8013812

WARD, JAMES REGINALD

THE RELATIONSHIP BETWEEN CLIENT CHARACTERISTICS AND THOSE NON-MEDICAL HOME CARE SERVICES AND COSTS ASSOCIATED WITH THE LONG-TERM CARE OF THE IMPAIRED AGED IN MICHIGAN

Michigan State University PH.D. 1979

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106 18 Bedford Row, London WC1R 4EJ, England

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs ______
2. Colored illustrations ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Print shows through as there is text on both sides of page ______
6. Indistinct, broken or small print on several pages ✓ throughout ✓
7. Tightly bound copy with print lost in spine ______
8. Computer printout pages with indistinct print ______
9. Page(s) ______ lacking when material received, and not available from school or author ______
10. Page(s) ______ seem to be missing in numbering only as text follows ______
11. Poor carbon copy ______
12. Not original copy, several pages with blurred type ______
13. Appendix pages are poor copy ______
14. Original copy with light type ______
15. Curling and wrinkled pages ______
16. Other ______

THE RELATIONSHIP BETWEEN CLIENT CHARACTERISTICS AND THOSE NON-MEDICAL HOME CARE SERVICES AND COSTS ASSOCIATED WITH THE LONG-TERM CARE OF THE IMPAIRED AGED IN MICHIGAN

By

James Reginald Ward

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Agricultural Economics

1979

ABSTRACT

THE RELATIONSHIP BETWEEN CLIENT CHARACTERISTICS AND THOSE NON-MEDICAL HOME CARE SERVICES AND COSTS ASSOCIATED WITH THE LONG-TERM CARE OF THE IMPAIRED AGED IN MICHIGAN

By

James Reginald Ward

A major problem in the provision of non-medical long-term care services to the chronically ill or disabled patient lies in making the assignment of appropriate types of services to match the condition or needs of the patient. A second problem lies in identifying key predictors of service costs to facilitate improved cost control.

Support services delivered in the client's home are primarily non-medical and in this study include light cleaning, meal preparation, home maintenance and repair, shopping, transportation, non-nursing personal services, and others. In 1975-76 these services were supplied to over 10,000 low income Michigan residents by the Adult Chore Service Program of the Michigan Department of Social Services. The services were proportedly assigned to clients on the basis of attributes such as the clients' functional status (e.g., mobility, dexterity, sensory perception, comprehension and ability to manage

home); socio-economic status (e.g., age, sex, location, income, relationship to provider and living arrangement); and medical status.

In part, this study was addressed to determining if patterned relationships could be detected between client attributes and the services assigned to compensate for client impairments. If such relationships were detected, it should be possible to infer that for any client profile, a range of services is appropriate. Knowing that range, it should then be possible, by analyzing costs, to identify the appropriate service or service mix of lowest cost. A cost analysis was conducted to determine which attributes significantly explained variation in cost.

The sample of 628 cases was randomly divided into two groups, a 'training set' of 428 cases and a 'test set.' An information theory based algorithm called entropy minimax was used in an attempt to detect patterned relationships between a client's attributes and the services assigned to them. For example, given a specific client profile, it was anticipated the probability of services assigned to that profile could be determined. No such relationships were discovered. These results could be explained by: 1) the failure of social workers to assign systematically services to clients on the basis of the client profile; or 2) no such relationship implies the lack of a systematic relationship between needed, and therefore, assigned services; or 3) inappropriate specification of the attributes and services may have obscured the the real relationship.

The results of a regression analysis using three models in which the endogenous variables were, cost per month of service, hours of service provided per month, and cost per hour respectively, tended to confirm the pattern detection results. Detailed functional status variables explained very little of the variation in costs and hours. More aggregated functional status indicators such as comprehension and ability to manage the home, explained more monthly cost variation than mobility, dexterity or sensory perception. The form of the functional status variable explaining over 40 percent of the variation in monthly costs was specified according to services. The most significant explanatory service categories were light cleaning, meal service and non-nursing personal service. Location was the most significant socio-economic status category explaining monthly cost. Categories of variables best explaining hours of service per month and service costs per hour were relationship of the provider to the client and the living management between them. Medical status was not a significant variable in any of the models.

The results of this research indicate further work is necessary on relating client attributes to services received. A shortcut to achieving this goal might be to classify patients according to service needs. The other attributes could be used as an accountability check on the program and its workers.

DEDICATION

To my wife Sheila, for her patience and encouragement,

and

my children Kerry and Laura, whose joyous young presence opportunely dovetailed with the discretionary times afforded by academic studies,

and

my mother, who despite her own judgement of what her son should be doing, has always been supportive of, if at a loss to explain, what he actually does.

ACKNOWLEDGMENT

A diverse group of people contributed to this work. Lois Libby induced me to become involved in the local health planning systems and the thorny issue of long-term care. Jackie Argyle and Diane Emling of the Michigan Department of Social Services were helpful in describing the Chore Service Program and providing data.

The initial conceptualization of the thesis was assisted by Dr. Paul Ginsburg. The assistance of Mr. Bill Schonbein and his associate, Dave Gift, were of enormous value in understanding the intricacies of the entropy mini-max procedures and their applications. Bob Stevens has always been supportive of my activities, both academic and otherwise, in the health care system.

Alan Schmid has had a pervasive influence on this research stemming from his insightful course on public policy and his active search for deeper insights into various schools of economic theory from an institutional perspective.

When the inevitable obstacles thwarted progress on this study, or methods and results were unclear, the thesis committee chairman, Lester Manderscheid, unfailingly and carefully came to my assistance. For such help and for his

fairness, integrity, and flexibility as my program advisor, I will always be in his debt.

Thanks are also due to Alyne Tennis and Sue Smith for the typing of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

LIST OF FIGURES

CHAPTER 1

PROJECT ORIENTATION

I. Problem Statement

During the last 5 to 10 years there has been increasing debate over the question of what constitutes appropriate long-term care services for the chronically ill and disabled. This debate stemmed, in part, from an intensifying concern about the increasing institutionalization of the aged and a concomitant rise in the costs of medical care. A great deal of uncertainty prevails about what constitutes the appropriate care for populations of long-term care patients with varying functional abilities and in varying socio-economic circumstances.

The general aim of this research is to shed light on the inter-relationships among characteristics of long-term care patients, the services they receive and the costs of those services. Specifically, it focuses on the long-term non-medical services delivered in the home. Since long-term care patients are relatively stable medically, particular interest is focused on specific patient attributes such as functional levels of impairment, socio-economic status and the non-medical types of services assigned to compensate for impairment. The study is conducted on a sample of chronically ill and disabled, predominantly aged people who are beneficiaries of a Michigan chore service program.

The problems associated with identifying appropriate long-term care services became apparent to the writer while serving on the project review committee of the Capital Area Comprehensive Health Planning Agency in Lansing, Michigan between 1973 and 1975. The committee received several requests from nursing home operators to approve the construction of new nursing homes. It was claimed by the applicants that all nursing homes in the area had an occupancy rate in excess of 99 percent of capacity and that there was a waiting list for future vacancies. A major question (of nationwide significance) surfaced in the hearings: how appropriate were nursing home and other long-term care services for those disabled and chronically ill persons using them? It became apparent the answer to the question was fraught with controversy. Third party reimbursement of long-term care services overwhelmingly favored institutionalized long-term care patients. It seemed placement decisions were being influenced by what was available (KNOX ET AL 1973). Since few substitute services were available that usually meant institutionalization. Little attention was being paid by either public or private financing agencies, such as government social service departments or private insurance companies, to financing the care of the chronically ill outside of institutions.

The problem for the project review committee was accentuated by the lack of agreed upon concepts and variables by which the various kinds of long-term care, and thus costs, could be assessed. Consequently, there were no means by which to adequately compare the various types of long-term care services with each other. Attempts to compare the costs of different types of long-term care programs proved futile because it was not possible to relate the attributes employed in classifying patients to the type of services assigned to them; a major objective of this research.

In general, a patient's profile of attributes can be broken down into 3 distinct types: medical status, socio-economic status, and functional status. The latter index either measures the patient's ability to perform various everyday tasks such as walking, bathing, cleaning, etc., or uses measurable proxies for impairments such as degree of mobility or degree of mental capacity and extent of sensory faculties. The two approaches may be combined. Without a common understanding and agreement on the form of patient profile to be used for assessing the needs of long-term care patients, it was not possible to examine the appropriateness of the various services nor compare their costs. This kind of information is required to learn how decisions are made with respect to the assignment of long-term care services to the chronically impaired. At issue, is not only whether appropriate attribute profiles exist, but also whether they would be used consistently by workers who assign long-term care services to the impaired.

II. Project Justification

The aged constitute a large proportion of the chronically ill and disabled population. Between 1960 and 1970, the population increase among persons aged 65 years and over amounted to 21.1%; among those 75 years and over the increase totalled 37.1%. These increases occurred during an overall population growth of 12.1%.

As people age, the likelihood of them suffering from one or more chronic diseases increases. Diabetes, heart problems, cancer, arthritic and mental disorders are often accompanied by disablement. Forty two percent of those aged 65 and over have long-term activity limitation due to chronic illness (NATIONAL CENTER FOR HEALTH STATISTICS 1973). It is estimated that three million of the nations 20 million aged are so seriously disabled that they need some form of personal care (CARO 1972).

Modern medical practice, particularly chemotherapy, has enabled the chronically ill to maintain reduced but relatively stable physical and mental capacities. They are thus living longer in various states of dependency. The multiplication of therapies employed to stabilize the chronically disabled patient serves to increase the cost of both medical services and non-medical support services such as cleaning, personal care or meal preparation. The latter derive from the need to physically, psychologically and socially care for the patient. Among the aged, the number of disability days now averages 34 per person annually.

