excluded because of their very different health care needs. The details of the selection and processing of individuals for use in the study is summarized in section 3.2.

The health care behavior of individuals is known to be a function of their age and sex. In order to attempt to obtain relatively homogeneous groups with respect to their health needs, the population which survived the selection process described in 3.2 was divided into ten age/sex cohorts. These are defined in table 2.2.3. It is probable that there are substantial differences in behavior within these cohorts but the data requirements again require that

the number of such groups be kept small. These ten groups were thought to be adequate to capture the majority of the age/sex related differences in health care behavior.

Table 2.2.3 Age and Sex Cohort Definitions

Group	Age	Sex
1	1	male
2	1	female
3	2-15	male
4	2-15	female
5	16-45	male
6	16-45	female
7	46-65	male
8	46-65	female
9	> 65	male
10	> 65	female

2.2.3 Examining the Assumptions

In light of the above definition of the state space, let us examine the Markov assumptions to determine whether or not they can reasonably be expected to hold a priori. Consider an individual who goes to a Winnipeg doctor on a particular day. The Markov assumption states that this individual's probabilities of being in any of the 24 states on the following day are independent of his/her past history, in particular independent of the area in which he/she was residing prior to his/her doctor's visit. This cannot be true! An individual who comes to the doctor in Winnpeg from Brandon is much more likely to return to Brandon (eventually) than the native of Winnipeg who goes to see his/her local doctor. Thus to have any chance that the Markov assumptions will hold, the model must be dissaggregated by region as well as by age and sex.

Another source of potential a proiri violations of the model assumptions is the tendency for health care demand by an individual to be episodic in nature, usually generating a burst of health care activity. Thus one would expect that the number of doctor's visits an individual has had in the past would seriously affect the probability of the individu-

al having another doctor's visit in the immediate future. However, one can anticipate that this "bursty" nature of an individual's requirements for care will be smoothed out by averaging over a large number of similar individuals in much the same way as demand for telephone circuits is smoothed. Much success has been had by applying Markov models to telephone switching problems [see for example Kleinrock 1971] despite the "bursty" nature of individual demand. The episodic nature of an individual's health requirements does not seem to be a serious barrier to the application of the theory, though we must be prepared to assess the nature of the impact on the conclusions drawn from the deviations from theoretical predicted behavior.

Chapter 3 The Computational Problems

3.1 Introduction

The application of the Markov process model introduced in section [2.2] to the study of the health care delivery system in Manitoba requires that some substantive computational problems be overcome. First, there is the problem of obtaining estimates for the parameters of the model, either the p^*_{ij} of the imbedded chain, or the rate matrix, Γ. Since there is a large number of parameters (24^2=576 per population group) a large amount of data must be available, or be collected, and then reduced to the parameter estimates. Fortunately, some of these parameters are known a proiri, and others are determined by known functions of other parameters, but there are still 342 independent parameters to be estimated for each population group of interest. The 342 can be derived by excluding from the 24^2=576 matrix elements all the diagonal elements, which are determined as either the negative of the sum of each row (for Γ) or by one minus the sum of each row (for p^*_{ij}), and all of the row elements for each of the 4 absorbing states and the redundant state "extended care hospital, out of province" for which the rates and probabilities are known by definition.

In order to estimate such a large number of parameters (over 13,000 in all) a massive amount of data must be processed, edited, verified and reduced to a form which permits the estimates to be calculated and displayed quickly and easily. This data processing problem of assembling and reducing the data to an appropriate form is the subject of section [3.3]. The computation of the derived estimates of the interesting properties of the model is discussed in section [3.4], and the drawing of statistical inferences

concerning the model parameters and assumptions is outlined in section [3.5]. Before we can consider these questions we must first determine what data is needed for the estimated process. It is to this that we now address ourselves.

3.2 Estimation: The Sufficient Statistics Derived

3.2.1 The Rate Matrix: Continuous Time Estimates

Our objective is to estimate the rate matrix Γ of an assumed time homogeneous Markov process by observing the trajectories of individuals over a fixed time period [0,T] (one year= 366 days in this case). Figure 3.2.1 is a graph of one such trajectory. The method of maximum likelihood is to be used as the criterion for selecting an appropriate $\hat{\Gamma}$.

For each individual under observation during [0,T] we observe a trajectory of the form illustrated in figure 3.2.1. The notation used means that the individual spent r^0 units of time in state s^0, then had a transition to state s^1, spent r^1 units of time in s^1 before transferring to s^2 and so on until time T at which time the individual had been in state s^n for δ time units. The number of transitions in [0,T] is, of course, a random variable, but we may still write the likelihood function for this trajectory as

$$L(\Gamma) = \Pr(s^0)\ [\Pr(r^0|s^0)\ \Pr(s^0 \to s^1)][\Pr(r^1|s^1)\ \Pr(s^1 \to s^2)]\ldots$$
$$[\Pr(r^{n-1}|s^{n-1})\ \Pr(s^{n-1} \to s^n)]\ \Pr(r^n > \delta | s^n) \qquad \text{(E3.2.1)}$$

If we take the initial distribution as fixed : $Pr(s^0)=1$,

$$L(\Gamma) = 1\cdot[\partial F_{s^0}(r^0;\gamma_{s^0})\ \gamma_{s^0 s^1}/\ \gamma_{s^0}]\cdot[\partial F_{s^1}(r^1;\gamma_{s^1})\ \gamma_{s^1 s^2}/\ \gamma_{s^1}]\cdots [\partial F_{s^{n-1}}(r^{n-1};\ _{s^{n-1}})\ \gamma_{s^{n-1}s^n}\ /\ \gamma_{s^{n-1}}]\ [1-\partial F_{s^n}(\delta;\gamma_{s^n})\ ] \quad \text{(E3.2.2)}$$

The r^i are simply the actual length of stay in s^i and are thus exponentially distributed [2.1.4], thus

$$L(\Gamma) = [\prod_{i=1}^{n-1} \exp(-\gamma_{s^i} r^i)\ \gamma_{s^i s^{i+1}}]\ [\exp(-\gamma_{s^n}\ \delta)] \quad \text{(E3.2.3)}$$

Taking logs and regrouping by states gives

$$\ln(L(\Gamma)) = \sum_{i=1}^{NS}\sum_{j\neq i}^{NS} a_{ij}\ln(\gamma_{ij}) - \sum_{i=1}^{NS}\gamma_i d_i - \gamma_{s^n}\ \delta \quad \text{(E3.2.4)}$$

where: a_{ij} = total number of (absolute) transitions from state i to state j in $[0,T]$.

d_i = total time units spent in state i in $[0,T-\delta]$

Observing N independent individuals in this manner permits us to add the log likelihoods to obtain

$$\ln(L(\Gamma)) = \sum_{i=1}^{NS}\sum_{j\neq i}^{NS} a_{ij}\ln(\gamma_{ij}) - \sum_{i=1}^{NS}\gamma_i d_i - \sum_{i=1}^{NS}\gamma_i\phi_i \quad \text{(E3.2.5)}$$

where: ϕ_i = total time spent in state i as a final state.
and a_{ij} and d_i are as above but pooled over all individuals.

Remembering that

$$\gamma_i = \sum_{i=1}^{NS} \gamma_{ij}$$

[E2.1.8] and defining $D_i = d_i + \phi_i$ = total time units spent in state i in [0,T] E3.2.5 becomes:

$$\ln(L(\Gamma)) = \sum_{i=1}^{NS} \sum_{j \neq i}^{NS} a_{ij} \ln(\gamma_{ij}) - \sum_{i=1}^{NS} D_i \sum_{j \neq i}^{NS} \gamma_{ij} \qquad \text{(E3.2.6)}$$

Differentiation gives:

$$\frac{\partial \ln(L(\Gamma))}{\partial \gamma_{ij}} = \frac{a_{ij}}{\gamma_{ij}} - D_i \qquad \text{(E3.2.7)}$$

Therefore the maximum likelihood estimates are

$$\hat{\gamma}_{ij} = a_{ij} / D_i \qquad i \neq j \qquad \text{(E3.2.8)}$$

FIGURE 3.3.1 A Trajectory : Continuous Time Observations

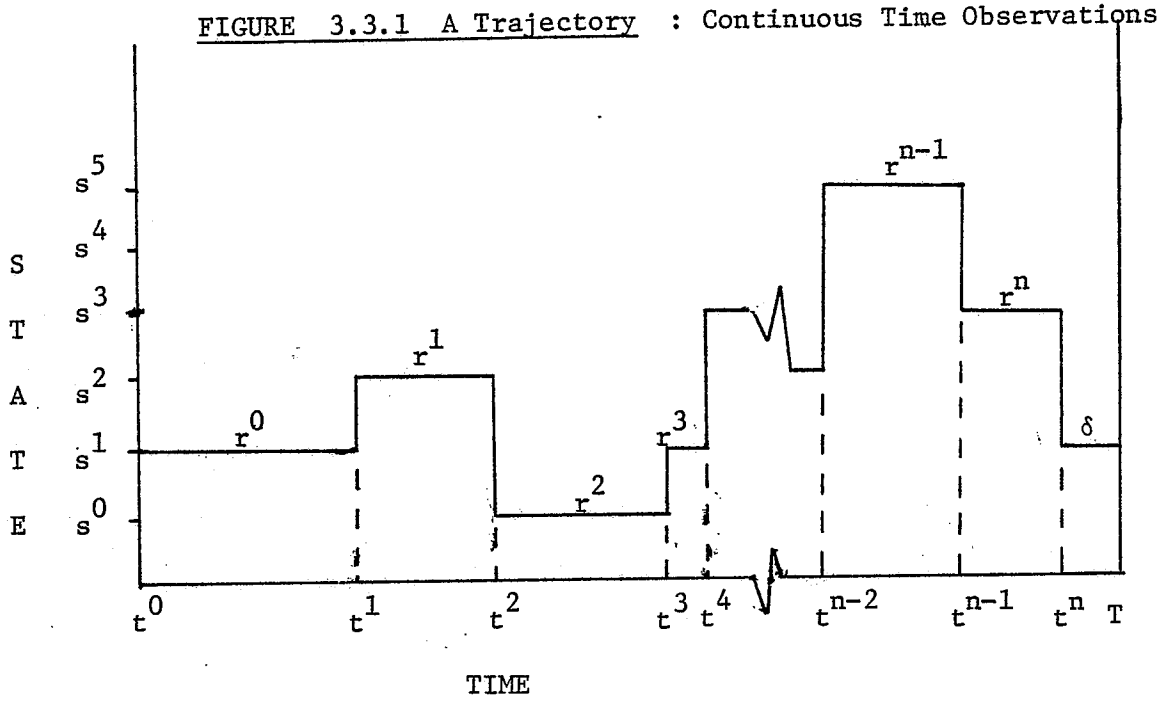
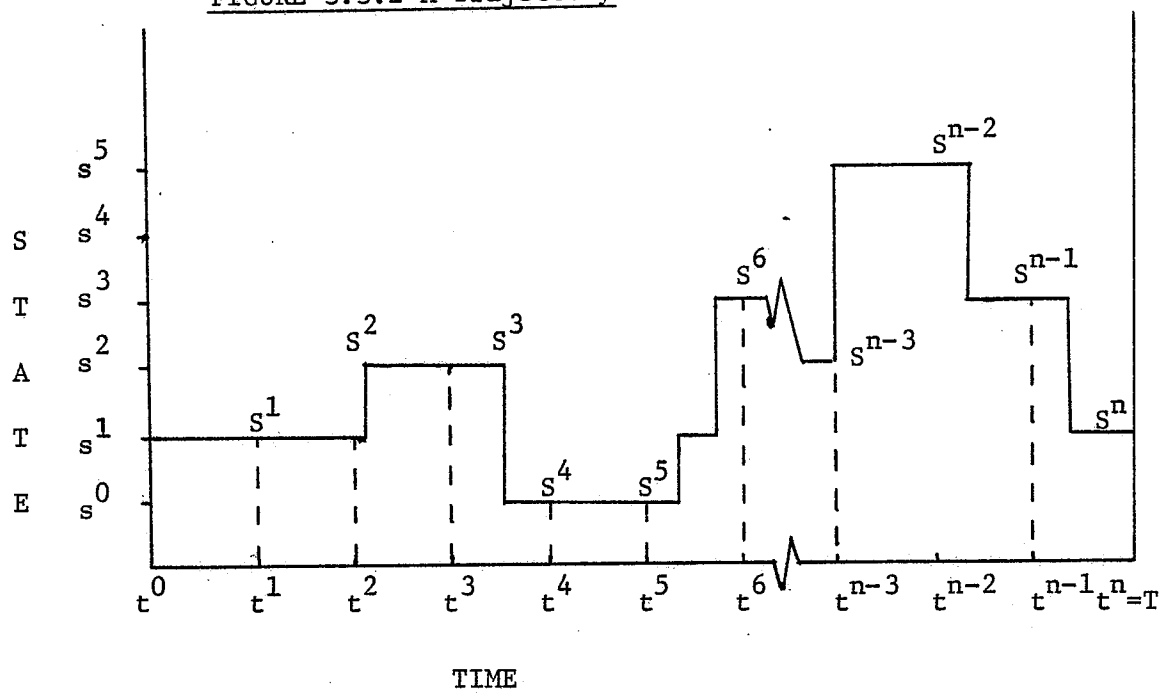


FIGURE 3.3.2 A Trajectory : Discrete Time Observations



The large sample (asymtotic) variance of these estimates can be obtained [Wilkes Ch 13.] from the results for maximum likelihood estimation as:

$$-\left\{E\left(\frac{\partial^2 \ln(L(\Gamma))}{\partial\gamma_{ij}\,\partial\gamma_{hk}}\right)\right\}^{-1} = \begin{Bmatrix} E(a_{ij})/\gamma_{ij}^2 & & o \\ & \ddots & \\ o & & \ddots \end{Bmatrix}$$

$$\begin{aligned} \text{var}(\hat{\gamma}_{ij}) &= \gamma_{ij}^2/E(a_{ij}) \\ \hat{\text{var}}(\hat{\gamma}_{ij}) &= \hat{\gamma}_{ij}/\,D_i \end{aligned} \qquad \text{(E3.2.9)}$$

The sufficient statistics for estimating Γ by the method of maximum likelihood are the absolute transition matrix $A=(a_{ij})$ and the total sojourn time vector D_i.

A more rigorous approach to this topic can be found in Billingsley [1961]. Further results were obtained by Doob [1953], Bartlett [1949] and Cuthbert [1973]. The results in these papers connect the continuous estimates derived above to the imbedded chain estimates discussed below.

3.2.2 The Rate Matrix: Discrete Time Estimates.

Zahl [1955] first presented the material summarized below as part of a paper advocating Markov process models for medical follow-up studies. In this approach, the observations are taken on the imbedded chain, at fixed equally spaced times t' in [0,T]. Observations tell only the state occupied at each of the time instants $t^i = it$, $i=0,1,\ldots,n$; that is, they consist of the (t^i, S^i) pairs of figure 3.2.2.

For such a path, the likelihood function can be written in terms of the transition probabilities $p_{ij}^* = p_{ij}(t')$ as:

$$L(P^*) = Pr(s^0)\, p^*_{s^0 s^1} p^*_{s^1 s^2} \cdots p^*_{s^{n-1} s^n} \qquad \text{(E3.2.10)}$$

where now n (the number of transitions) is fixed for all individuals. Neglecting the initial distribution as before [3.2.1] the likelihood of N such independent paths is:

$$L(P^*) = \prod_{i,j=1}^{NS} (p^*_{ij})^{n_{ij}} \qquad \text{(E3.2.11)}$$

where: n_{ij} = the total number of times over all individuals and all transitions in [0,T] that $S^k = i$ and $S^{k+1} = j$. Taking logs and differentiating yields the maximum likelihood estimates:

$$\hat{p}^*_{ij} = n_{ij} \Big/ \sum_{k=1}^{NS} n_{ik} \qquad \text{(E3.2.12)}$$

To convert these estimates of the transition probabilities of the imbedded chain to the estimates of the rate matrix of the overall process we note that the one dimensional analogue of E2.1.15 can be inverted by using the logorithm function. The power series expansion

$\ln(1+x) = x - (x/2) + (x/3) - (x/4) + (x/5) - \ldots$ is well known. Therefore, write $\tilde{P} = P^* - I$ and E2.1.15 becomes:

$$I + \tilde{P} = e^{\Gamma t'} \qquad \text{(E3.2.13)}$$

then:

$$\Gamma t' = \ln(I + \tilde{P}) = \tilde{P} - (\tilde{P}^2/2) + (\tilde{P}^3/3) - \ldots \qquad \text{(E3.2.14)}$$

This suggests that we estimate Γ as:

$$\hat{\Gamma} = \frac{1}{t'}\left[\hat{\underset{\sim}{P}} - (\hat{\underset{\sim}{P}}^2/2) + (\hat{\underset{\sim}{P}}^3/3) - \ldots\right] \qquad \text{(E3.2.15)}$$

where : $\hat{\underset{\sim}{P}} = \hat{\underset{\sim}{P}}^* - I = \hat{P}(t') - I$

Zahl [1955] shows that, if the series on the right hand side of E3.2.13 converges elementwise, then it is equivalent to E3.2.12. He further states that, if the series in E3.2.14 converges, then $\hat{\Gamma}$ is the unique maximum likelihood estimate of Γ. If one has control over the observation interval t', then Zahl shows that the convergence of E3.2.14 can be made as near to certain as desired. This is accomplished by choosing t' sufficiently small so that $p^*_{ii} >> 1/2$ for all states i, a sufficient (but not necessary) condition for convergence. In our circumstances however, we cannot get finer observations than those taken at one day intervals. This is not sufficient to ensure $p^*_{ii} > 1/2$; thus Zahl's estimates will be of questionable value.

Cuthbert [1973] has shown that, as one might expect, p^*_{ij} does not uniquely determine Γ in all cases. We must be wary of "false" convergence produced by machine roundoff when interpreting Zahl's estimates.

The sufficient statistics are n_{ij} transition counts for the imbedded chain. These are intimately related to the a transitions for the overall process. For t' sufficiently small the off diagonal entries approach the respective a_{ij}, and n_{ii} approaches D_i.

3.3 The Data Reduction Problems

Much of the data processing for the current study was

done three years ago as part of ancther project. This work, which was preliminary to the estimation of the Markov model, was done under the direction of A.N.Arnason and D. Roch for the Manitoba Health Services Commission. Their report [1974] is the basis of most of this section.

The problem which is of interest here is one of a large class of problems which usually face the model builder. It is typical of present day computing systems that the majority of the information gathered and processed arises from administrative requirements. The needs of a researcher are seldom given much thought at any stage of system design or implementation. Consequently, extracting information from these administrative files for the purpose of estimating model parameters, verifying results, and most other research functions often becomes rather like pulling teeth from a whale as it goes about it's normal business. Great caution must be exercised to interpret information in the context in which it was collected and maintained rather than in the research context. One must expect administrative data to be accurate only in those areas which critically affect the administrative functions. Other data, with little administrative value, will typically suffer from inaccuracies and may even be missing in a large percentage of cases, even though collected on the same form and in the same way as the administratively critical fields.

The data source which was used in this study is the administrative files of the Manitoba Health Services Commission (M.H.S.C). These files are used to control the payment of claims under the Manitoba Health Insurance scheme to doctors, hospitals, medical laboratories etc. The major requirement of the administrative process is to be able to identify whether or not a claim represents an insured service performed by the claimant for an insured person. At

the time of this study, premium fees were still collected from individuals and employers for insurance coverage, so a secondary requirement was that these bills be issued and their payment acknowledged.

With this structure in mind, the following sections describe the process by which the sufficient statistics for the estimates of the model parameters, and various verification statistics, were extracted from the files.

3.3.1 The Data Source

The M.H.S.C operates its file system in the following way. The primary file is the Population Master File. It contains the information on each insured person and is maintained on a daily basis with any additions, deletions and changes which may be reported to the Commission. Individuals are registered as a family group: that is, a registration number is assigned to a family, which may consist of one or more individuals. Within a registration number an individual is identified uniquely by his/ her first initial, year and month of birth, and sex. At six month intervals a copy of the current population master is saved. These "snapshots" occur Dec. 31 and June 30 of each year.

Information on the medical services delivered are maintained in separate files by point of service: physicians' claims in one file, hospital inpatients' in another, etc. The administrative process is basically to match each service (claim) record against the population master file to determine if the person to whom the service was rendered was entitled to coverage on the date of the service. This is relatively simple to do with this file structure.

For our purposes, however, it is necessary to know, for each service, its duration (for D_i), it's location (region

of service), and the region and facility that the individual receiving the service was in just prior to (or, equivalently, just after) the service was performed (for a_{ij}). This information, even if it appeared to be present on the administrative records, is superfluous to the administrative requirement and thus unlikely to be either complete or accurate. Thus some alternative method which relies heavily on the administrative requirements must be used to estimate D_i and a_{ij}.

The approach taken was to use the matching requirement to attempt to produce, for each individual, a list of all the medical services he obtained during the year and the dates on which he obtained them. This process does use only that information critical to the administrative function. However, there are some important difficulties to be overcome. First there is the shear volume of data involved. The population of Manitoba in 1972 was approximately 1 million. A million people generate a lot of health care transactions in a year. Sampling could be used to reduce this volume, but sampling requires that there be a suitable frame, in this case a list of individuals, to sample from. The population master file is not a suitable frame because of its organization by family. This induces many individuals to appear on the file more than once, duplicated as their family status changes through marriage, turning 18, adoption, and other more obscure causes. Also, the number of parameters in our model is large, approximately 342 parameters for each age/sex/region group for a total of over 13,000 parameters. Thus we must be very careful if sampling is not to make some of our estimates devoid of observations, and therefore, of meaning.

The computational problem of producing the merged history of an individual is complicated by the registration

number that occurs, and by the incompatibilities between the medical service claim files and the "snapshots" of the population master file. Although the administrative procedure requires that a medical service claim be matched against a record on the population master file, it does not require that such a match be found on yesterday's master file or on tomorrow's master file. Since we must examine all the services performed in the year, yet we only have the population master file as it existed on the day before the year began, in the middle of the year, and on the last day of the year, we are bound to have some matching problems.

The actual merge, which was done in 1974 by B.Beach of the University of Manitoba Department of Computer Science for the Arnason and Roch study, proceeded as follows:

1) Pass the Dec.31,1971 Population Master File and reject all those records "cancelled" on or before Dec.31,1971.

2) Pass the Dec.31,1972 Population Master File and reject all those records "cancelled" on or before Dec.31,1971.

3) Attempt to match each record in the 1972 file with a record in the 1971 file, merging them to produce a file of registrants who were "active" at either the beginning or the end of 1972. At the same time, attempt to trace those who have changed number in the year and map them into a single number.

4) Merge each individual's medical services from the medical service claim files with his master file record. This produces a file containing, for each individual, a vector of his characteristics (e^k in table 3.3.1) and a

list of vectors describing any services he had during the year (t -table 3.3.1).

Table 3.3.1 Merged Medical History Information

Vector Element	Meaning
e^1	registration number family seq. no.
e^2	month of birth: 2 digits (1-12)
e^3	year of birth: 2 digits if >72 assume 1800
e^4	sex: 0-male , 1-female
e^5	age class (1-7)=(1,2-15,16-45,46-65,66+,?)
e^6	municipality code res. as of Jan.1,1972
e^7	entry time: day of entry (0-367)
e^8	entry code: 1 alive on Jan.1, 2 new birth, 3 new entry, 4 partial hist.
e^9	municipality of residence Dec 31,1972
e^{10}	exit time: day left if known (2-367)
e^{11}	exit code: 0-unknown, 1 alive at end, 2-died,3-moved, 4-partial hist.
e^{12}	number of medical transactions
t^1	date i-th service began
t^2	facility number (1-9)
t^3	morbidity code (188 break), i-th service
t^4	location of service
t^5	duration of i-th service
t^6	next facility used (not always given)
t^7	number of surgeries (0-3) hosp. only
t^8	intensive care days (inaccurately reported)

1. Note $i=1,2,...e^{12}$
2. Facility codes : 1 - Drs. care, 2 - Acute Hospital
   3 - Extended H. 4 - Out-Patients
   5 - N.P.C. 6 - Healthy 7 - Dead
   8 - Psych. Hosp. 9 - Psych. Drs. care

At a later stage it was discovered that there was a problem with the reporting of outpatient contacts, particularly at Winnipeg General Hospital clinic. The reporting failure was due to the independence of the financial support of the clinic from the type and amount of care that it dispensed. The clinic was run by staff doctors on a budget. No extra money for the clinic or the doctors was generated by individual contacts. As a result, detailed record keeping was not enforced, and during 1972 very few patients who attended the Winnipeg General Outpatient facility had their visit reported to the M.H.S.C..

Some primary care contacts such as public health nursing, are not recorded in the data. Care provided by these nurses may be a substitute for physician or outpatient services especially in the rural areas.

The major problems with this file from the point of view of estimating the Markov model are related to the problem of tracing registration number changes from the two year-end "snapshots" of the population master file. Of the 46,524 records for which number changes were detected, 13,435 were not traceable to another number in the 1972 file. Arnason and Roch report that discrepancies in counts indicate that a further 8,654 records should have been linked, but no linkage was attempted due to missing information. These ∿22,000 partial histories could introduce a serious bias into the model estimates. The 13,435 detected records were thus discarded for the purposes of estimation with the hope that the reasons for the unsuccessful link are not related to their health care demand. If this is so the ignoring them will not introduce a bias. Of course, nothing could be done to eliminate the further 8,654 records which are surmised to have similar failings.

Another problem detected at this stage in the reduction process was an error in classification of services rendered into the appropriate facilities. In particular non doctor/

patient contacts (NPC) were distorted because doctor/patient contacts by doctors with billing identification numbers >3000 were classified as NPC. This distortion resulted in the deletion NPC as a health care state in the Markov model for the present study.

The next stage in the data reduction was to select the population of interest from this file of merged histories and to convert the selected histories into a form more useful for computing the sufficient statistics for the Markov model of section [2.2]. The population of interest is all those non-Indian people alive and resident in Manitoba on Jan.1,1972. Thus the records selected were those with $e^8=1$. This rejects the partial histories ($e^8=4$ or $e^{11}=4$), new births ($e^8=2$), and new entries ($e^8=3$). Indians were detected on the basis of their beginning or final muncipality code ($400 \leq e^6 \leq 500$ or $400 \leq e^9 \leq 500$). Finally those who could not be classified as to age were rejected ($e^5=7$ or 0). This selection procedure yielded a population of 995,769 histories to be used in computing the estimates.

These histories were then reduced to a vector containing the minimal amount of information necessary for the model by a two stage process. First,for each history, two vectors of length 367, A_i and F_i, were constructed to give the location (A_i) and health care facility used (F_i) on each day of 1972. The location information (A_i) was mapped from the municipality codes (and hospital numbers) (e^6,e^9,t^4) according to the mapping supplied by the M.H.S.C. Facility was determined from (t^2) by grouping psychiatric hospital care (8) with acute inpatients (2), psychiatric doctors' care (9) with doctors' care (1) and ignoring all NPC transactions. A person was assumed healthy for those periods not covered by a health care transaction. Multiple transactions on one day were eliminated by using the priority scheme: outpatient >

acute hospital or extended hospital > doctors' care. This means that an admission to hospital from an outpatient clinic on the same day will be recorded as one day outpatient care followed by one or more days of hospital care. Similarly doctors' transactions with patients in hospital will be ignored and the person will remain in the hospital care state.

The location (A_i) on the i-th day is the location in which the medical service took place, mapped from t_k^4, or the location of residence for the healthy state. Location of residence was mapped using e^6 and e^9. For those who changed location (moved between regions) during the year, the date of the move was approximated by finding the first set of five medical transactions in which three of the five were in the final location and selecting a day randomly between the days of the 2nd and 3rd medical transactions obtained in the final locations. If this was not possible, then the location change was assumed to have occurred on a date chosen uniformly between the last medical transaction and the end of the year.

Once A_i and F_i were defined for i= 2,3...367, the state occupied on each day was computed as

$$S_i = 6 * (A_i - 1) + F_i \qquad i=2,3,...,367 \qquad \text{(E3.3.1)}$$

and a record consisting of the age/sex group, the number of states occupied during the year (NCH), and the coded information about the state and day of entry to the state (L_k) constructed.

(E3.3.2)

$$L_k = S_i + 32 * i \qquad \text{for } i \text{ such that } S_i \neq S_{i-1}$$

This is the file which was used to collect the statistics used for estimation and verification of the model. The data flow described above is summarized in figure 3.3.1. At each stage of the reduction process information has been lost. The utility of the model will depend upon its robustness to this infomation loss.

Before continuing with the discussion of the model estimation, it is appropriate to comment here on the various problems encountered in this reduction process and how their impact might have been lessened or eliminated had the administrative problem been solved in a slightly different way.

There are two primary causes of the problems encountered. It is inexcusable that a file of this size be constructed without a unique primary key for each record. Families are not the logical grouping for the information, either from the administrative point of view or from the research viewpoint. Individuals demand and receive health care,not families. This would not have been as much of a problem in the present study if it were not for the second failing in the administrative file structure. Records are deleted immediately upon receipt of a cancellation, if such a deletion does not cause the family identification number to disappear.

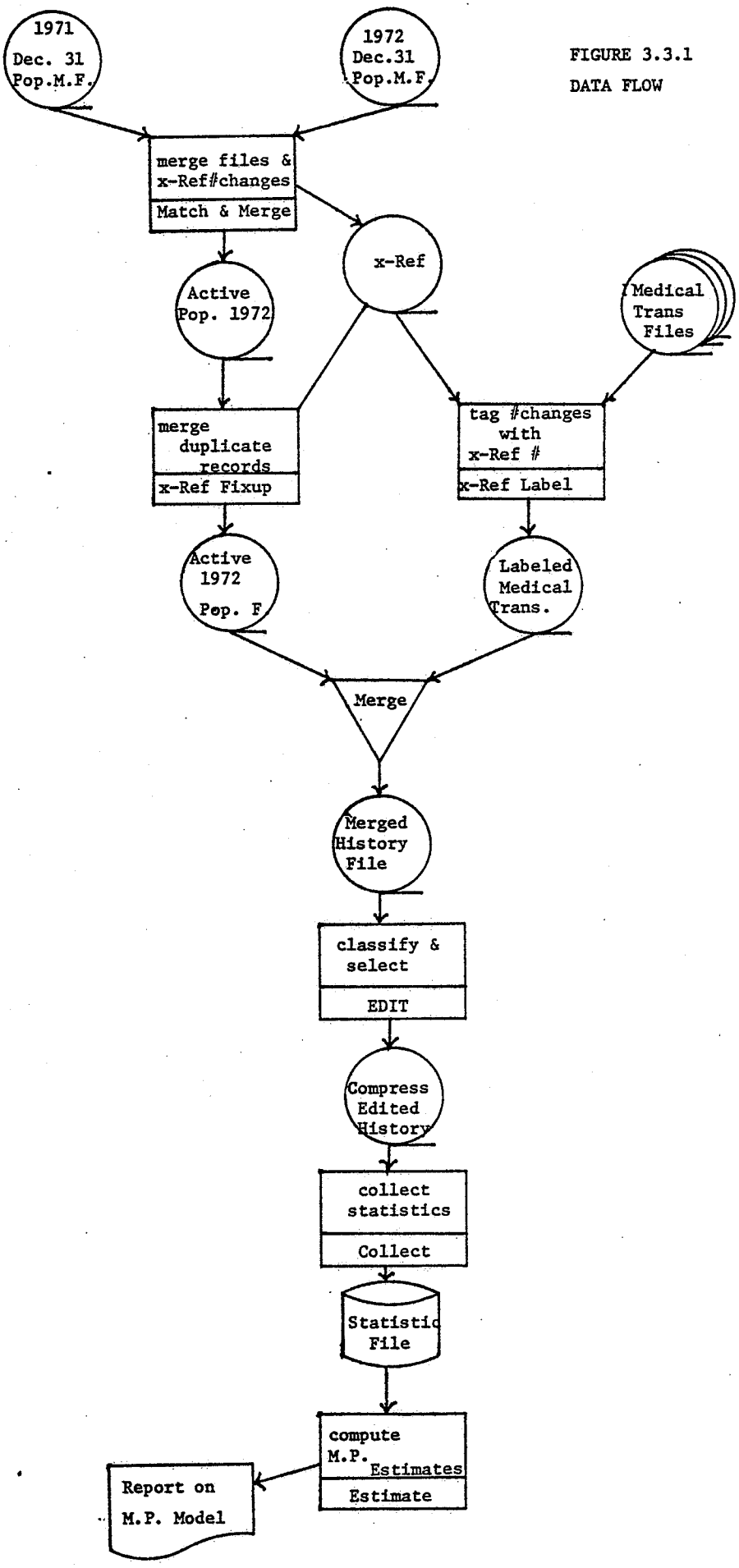


FIGURE 3.3.1

DATA FLOW

This makes the six month snapshots of the population master file woefully inadequate as representations of the population that existed and was insured between the snapshots. Thus individuals who change number are often untraceble, and therefore, unusable for our purposes. This practice must be expensive in terms of correspondence for the M.H.S.C. and has a simple solution: don't delete records until they are known to be inactive for a sufficient period of time, and delete only at the time of snapshots.

3.3.2 Sufficient Statistics Collection

Two more data processing problems remain before the Markov model can be used to measure the health care system. The sufficient statistics must be collected along with enough other information to validate the model's assumptions. Then the parameter estimates can be computed, and estimates of residency, daily and yearly probabilities, and other model determined statistics can be produced.

From data organized as the file of compressed and edited histories [3.3.1], the sufficient statistics could be produced in two ways. One could sort the output tape by group (region/age/sex) and process the sorted file to obtain the estimates. This approach requires little main storage, as only one set of statistics accumulators need be used, but requires a sort on 1 million records. This approach was rejected because the large amounts of relatively inexpensive main storage available on the University of Manitoba S370/168 permitted the statistics to be gathered simultaneously for the ten age/sex groups without the need for a sort.

Processing the statistics is logically done for each group at the completion of the data collecton. The decision not to sort the data thus implies a logical separation between the estimation and collection processes. Estimation requires data for one group at a time, while collection

(from an unsorted tape) requires data for all groups to be resident in main storage simultaneously. These tasks were therefore divided into two processes as indicated in figure 3.3.1, with the statistics collected in the first stage written to disk for later processing by the estimation program.

The sufficient statistics for the model are the total sojourn time in each state (D_i) and the absolute number of transitions from i to j (a_{ij}) for each pair of states. For model verification purposes, certain other statistics were collected as well. The standard deviation of the length of stay distribution is interesting because, if the model assumptions hold, it is equal to the mean length of stay [2.1.4]. The daily transitions matrix (n_{ij}) of the imbedded chain can be calculated by collecting the initial and final occupancy vectors (NI_i ,NF_i). The occupancy vector for the middle of the year (NM_i) was also collected, where the middle of the year was defined as day 185= July 2,1972. The effect of the truncation of observations as a result of the arbitrary one year time period was of interest. The total, mean and standard deviation for the untruncated (neither initial "stays" nor final "stays" of the year) length of stay were collected to assess this problem. The observed average residency times and their variance were also collected for comparison with the estimated values. These statistics were all collected simultaneously and then written with appropriate labeling to a sequential file for later processing by the estimation program (table 3.3.2).

In addition to the statistics passed to the estimation program, the collection process also generated some additional statistics. These were the distributions of the number of visits to each of the states in the model, conditional on the area of initial residence, and the means

and standard deviations of these distributions.

The collection process involved decompressing each medical history from the L_k vector of length (NCH) to two vectors S_i and DUR_i representing the state occupied prior to the i-th transition and the duration of the stay in that state:

$$S_k = L_k \bmod(32) \qquad k=1,2,\ldots,NCH$$

$$DUR_k = \lfloor (L_{k+1} - S_{k+1})/32 \rfloor - \lfloor (L_k - S_k)/32 \rfloor \qquad k=1,2,\ldots(NCH-1)$$

During this process the initial, final and mid-year states were identified. The length of stay and squared length of stay in each state was then calculated and accumulated for the total length of stay and untruncated total length of stay statistics. Transitions were counted by comparing S and S_i for i=2,...NCH. To be slightly more efficient, NCH=1, indicating no medical service during the year, was treated as a separate case. The distribution statistics were thus conditional on some change of state having occurred during the year, either through a location change or a medical transaction.