The financial burden of medical care falling on the chronically ill rose to such levels that indemnity insurance and outright government subsidization have become the chief sources of payment for medical services and for support services, principally in institutions. These developments slowly emerged from the awareness that, "Economic life in the health care sector is less subject to the influence of exchange as a means of organization." "Instead, the grant is rapidly becoming the instrument of political and economic organization" (BOULDING AND PFAFF 1972). Additional services required by the patient to compensate for his/her disability outside the institutional setting, e.g., cleaning, meal preparation etc., are often not covered in benefits received by the chronically ill. Exceptions to this generalization are increasing as in Michigan where patients in certain income categories can, through grants, secure non-medical assistance. It is important to note that reimbursement for institutional care in nursing homes and hospitals covers room and board in addition to professional medical services.

The financial strains imposed on third party payors have thus increased due to a larger population at risk, increasing medical intervention, and rapidly rising costs. Realizing this, third party payors have sought to gain information on appropriate types of medical and non-medical care for given patient characteristics. With this knowledge, they could then identify the lowest cost appropriate care and reimburse for it. In practice, patient characteristic profiles are not uniform, rendering comparisons difficult.

Institutions tend to emphasize the medical status of the patient and rarely supply sufficient detail on the functional status of the patient. Non-institutionalized patients are also predominantly characterized by their medical status. A population of the chronically ill and disabled in which this is not the case is to be found in Michigan, in the Adult Chore Service Program of the Michigan Department of Social Services. Patient functional status, socio-economic status and to a lesser extent, medical status have all been recorded.

III. Project Goals and Methodology

The allocation of funds by state and federal legislators among various long-term care programs is arbitrary to the extent it proceeds without significant insights into the relative merits and unit costs of those programs. This research is intended to examine one particular government program in Michigan and to develop a method whereby various long-term care programs can be compared with each other with respect to appropriateness of care and associated costs.

In comparing long-term care programs, it is important to have common indicators of patient status. Given common indicators we could then determine how they relate to the assignment of both medical and non-medical services to the patient. It is assumed there would be overlap between one type of service, e.g., home care, and another service such as nursing home care for some common patient profiles. In these instances, an incentive may be introduced or a cost ceiling imposed to induce the patient, or those assigning services, to choose the service of lowest cost.

This research has two primary objectives and an associated third objective. The first objective is to detail and test a methodology for determining whether a relationship can be detected between patient attributes and the long-term non-medical services assigned to them. If such relationships can be detected, the knowledge should be useful in determining the range of services deemed appropriate for patients with given attribute profiles. A knowledge of the alternative mixes of services appropriate for any given patient attribute profile when linked to associated costs should enable us to identify the appropriate service mix of lowest cost.

The hypothesis underlying the above objective is that given a specific patient attribute profile, a person assigning long-term care services will systematically base the choice of services on that profile. At a more mundane level, the question is whether or not Michigan Chore Service workers utilize the particular profile developed by the Michigan Department of Social Services to assign services to patients benefitting from the Chore Service Program.

The second major objective of this research is to determine to what extent the various attributes of the patient profiles are related to long-term care service costs. Here we are seeking to learn which particular attributes of the profile are the most useful predictors of service costs.

An associated third objective is to determine which long-term care services, as distinct from patient characteristics, are the most useful indicators of Chore Service Program costs.

To meet the first objective, a pattern detection algorithm called Entropy Minimax will be employed to determine if any systematic relationship exists between patient attributes and assigned services. If such a relationship is found to exist, regression analysis will be employed on a separate set of data to confirm the results. There are sufficient cases to randomly split the total sample of clients into 2 parts; the larger "training set" of at least 400 cases will be used to evaluate the effect of various specifications of the model; the smaller "test set" of approximately 200 cases will be employed to test the best results of the entropy minimax analysis.

In pursuing the second and third objectives, regression analysis will be used to estimate the costs of care and the hours of services provided. Since specification of the cost functions will involve many adjustments in the equations, the training set described above will be used to isolate the "best" specified equations. The test set will be used to test these models and ensure results from the "training data" are not contrived. It is anticipated this research will result in information and predictions that bear on policy-making in the Michigan Chore Service Program.

Of a more general nature, further research will be suggested that will enable comparisons of services among groups of patients receiving various types of long-term care services (e.g., nursing home care, home nursing, home chore services, and others).

Theoretical advances will include explicating procedures for identifying key variables within the context of a decision-making model and making them applicable to analysis of long-term care patients in general, and chore service beneficiaries in particular. Implicit in this procedure will be the identification of the specific key variables that influence assignment of services and costs.

IV. Description of Chapters

A more detailed analysis of the problems of long-term care are discussed in Chapter 2. There, assumptions and general hypotheses, about decision-making are explained together with a reference to the investigative boundaries of the research.

In Chapter 3, a brief survey of theoretical work is presented in the review of literature. A discussion of various approaches to conducting research on long-term care is also presented. In Chapter 4, a description of the Michigan Adult Chore Services Program is offered together with a description of the data.

Chapter 5 outlines the technical design underlying the research, detailing the models employed in the regression analysis of costs and hours of service. All variables are specified in Chapter 5 and entered into econometrically specified models. In Chapter 6, the variables are transformed for adaption to the pattern detection algorithm. This algorithm, entropy minimax, is of relatively recent origin and was developed from information theory; a relatively full explanation of the methodology and reasons for its use are provided. Its adaptation to the problem of estimating probabilistic relationships between sets of outcomes, e.g., services assigned, and sets of attributes is also described. In Chapter 7, further explanation of the entropy minimax procedure is supplied along with empirical results from its application.

Chapter 8 is devoted to the display and discussion of the regression results and analysis in which costs per month, costs per hour and hours per month constitute the endogenous variables. Each is estimated using various combinations of functional status characteristics, socio-economic status indicators, medical status and long-term care non-medical types of service.

Finally, the conclusions pertaining to the study are offered in Chapter 9. Further research is also suggested along with implications for policy making in non-medical long-term care services.

CHAPTER 2

NATURE AND SCOPE OF THE PROBLEM

I. Principal Dilemmas

The United States is one of the few countries in the world where the nursing home is viewed as a practical means of caring for the severely handicapped (TRACY 1974). Between 1966 and 1974 national expenditures on nursing homes increased from $1.41 billion to $7.45 billion (WORTHINGTON 1975). In 1974 the Michigan Medicaid program spent $250 million of state funds on nursing home services but less than $1 million on services delivered in the patient's home. Among Michigan residents the likelihood of a person entering a nursing home in his/her lifetime is now 1 in 4 (KASTENBAUM 1972).

Gerontologists are of the opinion that institutionalization of chronically ill and disabled patients often leads to traumatic effects among the aged and that home care would be more appropriate for purposes of maintaining a higher quality of life for many (BLENKNER 1974). Estimates of the proportion of people in nursing homes who could be better cared for outside of the institutional setting range between 20% and 50% (CARO 1972, BARNEY 1973, GENESEE 1970, KISTIN and MORRIS 1972, WAGER 1972, WHITE and TENBRUNSEL 1971, FLORIDA 1971,

WILLIAMS ET AL 1973, MICHIGAN 1975, DAVIS and GIBBON 1971, HUNDERT 1974, ROBINSON ET AL 1974).

Clearly there is considerable disagreement on what constitutes appropriate long term care for the physically and mentally disabled and the chronically ill. The disagreement on appropriateness of care is, in part, derived from disagreement on the issue of how to classify the long-term care patient. Indexes now used to classify patients can be broken down into three distinct areas; indexes describing the patient's medical status, socio-economic status, and functional status. Specifying medical status and socio-economic status, though somewhat problematic, is relatively straightforward compared to the confusion over how to define functional status.

Medical status is conventionally classified according to the International Classification of Disease Adapted (ICDA). However, physicians vary considerably in their assessment of pathology in an older person (SHANAS 1968). A possible problem here is that grouping of cases by diagnosis takes no account of their severity or complexity (EVANS 1971).

The ICDA index has had considerable importance in the designation of appropriate services, including non-medical services for the chronically ill. Since the patient requiring long-term care services is usually medically stable, the acute care intervention of a physician is often unnecessary. Monitoring and treatments by nurses or friends often constitute adequate care. This being the case, an important question is raised: how important is the patient's medical status in designating long-term care services for him/her? The socio-economic status of the patient is relatively clearly defined and the concepts involved generally agreed upon.

A consensus on the indexing of functional status has been difficult to achieve. It does not appear that one standard means of classifying functional status has been applied across the full range of long-term care patients. Several indexes have been used in various settings. Among them are the Activities of Daily Living (ADL) index (KATZ ET AL, 1963, 1970), Index of Functional Impairment (SHANAS 1971), the Townsend Scale (TOWNSEND 1963), the Minnesota State Periodic Medical Review (ANDERSON 1974) and the functional and socio-economic status assessment (Form DSS-3492) employed by the Michigan Adult Chore Service Program. The latter index is examined in this thesis.