STATISTICS COLLECTED TABLE 3.3.2

LABEL	DIMEN-SIONS	VARIABLE	MEANING
IN	S,S	a_{ij}	- number of transitions from state i to state j[2]
NO	S	$n_i(2)$	- occupancy vector : # individuals in state i on June 2,1972
NM	S	$n_i(185)$	- occupancy vector : # individuals in state i on day 185
NF	S	$n_i(367)$	- occupancy vector : # individuals in state i on December 31,1972
NNC	S	$\tilde{n}_i$	- # individuals who never leave state i during the year
TV[3]	A,S	$n_{aj.}$	- # visits by all individuals initially in area a to state j
SV[3]	A,S		- sum of squares of TV entries
NV[3]	A,S,0:9	n_{ajk}	- # individuals initially in area a who make k visits to state j
TRES	S,S	$e_{ij}(366)$	- sum of(time spent in state j by individuals initially in state i)
SRES	S,S	$e_{ij}(366)$	- sum of squares of TRES entries
TLS	S	D_i	- total sojourn time in state i over all visits to state i
SLS	S		- sum of squares of TLS entries
TSLT	S		- sum of lengths of stay in state i over all visits not the first or last visit for the individual
SLST	S		- sum of squares of TLST entries

1. Dimension Codes:
   S-states : 1-24
   A- areas or regions : 1-4

2. $a_{ii} = \sum_{j \neq i} a_{ij}$

3. not passed to estimates procedure

The statistics collected are summarized in table 3.3.2, which also indicates the form of the labeling used in passing the statistics to the estimation procedure. These labels are derived from the variable names used in the collection program. The collection program was run four times for the 24 state model, once for each of the Manitoba regions and once for the total population. The collection process for the total population required only slightly more than 11.1 min. of S370/168 CPU time and required a 384 K byte region of which approximately 188 K byte was used for statistics accumulation.

3.3.3 Computing the Estimates

To compute the maximun likelihood estimate for the instantaneous rates $\hat{\gamma}_{ij}$ [3.2.1] and for the imbedded chain transition probabilities $\hat{P}^{*}_{ij}$ [3.2.2] is a relatively simple procedure after the statistics collection has been done. $\hat{\Gamma}$ is calculated from E3.1.8 as (see table 3.3.2):

$$\gamma_{ij} = a_{ij} / D_i = \begin{cases} IN(I,J) & I \neq J \\ -IN(I,I) \ / \ TLS(I) & \text{otherwise} \end{cases}$$

To estimate P^{*}_{ij} we must first calculate the daily transition matrix A^{*} from the collected statistics. The off diagonal terms are just IN(I,J) , the number of times someone was observed to be in state J the day immediately following a day in which he was in state I. The diagonal terms can be calculated residually by noting that the number of times someone is observed to be in the same state I on successive days is simply the total number of days spent in state I minus the number of days which were "first days" in I. The number of "first days" in a state is the number of transitions into that state, therefore:

$$n_{ii} = D_i - \sum_{k\neq i}^{NS} a_{ki} = TLS(I) - \sum_{K\neq I}^{NS} IN(K,I)$$

$$\hat{p}^{*}_{ij} = n_{ij} \Big/ \sum_{k=1}^{NS} n_{ik} = \begin{cases} IN(I,J) \;/\; (TLS(I) - NF(I)) & I\neq J \\ (TLS(I) - \sum_{K\neq I}^{NS} IN(K,I)) / (TLS(I) - NF(I)) & \text{otherwise} \end{cases}$$

where : $TLS(I) - NF(I) = \sum n_{ik}$

$$= IN(I,I) + [TLS(I) - \sum_{K\neq I} IN(K,I)]$$

$$= TLS(I) - [\sum_{K\neq I} IN(K,I) - \sum_{K\neq I} IN(I,K)]$$

The estimated asymtotic variance of $\hat{\gamma}_{ij}$ can be calculated using E3.2.9 as:

$$\hat{var}(\hat{\gamma}_{ij}) = \hat{\gamma}_{ij} / D_i = a_{ij} / D_i^2 = IN(I,J)/TLS(I)^2 \qquad I\neq J$$

and:

$$\hat{var}(\hat{\gamma}_{ii}) = \sum_{j\neq i} \hat{var}(\hat{\gamma}_{ij}) = IN(I,I)/TLS(I)$$

since the covariances are asymtotically zero. The statistics actually reported were the standard deviation estimates, the positive square roots of the above estimated variances. Checks for zero divisions, indicating an unused state, were made and the respective estimates reported as zero for expectations, and -1 for standard deviations.

3.4 The Derived Estimates

The two most interesting properties of the Markov process model which can be derived from the estimates are the multistep transition probabilities, P(t), and the expected residency E(t). As described in section 2.1, these, theoretically at least, can be calculated from:

$$P(t) = e^{\Gamma t} = \frac{(\Gamma t)^n}{n!} = I + (\Gamma t) + \frac{(\Gamma t)^2}{2!} + \frac{(\Gamma t)^3}{3!} + \ldots \qquad \text{(E2.1.15)}$$

$$e_{ij}(T) = IT + \frac{\Gamma T^2}{2!} + \frac{\Gamma^2 T^3}{3!} + \frac{\Gamma^3 T^4}{4!} + \ldots \qquad \text{(E2.1.20)}$$

Unfortunately, these forms cannot be directly computed. This can be demonstrated by examining the simple 2 state process with rate matrix:

$$\Gamma = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

If we attempt to compute E (1yr. = 365 days), the one year residency matrix, using E2.1.20 we have as the n-th term of the series estimate of the expected residency:

$$E_n(366) = \begin{bmatrix} 366 & 0 \\ 0 & 366 \end{bmatrix} + \frac{366^2}{2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + \frac{366^3}{6} \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} + \frac{366^n}{n!} \begin{bmatrix} -a_n & a_n \\ a_n & -a_n \end{bmatrix}$$

where:

$$a_n = (-2)^{n-1}$$

The 51 st term of this series has elements of the form:

$$\left| \frac{(-2)^{50} \cdot \quad (366\ )^{51}}{51!} \right| \quad > 10^{62}$$

In a finite precision machine the difference (sum) of the 50th and 51st terms will be roundoff error with elements of order $\sim 10^{48}$ (assuming 16 decimal digits of floating point accuracy). There is no hope of computing the true residency estimate in this way. The formula presented by Zahl [1955] cannot be used to compute the residency matrix.

When this was discovered (the hard way I'm sorry to admit) alternatives were searched for in the literature. Chiang [1964, 1973] suggests that the solution be derived from the results for homogenous systems of differential equations. This gives:

$$e_{ij}(t) = \sum_{k=1}^{NS'} A_{ij}^{(k)} \frac{e^{\lambda_k t}}{\lambda_k} \Big/ \prod_{m\ k}^{NS'} (\lambda_k - \lambda_m)$$

$$p_{ij}(t) = \sum_{k=1}^{NS'} A_{ij}^{(k)} \quad e^{\lambda_k t} \Big/ \prod_{m \neq k}^{NS'} (\lambda_k - \lambda_m)$$

where $A_{ij}^{(k)}$ is the i,jth cofactor of the matrix $(\lambda_k I - \Gamma')$ $k=1 \ldots NS'$ are the eigenvalues of Γ', and Γ' is the transpose of the rate matrix for the non-absorbing states of the model. This is mathematically interesting, but would involve the evaluation of about 19^3, 18x18 determinants (in our case) for each age/sex/ region group. Accurate computation of determinates and cofactors is an almost totally unsolved problem in numerical methods. Thus this method was rejected after a few small scale attempts (in APL) at implementation.

The approach to computing the estimates of E(t) and P(t) which was finally selected was based on the well known theorems of linear algebra which relate a matrix to its

diagonal form through a similarity transformation. The basic result is that certain classes of matrices A are "similar" to a diagonal matrix in the sense that there exists a nonsingular matrix Q such that $Q^{-1}AQ = D_\lambda$ a diagonal matrix. If this is true then the elements on the diagonal of D_λ, are the eigenvalues of the matrix A, and Q is the matrix whose columns are the respective column eigenvectors of A. Given that we can find the similarity transformation of the rate matrix Γ which diagonalizes it, it is a relatively simple matter to estimate P(t) and E(t) since if $\Gamma = Q D_\lambda Q^{-1}$ then:

$$\Gamma^2 = Q D_\lambda Q^{-1} \; Q D_\lambda Q^{-1} = Q D_\lambda^2 Q^{-1}$$

and by induction:

$$\Gamma^n = Q D_\lambda^n Q^{-1} \qquad n=1,2,\ldots$$

Thus the matrix series for P(t) becomes:

$$P(t) = e^{\Gamma t} = \sum_{n=0}^{\infty} \frac{Q D_\lambda^n t^n Q^{-1}}{n!} = Q \sum_{n=0}^{\infty} \frac{D_\lambda^n t^n}{n!} Q^{-1} \qquad \text{(E3.4.1)}$$

$$= Q D_{e^{\lambda t}} Q^{-1}$$

and the residency matrix may be estimated as:

$$E(T) = \int_0^T P(t)\,dt = \int_0^T Q D_{e^{t}} Q^{-1} dt = Q \int_0^T D_{e^{t}}\,dt\, Q^{-1}$$

$$\text{therefore : } E(T) = Q \begin{bmatrix} \frac{e^{\lambda_1 T} - 1}{\lambda_1} & & 0 \\ & \frac{e^{\lambda_2 T} - 1}{\lambda_2} & \\ & & \ddots \\ 0 & & T \end{bmatrix} Q^{-1} \qquad \text{(E3.4.2)}$$

where the terms on the diagonal are $\frac{e^{\lambda_i T} - 1}{\lambda_i}$ if $\lambda_i \neq 0$, and T if $\lambda_i = 0$.

However, this decomposition assumes that the eigenvalues of Γ are all real and that the eigenvectors of Γ are linearly independent. In fact we really only need to require the second of these conditions if all of the calculation is done with complex arithmetic. Note that Γ, being a singular matrix, must have at least one zero eigenvalue. However, because the negative row sums are on the diagonal of Γ, we may invoke Gerschgorin's theorem [see Varga 1962 p 16] to guarantee that the real parts of the eigenvalues will be less than or equal to zero. This implies that if the decomposition works, E3.4.1 and E3.4.2 will be computationally useful, for now large values of T produce terms which are very small instead of very large.

The conditions under which Γ can be diagonalized in this manner are not known. If Γ were symmetric then we are guaranteed to have real, distinct eigenvalues and linearly independent eigenvectors, but the rate matrices with which we are concerned are far from symmetric in theory, let alone after estimation. However, there seemed to be empirical support for the view that the decomposition should work. In attempting to apply the decomposition to the estimated rate matrices for the 24 state health care model, only 6 of 33 matrices showed any sign of complex eigenvalues, and in none of the 27 successful cases did the eigenvectors turnout to be computationally singular.

The method employed was to extract the eigenvalues and eigenvectors of $\hat{\Gamma}$ using the IMSL [1974] subroutine library program EIGRF. The matrix of eigenvectors was then inverted using another IMSL routine, LINV3F. Direct computation of P(t) and E(t) using E3.4.1 and E3.4.2 for various values of t is then straightforward. The computational effort involved in the process is substantial. The matrix must first be balanced, then reduced to Hessenburg form, the

eigenvalues and eigenvectors extracted via the QR algorithm [see Wilkinson 1968]. The matrix of eigenvectors must be inverted. Two matrix multiplications for each value of t of interest then produces the result.

3.5 Inference

3.5.1 Goodness of Fit.

Formal statistical tests of the model assumptions are a necessary part of the model verification procedure. Since the models with which we are concerned are stationary Markov processes, there are basically two areas which require testing:

1) The Markov property: The probability that an individual who is in state i at time t is subsequently in state j at time $t' > t$ depends on the states, i and j, in question, but not on any previous state or sequence of states the individual may have occupied at times prior to t (E2.1.2).

2) The Stationary assumption: The probability that an individual who was in state i at time t is subsequently in j at time $t' > t$ is a function of the length of time between t' and t, $\Delta t = t' - t$, but not of the time t at which the interval began.

A test of (1) suggested by Zahl [1955] is based on examining the number of transition pairs of the form $h \rightarrow i \rightarrow j$, that is the number of times, n_{hij}, a transition from state h to state i is immediately followed by a transition from state i to state j. If the process is Markovian then the expected values of the proportions $n_{hij}/n_{.ij}$ of such transitions to all transitions from state i to state j should be the same for all states h.

$$E(p_{hij}) = E\ \frac{n_{hij}}{n_{\cdot ij}} = c_{ij} \quad \text{for all } h = 1,2,\ldots,NS$$

Thus a χ^2 test for homogeneity of proportions may be used as a test of property 1) by computing

$$\chi^2_m = \sum_h \sum_i \sum_j (n_{hij} - E(n_{hij}))^2 / E(n_{hij}) \qquad \text{(E3.5.1)}$$

where

$$E(n_{hij}) = (\sum_j n_{hij})(\sum_h n_{hij})/(\sum_h \sum_j n_{hij})$$

which is asymtotically χ^2_m with $m=NS(NS-1)$ degrees of freedom.

Alternatively we may use the property derived from the Markov assumptions that the length of stay is distributed according to the exponential distribution [2.1.4] and use any of the many goodness of fit tests. An intuitive and simple approach is to remember that the exponential distribution has the property that its mean μ and standard deviation σ are equal. Therefore the ratio $\frac{\bar{x}}{s}$ of the sample mean and sample standard deviation should be close to one for large samples. In fact, in view of the central limit theorem it is not unreasonable to expect:

$$F\left(\frac{\bar{x}}{s}\right) \overset{L}{\to} N(1;\ 1/\sqrt{n})$$

independent of the parameter of the distribution involved. Sampling experiments have confirmed that this is not a bad approximation even for relatively small sample sizes.

Testing for property (2) can be reduced to testing for homogeneous rates by separating the period of observation

[0,T] into two or more pieces and testing whether or not the rates are the same in each of the periods. Thus the problem of testing for stationarity can be solved in the same way as the problem of testing for homogeneous rates in two different age/sex cohorts or in two disjoint regions.

3.5.2 Homogeneity Tests

3.5.2.1 Likelihood Ratio Test

The likelihood ratio test [Wilkes ch 13] for homogeneous rates in two populations, A and B, can be developed as:

under H_0: $\Gamma^A = \Gamma^B$, the populations have the same rate matrix :

$$\ln(\hat{L}_0) = \sum\sum (a_{ij}^A + a_{ij}^B) \ln\left(\frac{a_{ij}^A + a_{ij}^B}{D_i^A + D_i^B}\right) - \sum\sum (D_i^A + D_i^B)\left(\frac{a_{ij}^A + a_{ij}^B}{D_i^A + D_i^B}\right)$$

$$= \sum\sum (a_{ij}^A + a_{ij}^B) \ln\left(\frac{a_{ij}^A + a_{ij}^B}{D_i^A + D_i^B}\right) - \sum\sum (a_{ij}^A + a_{ij}^B)$$

: under H_1: $\Gamma^A \neq \Gamma^B$, different rates :

$$\ln(\hat{L}_1) = \sum\sum a_{ij}^A \ln\left(\frac{a_{ij}^A}{D_i^A}\right) - \sum\sum a_{ij}^A$$

$$+\sum\sum a_{ij}^B \ln\left(\frac{a_{ij}^B}{D_i^B}\right) - \sum\sum a_{ij}^B$$

therefore the likelihood ratio is :

$$-2\ln(\lambda) = 2[\ln(L_1) - \ln(L_0)] =$$

$$2\sum\sum a_{ij} \ln\left(\frac{a_{ij}^A}{a_{ij}^A + a_{ij}^B} \cdot \frac{D_i^A + D_i^B}{D_i^A}\right) + a_{ij} \ln\left(\frac{a_{ij}^B}{a_{ij}^A + a_{ij}^B} \cdot \frac{D_i^A + D_i^B}{D_i^B}\right)$$

$$= 2\sum\sum a_{ij}^A \ln\frac{\hat{\gamma}_{ij}^A}{\hat{\gamma}_{ij}^{A\cup B}} + a_{ij}^B \ln\frac{\hat{\gamma}_{ij}^B}{\hat{\gamma}_{ij}^{A\cup B}}$$

which is asymptotically χ_m^2 (m=NS(NS-1) degrees of freedom). Corrections which improve the asymptotic convergence of this

and other similar tests for the parameters of Markov processes can be found in Sharp[1975].

3.5.2.2 Zahl's Test for Homogeneous Rates

An alternative to the above likelihood ratio test was proposed in S.Zahl's 1955 paper. Consider the statistic :

$$U_{ij} = \frac{|\gamma_{ij}^{A} - \gamma_{ij}^{B}|}{\sqrt{var(\gamma_{ij}^{A}) + var(\gamma_{ij}^{B})}}$$

Under H_0: $\Gamma^A = \Gamma^B$ or H_0': $\gamma_{ij}^{A} = \gamma_{ij}^{B}$, U_{ij} is asymptotically distributed as $N(0,1)$. Therefore a test for $H_0: \Gamma^A = \Gamma^B$ can be constructed using :

$$U^2 = \sum_{i=1}^{NS} \sum_{j \neq i}^{NS} U_{ij}^2$$

which, under H_0, is approximately χ_m^2 with $m=NS(NS-1)$ degrees of freedom.

Chapter 4 Examining the Statistics

4.1 Without the Model

The sufficient statistics as described in [3.1] and the auxillary statistics gathered with them have value as measurements of the performance of the health care delivery system which does not depend on the assumptions of the Markov model. Care must be exercised to recognize that the context of their collection, the Markov model, has a definite impact on their meaning. The decisions taken during the data collection process were primarily influenced by the requirements of the Markov model; thus it was inevitable that certain distortions in the statistics would result. However, much can be learned by examining the statistics and interpreting them in the light of these distortions.

4.1.1 Absolute Transitions

Table 4.1.1 presents the absolute transition matrix (a_{ij}) for the 24 state model for the entire population. The diagonal terms are the row sums, and represent the number of separations from each state. Thus there were 83,983 seperations from Winnipeg acute care hospital in 1972 by those people in the subject population, a measure of utilization of the facility. Note that this usage does not include those out of province who spent time in Winnipeg hospital during the year, nor does it include use of the facilities by those who were born in 1972, or who moved to Manitoba in 1972. Further distortion of the statistics results from the exclusion of Indians from the subject population.

Separations from Drs. Care must not be interpreted as doctor's visits for several reasons. First, visits to the doctor (not necessarily the same doctor) on successive days are not included. Multiple visits on the same day have also

been eliminated by the editing process. Because of the necessity of classifying a person in a unique state at each time, doctor's visits while the person receiving care is in the hospital are also lost. Thus separations from Drs. care represent a severe under-estimate of the total physician load imposed by the subject population, which is an under-estimate of the total population served by physicians for the reasons mentioned above for hospitals.