The importance of a patient profile for this analysis lies not so much in the accuracy with which it describes the patient _per se_ but in its utility in predicting appropriate services (and associated costs) required by the patient. We are concerned about finding those key attributes of the profile which account for the assignment of various types of services designated to be appropriate for the patient. Without some understanding of the impact of the various patient attributes on service determination, it is questionable whether the estimates of costs for the various patient profiles can be of value to policy makers in predicting future costs. The estimation of relationships between patient attributes and services provided would facilitate the use of administrative rule-making such as choice of patient eligibility criteria and the designation of fitting services. There is agreement within States, if not among them, as to what constitutes a patient profile that justifies institutionalization in a nursing home or hospital. Justification implies some go/no-go index which enables institutional administrators to determine whether or not their services will be reimbursed by third-party payors. It should be noted that justification for the provision of care does not necessarily imply provision of one specific type of appropriate care. Alternative forms of care might also be appropriate for the nursing home resident as was noted above.

Since no clearly established deterministic relationship between patient attributes and appropriateness of care has yet been discovered, the latter tends to be prescribed on an adhoc basis by the interplay of the concerned "publics" as they engage in transactions, rule-making and legislation. In order to gain insights into the consequences of the interplay of these four "publics" on the determination of appropriate care, it is useful to make some assumptions about their objective functions and assess their implications for the selection of appropriate care.

A) Patients. We might presume the objective functions of patients to be characterized by:

maximum independence and freedom of choice in the long-term care setting; home based care (BUSSE and PFIEFFER 1969, MICHIGAN 1975, WAGER 1972); and minimum out-of-pocket cost of service. Such an objective function would place institutionalized care low in priority for many. Appropriate care from the patient's point of view would be best provided in their own residence or that of relatives by family or friends and health professionals. The patient might thus push for those rules which provided for a wide choice of optional services for any given patient profile, the emphasis being on home care services.

B) Patient's Family. In addition to empathizing with the patient and approving of his/her objective function, the family might also be concerned about minimizing family members' opportunity costs. The cost of looking after a patient might be foregone employment. This might be offset by hiring outside help if the cost of doing so is less than benefits enjoyed in a job. The family might also favor institutionalization if it results in not only removing the inconvenience of looking after the patient in the family home, but also results in reimbursement for the patient's room and board while in the institution. Families thus might favor rules that resulted in reimbursement to themselves in order to offset the opportunity costs (at least in part) of foregone employment. Barring such rules they might then favor those alternatives that liberally allow for reimbursement of nursing home service charges by third party payors. This would have the effect of freeing them from the necessity of providing care and would also result in lower household costs since the family would no longer have to bear room and board expenses of the patient.

C. Service Administrators Including Owners. Since many long-term care services are provided by proprietary firms whose major objective is assumed to be profit maximization, it follows that such providers would seek to keep costs low, maintain high occupancy rates and produce as wide a range of reimbursable long-term care services as possible. Providers with institutions to maintain have fixed costs of a higher magnitude than those providers of in-home services. It follows that they would have strong incentives to increase the size of their firms e.g. supply more beds or, assuming current high occupancy levels, at least maintain the status quo. Small rule changes that would liberalize reimbursement criteria of non-institutionalized long-term care services could result in substantial losses to nursing home owners, at least in the short-run. They thus have a large vested interest in the nature of the rules as they apply to the relationship between patient attributes and the appropriateness of services provided.

D) Third Party Payors. Let us presume the objective function of third party payors is to minimize costs and avoid expenditures in excess of the amounts budgeted. Here the interest lies in determining the appropriate care at lowest cost. It might even extend to mandating the shift of costs to others (e.g. the family, if the family feels obligated to live with or near the patient). To facilitate this a rule might be promulgated to the effect that no patient below a given level of disability is eligible for reimbursement services if they live with their family. Or, as in many cases, it could ensure that only those patients who are institutionalized would have services reimbursed.

It is perhaps of interest that historically few nursing homes have existed. It is possible that the increase in nursing home services in the last decade has been nurtured by the growth of third party reimbursement for long-term care in institutionalized settings.

II. Research Priorities

From the foregoing overview of the problem, it follows that a major research priority lies in identifying and utilizing a methodology for detecting patterned relationships between a patient's profile of attributes and services designated appropriate. Then, given the range of what is customarily identified as appropriate services for given profiles, it is important to determine from among appropriate services those services of lowest cost. Figure 2.1 illustrates the components involved in the decision making process by which services are designated for long-term care patients.

When patients are recognized to be in potential need of long-term care services, the nature and degree of needs have largely been assessed according to the patient's medical status by a health professional. Rarely is the health professional left out of the decision making process. The patient's socio-economic status is appraised usually by agents (e.g., social workers) of the third-party payor. The assessment and categorizing of a patient's functional status is not consistently detailed according to a common index across long-term care programs, but may be consistent within a specific program.

Each index is generated and related to the constraints that are peculiar to that index. In addition, constraints with respect to permissable services, and the time period of such services, may be proposed together with possible ceilings

PATIENT PROFILE	CONSTRAINT	DECISION MAKERS	SERVICES PROVIDED	
		HEALTH PROFESSIONALS		CHORE SERVICES
MEDICAL STATUS			HOME SERVICES	
				PROFESSIONAL MEDICAL SERVICES
FUNCTIONAL STATUS	THIRD PARTY PAYOR RULES	MICHIGAN DEPARTMENT OF SOCIAL SERVICES WORKERS		
SOCIO-ECONOMIC STATUS			NURSING HOME (MEDICAL SERVICES ROOM BOARD)	
			NONE	
		PATIENT & FAMILY	OTHER, e.g., DAY CARE HOSPITALIZATION	

FIGURE 2.1.--Decision processes as to the disposition and assignment of long-term care patients.

on unit or total costs. Taking into consideration the patient's attributes and third-party constraints on patient eligibility according to both attributes and services, a package of appropriate long-term care services can be assembled and made available to the patient. In some cases, services are offered in one undifferentiated lump such as in hospitals and nursing homes which include room and board. In the case of non-institutionalized services, benefits can be disaggregated into discrete types and amounts. The health component of services rendered can be delivered (if necessary) separately from personal care. With respect to home delivered long-term care services, cost centers can also be identified and charges made only for those specific services received. The primary thrust of this thesis is to analyze non-medical services delivered in the home setting. There is considerable justification for this thrust in so far as the demand for such care becomes more intensely articulated.

It is considered axiomatic that people suffering from chronic disabilities wish to maintain their independence for as long as possible. Surveys of the chronically ill themselves attest to their desire to live in familiar surroundings with familiar people. Long-term care services are valued to the extent they facilitate the patient's independence (BUSSE and PFEIFFER 1969, BARNEY 1973 and 1975). It is generally acknowledged that institutionalization leads to virtual denial of independent decision making and to unwanted regimentation. Gerontologists have observed the positive differences in attitudes and outlook on life that favor those who live in independent as opposed to institutional settings.

It is postulated here that the major aim of describing a patient profile is to facilitate the matching of a patient's attributes to services required to compensate for the patient's disabilities. It is further assumed that it is desirable to keep the profile to a minimum, thus reducing sources of error, facilitating the training of those constructing profiles, and cutting the costs of processing and administering the profile.

An important assumption in this thesis is that different and mutually exclusive types of appropriate services are available (not necessarily in one geographic location) for groups of patients characterized by virtually identical profiles. There are probably extremes in this range of service types which would not be considered alternatives by a large proportion of patients or those making decisions on appropriate services. The full range of long-term care now being delivered to patients can be statistically determined. Statistical analysis can help define those services. It is the bulk of the services designated "appropriate" that we wish to examine with respect to their costs. It is anticipated that some patterns of patient attributes will yield a wider range or variety of alternative acceptable services than others. For example, a person who is relatively immobile may justify the provision of cleaning services, or meals, or laundry services, or yardwork or non-nursing personal care services or any combination of those services. Whereas a person who is mentally confused to some degree may only justify the provision of financial management services.

III. Investigative Boundaries

Theoretically, this thesis addresses the whole range of long-term care services together with those patient attributes deemed appropriate indicators of service needs. In part, this research also examines what may be loosely termed the "production function" of long-term care where the patient profile is viewed as a set of inputs and the services rendered as outputs. Services rendered are a proxy for the output that is viewed here as the maintenance of an independent lifestyle.

Empirical research here is limited to analyzing statistically a government financed population of people with low incomes. Neither professional medical care nor institutional care are empirically analyzed in this thesis. Medical status has been neither sufficiently clearly specified in the data used here nor consistently recorded in the patient profiles, therefore, little weight should be placed on findings with respect to the significance of the patient's medical status. Since the thesis is concerned with persons who are not necessarily receiving medical services, these persons will from now on be referred to not as patients but as clients.

The functional status indicator analyzed here was specifically developed for the Michigan Department of Social Services Adult Chore Service Program. Generalization of the results to other populations should, therefore, be viewed with circumspection.

IV. Problems Associated with the Study

The research conducted here was directed into two major areas: (i) the analysis of the relationship between client profiles and the services provided to clients; and (ii) the estimation of the relationships between client profile and costs. With the exception of a few studies cited in Chapter 3, there has been little attention devoted to providing theoretical structure nor empirical work to suggest the nature of specific relationships between client attributes and the kinds of services received by clients with given attributes. The problem of analysis is accentuated by the presence of a relatively large number of endogenous binary variables (i.e. services) which serve to render more common forms of econometric analysis unsuitable in this case. This issue is examined in more detail in Chapter 6. The employment of proxies for aggregations of services such as service costs or hours of service provided was considered but would not meet the specific objectives of determining the relationship between specific attributes and specific services. A broad hypothesis adopted here is that client attributes result in appropriate assigned services. If this is the case there should be some detectable patterned relationship between client attributes and services received. A procedure, namely, entropy minimax, was utilized to attempt the detection of patterned relationships. This procedure is detailed in Chapter 6. If such relationships are observed, they would serve as more conventional hypotheses for confirmation of the results utilizing more common estimation procedures such as regression analysis.