It would be very nice if the separations from the healthy state represented the number of "illnesses" experienced by the subject population, but this is not the case. Physicians perform large numbers of services for people who are healthy, and each of these will probably generate a separation from the healthy state. More severe distortion is introduced by the episodic nature of illness. The number of health car transactions generated by a single "episode" of ill health is unknown. [See J. Ashford and R.G. Hunt 1974]. Each of these transactions can cause a healthy state separation [and usually does]. These factors make healthy state separations very difficult to interpret.

The out-patient transitions are distorted for Winnipeg by the reporting problem previously described in section 3.3.1, as well as the lack of measurement of part of the population serviced, noted above.

TABLE 4.1.1 ABSOLUTE TRANSITION MATRIX

FROM \ TO		WINNIPEG DR.CARE	ACUTE	EXTEND	OUT-PAT	HEALTHY	DEAD
WINNIPEG	DR.S CARE	1985136	0	0	10307	1741219	1
	ACUTE HOSPITAL	1495	83983	0	644	62610	2088
	EXTENDED HOSP.	10	0	371	29	143	119
	OUT PATIENTS	9127	16746	23	145642	107128	221
	HEALTHY	1740674	51965	277	120273	2004803	747
BRANDON	DR.S CARE	172	0	0	21	3825	0
	ACUTE HOSPITAL	10	0	0	1	110	7
	EXTENDED HOSP.	0	0	0	0	10	1
	OUT PATIENTS	14	1	0	6	392	0
	HEALTHY	2893	390	0	308	816	0
REST MAN.	DR.S CARE	1623	0	0	164	19146	0
	ACUTE HOSPITAL	220	7	0	163	1431	27
	EXTENDED HOSP.	0	0	0	0	5	2
	OUT PATIENTS	529	599	2	147	6495	10
	HEALTHY	226939	14372	70	13235	12163	1
OUT PROV.	DR.S CARE	178	0	0	78	18981	1
	ACUTE HOSPITAL	17	2	0	22	1086	27
	EXTENDED HOSP.	0	0	0	0	0	0
	OUT PATIENTS	59	8	0	21	3732	0
	HEALTHY	973	47	0	65	15	0

		BRANDON DR.CARE	ACUTE	EXTEND	OUT-PAT	HEALTHY	DEAD
WINNIPEG	DR.S CARE	121	0	0	4	2911	1
	ACUTE HOSPITAL	10	0	0	8	399	5
	EXTENDED HOSP.	0	0	0	0	1	0
	OUT PATIENTS	25	1	0	5	286	0
	HEALTHY	3837	124	12	405	670	0
BRANDON	DR.S CARE	134379	0	0	586	87850	1
	ACUTE HOSPITAL	57	6618	0	39	4059	93
	EXTENDED HOSP.	23	0	623	14	218	49
	OUT PATIENTS	528	68	20	14519	10758	7
	HEALTHY	87866	4179	279	10719	116796	52
REST MAN.	DR.S CARE	244	0	0	38	3179	0
	ACUTE HOSPITAL	70	4	0	18	312	3
	EXTENDED HOSP.	0	0	0	0	0	1
	OUT PATIENTS	167	131	7	43	1078	0
	HEALTHY	41242	2101	305	2596	2801	0
OUT PROV.	DR.S CARE	3	0	0	1	781	0
	ACUTE HOSPITAL	4	1	0	0	66	0
	EXTENDED HOSP.	0	0	0	0	0	0
	OUT PATIENTS	0	1	0	0	251	0
	HEALTHY	174	15	0	15	6	0

TABLE 4.1.1 (CONTINUED) ABSOLUTE TRANSITION MATRIX

REST OF MANITOBA

FROM	TO	DR.CARE	ACUTE	EXTEND	OUT-PAT	HEALTHY	DEAD
WINNIPEG	DR.S CARE	972	0	0	424	227987	1
	ACUTE HOSPITAL	173	2	0	240	15786	432
	EXTENDED HOSP.	1	0	0	2	40	25
	OUT PATIENTS	146	10	0	73	11695	21
	HEALTHY	19116	1378	6	6798	8659	0
BRANDON	DR.S CARE	214	0	0	232	41295	0
	ACUTE HOSPITAL	24	0	0	24	2115	62
	EXTENDED HOSP.	6	0	0	8	191	102
	OUT PATIENTS	20	11	0	16	2655	4
	HEALTHY	3187	308	1	1068	1602	0
REST MAN.	DR.S CARE	781196	0	0	8386	748090	3
	ACUTE HOSPITAL	1046	63438	1	616	58327	1167
	EXTENDED HOSP.	16	1	576	5	496	50
	OUT PATIENTS	5196	6334	29	135713	114749	141
	HEALTHY	750699	55490	540	117741	1275816	370
OUT PROV.	DR.S CARE	38	0	0	8	13197	0
	ACUTE HOSPITAL	14	0	0	1	1106	21
	EXTENDED HOSP.	0	0	0	0	0	0
	OUT PATIENTS	34	4	0	11	2752	0
	HEALTHY	219	13	0	41	21	0

OUT OF PROVINCE

		DR.CARE	ACUTE	EXTEND	OUT-PAT	HEALTHY	DEAD
WINNIPEG	DR.S CARE	131	0	0	21	1036	0
	ACUTE HOSPITAL	18	0	0	3	70	0
	EXTENDED HOSP.	1	0	0	0	0	0
	OUT PATIENTS	64	1	0	5	65	0
	HEALTHY	18980	1160	0	3810	25912	0
BRANDON	DR.S CARE	4	0	0	2	177	0
	ACUTE HOSPITAL	1	0	0	0	16	0
	EXTENDED HOSP.	1	0	0	0	0	0
	OUT PATIENTS	0	0	0	1	18	0
	HEALTHY	787	68	0	243	2030	0
	DEAD	0	0	0	0	0	0
REST MAN.	DR.S CARE	49	0	0	38	236	0
	ACUTE HOSPITAL	7	0	0	0	19	0
	EXTENDED HOSP.	0	0	0	0	0	0
	OUT PATIENTS	11	0	0	5	40	0
	HEALTHY	13208	1157	0	2747	18039	0
OUT PROV.	DR.S CARE	39686	0	0	343	6077	0
	ACUTE HOSPITAL	56	2918	0	40	455	0
	EXTENDED HOSP.	0	0	0	0	0	0
	OUT PATIENTS	268	79	0	8190	970	0
	HEALTHY	6100	464	0	933	9101	0

The off-diagonal flows within a region but between health care states are difficult to interpret. Notice that no transitions are reported from Drs. Care to Acute Hospital care. This does not indicate that Manitoba doctors did not send anyone directly to hospital after seeing them, it means only that during the editing process a patient was considered to be in the Acute Hospital care state on his discharge day but not on his admit day. Despite this, doctors' contacts on his admit day were assumed to be given in the hospital, and thus were eliminated to be consistent with the priority scheme[3.3.1]. Thus direct referrals to hospital are recorded in the model statistics as a transition from Drs. Care to Healthy, and then a transition from healthy to Acute Hospital. Thus these transitions in no way reflect the nature of the referral process being practiced by the medical profession. The most interesting and readily interpretable of the within-region flows are the transitions to the various "death" states. Because death is an absorbing state (the ultimate absorbing state), transitions into a death state represent an individual's death. Thus the absolute transitions measure the number of deaths from each of the health care facilities. For example, 2,088 people died while in acute care hospitals in Winnipeg, a little more than 64% of those Winnipegers who died in 1972. Table 4.1.2 summarizes these mortality statistics in terms of numbers of deaths in each facility, percentage of total deaths, and death rates per 1,000 population. These agree fairly well with the vital statistics for 1972 [Statistics Canada publication CS84-206], in the aggregate at least. Mortality rates such as these have traditionally been used as a measure of a population true health or of its need for additional health care. The weaknesses of the approach (and various other approaches) have been pointed out in the

literature by Sanders [1964], Moriyama [1968] and Goldsmith [1972] among others. Nevertheless, mortality rates are surely a direct and important measure of the perfomance of the health care systems, and rates which are specific to medical facility should give further insight into this problem.

The cross regional flows as measured by the absolute transition matrix are of primary interest when addressing the question of equality of access to health care raised in Chapter 1. The form presented in table 4.1.1 is not the best for examining these flows. What we are interested in is the total flow from one region, say Brandon, to a particular health care state in another region, say Drs Care in Winnipeg. A better form for examining these flows is that of table 4.1.3.1, in which the transitions have been summed over health care states to give flows from all states in each region to the Dr.'s care facilities in each region. In addition to the problems mentioned above for separations from Doctor's care this table suffers one additional problem which influences its interpretation. The flows represent the location of the individual on the day before his/her visit(s) to the doctor, not the location of that individual's residence. Thus an individual who lives in Brandon and goes to an out-patient facility in Winnipeg on a particular day, followed on the next day by a visit to his/her local physician, will be counted as a Winnipeg to Brandon flow. However, the probability of such a set of transactions on successive days is so small relative to the probabilty of an "actual" flow that this effect can be ignored. Tables 4.1.3.2 and 4.1.3.3 give similar regional flows for the facility acute care hospitals and out-patient clinic respectively.

MORTALITY STATISTICS TABLE 4.1.2

DEAD IN: / FROM:	DRS. CARE	ACUTE HOSPITAL	EXTENDED HOSPITAL	OUT PATIENT	HEALTHY	TOTAL
WINNIPEG	2	2149	122	231	748	3252
row %		66.1	3.8	7.1	23.0	
/1000P	0.00	4.00	0.23	0.43	1.39	6.04
BRANDON	2	101	50	7	52	212
row %	0.9	47.5	23.6	3.3	24.5	
/1000P	0.06	2.85	1.41	0.20	1.47	6.00
REST MAN	4	1682	177	166	370	2399
row %		70.1	7.4	6.9	15.4	
/1000P	0.01	4.00	0.42	0.39	0.88	5.69
TOTAL	8	3932	349	404	1170	5863
row %		67.1	6.0	6.9	20.0	
/1000P	0.01	3.95	0.35	0.46	1.18	5.90

DR.S'CARE FLOWS TABLE 4.1.3.1

FROM	TO DR.S'CARE IN : WINNIPEG	BRANDON	REST MAN	OUT PROV	TOTAL
Winnipeg	1751306	3993	20408	19194	1794901
row %	97.6	0.2	1.1	1.1	
col %	88.4	3.0	2.6	48.4	61.0
/capita	3259	7.4	37.9	35.6	3.34
Brandon	3089	88474	3351	793	95707
row %	3.2	92.5	3.5	0.8	
col %	0.2	65.7	0.4	0.2	3.3
/capita	872	2495	946	224	2.70
Rest Man	229311	41723	756957	13275	1041266
row %	22.0	4.0	72.5	1.3	
col %	11.1	31.0	96.8	33.4	35.4
/capita	545	991	1798	314	2.48
Out Prov	1227	181	305	6424	8137
row %	15.1	2.2	3.7	79.0	
col %	0.1	0.1	0.0	1.6	0.3
/capita					
Total	1984933	134371	781021	39686	2940111
row %	67.5	4.6	25.6	13.5	

ACUTE CARE HOSPITAL FLOWS (ADMISSIONS) TABLE 4.1.3.2

FROM	TO ACUTE HOSPITAL CARE IN : WINNIPEG	BRANDON	REST MAN	OUT PROV	TOTAL
Winnipeg	68711	125	1390	1161	71387
row %	96.4	0.2	1.9	1.6	45.4
col %	81.8	1.9	2.2	39.7	
/1000P	126.3	0.2	2.6	2.2	132.2
Brandon	391	4247	319	68	5025
row %	7.8	84.4	6.4	1.4	
col %	0.5	64.0	0.5	2.3	3.2
/1000P	11.1	119.9	9.0	1.9	142.0
Rest Man	14978	2236	61824	1157	80195
row %	18.7	2.8	75.9	1.4	
col %	17.8	33.7	97.3	39.5	51.0
/1000P	35.6	5.1	147.0	2.5	190.5
Out Prov	57	17	17	543	634
row %	8.9	2.7	2.7	85.6	
col %	0	0	0	18.5	0.4
/1000P					
Total	84137	6625	63550	2929	157241
	53.5	4.2	40.4	1.9	

OUT-PATIENT FLOWS TABLE 4.1.3.3

FROM	TO OUT-PATIENT CARE IN : WINNIPEG	BRANDON	REST MAN	OUT PROV	TOTAL
Winnipeg	131253	422	7537	3839	143051
row %	91.7	0.3	5.3	2.7	
col %	90.2	2.9	5.6	24.9	46.0
/1000P	243.0	.8	13.9	7.1	265.9
Brandon	336	11358	1348	246	13288
row %	2.5	85.5	10.1	1.9	
col %	0.2	78.4	1.0	1.6	4.3
/1000P	9.5	322.0	38.2	7.0	375.4
Rest Man	13709	2695	126748	2790	145942
row %	9.4	1.8	86.8	1.9	
col %	9.4	18.6	93.4	18.1	46.9
/1000P	32.6	6.4	301.0	6.6	346.6
Out Prov	186	16	61	8563	8826
row %	2.1	0.2	0.7	97.0	
col %	0.1	0.1	0.0	55.6	2.8
/1000P					
Total	145484	14491	135694	15438	311107
row %	46.8	4.7	43.6	5.0	

POPULATION MOVEMENTS TABLE 4.1.4

FROM	TO HEALTHY STATE IN : WINNIPEG	BRANDON	REST MAN	OUT PROV	TOTAL
Winnipeg		670	8659	25912	35241
row %		1.9	24.6	73.5	
/1000P		1.2	16.1	48.2	65.5
Brandon	816		1602	2030	4448
row %	18.4		35.2	45.8	
/1000P	23.1		45.3	57.3	125.6
Rest Man	12163	2801		18039	33003
row %	36.8	8.5		54.7	
/1000P	28.9	6.7		42.8	78.4
Total	12979	3471	10261	45981	72692
row %	17.9	4.8	14.1	63.3	

Examining these tables carefully we discover some mildly surprising things. Most surprising, perhaps, is the usage of Brandon Dr.'s care by those in Winnipeg exceeds, in absolute terms, usage of Winnipeg Dr.'s care by Brandon residents (3,993 to 3,089). This is a population effect, as can be seen by examining either the percentage of total doctor's care this represents (3.2% to 0.2%) or the rates per 1,000 population (87.2 to 7.4) also contained in the table. Thus Brandon residents do use Winnipeg doctors more heavily than Winnipeg residents use Brandon doctors. However, comparing the usage of Brandon residents of doctors in the Rest of Manitoba with that of Winnipeg doctors (3.5 to 3.2%) would seem to indicate that this rate of use is a function of the characteristics of Brandon's population and is not due to lack of access to health care facilities.

In contrast with this behavior are the rather massive flows observed from the rest of Manitoba region. This population appears to get about 22% of its physician's care, 19% of its acute hospital care and 10% of its out-patient care from facilities and doctors located in Winnipeg. The doctor's care figure may be somewhat inflated because doctors who normally practice in Winnipeg do travel to Northern Manitoba to treat patients. Since the location of a doctor service was determined by the location of the doctors' normal practice these transactions will be represented as flows from the rest of Manitoba to Winnipeg doctors' care. However, even if this problem does distort the direction of the flow, it still seems to indicate a lack of locally available physicians in the region outside of Winnipeg and Brandon. These figures will also be distorted somewhat by the omission of data on public health nursing care. A finer regional breakdown of these substantial flows would be interesting, as it would enable the separation of flows generated close to Winnipeg (result-

ing from a failure of our definition of Winnipeg to correspond to the health care catchment area of Winnipeg) and those from substantial distances away.

Another interesting property to be learned from these tables is the difference in the overall usage pattern for those in the rest of Manitoba as compared to Winnipeg. Doctor's care is a much more important component of the health service system to those who live in Winnipeg than to those in the rest of Manitoba, as indicated by the per capita utilization rates (Winnipeg-3.34, Rest of Man.-2.48). But the reverse is true for acute care hospital admissions (132 per 1,000 in Winnipeg to 191 per 1,000 in Rest of Man.) Thus there appears to be a substitution of hospital care by those in the Rest of Manitoba region for services which are performed by doctors for those living in Winnipeg. This is supporting evidence for the hypothesis that the Rest of Manitoba lacks the access to physicians that is enjoyed elsewhere in the province. However, it is unreasonable to expect that this would not be so. As long as the relative population densities of these two regions are so vastly different, it must be expected that specialists services will tend to be concentrated in Winnipeg. This centralism is to be expected for all kinds of services, not just health care. The question of interest is whether or not the health care delivery system in Manitoba exhibits this central tendency to an excessive degree.

The movement of people (changes in residence) should reflect the degree of centralism to be expected in the economy as a whole. Table 4.1.4 gives the healthy to healthy state transitions and their respective rates per 1,000 population. Notice that the net flows are into Winnipeg as expected, but the relative flow rate for Rest of Manitoba/Winnipeg are not as large as those for the various

health care facilities. This would seem to indicate a greater degree of centralism in health care delivery than in the economy of the province as a whole. This interpretation depends, of course, on the algorithm used to determine movements [3.3.1], which is suspect due to late reporting (and non-reporting) of movements as well as the elimination of multiple moves and transient population. Further complicating this interpretation is the high technology of health care in comparison to the rest of the economy. Some degree of greater concentration can be expected for industries and services where high technology, and therefore, a large amount of capital input, is used. However, there is certainly evidence that the two regions have different levels of access to the various health care facilities. If equality of access [1.2] is to be a goal of the health care system then further study seems indicated to determine the real extent of the problem and determine methods to improve access in the Rest of Manitoba region.

4.1.2 Length of Stay

Table 4.1.5 reports the observed average length of stay in each of the non-absorbing states, and the standard deviation of the length of stay, for each age/sex cohort, over all stays in each state in 1972. These figures suffer from the same kinds of interpretation problems as those mentioned above for absolute transitions. The subject population is not the serviced population. Indeed, those born in 1972 or who moved to Manitoba in 1972, and transients, have been removed prior to the calculation of these statistics. In addition, the definitions of the health care state "Drs. Care" and "Out-patient" make the length of stay in those states essentially deterministic with mean one day. This problem also distorts the "Healthy" state length of stay distribution by counting people as healthy who are simply between doctor's visits, waiting to get into hospital, and, in general, <u>not</u> healthy. However,

the two hospital care states should have defined and measured lengths of stay distribution, and the remaining statistics are useful for examining the assumptions of the Markov model [2.1.4].