Regression analysis is employed in meeting the second objective of this thesis namely the estimation of relationships between specific attributes and costs of specific services. Using regression analysis, it is possible to determine which attributes affect costs in a statistically significant way.

The reader might be puzzled as to why, if cost estimation is of ultimate concern, we do not forego the problem of estimating the relationships between attributes and services since cost estimates can be obtained directly from both attributes and services. The direct estimation of costs is not of great value to the decision maker who is seeking, for any given patient profile, the range of services appropriate and that service-mix of lowest cost that meets the needs of a specific client. Direct estimation of costs does not indicate which services are appropriate for which clients. From another perspective, the services assigned might be appropriate when considering decision rules that award grants to the clients who then choose and purchase their optimal mix of services.

V. Variables

The specific designation of the variables employed, e.g. endogenous or exogenous, depends in part on the type of analysis used. The endogenous variables to be employed are: services (where they are designated as outcomes in the entropy minimize algorithim); costs of services per month; hours of services provided to clients monthly; and costs per hour.

In the regression analysis, services are treated as an exogenous variable, the categories of which are: laundry, shopping, heavy cleaning, light cleaning, meal preparation, non-nursing personal care, transportation, financial management, attendant, home maintenance and repair, and yard work. Medical services are considered in this study but are given little credence because of the poor quality of the data used. The other exogenous variables or patient profile attributes are divided into the three major areas: medical status, socio-economic status, and functional status.

The medical status variable is indicated by the major medical problem judged to characterize the patient. Medical attributes are: MI (mental illness), MR (mental retardation), heart problems, respiratory system problems, mobility, diabetes, cancer, recuperation, and other.

The socio-economic attributes are characterized by categories of: age, sex, location, income, living arrangement, and the relationship between the provider of services and the patient. The living arrangement category is subdivided into: owns own home, rents own home, rents apartment, rents room, rooms with parent, rooms with child, rooms with non-relative, non-householder and other. The relationship of the primary provider to the patient is broken down in the following way: parent, daughter/son, other relative, non-relative, (each of the foregoing categories is subdivided according to whether the service is in the home of that category or not), and other. The living arrangement and relationship variables were aggregated into 4 groups: client lives in same home as provider who is a relative; client lives separately from the relative providing services; client lives with non-relative service provider; client lives separately from non-relative provider.

The functional status attributes are characterized by: mobility, dexterity, sensory perception and comprehension faculties and home management abilities. Each of these designations is divided into five levels from zero to high degrees of disability or impairment.

The treatment of these variables in the context of long-term care services has been seriously attempted by few other analysts as is evident in the next chapter which surveys studies in long-term care services delivered in the home. The variables, models and data base are developed in more detail in Chapters 4 and 5.

CHAPTER 3

LITERATURE REVIEW

I. Introduction

This chapter contains a review of the various types of cost analysis applicable to long-term care of the chronically ill and disabled. In addition, an overview of a specific model is presented in which features of particular relevance to this study are emphasized. Clustering of services and their relationship to patient attributes will also be discussed together with references to various indexes of patient attributes. Finally, estimation procedures used in a major relevant empirical study will be briefly considered followed by a brief discussion of health industry related cost functions.

II. An Overview of Approaches to Cost Analyses

Assuming policy makers seek to learn which long-term care programs are most appropriate in this society, one basis for selection is to conduct a cost-benefit analysis of the various alternative programs. Such an approach implies identification of all cost factors and all benefits quantified in monetary units where possible. To date, the sole identified exercise in cost/benefit analysis in long-term care, identified costs and benefits but failed to monetarize the latter (WAGER 1972). Patient satisfaction was identified as the chief benefit in the sense that most elderly people asserted they would prefer to maintain their independence in familiar settings as opposed to being institutionalized.

Since quantifying benefits is problematic, a second means of analyzing costs is to conduct a "cost effectiveness" analysis. Given certain goals, e.g. minimizing costs or maximizing client satisfaction with respect to ameliorating the effects of physical and mental impairments, we can compare costs of meeting these goals for specific groups of clients and attempt to identify the lowest cost approach (ANDERSON 1974). It is assumed under this approach that patient profile formats are common across alternative programs being offered. The lack of such uniformity of data makes comparisons of appropriateness of care and cost of care very difficult.

A third means of estimating costs of care and the one most commonly used in the health research literature, is to analyze the costs of service directly without taking into consideration the purpose of the services.
No benefits are considered here, merely implicit costs of components that explain variation in the cost of a specified unit of service such as a patient day of care or an episode of illness. There have been a few attempts at estimating costs on this basis in the area of long-term care, principally

with respect to institutions such as nursing homes (BERGER 1970, HURTADO ET AL 1972, SPROAT 1972, SAGER, 1977, BURTON ET AL 1974, CONGRESSIONAL BUDGET 1977). However, most of the work using this approach has been conducted in the analysis of hospital costs (LAVE, LAVE and SILVERMAN 1972, EVANS 1971).

Less systematic studies of homemaker and home health aide costs have been conducted but they tend to be anecdotal and heterogeneous with respect to specification of client attributes and services received. "Single agencies and some multiple-entity programs such as state agencies generate statistics on costs and benefits, however, such statistics are so diverse that there is little or no basis for comparison," (ROBINSON 1974, pp. 9). There were seemingly no systematic methodologies developed to enable comparisons among various types of long-term care programs.

The only evident empirical cross-sectional analysis of long-term home care was conducted by Greenberg who is extending the project (ANDERSON 1974). This is discussed more fully in the section on empirical findings in this chapter.

Two major emphases have been identified. One recommends that studies of various long-term care programs take into consideration all costs including opportunity costs of providers such as close relatives who offer their services for no remuneration (POLLACK 1973). The problems associated with this approach were noted above. The second thrust advocated by Greenberg goes a step further and recommends emphasis be placed on estimation of the relationship between costs and what he considers to be the most important aspect of the patient profile, namely the functional status attributes of the client (ANDERSON 1974).

III. A Model for the Analysis of Long-Term Care

Pollack was concerned about ensuring that cost comparisons between various types of long-term care be based on total social costs, the common denominator being costs per client per unit of time. Because this data is difficult to collect and quantify, a 'second best' approach is adopted which analyzes on some components of social cost such as government expenditures on long-term care services.

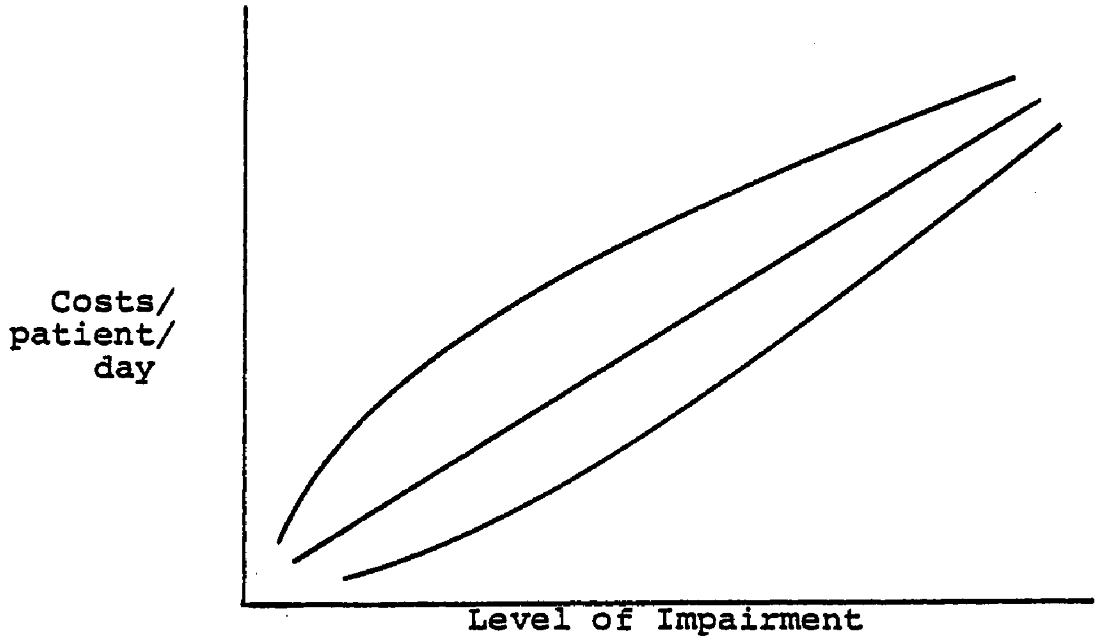
Real social costs according to Pollack depend on: (i) the client's functional status; (ii) the family status of the client; and (iii) the quality of care provided. Quality of care is usually unspecified but is assumed appropriate for whatever type of service is provided. This does not imply that the quality of one type of care is equivalent to another even though both types of care may be deemed appropriate for a particular client.

In addition to considering the comprehensiveness of cost categories and the completeness of costs within categories, Pollack stressed the importance of identifying and measuring the cost of non-market or quasi-market inputs. To assure comprehensiveness, he recommended measuring costs of housing, nutrition, supervision, personal care, transportation, recreation, medical care, environmental hygiene and other relevant miscellaneous items. Further completeness of costs may be ascertained by detailing administrative costs, and accounting for subsidies on such items as rent, parking, etc. Non-market and quasi-market inputs would include such items as travel costs of providers, shadow prices of opportunity costs of volunteers (including family members) and foster families. Pollack stressed this point because future projections of eligible need may involve expansion of services to such an extent that volunteers cannot be relied upon to provide the additional services.

In Pollack's opinion, family status is a very significant category of the socio-economic status variable. He hypothesizes that costs of the personal care of clients with identical functional status attributes will differ substantially depending on their family status. For example, for a given client, nursing home care may cost $22 per day, in-home care with family $40 per day, and home care with a retired spouse $4 per day. He also hypothesizes ceteris parabus that as functional status decreases, costs will rise as illustrated in Figure 3.1.

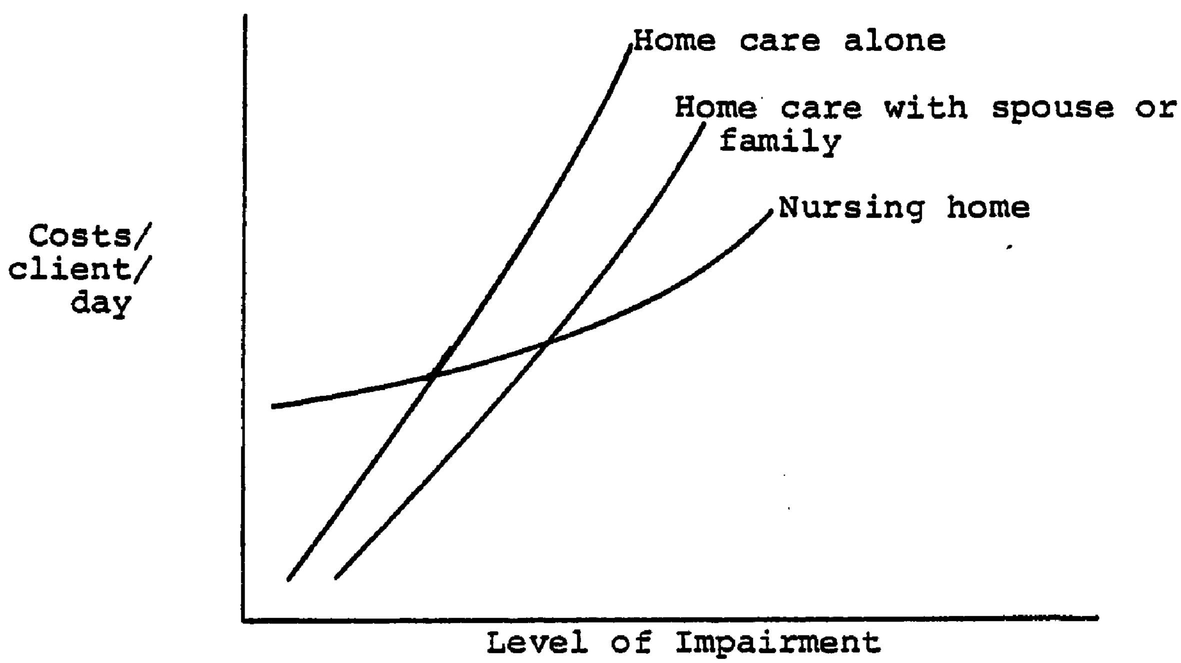
When comparing two services, say home care and nursing home care, Pollack hypothesized the relationships would be characterized by the graph in Figure 3.2.

Here it is clear an able spouse of family can delay entry into an institution for greater degrees of client



Source: Pollack W. Costs of Alternative Care Settings for the Elderly. Working Paper 963-11 (Draft), Urban Institute, Washington, D.C., March 12, 1973.

FIGURE 3.1.--Cost of patient care according to change in impairment. Alternative relationships.



Source: Pollack W. Costs of Alternative Care Settings for the Elderly. Working Paper 963-11 (Draft), Urban Institute, Washington, D.C., March 12, 1973.

FIGURE 3.2.--Costs of different kinds of care at differing levels of impairment.

impairment than is possible with a client living alone. Clearly, over some impairment ranges costs are less for one specific type of care than another but for other ranges the relative costs may be reversed. In the case of nursing home care, costs may not inevitably rise with the degree of impairment. Persons immobile in their beds with little comprehension of their surroundings possibly require less care than the partially mobile, alert patient (HUNDERT 1974). From this viewpoint, a case could be made for the peaking of the cost impairment curve as illustrated by this writer in Figure 3.3.

Implicit in Figure 3.3 is the observation that for any one individual more than one mode of long-term care may be appropriate. Table 3.1 illustrates an example provided by Pollack.

TABLE 3.1.--Appropriate care settings for different patients.

	Home Care	Foster Care	Nursing Home	Hospital Extended Care
Mrs. Mayer	X	X		
Mr. O'Neill	X			
Mrs. Cooke	X		X	X
Mrs. Jones		X	X	

Source: Pollack, W. Cost of Alternative Care Settings for the Elderly. Working paper 963-11 (Draft), Urban Institute, Washington, D.C., March 12, 1973.

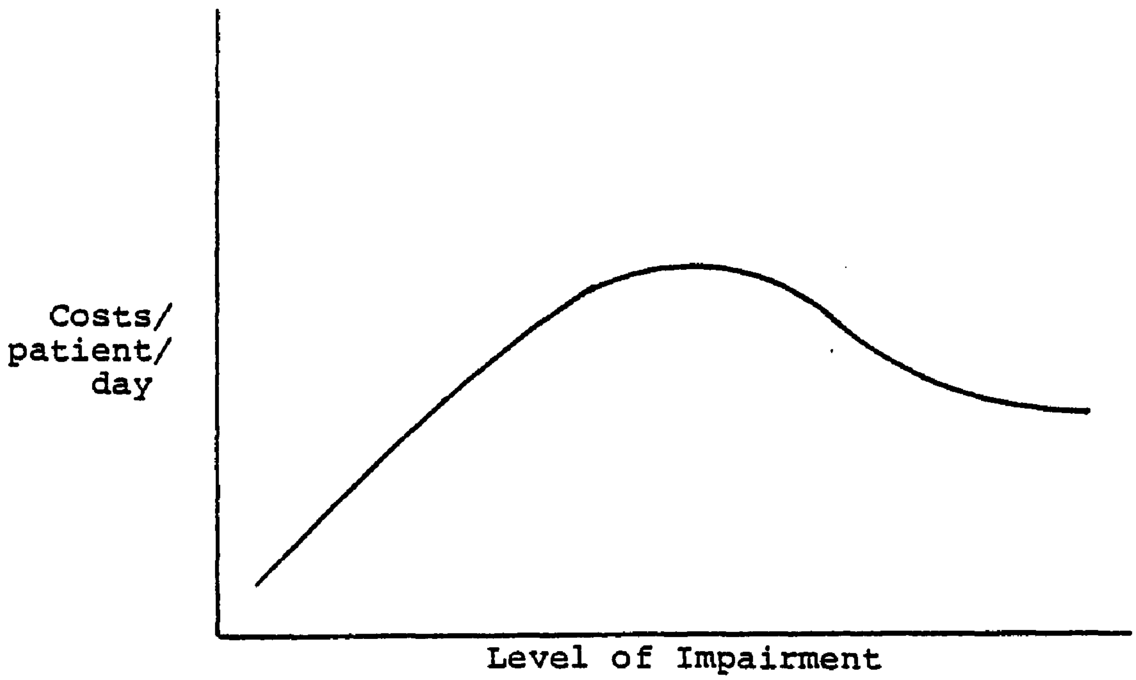


FIGURE 3.3.--Hypothetical costs of care for different levels of impairment for patients in nursing home.

Given these ranges of appropriate types of care for different individuals it is possible to identify those of either lowest social cost, or lowest cost to third party payors such as governments.

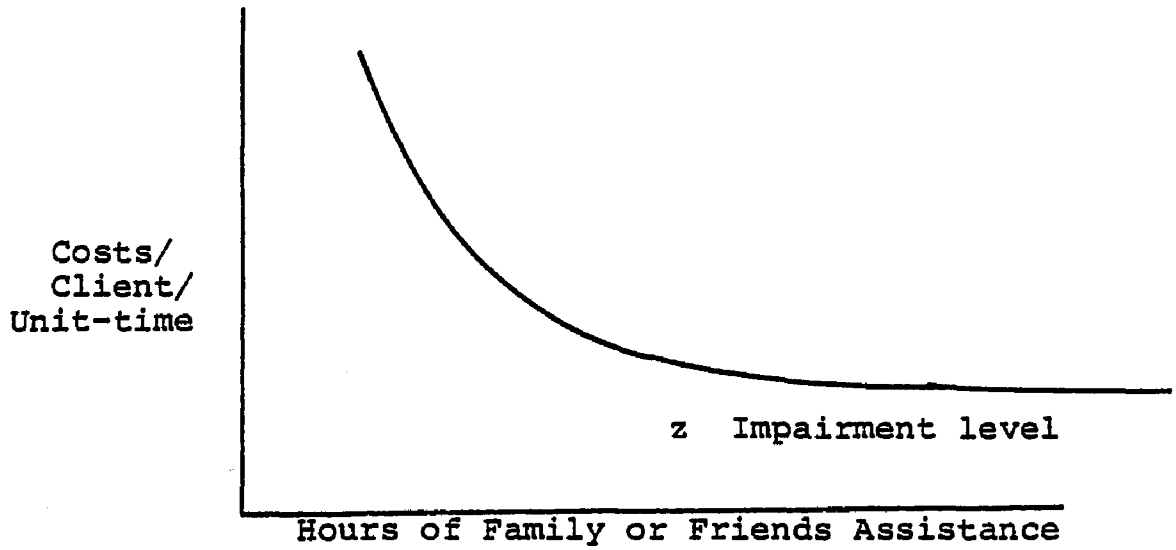
In his cost analysis of long-term care, Greenberg examines not only patient attributes but also service provider characteristics including geographic and demographic features. He also takes into consideration the degree of population dispersal and the degree of integration of long-term services.