The age and sex variation of mean length of stay in each of the health care states follows the expected trends. In general, the older the populaton the greater its need for health care, and, therefore, the longer the average length of stay. A notable exception are the very young, here those between one and two years of age, who require significantly more care than the 2 to 15 year olds and thus have longer lengths of stay in hospital on average. The sex effects are less uniform. Females in the child bearing years have shorter average length of stay in hospital, probably because of the large number of maternity related admissions which will have relatively short stays associated with them. Females over 65 years of age spend a longer time on average in the hospital than males counterparts while the opposite is true in the 46-65 year old cohort. The younger cohcrts do not seem to have consistant sex effects across all regions, and the effect appears small in comparison to the older groups.

Tables 4.1.5 and 4.1.6 are calculated using all stays in each state which occured in 1972, including those which started in 1971, and those which extended into 1973. Thus some of the stays are truncated by the failure to include those parts which did not occur in 1972.

LENGTH of STAY DISTRIBUTIONS (MEAN/STANDARD DEVIATIONS) TABLE 4.1.5

FACILITY		MALES					
		1	2-15	16-45	46-65	65+	MANITOBA
WINNIPEG	DRS. CARE	1.02	1.03	1.06	1.14	1.49	1.09
		.510	.987	.630	2.34	66.79	2.07
	ACUTE	6.03	5.510	9.19	13.45	18.93	10.88
	HOSPITAL	7.21	8.23	13.3	16.93	23.65	16.43
	EXTENDED	-	120.0	122.0	25.32	61.45	55.65
	HOSPITAL	-	159.4	134.9	36.46	73.45	63.79
	OUT	1.02	1.02	1.04	1.09	1.11	1.04
	PATIENTS	.166	.189	.359	.560	.523	.347
	HEALTHY	49.70	100.3	99.0	73.22	54.43	75.47
		56.4	107.2	115.5	102.33	84.64	97.87
BRANDON	DRS. CARE	1.02	1.02	1.05	1.05	1.24	1.10
		.332	.256	.502	.525	3.0	1.81
	ACUTE	5.48	3.79	6.83	11.31	13.80	8.51
	HOSPITAL	4.91	3.84	7.52	10.57	11.20	8.70
	EXTENDED	-	-	43.9	58.02	63.91	66.37
	HOSPITAL	-	-	46.4	68.81	65.58	68.54
	OUT	1.01	1.03	1.08	1.23	1.25	1.09
	PATIENTS	.113	.297	.898	2.86	1.28	1.10
	HEALTHY	52.75	112.3	108.6	82.22	58.27	83.76
		63.86	113.9	118.8	108.52	87.76	103.2
REST MAN	DRS. CARE	1.06	1.03	1.05	1.06	1.18	1.06
		.997	.414	1.48	1.60	3.77	1.83
	ACUTE	6.30	4.37	4.96	7.91	11.44	7.08
	HOSPITAL	5.21	5.23	5.75	8.79	13.84	9.29
	EXTENDED	-	-	14.5	29.34	35.59	38.27
	HOSPITAL	-	-	15.4	46.45	49.35	57.66
	OUT	1.05	1.03	1.09	1.13	1.17	1.08
	PATIENTS	.343	.284	.574	.844	2.76	.878
	HEALTHY	57.66	120.4	109.6	86.16	63.34	87.37
		71.67	122.4	120.9	111.84	92.80	107.68

LENGTH of STAY DISTRIBUTIONS (MEAN/STANDARD DEVIATIONS) TABLE 4.1.5

FACILITY		FEMALES					
		1	2-15	16-45	46-65	65+	MANITOBA
WINNIPEG	DRS. CARE	1.03	1.03	1.06	1.07	1.10	1.09
		.437	.344	.865	1.84	1.50	2.07
	ACUTE	6.837	5.58	6.52	12.91	20.39	10.88
	HOSPITAL	7.83	8.72	8.36	15.83	27.22	16.43
	EXTENDED	-	-	59.2	53.09	56.15	55.65
	HOSPITAL	-	-	92.5	52.0	54.42	63.79
	OUT	1.02	1.02	1.03	1.04	1.02	1.04
	PATIENTS	.186	.163	.289	.302	.181	.347
	HEALTHY	52.17	102.4	65.8	59.57	50.01	75.47
		58.7	108.6	85.7	85.42	77.26	97.87
BRANDON	DRS. CARE	1.03	1.02	1.07	1.09	1.31	1.10
		.698	.175	1.80	2.76	2.10	1.81
	ACUTE	5.36	3.31	6.06	10.88	12.80	8.51
	HOSPITAL	4.16	3.77	5.36	9.38	9.84	8.70
	EXTENDED	-	-	23.3	58.07	78.15	66.37
	HOSPITAL	-	-	21.8	65.88	74.04	68.54
	OUT	1.03	1.03	1.06	1.13	1.06	1.09
	PATIENTS	.210	.258	.660	.622	.510	1.10
	HEALTHY	54.15	117.2	73.0	67.01	54.79	83.76
		60.90	116.4	89.6	91.17	81.09	103.2
REST MAN	DRS. CARE	1.06	1.02	1.03	1.05	1.16	1.06
		1.05	.325	.491	1.03	3.91	1.83
	ACUTE	6.84	4.22	4.94	7.73	12.09	7.08
	HOSPITAL	7.62	4.56	3.95	8.42	15.47	9.29
	EXTENDED	-	-	18.2	23.93	50.58	38.27
	HOSPITAL	-	-	23.9	46.50	69.37	57.66
	OUT	1.03	1.03	1.08	1.10	1.06	1.08
	PATIENTS	.390	.276	.656	.779	.432	.878
	HEALTHY	62.43	122.6	71.89	66.43	56.80	87.37
		77.10	123.1	90.90	90.73	82.29	107.68

LENGTH OF STAY BY REGION OF ORIGIN : ACUTE HOSPITALS

TABLE 4.1.6

REGION of ORIGIN	REGION OF SERVICE WINNIPEG	BRANDON	REST MAN	OUT PROV
Winnipeg	10.60 (16.68) 65750	7.58 (7.06) 176	5.57 (6.53) 2103	7.86 (9.89) 1427
Brandon	11.79 (15.37) 466	8.29 (8.67) 3979	5.67 (6.66) 479	8.83 (10.94) 92
Rest Man	11.88 (15.45) 17912	8.92 (8.84) 2470	7.14 (9.39) 60963	7.65 (9.75) 1378
Manitoba	10.88 (16.43) 84137	8.51 (8.70) 6625	7.08 (9.29) 63551	7.78 (9.96) 2929

Rows are : mean - length of stay

(std. dev.)- length of stay

n - number of stays

NOTE: Region numbers do not add to Manitoba figures because of the ommission of Out of Province intial region.

There does appear to be a strong regional effect on hospital length of stay. Table 4.1.6 demonstrates this effect clearly by separating the length of stay distribution statistics by region of residence on Jan 1, 1972 and by the region where the care was obtained. Here we can notice that the average length of stay in Winnipeg acute care hospitals is much greater than that in either Brandon or Rest of Manitoba hospitals, regardless of the original residence of the patient. Similarly, Brandon acute care hospitals exhibit longer average length of stay than those in the Rest of Manitoba, but the differences are less severe. This seems to indicate that medical practices, as they effect the duration of hospital stays, vary with the location of the hospital and its characteristics, rather than with the characteristics of the population served. Alternatively one could interpret table 4.1.6 as evidence of greater demand for hospital space in rural areas, generating a need for faster turnover of beds and thus shorter average lengths of stay. Given the already noted [4.1.1] greater per capita utilization of acute care hospitals by those in the Rest of Manitoba region and lower uilitization of Drs. Care a combination of differences in treatment patterns and greater pressure for beds probably explains this difference.

One possible use for the mean length of stay in the healthy state is as a measure of the overall health of the population. It represents the average time between health care transactions (or moves) which are of different types or are non-consecutive. The major problem in using these figures as an index of a populations' health is that a large number of stays in the healthy state were truncated, that is begun before the beginning of the year or extended beyond the end of the year or both. This plus the problem of

defining healthy as not experiencing a health care transaction on this day, both cause the calculated average length of stay in the healthy state to be much smaller than it is in reality. If the effects of these two problems are uniform across the populations we wish to compare then the statistics would still be of use as an index of relative health. Table [4.1.7] gives insight into the reliability of the assumption of uniform effect by examining the age/sex/region distribution of people who had no health care transactions in 1972, and did not move between regions in that year. If the truncation effect is to be uniform then these health care inactive populations should represent a fixed proportion of the respective total cohort population. Unfortunately, there appears to be strong variations in the proportion of the population inactive in all three of the separating variables: age, sex, and region. Thus there is little hope that the effect of truncation on the average length of stay in the healthy state is uniform enough for it to be useful as a measure of the relative health of the populations. For these reasons, the use of mean length of stay in the healthy state as an indicator of population health is totally inappropriate.

One further note on table 4.1.7. The larger proportion of inactives in the Brandon and Rest of Manitoba regions is evidence in support of the hypothesis that people in those regions have less access to health care. Of course, the same effect would result if the populations in the rural areas were basically "healthy", needing less care than their Winnipeg counterparts.

HEALTH CARE INACTIVE POPULATIONS TABLE 4.1.7

MALES	AGE 1	2-15	16-45	46-65	65	TOTAL
Winnipeg	186	15645	33745	13699	4623	67898
of	4628	68475	116507	51777	22375	263762
%	4.0	22.8	29.0	26.5	20.6	25.7
Brandon	19	1211	2315	984	351	4880
of	297	4705	7569	3441	1662	17674
%	6.4	25.7	30.6	28.6	21.1	27.6
Rest Man.	284	21155	26869	13040	5098	66446
of	3872	63655	86075	42256	21510	217368
%	7.3	33.2	31.2	30.9	23.7	30.6
FEMALES						
Winnipeg	192	15335	15285	10340	5178	46330
of	4472	65701	117410	58017	29565	275165
%	4.3	23.3	13.0	17.8	17.5	16.8
Brandon	13	1236	995	735	387	3366
of	292	4419	7267	3657	2106	17741
%	4.5	28.0	13.7	20.1	18.4	19.0
Rest Man.	363	20554	12424	8031	3985	45357
of	3743	60906	78820	39523	21006	203998
%	9.7	33.7	15.8	20.3	19.0	22.2

Returning to the length of stay distributions [table 4.1.5] it is important to note that the truncation problems discussed above for the healthy state length of stay are likely to distort the calculated lengths of stay for any state where the actual length of stay is large with respect to the period of observations, one year in this case. Thus the lengths of stay in extended care hospitals are probably somewhat larger than those reported.

One last set of useful statistics closely related to the length of stay distribution statistics are the per capita "utilization" or "participation" rates of the various health care facilities. These are obtained by dividing the total sojourn time figures, D, by the number in the subject population. Since these rates permit forecasts of populaton to be converted into forecasts of health care demand, these rates are particularly interesting. Table 4.1.8 gives these utilization rates for each of the health care facilities in each region, in days per capita. The use of these rates as a forecasting or planning tool will be discussed in chapter 5, and an alternative means of estimating these rates using the Markov model in section 4.2. A full discussion of these numbers is postponed until later. It suffices at this point to note that, since almost everyone is in the "healthy" state at the beginning of the year, these participation rates correspond almost exactly to the healthy state residency times[2.1.5].

TABLE 4.1.8 PER CAPITA FACILITY USAGE

		DOCTORS			HOSPITALS			OUT-PATIENT		
		WPG	BRDN	ROM	WPG	BRDN	ROM	WPG	BRDN	ROM
WPG AGE 1	M	5.0261	.0080	.0983	1.1409	.0028	.1054	.6221	.0013	.0227
	F	4.7838	.0197	.1221	.9680	.0018	.0783	.5143	.0020	.0268
2-15	M	2.1207	.0041	.0339	.3207	.0004	.0112	.3202	.0007	.0153
	F	2.1196	.0041	.0322	.2808	.0003	.0111	.2419	.0005	.0127
16-45	M	2.2869	.0097	.0542	.5115	.0019	.0137	.2249	.0022	.0266
	F	4.0528	.0139	.0707	1.1190	.0023	.0245	.1748	.0009	.0154
46-65	M	3.7617	.0071	.0334	1.9299	.0011	.0199	.3119	.0013	.0171
	F	4.9464	.0082	.0363	1.6213	.0032	.0133	.1702	.0002	.0105
OVER 65	M	6.8633	.0139	.0638	5.8963	.0126	.0415	.5791	.0008	.0129
	F	5.9952	.0190	.0928	4.4052	.0086	.0732	.1873	.0003	.0104
BRDN AGE 1	M	.1616	4.0370	.3367	.0572	.9495	.3872	.0337	.5859	.1077
	F	.1164	3.9897	.4315	.5342	.6267	.2329	.0342	.5103	.0822
2-15	M	.0784	1.3256	.1207	.0425	.2457	.0546	.0125	.3522	.0385
	F	.0864	1.3299	.1184	.0459	.1598	.0247	.0104	.2849	.0253
16-45	M	.1258	1.4810	.1222	.0958	.3532	.0240	.0160	.3977	.0526
	F	.1924	2.9261	.2006	.1031	1.0076	.0731	.0113	.2954	.0380
46-65	M	.1192	2.7422	.1017	.4162	1.3676	.0942	.0209	.3595	.0253
	F	.1444	3.9357	.1184	.1952	1.3227	.0514	.0082	.3112	.0238
OVER 65	M	.1606	4.7509	.1594	.4434	3.3448	.2605	.0138	.3544	.0656
	F	.1377	5.2896	.1467	.2673	2.6462	.2422	.0038	.1880	.0266
ROM AGE 1	M	.9367	.1377	3.0850	.4217	.0455	2.5181	.0876	.0142	.5251
	F	.8055	.1419	2.9073	.3780	.0278	2.2533	.0799	.0139	.4170
2-15	M	.3924	.0619	1.1254	.1510	.0140	.3817	.0347	.0053	.2755
	F	.3880	.0598	1.1360	.1213	.0082	.3343	.0293	.0050	.2223
16-45	M	.4934	.0644	1.2394	.2975	.0205	.3227	.0435	.0115	.3590
	F	.9407	.1573	2.2955	.4455	.0529	.9716	.0334	.0086	.3329
46-65	M	.6641	.1248	1.9550	.8305	.0773	1.0790	.0580	.0084	.3888
	F	.9523	.1897	2.8623	.7789	.0885	1.3213	.0275	.0066	.3482
OVER 65	M	.8289	.2564	3.2063	1.8325	.2172	3.9511	.1026	.0153	.4097
	F	.9812	.3870	3.8332	1.2724	.1418	4.0699	.0318	.0073	.2817

4.2 With the Model

4.2.1 Examining the Assumptions

As noted before [2.1], in order to use the model as a tool for examining health care delivery, we must first determine the extent to which the Markov assumptions are true, and the impact of any deviations from these assumptions on the important statistics. As noted in section 2.1.7, if we are observing a stationary Markov process then the length of stay distributions are exponential distributions with fixed time-independent parameters. This implies that the sample mean and standard deviation have the same expected value and that their ratio should be close to one. Examining table 4.1.5 we observe that the ratio $(\frac{s}{x})$, the coefficient of variation, varies from about 0.16 (female age 1 outpatient care in Winnipeg) to more than 4.0. Without applying any formal statistical procedure, it is clear that the observed length of stay distributions are not exponential distributions as the model predicts they should be.

Most of the problem probably stems from the failure of the state definitions to fully summarize the information available in a person's health care history. The eposidic nature of sickness, referral practices, and cyclical or periodic treatment patterns will, for example, cause deviations from the model assumptions and distort the exponential lengths of stay they imply. Lack of homogeneous groups with regard to their health care behavior, and a certain amount of lack of independence between individuals (for example within family) may also be a cause of the observed deviations from Markov behavior. Finally, it is certainly possible that the Markov type model is totally inappropriate as a model of health care delivery in Manitoba. However, it is hoped that the following material will demostrate that,

despite these obvious violations of the Markov assumptions, useful and interesting conclusions may be drawn from the model and its related statistics.

4.2.2 Daily Transition Probabilities

The daily transition probabilities of the imbedded Markov chain can be estimated in two ways: by the maximum likelihood method suggested by Zahl [1955] [3.2.2], or from the series approximation [E2.1.15] using the estimated matrix, $\hat{\Gamma}$, with t=1. In the first case, the estimates are simply the proportion of the total number of days spent in a state i where the state occupied on the following day is j. Thus they correspond with the intuitive notion of the conditional probabilities of being in a particular state j tomorrow given that the individual occupies state i today. The second method does not have as clear an intuitive meaning, and depends heavily on the estimated matrix and on the model assumptions.

Table 4.2.1.1 gives the within Winnipeg transition probabilities estimated by Zahl's method, and table 4.2.1.2 the estimates computed via equation E2.1.12 and the similarity transformation of section 3.4. Notice that these matrices are quite different, especially for the Doctor's care and Outpatient care states. The problem is with the probabilities estimated from the rates and arises from the nearly deterministic length of stay in these states. An exponentially distributed random variable with a mean of one day has a probability of exceeding one day, and thus being in the same state on successive days of $(e^{-1})=.369$. The maximum $\hat{\gamma}_i$ rate that can be estimated by equation E3.2.8 given daily transition data is 1.0 because the total sojourn time in a state must be at least as large as the number of transitions into that state. Therefore the minimum one day transition probability between a state and itself that can be estimated from the rate matrix is approximately .369,

while the actual transition probabilities probably correspond more closely to the ~.08 and ~.04 estimate obtained from Zahl's method. This problem, when combined with the row sum constraint, distorts the entire matrix of probability estimates. To correct this problem, the data would have to indicate the time of day of the medical transaction, thus permitting more realistic lengths of stay in short duration states such as doctors care to be estimated, and permitting multiple medical transactions in a single to be recorded. The number of separations from a state could then exceed the total number of days spent in that state. Rates greater than one would be possible and more realistic probabilities would likely result from the series approximation. Our data did not permit this to be done. Greater ease of interpretation for many of the statistics would have resulted if multiple transactions on a single day had not been eliminated, thereby greatly increasing the utility of the model.