In his theoretical analysis of social characteristics, Greenberg appears to ignore Pollack's concern for family member opportunity costs, hypothesizing that cost per client per unit time decreases to a constant level as hours of assistance from family and friends increases. This is shown in Figure 3.4.

If we let the impairment level of the client case illustrated in Figure 3.4 be z, then we can transform the graph into a relationship between cost and level of impairment as pictured in Figure 3.5.

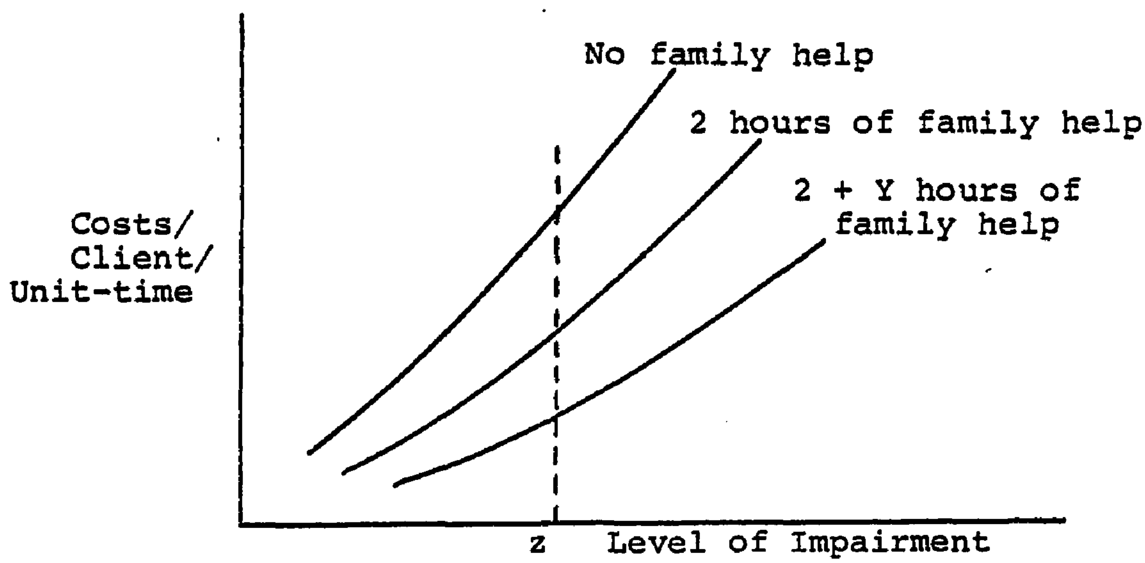
As shown in Figure 3.5, third party costs of client care decrease with increased time offered by family members for any given level of impairment.

From the foregoing, it is clear the identification of impairment level, i.e., functional status, is important for policy makers in their quest to control costs. To quote Greenberg "...if the study does not control for disability level at all, then there is no way to distinguish cost savings due to place of service from cost savings due to treating healthier people" (ANDERSON 1974, p. 9). The underlying assumption behind this assertion is that there is a relationship between the functional status of the client and the



Source: Anderson, N.N. "A Planning Study of Services to Non-Institutionalized Older Persons in Minnesota," The Governor's Citizens Council on Aging, State of Minnesota, 1974.

FIGURE 3.4.--Variation in costs of client care according to the extent of family assistance for specified impairment level.



Source: Anderson, N.N. "A Planning Study of Services to Non-Institutionalized Older Persons in Minnesota," The Governor's Citizens Council on Aging, State of Minnesota, 1974.

Figure 3.5.--Cost of care according to the length of time the client receives assistance from family members or friends.

services delivered to compensate for those impairments. In addition, is is assumed decisions as to services assigned to clients are based at least in part on functional status. A small amount of research is available on these relationships (MEDICUS 1975, KATZ ET AL 1966, KATZ ET AL 1972, STROUD 1978). However, at present, little knowledge is available about these relationships for most long-term care conditions.

IV. The Detection of Patterned Relationships Between Patient Profiles and Services

Among the literature reviewed by this author, a statistical analysis of the relationship between client attributes and services has not been attempted in the area under discussion here. The nearest approach to it has been conducted with respect to assessing efficacy of medical tests for acceptable X-ray diagnosis of complaints associated with the head (SCHONBEIN ET AL 1974, 1978). The researchers in this instance were attempting to discover if a specific test ordered by a physician made any difference to the health outcome of the patient. Using the information theory-based entropy minimax pattern discovery algorithm, the study took patient profiles and attempted to find if there were any patterned relationships between patient attributes together with the use or non-use of the test by the physician and the patient's health outcome. Should there be no difference between the patient's health outcome when using or when not using the test and should the profiles be comparable, then it would be concluded the test was of no value. But should there be such a relationship determined, then the study would enable the researchers to determine which patients (classified according to profile) were most likely to benefit from the test, and, more importantly, which patients with given profiles were unlikely to benefit from the test. If such profiles can be identified, it should be possible to cut down on patient care costs by foregoing certain tests on patients manifesting those profiles. For the purposes of the pattern detection research involving efficacy of tests, the result of the test itself can be considered a patient attribute. If the result is positive, the patient outcome should be different from the outcome where the test is negative. The range of outcomes with and without the test may be greater than two, numbering up to five or six. Beyond this number of outcomes, the reliability of the technique drops. However, it is precisely because there is a possibility of relating multiple outcomes to a set of attributes that makes the entropy minimax procedure appropriate for the purposes of research on the topic of this thesis. In our case, we are not seeking the efficacy of a test but the relationship, if any, between a set of service outcomes and client attributes.

To the extent there is a relationship between service outcome mixes and client attributes, we should be able to "distinguish between cost savings due to place of service and cost savings due to treating healthier people."

V. Empirical Findings

In attempting to compare costs of home services with institutional type services, indirect costs such as housing costs, food costs, laundry costs, etc. must be accounted for with the home-bound patient. Since the population examined by this thesis is receiving care exclusively in non-institutional settings, such costs can be ignored as constant for any client.

For our purposes, the salient components of the empirical research conducted by Greenberg include functional status, socio-economic status and medical status variables together with a specification of services rendered to the patient receiving strictly chore-type services. Dependent variables employed by Greenberg included hours of homemaker aid per week, global rating of patient functional ability and cost of homemaker/home-health aid per month. Only a few of Greenberg's categories of medical diagnosis or functional status are compatible with those adopted in this thesis. Functional status categories included: ambulation, bathing, transfer (moving from one room to another), eating, communication, toileting and mental status. Greenberg analyzed only 47 cases drawn from a sample of 139 persons 65 and over, all of whom, for various reasons, no longer received services. Using the Freedman/Tukey T^2 statistic because of low cell counts, he found through partial correlation analysis, no association between hours of service and medical diagnosis. Among the exogenous variables only three functional status categories, ambulation, bathing and transfer were found to be significantly correlated at the .05 level with the endogenous variable hours of homemaker/home-health aid services. Employing regression analysis to estimate average cost as a function of various independent variable categories for nursing homes, Greenberg found the log specification of endogenous and exogenous variables yielded a slightly larger explanation of variance than a linear specification. The only two variables to prove significant were the percentage of fully ambulatory patients and type of ownership of the nursing home. The R^2 varied from .50 to .54.

VI. Conclusion

From this and previous chapters it is clear that research has only just begun to focus on the subject of long-term care, especially that provided in the home. The theoretical groundwork has been laid by Pollack and Greenberg with respect to identifying the pertinent variables.

Considerable work has been devoted to constructing indexes of functional status but little effort made to statistically relate the indexes to patterns of services received. Schonbein's efforts in assessing the efficacy of specific diagnoses pointed the way to determining the utility of functional status indexes with respect to judging appropriateness of long-term care services.

The next chapter describes the specific services analyzed in this thesis as well as the data base.

CHAPTER 4

DESCRIPTION OF ADULT CHORE SERVICE PROGRAM*

I. Overview of Chore Service Program

Started by the Michigan Department of Social Services in 1972, the Adult Chore Service program was designed to meet the strictly non-medical needs of the chronically ill and disabled, hereafter referred to as clients. The principal purpose of the program was to enable clients to maintain a relatively independent life-style in familiar surroundings. "Community services must no longer be thought of simply as replacing the services of the family or substituting for them when they cannot exist - though this is one primary role. They must also increasingly be thought of as supplementing or complimenting the service of the family" (MICHIGAN HEALTH COUNCIL 1965, p. 113). To the extent this purpose is achieved, the necessity of institutionalization of the disabled and chronically ill in nursing homes or homes for the aged is forestalled or eliminated.

---

*Much of the information in this chapter is taken from a Michigan Department of Social Services study, "Adult Chore Services" authored by Diane Emling (EMLING 1976). The data used by Emling were also employed in this thesis.

In January 1976, 10,280 clients were beneficiaries of the Chore Service program at a cost of $1.7 million per month, an average of roughly $165 per client per month. By September 1976, the number of clients had grown to 11,180, costing $2 million per month or $179 per client.

The observations on each client were collected by MDSS workers who also made the service assignments. All the raw information used in this dissertation was recorded for each client on form DSS-3492 (Rev. 2-75) and reproduced in Appendix A. Details on the specifications of the variables are provided together with the theoretical framework in which the data is analyzed in Chapters 5 and 6.

Services offered to eligible clients included: light cleaning, heavy cleaning, minor household repair, meal preparation, laundry, shopping and/or other errands, guide dog, assistance in financial management, transportation, an attendant, yardwork such as lawn care or snow removal, non-nursing personal care, interpreter, and unspecified miscellaneous services. The services are provided to the client by a provider who is usually chosen and hired by the clients. Sometimes the social worker helps the client locate and hire a provider. The program is unusual in the sense that the client is provided with the resources to hire a provider rather than the government reimbursing the provider directly. Since the state government is not the employer of the provider, it need not pay the statutory minimum wage.

To enter the program, prospective clients must be eligible; that is, they must be recipient of or eligible to receive Supplemental Security Income (SSI). In 1975, the maximum social assistance payments to SSI recipients was \$170 per month if they were single and \$275 per month for couples. Supplemental Security Income is provided to persons who meet specific federal criteria with respect to age, blindness or other disabilities and who have savings of less than \$1500 in the case of single people and \$2500 in the case of couples. Further eligibility criteria are mentioned in the Emling study.

Referring more directly to the findings of Emling, 92 percent of the clients in the study had incomes of less than \$250 per month. This included SSI payments, and income from other sources. Eighteen percent had incomes of less than \$150 per month.

In order to assess a potential client's eligibility the local county Social Services Department would send a social worker to draw up a profile on the client according to a predetermined format [form DSS-3492 (Rev. 2-75)] (see Appendix A).

The social worker determines the impairment levels or functional status of the individual and records the details of various indicators of socio-economic status. The kinds of services required are then agreed upon and their costs determined. Payments for costs incurred by the client are forwarded to the client or to his/her agent in cases where the client is mentally incompetent. The client then seeks to employ the services of a provider for the hours, tasks and wages agreed upon. The average rate of pay in 1975 was $1.78 per hour, the mode being $2.00 per hour while the statutory minimum wage was $2.25 per hour. Further details on wage rates are given in the County Rate Schedule in Appendix B.

A ceiling of $270 was placed on the total monthly payments a client may receive from the state government. The percentage distribution of payments to providers is illustrated in Figure 4.1 (EMLING 1976, p. 36). It should be noted that only 23 percent of those providers in the top range of payments, namely $240 to $270 per month, have their expenses paid entirely out of program dollars.

The monthly payment received by providers breaks down into an hourly rate as shown in Figure 4.2.

The hours of services provided by providers breaks down as shown in Figure 4.3. The figure shows only hours of service of what Emling refers to as the Primary Provider. In only 5 percent of cases did the client receive services from providers in addition to the Primary Provider. The minimum age for an eligible client was 18 years old. Two-thirds of the sample were over 60 years old and 6 percent over 90 years. Males constituted 25 percent of the client population.

Among the clients, 21 percent owned their own homes, 19 percent rented homes, and 21 percent rented apartments.

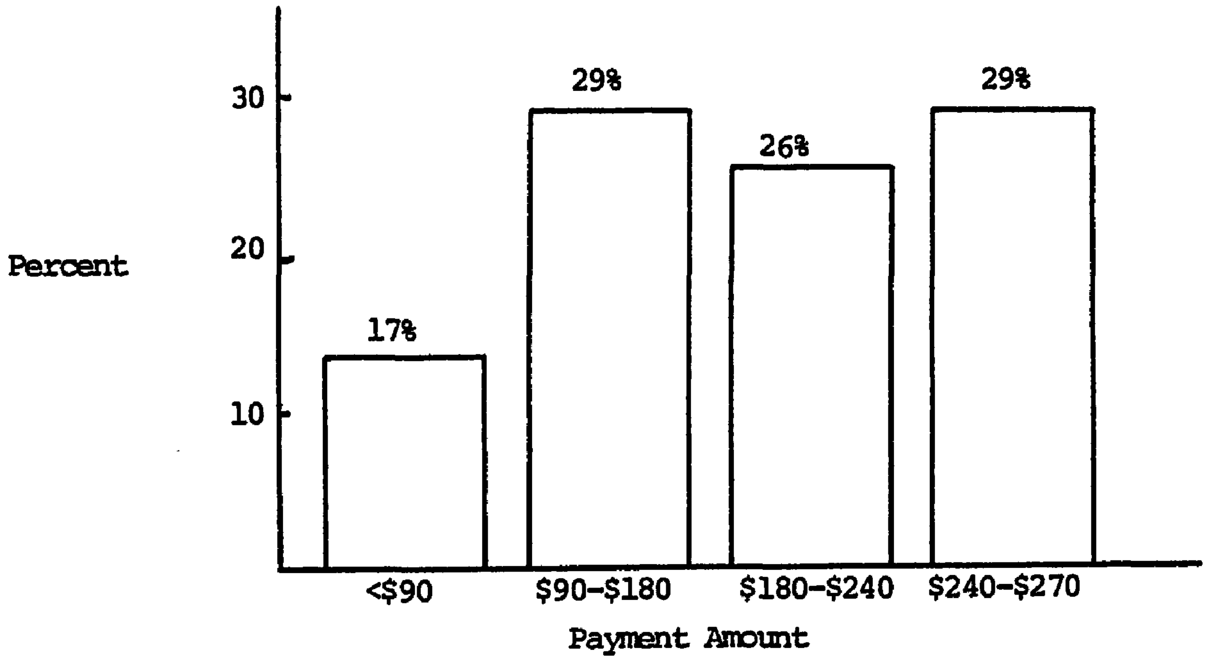


FIGURE 4.1.—Rate of provider payment.

Source: Emling, Diane. Adult Chore Service. A Profile of In-Home Assistance. Studies in Welfare Policy No. 10, Michigan Department of Social Services, November 1976.

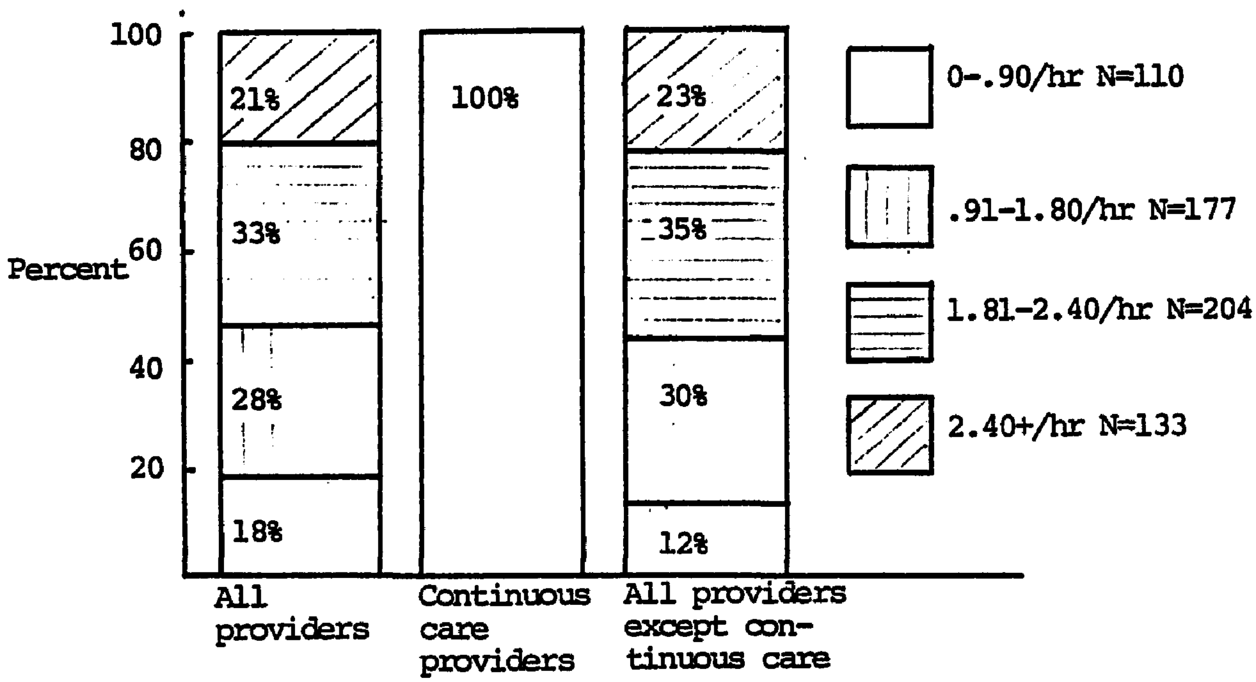
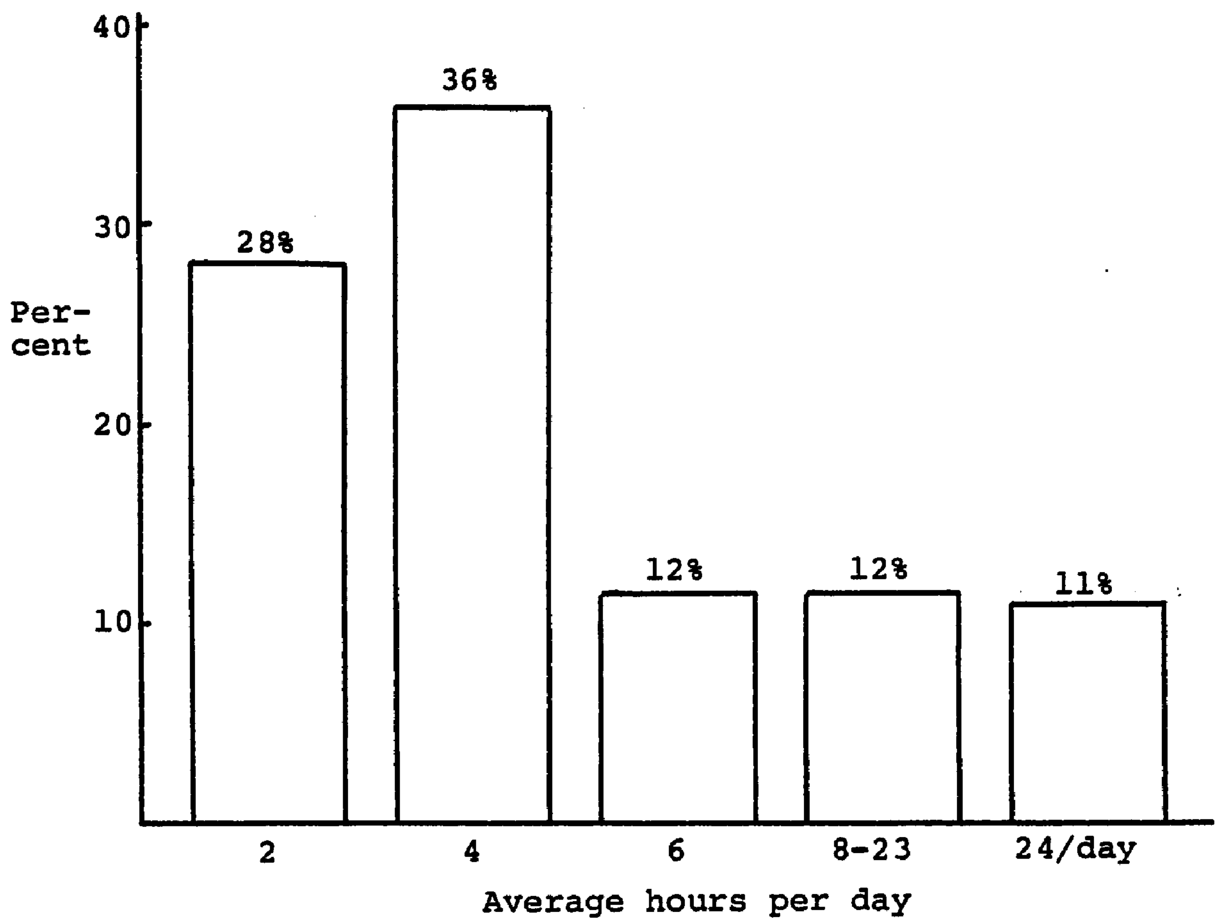