TABLE 4.2.1.1 One Day Transition Probabilities

M.L.E. ESTIMATES: WINNIPEG ONLY

TO: / FROM:	DR.S' CARE	ACUTE HOSPITL	EXTEND HOSPITL	OUT-PATIENT	HEALTHY	DEAD
DR.S CARE	.0847	-	-	.0049	.9092	-
ACUTE H.	.0018	.9058	-	.0008	.0882	.0030
EXT. HSP.	.0004	-	.9822	.0017	.0086	.0070
OUT-PAT.	.0616	.1074	.0002	.0443	.7823	.0016
HEALTHY	.0092	.0003	-	.0006	.9895	-

Since the series formula with t=1 gives such poor results, the same problem can be expected for larger values of t as well. Thus t day transition probabilities are better estimated by powering the one day matrix, P(1),

TABLE 4.2.1.2 One Day Transition Probabilities

SERIES ESTIMATES: WINNIPEG ONLY

TO: / FROM:	DR.S' CARE	ACUTE HOSPITL	EXTEND HOSPITL	OUT-PATIENT	HEALTHY	DEAD
DR.S CARE	.4027	.0002	-	.0021	.5941	-
ACUTE H.	.0014	.9102	-	.0005	.0847	.0028
EXT. HSP.	.0003	-	.9823	.0011	.0091	.0070
OUT-PAT.	.0263	.0657	.0001	.3842	.5210	.0012
HEALTHY	.0060	.0003	-	.0004	.9929	-

obtained from E3.2.12 :

$$P(t) = (P(1))^{t} \qquad \text{(E4.2.1)}$$

However, since for large values of t, such as t=366 or one year, the transition probabilities can be expected tc be close to the limiting overall proportions of the population in each state on an average day, there is some hope that the series approximation may give reasonable results for large t in spite of the inaccuracies involved in the estimates of P(1).

In interpreting the daily transition probabilities, P(1), many of the same observations made earlier about the absolute transition matrix, in particular the per capita rates derived there from [4.1.1] can be made again. Here the information is already in per capita form, but knowledge of the number of people is not required to compute the estimates. One may examine inter-region flows by examining the conditional probabilities of using a facility in a particular area, given that the facility is to be used. These probabilities are easily calculated from the one day transition probabilities by summing the probabilities of

being in the facility tomorrow over all regions and taking the ratio of each of the respective regional probabilities to this sum. For example the conditional probability that an individual who is healthy in the rest of Manitoba today will go to a doctor in Winnipeg tomorrow, given that he will go to the doctor somewhere tomorrow can be computed as:

$$\Pr(\text{Rest of Man.\&Healthy} \rightarrow \text{Winnipeg Doctor} \mid \text{goes to a doctor})$$

$$= \Pr(\text{Rest of Man.\&Healthy} \rightarrow \text{Winnipeg Doctor}) \div$$
$$[\Pr(\text{Rest of Man.\&Healthy} \rightarrow \text{Winnipeg Doctor}) +$$
$$\Pr(\text{Rest of Man.\&Healthy} \rightarrow \text{Brandon Doctor}) +$$
$$\Pr(\text{Rest of Man.\&Healthy} \rightarrow \text{Rest of Man. Doctor}) +$$
$$\Pr(\text{Rest of Man.\&healthy} \rightarrow \text{Out of Prov. Doctor})] \doteq 0.23$$

Which says that between one in four and one in five doctors visits made by those in the Rest of Manitoba are to Winnipeg doctors. This agrees with the large flows noted earlier. However, because the only advantage to this approach over examining the absolute transitions is that the populations need not be known, and because of time limitations, this approach was not pursued further.

4.2.3 The Rates Γ

The rate estimates computed from the continuous estimation procedure [3.2.1] are essentially per capita flow rates based on the population (average) of the particular state from which the flow originates. They can be examined for relative size in much the same way as the per capita absolute transitions [4.1.1] and daily transitions probabilities [4.2.2], and the same conclusions drawn. Again, no knowledge of the total populations is required to compute the rates or draw conclusions about their relative sizes.

If we assume that the model assumptions hold sufficiently that transitions from a state are essentially multinomially distributed [2.1.4] then the conditional probabilities (γ_{ij} / γ_i) that are individuals next state will be j given his current state in i may be computed and interpreted. With

our data, and state definitions, no surprises occurred here. Drs. Care is the most likely next state for a healthy person, followed by outpatient care, acute hospital care, extended hospital care, and death, in that order. Care received within some region is preferred to that obtained elsewhere. Those in any health care state are not likely to be "Healthy" next, this being true almost by the definition of the healthy state. One possibly interesting result of examining these conditional probabilities is that extended care hospitals appear to have different functions in Winnipeg as compared to the Rest of Manitoba. A person in a Winnipeg extended care facility is almost as likely to have "death" as his next state as he is to have "Healthy". This seems to be different from the behavior of extended care facilities in the Brandon and Rest of Manitoba regions. However, the number of observations involved is quite small and the possiblities of reporting differences between regions is quite large, so this may be a spurious conclusion.

Care must again be exercised in interpreting the magnitudes of the rates. The process of estimation using (E3.2.8) and the definitions of the states have constrained the rates such that the row sum of the off diagnal elements, (γ_i), must be less than one. This is a consequence of the fact that multiple transactions on one day were eliminated, so that the minimum length of stay in any state is one day. Thus the total number of separations from a particular state cannot be greater than the number of days spent in that state. The impact of this constraint on the estimates of the daily transition probabilities was discussed in the previous section [4.2.2]. Of course, all of the inaccuracies discussed in [4.1] with regard to the absolute transitions and length of stay statistics are inherited by the rate estimates as well.

4.2.4 Residency

Perhaps the most useful statistic which can be derived from the model are the expected residency times [2.1.5]. With t=366 days or one year, the estimated expected residency gives a concise picture of how much health care each individual in the society can expect to use in that year and where he is likely to receive it. Also of potential interest are the effects of an individual's initial health care status on his own use of the various health care services in the coming year.

Both observed residency and its estimate derived from the rate matrix were computed, along with the standard deviation of the observed residencies. Table [4.2.2.1] gives the observed residency matrix and the standard deviations. Jan. 1 was an unfortunate choice as an intial day, as far as interpretation of these residencies is concerned, because it is very different from a normal day from the point of view of the use of health facilities. This choice, and an error in classification of those in hospital on the first day of the year, caused a severe distortion in the initial state occupancy as compared to its everyday value [table 4.2.3]. As a result, no one was recorded as being in a hospital care state (either acute or extended) on the initial day, and the residency times for these states could not be estimated. Similarly, the Drs. Care and Outpatient care facilities are almost unused on Jan. 1, compared to their average occupancy, and the observed residency figures are thus distorted. One must be very sick to see a doctor or go to a hospital on New Year's Day.

TABLE 4.2.2.1 OBSERVED RESIDENCY MATRIX

		DOCTORS' CARE	ACUTE HOSPITAL	WINNIPEG EXTENDED HOSPITAL	OUT PATIENTS	HEALTHY	DEAD
WINNIPEG	DR.S CARE	79.327	5.899	0.000	3.283	222.514	5.549
		129.7201	16.3599	0.0000	8.7579	161.3666	44.9982
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	6.150	3.830	0.035	2.489	307.132	6.401
		17.8063	12.6787	0.6983	8.7811	117.4579	43.0109
	HEALTHY	3.411	1.287	0.030	0.244	347.283	0.842
		5.5779	8.2085	2.0834	1.0193	58.2760	15.6209
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	0.520	0.640	0.000	0.080	20.860	0.000
		1.5905	3.6153	0.0000	0.5600	76.1567	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.271	0.792	0.000	0.042	22.583	0.000
		0.7286	5.4274	0.0000	0.1998	87.4661	0.0000
	HEALTHY	0.131	0.154	0.000	0.013	4.752	0.000
		1.0143	2.7622	0.0000	0.1866	34.6040	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	0.796	1.318	0.000	0.078	23.310	0.000
		2.2386	5.2904	0.0000	0.3577	83.1811	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	1.029	0.737	0.000	0.101	25.068	1.182
		4.7457	3.7545	0.0000	0.5342	87.2040	20.7071
	HEALTHY	0.655	0.504	0.010	0.041	6.447	0.007
		2.5033	4.7678	1.0984	0.4843	41.6035	1.1565
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
OUT PROV.	DR.S CARE	5.893	1.214	0.000	1.000	194.321	12.607
		18.7984	4.8870	0.0000	2.0529	160.0946	65.5086
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	2.652	0.391	0.000	0.261	192.261	0.000
		3.7603	1.3429	0.0000	1.0307	173.1308	0.0000
	HEALTHY	0.400	0.000	0.000	0.000	0.000	0.000
		1.2000	0.0000	0.0000	0.0000	0.0000	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 4.2.2.1 OBSERVED RESIDENCY MATRIX (CONTINUED)

BRANDON

		DOCTORS' CARE	ACUTE HOSPITAL	EXTENDED HOSPITAL	OUT PATIENTS	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.007	0.000	0.000	0.002	0.005	0.608
		0.1165	0.0000	0.0000	0.0412	0.0922	14.7224
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.010	0.000	0.000	0.000	0.000	0.000
		0.1995	0.0000	0.0000	0.0000	0.0000	0.0000
	HEALTHY	0.010	0.002	0.003	0.001	0.245	1.2E-04
		0.2757	0.2167	0.8322	0.0469	7.8444	0.0900
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	25.460	2.300	0.000	0.360	187.080	6.300
		58.5799	5.5435	0.0000	0.6248	173.8727	48.4133
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	5.917	2.021	0.000	7.750	296.458	0.000
		8.0463	4.0285	0.0000	21.3420	125.3816	0.0000
	HEALTHY	2.502	0.928	0.434	0.322	335.159	0.763
		5.0371	4.5991	7.6184	1.5090	78.8458	14.5345
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	0.115	0.067	0.911	0.022	2.098	0.000
		0.9778	0.8944	12.2078	0.3250	26.1998	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.179	0.036	0.000	0.013	6.597	0.000
		1.1526	0.4580	0.0000	0.1132	47.1699	0.0000
	HEALTHY	0.126	0.052	0.057	0.008	1.542	0.004
		1.3947	1.0833	2.9375	0.1530	20.6329	1.0078
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
OUT PROV.	DR.S CARE	0.000	0.000	0.000	0.000	0.929	0.000
		0.0000	0.0000	0.0000	0.0000	4.8250	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.000	0.000	0.000	0.000	0.609	0.000
		0.0000	0.0000	0.0000	0.0000	2.8550	0.0000
	HEALTHY	0.300	0.000	0.000	0.000	0.000	0.000
		0.9000	0.0000	0.0000	0.0000	0.0000	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 4.2.2.1 OBSERVED RESIDENCY MATRIX (CONTINUED)

REST OF MANITOBA

		DOCTORS' CARE	ACUTE HOSPITAL	EXTENDED HOSPITAL	OUT PATIENTS	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.240	0.244	0.000	0.072	39.951	1.090
		1.5927	2.1973	0.0000	0.8589	107.9568	15.0221
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.150	0.180	0.000	0.045	34.663	0.000
		1.0487	3.1572	0.0000	0.5218	103.5641	0.0000
	HEALTHY	0.052	0.021	8.4E-04	0.017	2.386	0.002
		0.6566	0.6297	0.3107	0.2924	23.5941	0.5605
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	0.460	0.420	0.000	0.040	117.020	0.000
		1.4030	2.2811	0.0000	0.1960	159.3479	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.042	0.083	0.000	0.104	29.917	0.000
		0.2857	0.5713	0.0000	0.3055	99.2322	0.0000
	HEALTHY	0.142	0.076	0.002	0.038	7.769	0.013
		0.9604	1.3962	0.3033	0.4771	43.8108	1.5452
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	19.427	6.883	0.193	0.757	290.994	7.095
		60.4228	19.3917	3.2458	2.4845	122.5415	45.4082
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	4.120	3.032	0.000	4.429	303.442	5.003
		7.4600	7.7169	0.0000	22.4320	115.2524	39.7573
	HEALTHY	1.874	1.027	0.051	0.321	342.837	0.889
		3.8510	5.5681	2.6469	1.4009	69.0293	15.6299
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
OUT PROV.	DR.S CARE	0.786	1.857	0.000	0.214	46.857	4.786
		3.0042	9.6500	0.0000	1.1135	110.2253	24.8673
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.478	0.087	0.000	0.174	109.739	0.000
		0.9722	0.4079	0.0000	0.8157	165.9390	0.0000
	HEALTHY	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 4.2.2.1 OBSERVED RESIDENCY MATRIX (CONTINUED)

		DOCTORS' CARE	ACUTE HOSPITAL	OUT OF PROVINCE EXTENDED HOSPITAL	OUT PATIENTS	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.019	0.000	0.000	0.017	7.174	0.000
		0.1588	0.0000	0.0000	0.1420	41.2681	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.202	0.025	0.000	0.020	4.668	0.000
		1.9740	0.3661	0.0000	0.1988	31.8167	0.0000
	HEALTHY	0.049	0.021	0.000	0.009	10.085	0.000
		0.6243	0.7201	0.0000	0.1503	50.1416	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	0.020	0.280	0.000	0.020	4.240	0.000
		0.1400	1.9600	0.0000	0.1400	29.6800	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.000	0.000	0.000	0.021	0.000	0.000
		0.0000	0.0000	0.0000	0.1428	0.0000	0.0000
	HEALTHY	0.037	0.023	0.000	0.010	12.731	0.000
		0.5064	0.8399	0.0000	0.2738	56.7821	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	0.034	0.034	0.000	0.003	11.866	0.000
		0.2217	0.5379	0.0000	0.0528	52.3224	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.032	0.159	0.000	0.088	10.753	0.000
		0.2256	2.5693	0.0000	1.2069	49.4920	0.0000
	HEALTHY	0.041	0.025	0.000	0.008	9.474	0.000
		0.4761	0.7750	0.0000	0.1409	49.4082	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
OUT PROV.	DR.S CARE	7.786	1.143	0.000	0.464	86.143	0.000
		15.6327	5.0336	0.0000	1.8800	140.0300	0.0000
	ACUTE HOSPITAL	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	EXTENDED HOSP.	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	OUT PATIENTS	0.217	0.000	0.000	1.043	58.087	0.000
		0.4125	0.0000	0.0000	0.2039	125.5928	0.0000
	HEALTHY	0.700	18.700	0.000	0.100	345.800	0.000
		1.0050	27.6516	0.0000	0.3000	26.9177	0.0000
	DEAD	0.000	0.000	0.000	0.000	0.000	0.000
		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

TABLE 4.2.2.2

ESTIMATED EXPECTED RESIDENCY
TOTAL POPULATION - COMPOSITE ESTIMATES

WINNIPEG

		DR.S'CARE	ACUTE HSP.	EXTEND HSP	OUT-PAT.	HEALTHY	DEAD
WINNIPEG	DR.S CARE	4.5653	1.2606	0.0258	0.2527	345.7952	1.0598
	ACUTE HOSPITAL	3.2958	11.8150	0.0242	0.2417	326.0599	12.2218
	EXTENDED HOSP.	1.8178	0.7530	56.0480	0.2241	177.8238	123.1152
	OUT PATIENTS	3.5146	2.4387	0.0347	1.2912	342.7203	2.9445
	HEALTHY	3.4844	1.2584	0.0258	0.2479	346.9891	1.0561
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	366.0000
BRANDON	DR.S CARE	0.1368	0.1530	0.0000	0.0134	5.2873	0.0000
	ACUTE HOSPITAL	0.1311	0.1463	0.0000	0.0127	4.9829	0.0000
	EXTENDED HOSP.	0.0942	0.1071	0.0000	0.0093	3.4623	0.0000
	OUT PATIENTS	0.1377	0.1541	0.0000	0.0134	5.3456	0.0000
	HEALTHY	0.1360	0.1529	0.0000	0.0133	5.2156	0.0000
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	0.6594	0.4916	0.0089	0.0418	7.0156	0.0092
	ACUTE HOSPITAL	0.6374	0.4787	0.0085	0.0427	6.4599	5.2790
	EXTENDED HOSP.	0.5381	0.4012	0.0071	0.0341	5.5936	0.0068
	OUT PATIENTS	0.6587	0.5301	0.0093	0.0423	7.0589	0.0164
	HEALTHY	0.6578	0.4918	0.0089	0.0416	6.8650	0.0092
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

BRANDON

		DR.S'CARE	ACUTE HSP.	EXTEND HSP	OUT-PAT.	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.0099	0.0025	0.0024	0.0011	0.2734	0.0007
	ACUTE HOSPITAL	0.0093	0.0024	0.0023	0.0011	0.2610	0.0005
	EXTENDED HOSP.	0.0050	0.0013	0.0012	0.0006	0.1317	0.0002
	OUT PATIENTS	0.0099	0.0026	0.0024	0.0011	0.2762	0.0005
	HEALTHY	0.0098	0.0025	0.0024	0.0011	0.2702	0.0005
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	3.6011	0.9076	0.3668	0.3339	332.8909	0.9769
	ACUTE HOSPITAL	2.4229	9.1686	0.3505	0.3239	318.9966	9.1006
	EXTENDED HOSP.	1.8485	0.6449	56.7605	0.2543	237.1344	50.8537
	OUT PATIENTS	2.5551	0.9391	0.4034	1.4358	332.4879	1.2420
	HEALTHY	2.5284	0.9114	0.3684	0.3295	334.3932	0.9784
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	366.0000
REST MAN.	DR.S CARE	0.1264	0.0513	0.0464	0.0085	1.6464	0.0028
	ACUTE HOSPITAL	0.1229	0.0501	0.0446	0.0084	1.5984	0.0027
	EXTENDED HOSP.	0.1028	0.0418	0.0370	0.0068	1.2730	0.0020
	OUT PATIENTS	0.1282	0.0596	0.0488	0.0088	1.8011	0.0033
	HEALTHY	0.1262	0.0513	0.0465	0.0084	1.6179	0.0028

TABLE 4.2.2.2

ESTIMATED EXPECTED RESIDENCY (CONTINUED)
TOTAL POPULATION - COMPOSITE ESTIMATES

REST OF MANITOBA

		DR.S'CARE	ACUTE HSP.	EXTEND HSP	OUT-PAT.	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.0562	0.0232	0.0007	0.0179	3.1597	0.0057
	ACUTE HOSPITAL	0.0542	0.0225	0.0007	0.0187	3.1398	0.0434
	EXTENDED HOSP.	0.0286	0.0122	0.0004	0.0127	1.5575	0.0033
	OUT PATIENTS	0.0568	0.0233	0.0007	0.0182	3.2177	0.0128
	HEALTHY	0.0555	0.0230	0.0007	0.0177	3.0466	0.0056
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	0.1528	0.0807	0.0014	0.0407	9.6678	0.0188
	ACUTE HOSPITAL	0.1462	0.0775	0.0013	0.0407	9.2485	0.1076
	EXTENDED HOSP.	0.1054	0.0566	0.0009	0.0357	6.3639	1.1533
	OUT PATIENTS	0.1523	0.0809	0.0014	0.0405	9.5822	0.0198
	HEALTHY	0.1506	0.0796	0.0014	0.0400	9.2848	0.0185
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	2.9421	1.0167	0.0461	0.3342	341.9794	0.9941
	ACUTE HOSPITAL	1.8322	8.1360	0.0450	0.3221	329.9349	7.7363
	EXTENDED HOSP.	1.5752	0.8464	38.0825	0.2749	281.0716	29.6691
	OUT PATIENTS	1.9153	1.3557	0.0546	1.4096	340.5983	1.7641
	HEALTHY	1.8871	1.0167	0.0462	0.3237	343.2224	0.9902
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	366.0000

OUT OF PROVINCE

		DR.S'CARE	ACUTE HSP.	EXTEND HSP	OUT-PAT.	HEALTHY	DEAD
WINNIPEG	DR.S CARE	0.0484	0.0201	0.0000	0.0087	9.4100	0.0000
	ACUTE HOSPITAL	0.0457	0.0189	0.0000	0.0082	8.7129	0.0000
	EXTENDED HOSP.	0.0285	0.0102	0.0000	0.0045	4.4204	0.0000
	OUT PATIENTS	0.0485	0.0200	0.0000	0.0087	9.3566	0.0000
	HEALTHY	0.0485	0.0202	0.0000	0.0087	9.4251	0.0000
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BRANDON	DR.S CARE	0.0359	0.0219	0.0000	0.0103	11.3021	0.0000
	ACUTE HOSPITAL	0.0342	0.0209	0.0000	0.0098	10.6771	0.0000
	EXTENDED HOSP.	0.0246	0.0151	0.0000	0.0071	7.0681	0.0000
	OUT PATIENTS	0.0358	0.0219	0.0000	0.0104	11.3405	0.0000
	HEALTHY	0.0359	0.0219	0.0000	0.0103	11.3299	0.0000
	DEAD	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
REST MAN.	DR.S CARE	0.0402	0.0242	0.0000	0.0083	8.5067	0.0000
	ACUTE HOSPITAL	0.0387	0.0233	0.0000	0.0080	8.0995	0.0000
	EXTENDED HOSP.	0.0326	0.0196	0.0000	0.0083	6.3769	0.0000
	OUT PATIENTS	0.0400	0.0241	0.0000	0.0083	8.4646	0.0000
	HEALTHY	0.0404	0.0243	0.0000	0.0083	8.5137	0.0000

TABLE 4.2.3 State Occupancy Vectors

		INITIAL	DAY 185	FINAL
WINNIPEG	DR.S CARE	587	694	384
	ACUTE HOSPITAL	0	2356	154
	EXTENDED HOSP.	0	59	1
	OUT PATIENTS	401	446	243
	HEALTHY	537939	538417	512443
	DEAD	0	0	3252
BRANDON	DR.S CARE	50	49	42
	ACUTE HOSPITAL	0	150	7
	EXTENDED HOSP.	0	128	0
	OUT PATIENTS	48	47	20
	HEALTHY	35317	35631	34147
	DEAD	0	0	212
REST MAN.	DR.S CARE	358	382	283
	ACUTE HOSPITAL	0	1068	113
	EXTENDED HOSP.	0	72	1
	OUT PATIENTS	308	484	289
	HEALTHY	420700	414172	395647
	DEAD	0	0	2399
OUT PROV.	DR.S CARE	28	41	28
	ACUTE HOSPITAL	0	66	11
	EXTENDED HOSP.	0	0	0
	OUT PATIENTS	23	62	24
	HEALTHY	10	1445	46069

Since most everyone is classified as healthy, especially on Jan. 1, it is the healthy state residencies which are of particular interest. They may be interpreted directly as "utilization" or "participation" rate for the various health care facilties by the subject population. Thus an individu-

al who is healthy in Winnipeg on Jan. 1, 1972 could expect to spend 3.411 days seeing a doctor in Winnipeg (at least once each day) and 1.287 days in a Winnipeg acute care hospital, while his counterpart in the Rest of Manitoba region could expect to spend only 1.874 days seeing his local doctor and 1.027 days in his local hospital. The use of these participation rates as tools for forecasting health care demand is discussed in chapter 5.

The estimated expected residencies in table [4.2.2.2] are a composite of three separate estimates, one for each region, using the computational procedure described in 3.3.3. Comparing these estimates with the actuals we notice that they are quite different for all the initial states except the "Healthy" state. This is because of the distortions, noted above introduced by the special nature of Jan. 1 in our society, and the time homogeneity assumption of the model which does not permit any "special" days. So the model attempts to estimate the average amount of time an individual would expect to spend in each of the other states in the year following a typical visit to the initial state, while the observed resdencies are the average amount one would expect to spend in the year following Jan. 1, given that one visited the initial state on that particular day. These are quite different in meaning, especially for health care initial state. The differences are caused by the failure of the time homogenous rate assumptions of the Markov model to hold in practice, at least at certain special times of the year, one of which is clearly Jan. 1.

The estimates of the healthy state residencies are remarkably accurate, considering the already noted violations of the Markov assumptions, and the high variability of the observed figures. For the most part, the estimates are accurate to within a few percent of the observed average

residencies for the Healthy state. This surprising accuracy is probably a result of the averaging or smoothing effect on the statistic of the large time parameter (t=366) used in computing the estimates. The observed robustness of these statistics shows that there is some hope that the model can be usefully applied, despite all of the problems with the failure of assumptions.

Another application for the residency matrix, suggested by C.L. Chiang[1973] for a similar model of health care behavior, is to use the expected time spent in the "Healthy" state as an overall index of a populations' health. The problems with this_index are that it is relatively difficult to compute, and that it does not take into account the severity of an illness, only its duration. In our case it is not even the duration of the illness which is involved. but only the duration of treatment for illness. Table 4.2.4 presents these healthy state residency figures for the various regions along with the death rates per 1000 population and the average length of stay in the healthy state, for comparison as indices of health. The separation between population for this proposed index does not seem to be as good as for the other two indices in the table. Nor do the three indicies agree consistently on which population is the "healthier". Increasing death rates are associated with longer average residency in the "Healthy" state in some cases. Thus there does not seem to be a safe way to interpret this index as a measurement of population health.

INDICES OF POPULATION 'HEALTH' TABLE 4.2.4

AGE/SEX GROUP	WINNIPEG HEALTH DAYS	WINNIPEG DAYS BTN CARE	WINNIPEG DEATH RATE	BRANDON HEALTH DAYS	BRANDON DAYS BTN CARE	BRANDON DEATH RATE	REST OF MANITOBA HEALTH DAYS	REST OF MANITOBA DAYS BTN CARE	REST OF MANITOBA DEATH RATE
MALES 1	338	49.7	1.08	321	52.8	----	333	57.8	1.81
2-15	347	100.6	0.48	332	115.0	0.42	348	121.0	0.50
16-45	347	99.7	0.63	331	110.7	0.66	339	110.3	0.94
46-65	352	73.4	8.21	345	82.3	6.99	352	86.3	8.14
65 +	335	54.7	52.5	333	58.8	49.5	338	63.4	46.4
FEMALES 1	337	52.2	0.90	319	53.8	----	334	62.7	2.14
2-15	347	102.7	0.26	334	120.0	0.23	348	123.1	0.23
16-45	346	66.1	0.46	329	74.2	0.55	335	72.3	0.57
46-65	353	59.6	5.10	348	67.0	2.47	351	66.5	4.48
65 +	342	50.2	39.5	343	56.1	37.0	341	56.9	33.4
TOTAL ---	347	75.7	6.03	334	84.7	5.82	343	87.7	5.73

NOTES :

Healthy Days: estimated expected residency in local " Healthy "state, given initially healthy

DAYS BTN CARE: average Length of Stay in local " Healthy" state

DEATH RATE: crude death rate per 1000 individuals

Once again, the residency estimates can be examined as measurements of the flows of people between regions to obtain health care. That is, if there is a health care need which cannot be met locally, then this should be reflected in the residency matrix as relatively large expected residencies in health care states in other regions. This is certainly noticeable for the Rest of Manitoba region's residencies in Winnipeg medical states, further corroborating the already noted [4.1] apparent lack of access to facilities in that region. Table 4.2.5 gives the overall population healthy state residencies for the three health care facilities Drs. Care, Acute Hospital Care, and Outpatient Care. Note that a person, who is healthy in the Rest of Manitoba, can expect to spend ~2.7 days seeing a doctor, with almost 1/4 of that time (.657 days) spent seeing a doctor in Winnipeg. Surely this is an indication of a need for more doctors in the rural areas of Manitoba.

Note also in table [4.2.5] the variation in utilization of the various facilities by the populations in different regions. Winnipeg residents use much more doctors' care than the Rest of Manitoba counterparts on average, while the reverse is true for acute hospital care. Thus there appears to be some substitution of acute hospital facilities, at high tax payer cost, for services performed normally by doctors for those in Winnipeg. These conclusions are in complete agreement with those drawn without the use of the model in section 4.1.

HEALTHY STATE RESIDENCY TABLE 4.2.5

STATE		WINNIPEG	BRANDON	REST MAN
DRS. CARE	WINNIPEG	3.411	.131	.655
	BRANDON	.010	2.502	.120
	REST MAN	.052	.142	1.874
	OUT PROV	.049	.037	.041
	TOTAL	3.532	2.812	2.690
ACUTE HOSPITAL	WINNIPEG	1.287	.154	.504
	BRANDON	.002	.928	.052
	REST MAN	.021	.076	1.027
	OUT PROV	.021	.023	.025
	TOTAL	1.330	1.181	1.608
EXTENDED HOSPITAL	WINNIPEG	.030	.000	.010
	BRANDON	.003	.434	.057
	REST MAN	.001	.002	.051
	TOTAL	.034	.436	.118
OUT PATIENTS	WINNIPEG	.224	.013	.041
	BRANDON	.001	.322	.008
	REST MAN	.017	.038	.321
	OUT PROV	.009	.010	.008
	TOTAL	.251	.383	.378

Based on total population estimated residency in healthy state.

4.3 In Summary

Chapter 4 has attempted to present the information obtained during the execution of the computational procedure described in chapter 3 in two primary areas. The first area was the knowledge extracted during that process about the functioning of the Manitoba Health Care delivery system in 1972. The second area involves the degree to which the Markov process model of chapter 2 can be used in examining the health care delivery system and assessing the models "goodness of fit" to the actual process.

In the first area the major results can be listed as follows:

1) There appears to be evidence of a lack of access to health care outside of the cities of Winnipeg and Brandon, particularly with regard to physicians' care.

2) There is no similar evidence of a shortage of health care in either of the urban areas which would not be explained by their relative sizes.

3) In order to get the care they require, Manitobans outside Winnipeg and Brandon travel to Winnipeg and the extent to which this travel occurs can be quantified precisely.

4) Some health care which is provided to the urban population in Winnipeg and Brandon by physicians appears to be provided by acute care hospitals elsewhere in the province.

In the second area the results are mostly negative:

1) The length of stay distributions are not exponential distributions as they should be for a time homogeneous Markov process.

2) The rates do not appear to be time homogeneous, especially close to days with "special" significance such as statutory holidays.

3) Almost by definition, the pertinent information about individuals' health care history, as it affects his future health care needs, is not summarized sufficiently by knowledge of current state as defined in chapter 2.

4) Despite these massive viclations of the model assumptions, the computed long term residency estimates appear to correspond well with the actual usage, especially when the conditioning state is the "Healthy" state. However, other derived statistics such as the transition probabilities are less robust to the violations of the model assumpticns.

Chapter 5 Applications of the Model

The purpose of this chapter is to discuss, in the light of the computational and theoretical difficulties encountered, how the Markov model might be of use to the health care planner or manager. The major emphasis is on the application of those statistics which showed themselves to be robust to both violations of the model assumptions [ch. 4] and to the computational problems [ch. 3]. In addition, some comments are made on how the model might have been used had time permitted some of these problems to be resolved. Directions for further study which were identified during the course of the study are suggested.

5.1 Forecasting Demand by Location

Health care planners must consider the level of future health care demand that will result if current utilization and flow rates persist. These rates are determined by a complex interaction of population preferences, needs, and availability of services, which will change as services develop and, even when services remain fairly static, will change as access, attitudes and health related attributes of the population change. Forecasting demand, conditional on current rates, is thus a means of gaining insight into what would happen to future demand, in response to population changes (changes in age/sex/regional structure due to growth and flows), and is of value mainly in the short to medium term (5 to 10 years), since it presumes little change in utilization behaviour.

The critical determinants of the aggregate amount of health care demanded of a facility are the population served by the facility and the characteristics of that population as they affect individual health. Thus, to forecast the demand for care from a particular facility, in a particular location, the most obvious approach is to attempt to forecast the growth of the serviced population, and then convert these population projections into health care facility requirements through per capita participation or utilization rates. These rates, usually age and sex specific, describe the variations in demand of the various groups and

are obtained by observing the past and present use of the facility. Many examples of this type of approach exist in the literature, particularly with regard to hospital bed requirements [Roch 1974].

The major problem with this approach is that the definition of the serviced population, usually referred to as the "catchment area" of the facility, is an almost impossible task. In the case of Manitoba, for example, the catchment area for Winnipeg Hospitals is not just metropolitan Winnipeg. About 18% of admissions to Winnipeg hospitals are resident outside the Winnipeg area [table 4.1.3.2]. It would be folly to attempt to forecast the future demand on Winnipeg hospitals without including this significant requirement for care generated by the population of the remainder of the province. However, in order to produce population base figures for both forecasting and the construction of the participation rates, it is necessary to divide the population geographically into regions. In fact, if accurate base population figures are to be available, the boundaries of the regions must be restricted to those defined by the census takers. It is inevitable that the geographically appropriate definition will fail to represent the actual population served, and the participation rates will be artificially inflated by usage from outside the defined region. If the relative sizes of the regions are not changing this may not be a serious problem for the aggregate forecasts. But with the continuing rural-urban migration occurring in Manitoba, forecasts of facility usage may be distorted significantly by this failure to adequately define and forecast the population of the catchment area.

The regional flow model concept eliminates the need to define catchment areas. This is accomplished by collecting and using information on the region of origin of health care

demands. Knowing the region of origin of each person who demands care at a particular institution, it is possible and natural to construct participation or utilization rates which are inter-regional as well intra-regional. Thus the amount of hospital care demanded by residents of Brandon which is delivered in Winnipeg hospitals can be converted into a per capita rate whose base population is the population of the Brandon region, a well defined quantity. The projections of total demand for a particular facility in a particular location computed by adding the demands on that facility over all of the regions, can be based upon any division of the area served which is convenient for population projection, without distortion of the participation rates.

The regional flow model approach completely eliminates the catchment area problem only if the population served is a closed population; that is , one which can be divided into regions which exhaust the served population and which each define a relevant and "forecastable" population group. In reality this will not always be possible. The Out-of-Province catch-all in our model is really an exogenous influence on the system. Growth in its population may occur in ways which are irrelevant to the demand placed on Manitoba's health care facilities (for example , the Chinese population). However, the majority of the care delivered is received by those residents in the province, and the regional flow approach permits the division of the province into arbitrary regions without much danger of ignoring significant demands, yet providing a more accurate picture of demands on the health care delivery system. For a complete forecast, some correction for the exogenous demands on the system (i.e. to account for use by transients,new births, Indians and other excluded elements) would be required, using some other methodology.

To explain more fully the regional flow model approach to the forecasting problem, the Manitoba model will be a useful example. The population of Manitoba is essentially stable at present, at least in terms of the total number of people in the province. This does not mean that no forecasting is necessary or that new facilities will not be needed above the replacement levels. The age structure of the population is changing dramatically as the post-war "baby-boom" ages, and fecundity and morbidity rates decline. Figure 5.1.1 illustrates the expected movement of the age structure in the classic demographic pyramid diagram. The population is becoming older, on average, and as this aging process continues it will affect the demand for health care.

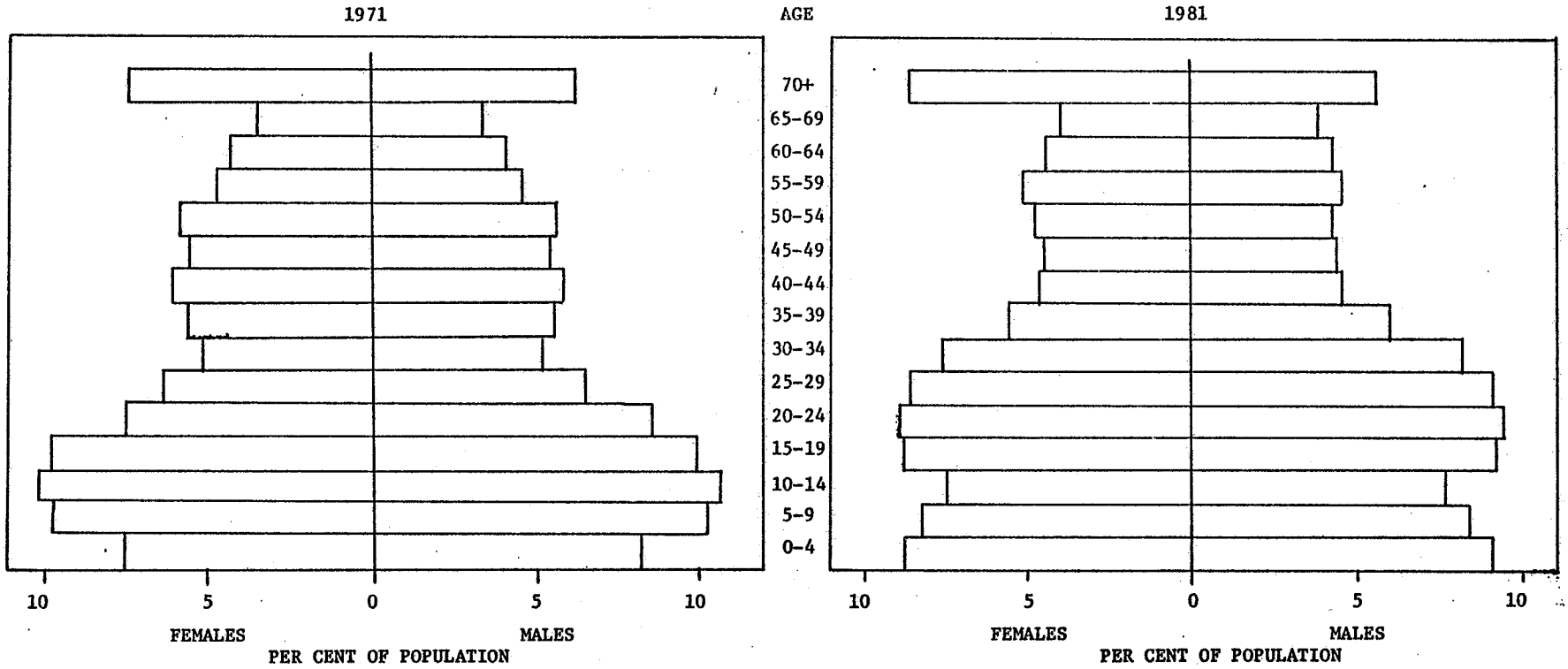
The nature of the forecasting process using the regional flow model approach is as follows. Regions are defined based upon administrative criteria, but keeping in mind the need for accurate population figures. Participation or utilization rates are then produced, measuring the expected per capita demand for the facility (or set of facilities) by the population of sex 's', in age cohort 'a', residing in region 'k'. Note that the sets of facilities for which forecasts are desired may be grouped into administratively useful groups which may or may not relate to the regions into which the serviced population has been split. The generation of the forecast of total usage for the i-th facility group then involves only the production of age and sex cohort population forecasts for each region, C_{ask}, and the computation:

$$U = \sum_a \sum_s \sum_k u_{ask} C_{ask} \qquad \text{(E5.1.1)}$$

TABLE 5.1.1 HEALTH CARE FACILITY USAGE : ESTIMATED EXPECTED RESIDENCY

			DOCTORS			HOSPITALS			OUT- PATIENT		
			WPG	BRDN	ROM	WPG	BRDN	ROM	WPG	BRDN	ROM
WPG	AGE-1	M	5.015	.008	.099	1.135	.003	.106	.618	.001	.023
		F	4.751	.020	.131	.950	.002	.085	.511	.002	.028
	2-15	M	2.117	.004	.036	.317	.000	.012	.319	.001	.016
		F	2.116	.004	.034	.277	.000	.012	.241	.001	.013
	16-45	M	2.280	.010	.054	.509	.002	.014	.222	.002	.027
		F	4.037	.014	.078	1.101	.002	.027	.174	.001	.016
	46-65	M	3.750	.007	.035	1.864	.001	.020	.311	.001	.017
		F	4.927	.008	.038	1.566	.003	.014	.169	.000	.011
	OVER 65	M	6.790	.014	.066	5.567	.012	.042	.573	.001	.013
		F	5.948	.019	.094	4.145	.008	.073	.185	.000	.010
BRDN	AGE 1	M	.164	3.956	.341	.058	.952	.392	.034	.560	.102
		F	.117	3.966	.407	.538	.621	.234	.034	.510	.083
	2-15	M	.081	1.325	.127	.043	.244	.056	.013	.351	.040
		F	.088	1.329	.125	.046	.159	.026	.011	.285	.027
	16-45	M	.130	1.477	.100	.096	.348	.025	.016	.395	.055
		F	.200	2.907	.215	.104	.989	.077	.012	.292	.040
	46-65	M	.120	2.758	.106	.400	1.325	.094	.021	.356	.026
		F	.143	3.801	.118	.194	1.324	.051	.008	.311	.024
	OVER 65	M	.166	4.715	.155	.428	3.210	.254	.014	.350	.065
		F	.138	5.222	.145	.268	2.646	.243	.004	.189	.027
ROM	AGE 1	M	.939	.139	3.079	.414	.045	2.481	.088	.014	.524
		F	.807	.142	2.897	.370	.027	2.215	.080	.014	.416
	2-15	M	.394	.062	1.125	.149	.014	.379	.035	.005	.275
		F	.389	.060	1.134	.119	.008	.331	.029	.005	.222
	16-45	M	.499	.065	1.242	.292	.020	.321	.044	.012	.359
		F	.945	.163	2.289	.440	.058	.961	.034	.009	.332
	46-65	M	.664	.125	1.951	.848	.075	1.058	.058	.008	.387
		F	.957	.190	2.854	.751	.086	1.295	.028	.007	.347
	OVER 65	M	.828	.256	3.192	1.746	.210	3.830	.102	.015	.408
		F	.979	.386	3.812	1.205	.138	3.932	.032	.007	.280

FIGURE 5.1.1 MANITOBA POPULATION AGE DISTRIBUTION : 1971 & 1981



In the context of our study the production of these specific utilization rates can be performed in two ways. The most obvious way is to divide the total sojourn times for each population in each facility by the respective base population to obtain per capita usage [table 4.1.8]. By using the Markov model and noting that almost everyone is healthy on Jan. 1, 1972, one may also estimate the participation rates by selecting the healthy state rows from the expected residency matrix. As already noted, the one year estimated expected residency statistics appear to be extremely robust to the violations of the Markov assumptions which were encountered. This robustness was particularly strong for the healthy state. Table 5.1.1 contains the participation rates assembled from the various estimated expected residency matrices. Since time did not permit the problem of the 6 cases of complex eigenvalues to be resolved [3.4], observed expected residency in the healthy state were used to complete the table. The similarity of the two sets of participation rates is remarkable, given the violations of the model assumptions which intimately affect the derivation of the formula used to compute table 5.1.1. Note that, as products of the model, only the sufficient statistics for estimation of the model (a_{ij} and D_i see section 3.2) are needed to compute table 5.1.1. Knowledge of the number of people involved is not required.

POPULATION PROJECTIONS FOR MANITOBA TABLE 5.1.2

			1971	1972	1976	1981
Winnipeg	Age 1:	Males	4087	4628	4933	5528
		Females	3855	4472	4726	5295
	2-15:	Males	68089	68475	70108	70659
		Females	65665	65701	67056	67927
	16-45:	Males	111136	116507	129618	147139
		Females	117327	117410	128518	142402
	46-65:	Males	49108	51777	52726	52824
		Females	55182	58017	60724	60881
	65:	Males	20493	22375	23765	27344
		Females	27773	29565	31613	37204
	total:	Males	252913	263762	281150	303494
		Females	269802	275165	292637	313709
Winnipeg	Total:		522715	538927	573787	617203
Brandon	Age 1:	Males	307	297	309	334
		Females	301	292	291	313
	2-15:	Males	4955	4705	4834	4601
		Females	4881	4419	4436	4165
	16-45:	Males	7894	7569	6896	7315
		Females	8260	7267	7883	8363
	46-65:	Males	3502	3441	3198	2850
		Females	3736	3657	3780	3729
	65:	Males	1788	1662	1470	1548
		Females	2186	2106	1912	2015
	total:	Males	18446	17674	16707	16648
		Females	19364	17741	18302	18585
Brandon	Total:		37810	35415	35009	35233
Rest Man	Age 1:	Males	3421	3872	3543	3826
		Females	3274	3743	3351	3617
	2-15:	Males	58612	63655	54161	50828
		Females	55978	60831	50587	47222
	16-45:	Males	72328	86075	84348	89570
		Females	67645	78820	73585	77160
	46-65:	Males	39067	42256	39262	36833
		Females	36661	39523	36441	34505
	65:	Males	18757	21510	21150	23074
		Females	18155	21006	20974	23475
	total:	Males	192185	217368	202464	204138
		Females	181713	203923	184938	185979
Rest Man	Total:		373898	421291	387402	390117

Returning to the forecasting process, table 5.1.2 presents the necessary population projections by age, sex and region. The 1972 figures in the table are the actual counts of individuals in each age/sex cohort in each region which is used in the estimation process of chapter 3. The remaining columns were derived from a set of Leslie model projections adjusted for migration which were produced in 1971 by A.N. Arnason [1971].

Applying E5.1.1 to the participation rates derived from the estimated expected residence [table 5.1.1] and these population projections [table 5.1.2] yields table 5.1.3. No claim is made that these projections represent reality as it was in 1976 or will be in 1981. Many changes in the health care delivery system which would have large impacts on these facilities have occurred since 1972. The introduction of Nursing Home care, continued expansion of extended care facilities, and various programs aimed at improving care in Northern Manitoba have severely altered the context in which the participation rates were estimated. Thus these projections represent more what might have been in the absence of these changes. They do serve to illustrate the technique, however, and, since the style of health care delivery appears to be stablizing, a re-estimation of the participation rates might prove a valuable forecasting tool. At the same time, some of the problems noted in this study could be corrected.

FORECASTS: HEALTH CARE UTILIZATIONS TO 1981 TABLE 5.1.3

FACILITY		1971	1972	1976	1981
WINNIPEG	HOSPITALS	831	886	917	992
	DOCTORS	2059	2158	2261	2428
	OUT-PATIENT	145	151	158	169
BRANDON	HOSPITALS	55.2	55.7	53.8	55.2
	DOCTORS	147	148	144	147
	OUT-PATIENT	16.1	15.8	15.4	15.7
REST MAN	HOSPITALS	394	445	422	440
	DOCTORS	738	829	777	794
	OUT-PATIENT	131	147	138	140

UNITS:

Hospitals-10^3 Bed-Days

Doctors -10^3 Patient Contact-Days

Out-Patients-10^3 Patient Contact-Days

The forecasts are interesting in the way that they illustrate the effect of aging on health care needs. Brandon is projected to experience an absolute decline in population, yet its increasing average age keeps the demand on its health care facilities relatively stable. The anomalous results listed for 1972 are a result of the incompatibilities between the population projections, based on census information, and the population derived from the M.H.S.C files. These incompatibilities and failure to apply any correction for usage by the exogenous population (those who were born in 1972, who entered Manitoba in 1972, Indians, and transients) probably imply that these forecasts are too small in most cases. A useful forecasting procedure would have to resolve these problems, but the regional flow model approach should assist in the process of monitoring and forecasting the bulk of the health care demanded.

5.2 Simulation Studies

One of the most attractive properties of a Markov process is the ease with which it may be simulated. Thus if a Markov model can be found which has some or all of its interesting derived statistics in close correspondence to reality, relatively cheap experiments into changes in the structure of the system can be performed through simulation. Unfortunately, it is often difficult to relate the desired changes in the real system to changes in the Γ or P^* matrices which are required for a simulation experiment.

For example, the discussion in chapter 4. concluded that there was evidence of a lack of access to physicians for those living outside of Winnipeg and Brandon. In planning to correct this situation such policies as the provision of additional clinics or the training of additional doctors might be considered. It would be helpful in

assessing the policy alternatives if the impact of a particular policy on the entire health care delivery system could be estimated. In order to use the Markov regional flow model in this context, the impact the policy might have on either the one day transition probabilities or the rates (γ_{ij}) must be hypothesized. The process of formulating this hypothesis is assisted by the interpretation which can be placed upon these rates [2.1.4]. One could anticipate,for instance, that the opening of a new clinic in Thompson should decrease the overall probability of going to Winnipeg for doctor's care for those in the rest of Manitoba, and increase the respective probability of attending a local physician. Quantifying the amount of this change would be difficult, and is probably best accomplished by a trial and error type search using the simulation process. Constraints on this search procedure exist in the maintenance of the theoretical properties of the rate matrix (that its row sums are zero) and of the estimated population characteristics, such as lengths of stay.

Care must be taken to restrict attention in the simulation study to those properties of the model which have shown robustness to the observed failures in the model assumptions. In the present case this means that the only really useful output of a simulation would be the long-run estimated expected residencies. As a result , and because of time pressures, no simulation results will be reported here, although some preliminary work on simulation was done to provide realistic test data for the software used in the data reduction process of chapter 3. For a discusion of a method of simulation for this class of compartment models see Arnason[1975].

5.3 Problems and Directions for Further Study

Most of the major problems with the application of the Markov flow model to the Manitoba health care delivery

system were the result of two factors. Firstly, there were a series of critical decisions made during the data reduction process which, in retrospect, turned out to have unfortunate consequences for the use and interpretation of the model. Briefly these were :

1) The choice of observation interval as beginning on Jan. 1 was unfortunate, as this is a peculiar day for the health care system.

2) The elimination of multiple contacts on one day. This resulted in a loss of information wich could have been avoided by permitting lengths of stay under 1 day, and ordering arbitrarily the day's transactions.

3) The exclusion of the Indian populations, and possibly also the transient population caused a similar loss of information.

4) Too few regions were defined, and the information necessary to alter the region definitions was not retained.

5) Flexibility with regard to age cohort boundaries was not maintained and the chosen boundaries did not conform to standard demographic practice.

The elimination of these mistakes would be relatively simple were the study to be repeated or a similar study performed. The basic mistake in most of the above was a failure to retain as much information and as much flexibility as possible for as long as possible during the data reduction process.

Much more difficult to avoid would be the violations of the Markov assumptions which were noted. In order to correct this problem, more realistic mechanisms for deter-

mining the length of stay for such facilities as doctors and outpatient clinics would be necessary. Even then, there is probably a requirement for more complexity in the definitions of the states than is the case with our simple, administratively convenient, state definitions. Higher order Markov process models, and/or semi-Markov models could be profitably explored as possible alternatives, though the relevant theory is less well known. Much further work could easily be done on development of a model which would more closely represent the functioning of the health care system.

The computational problem [3.4] of actually computing the estimated expected residency and transition probability matrix functions was not completely solved. There may well be a more complete, if not more elegant, solution than the one used here. There may even be a theorem which would guarantee real eigenvalues for some class of rate matrices. The problem of what form the solution should take, should non-distinct eigenvalues result in linearly dependent eigenvectors, should also be explored more fully.

Another interesting direction for further study uncovered by the present work was the need to quantify the robustness of the various Markov derived properties to systematic violations of the Markov assumptions. A simulation study of the classes of violation which might be expected in practice would certainly enhance the confidence with which this type of model could be applied. In particular, a simulation study of the sensitivity of the estimated expected residency to sample size variations would directly extend the present results, which were based on census type information. Cost savings could be quite easily realized should sampling variations not severely affect the statistics.

BIBLIOGRAPHY

1)Arnason, A.N. (1975) : Computer Simulation of Multi-Dimensional Discrete Markov Processes. Proceedings of the Fifth Manitoba Conference on Numerical Mathematics, University of Manitoba, Winnipeg, Manitoba.

2)Arnason , A.N. and Roch D.J. (1974) : A Critical Examination of the Data Structure of the Manitoba Health Services Commission Tape Files and their Potential for use in Research on Health Care Delivery . Report to the Manitoba Health Services Commission. (unpublished).

3)Arnason , A.N. (1971) : Statistical Projection of the Manitoba Population in Age Cohorts by Towns, Municipalities, and Regions with estimates of Migration. University of Manitoba Department of Computer Science Scientific Report No. 39, Winnipeg Manitoba.

4)Ashford, J.R. and Hunt, R.G. (1974) : The Distribution of Doctor-Patient Contacts in the National Health Service. Journal of the Royal Statistical Association Ser. A, 137 p347

5)Bartlett, M.S. (1949) : Some Evolutionary Stochastic Process. Journal of the Royal Statistical Association, Ser.B 11 p211.

6)Bhat, U.N.(1972) : ELEMENTS OF APPLIED STOCHASTIC PROCESSES. John Wiley, New York.

7)Bharucha-Reid, A.T. (1961) : ELEMENTS OF THE THEORY OF MARKOV PROCESSES AND THEIR APPLICATIONS. Magraw-Hill, New York.

8)Billingsley, P. (1961) : STATISTICAL INFERENCE FOR MARKOV PROCESSES. University of Chicago Press, Chicago.

9)Chiang C.L.(1973) : How to Measure Health: A Stochastic Model for an Index of Health. International Journal of Epidemiology 2 p7.

10)Chiang C.L. (1968) : INTRODUCTION TO STOCHASTIC PROCESSES IN BIOSTATISTICS. John Wiley, New York.

11)Chiang C.L. (1964) : A Stocastic Model of Competing Risks of Illness and Competing Risks of Death in STOCHASTIC MODELS IN MEDICINE AND BIOLOGY. J. Gurland ed., University of Wisconson Press, Madison.

12)Cox,D. R. and Miller, H.D. (1965) : THE THEORY OF STOCHASTIC PROCESSES. John Wiley, New York.

13)Cuthbert,J.R. (1973) : The Logarithm Function for Finite State Markov Semi Groups. London Mathematical Society Journal 6 p524.

14)Doob, J.L. (1953) : STOCHASTIC PROCESSES. John Wiley, New York.

15)Feller,W (1940) : On the Integrodifferential Equation of a Completely Discontinuous Markov Process. Trans. American Mathematical Society 48 p488.

16)Goldsmith, S.B. (1972) : The Status of Health Status Indicators. Health services Reports 87 p212.

17)International Mathematical and Statistical Library(IMSL) Reference Manual version 1.5. IMSL Inc. Houston Texas (1974).

18)Kleinrock,L. (1975) : QUEING SYSTEMS Vol 1 and 2. John Wiley,New York 1975-76.

19)Moriyama, I.M. (1968) : Problems in the Measurement of Health Status in INDICATORS OF SOCIAL CHANGE. E.Sheldom and W. Moore ed. Russel Sage Foundation, New York 1968.

20)Navarro,V. (1970) : Planning Personal Health Services :A Markovian Model. Department of Medical Care and Hospitals, Johns Hopkins University, Baltimore Md.

21)Parzen E. (1962) : STOCHASTIC PROCESSES. Holden-Day, San Fransisco.

22)Roch, D.J. (1974) : A Comparitive Analysis of Hospital Supply Planning Models. M.A. Thesis in Economics, University of Manitoba, Winnipeg.

23)Sanders,B.S. (1964) : Measuring Community Health Levels. American Journal of Public Health 54 p1063.

24)Sharp, S.A.(1975) : Corrections to Likelihood Ratio Tests of Hypotheses Concerning the Parameters of Markov Chains. Biometrika, 62 p595.

25)Varga, R.S. (1962) : MATRIX ITERATIVE ANALYSIS.Prentice Hall, New Jersey.

26)Statistics Canada : VITAL STATISTICS :1972. Statistics Canada Publication No. CS84-206/1972.

27)Wilkinson, J.H.,Bowdler H.,Martin,R.S.,and Reinsch, C. (1968) : The QR and QL Algorithms for Symetric Matrices. Numerishe Mathematic 11 p293.

28)Wilkes, S.S. (1972) : MATHEMATICAL STATISTICS. John Wiley, New York.

29)Zahl, S. (1955) A Markov Process Model for Follow-Up Studies. Human Biology 27, p90.