FIGURE 4.2.—Percentage distribution of hourly rate of pay by provider type.

Source: Emling, Diane. Adult Chore Service. A profile of In-Home Assistance. Studies in Welfare Policy No. 10, Michigan Department of Social Services, November 1976.



Source: Emling, Diane. Adult Chore Service. A profile of In-Home Assistance. Studies in Welfare Policy No. 10, Michigan Department of Social Services, November 1976.

FIGURE 4.3.--Percentage distribution of hours of service provided by primary providers.

Of the remaining clients, 11 percent lived with a child, 13 percent with a parent, and 14 percent in other households.

Broadly defined, medical problems were categorized as follows: mentally disabled (9 percent of clients), heart problems (23 percent), mobility problems (27 percent), diabetes (8 percent), other problems (33 percent). Post hospital recuperation accounted for 4 percent of the clients.

On the "average," clients received six of the above mentioned services, the most common ones being laundry (received by 91 percent of clients), shopping and errands (87 percent), heavy cleaning (82 percent), light cleaning (77 percent) and meal preparation (75) percent.

Providers of services were predominantly females over the age of 40, at least 50 percent of whom were unrelated to the client. Thirty-nine percent of all providers lived with the client.

Whether or not the state should be paying the expenses of relatives caring for clients has raised questions about the property right obligations of one family member to another. The issue continues to be debated. One suggested approach has been to render clients ineligible if they live in households where the collective income exceeds a predetermined limit. A major problem appears to be that an adult cannot be held legally responsible for another adult to whom he/she is not married while a resident in Michigan.

II. Data Base

The sample of 628 Michigan clients which constitutes the data base in this thesis were randomly drawn from a total state chore service population of between 10,100 and 10,500 clients. The range in client population stems from the extended period of selection from October 1975 to February 1976.

The cases were drawn from those 23 counties, each harboring 50 or more chore service clients. Due to some clients moving subsequent to their selection, the actual number of counties represented was 31. Of these, 17 encompassed Standard Metropolitan Statistical Areas and were designated as urban for the purposes of this study. The other 14 counties were designated as rural areas and served 81 of the 628 chore service clients. All the data employed in the study were extracted from form DSS-3492 (Rev. 2-75).

Though justified in Chapters 5 and 6, it should be pointed out that largely for purposes of testing the empirical results of the entropy minimax algorithm the data were divided into 2 sets; one set being a "training set" of 428 cases on which manipulation of the various variables was extensive, the other set being a "test set" of 200 cases on which the hypotheses to be generated from the pattern detection algorithm (see Chapters 6 and 7) could be tested. In addition, the results of the regression analysis of costs could also be tested.

The variables employed in this study, together with classification and specification of their categories and sub-categories are listed and detailed in Chapter 5.

CHAPTER 5

TECHNICAL DESIGN AND SUBSTANCE OF COST ANALYSIS

I. Introduction

The principal purpose of this chapter is to clarify and specify the major research problems and to construct testable hypotheses. Variables employed in the analysis will be defined and hypotheses posited with respect to the impact of the exogenous variables on the endogenous variables.

The division of chapters in this thesis is somewhat unusual in that this chapter will describe the variables employed in Chapters 6, 7 and 8. It will also describe the hypotheses, economic model and econometric specification with respect to the empirical analysis of models estimated in Chapter 8 only. These models describe the behavioral relationships between, on the one hand, costs of services and hours of service and, on the other, patient attributes such as functional status, socio-economic status and medical status.

Chapters 6 and 7 will address, respectively, the theory and empirical results associated with the pattern detection algorithm. Should distinct patterns of relations between client attributes and services be observed and, upon further testing, be confirmed, the outcomes may have an effect on the forms of the models discussed in this chapter as they are empirically analyzed in Chapter 8. To determine whether any significant patterns were detected, the reader should advance to the last section of Chapter 7.

II. Technical Design

The major emphasis of this research focuses on estimating the relationships between total service costs per unit of time per client and client attributes. Two related relationships to be examined are: those between patient total service hours per unit of time and patient attributes; and those between service cost per hour and patient attributes. These are not technical cost functions featuring factors of production. They are behavioral cost functions typical of cost prediction studies in health care (EVANS 1971). The major aim is to identify variables which will explain a significant proportion of the variation in costs. This approach obviates the necessity of clearly specifying the output.

With much economic analysis of the production of health, definitions of output that lend themselves to econometric manipulation of cost functions are rare. In the case of chore services, there is no readily agreed upon measure of output. Qualitatively, chore service output might be described as the maintenance of independent lives for individuals who otherwise could not live independently. There is both a wide variety of services and service mixes available to chore service clients seeking to maintain their independence in familiar settings. Within each service category, the quantity of service offered may vary. The proxies used in this analysis to represent both the service mix and intensity of service are three measurable entities: Service Cost Per Client Per Month (Total Cost per Month abbreviated TCM); Hours of Service Per Client Per Month (Total Hours per Month abbreviated THM); and Service Cost Per Client Per Hour (Cost per Hour abbreviated CH). These proxies can be related to a variety of client attributes which are listed and described later in this chapter. The crucial variable among these attributes is functional status. Together with the variable socio-economic status, functional status brings in a level of detail found in few other studies.

To further the understanding of client attributes and their relationship to costs, we will posit three mathematical models, Model A, Model B and Model C, that lend themselves to simplification of the relationships and, in addition, to the construction of hypotheses and statistical analysis. In general terms, it is hypothesized that the dependent variable, e.g., Total Costs Per Client Per Month (TCM), is a function of the client's Functional Status (FS), Socio-Economic Status (SES), and Medical Status (HS).

$$\text{Cost/Month (TCM)} = f\ (\text{FS, SES, HS}) \qquad \text{Model A}$$

Similarly, the same type of relationships is presumed for the other two endogenous variables, Total Hours

Per Month worked by the provider(s) (TCM), and Cost Per Hour (CH), the payment received by the provider.

Hours/Month (THM) = f (FS, SES, HS) Model B

Cost/Hour (CH) = f (FS, SES, HS) Model C

The variables employed in the analysis have been selected from among classifications of data obtained from the form DSS-3492 (revised Feb. 1975) which is reproduced in Appendix A.

Following is a description of each variable and its categories and why it was selected, and in the case of the exogenous variables and their categories, the hypothesized relationship between them and the three endogenous variables.

A. Endogenous Variables. All three endogenous variables are continuous.

1. Total Cost per Month (TCM). Total cost refers to the total cost of providing chore services to a client over a period of one month. Since the clients under consideration had long-term disabilities that were not subject to rapid change, costs assessed over periods of one month remained stable. The cost of service paid for by the Michigan Department of Social Services was limited to a maximum of $270 per month, an additional justification for the chosen time unit. An alternative means of measuring

cost for purposes of analyzing medical services is to conduct it on a case basis. This approach was not deemed practical because services for the disabled and chronically ill are often provided for the duration of the patient's lifetime. Government rules, regulations and modes of rationing services inevitably undergo change, thus rendering the clear interpretation of results difficult if not impossible. Also, price changes may result from inflation. For this reason, the highly desirable time-series analysis of long-term care problems is likely to be fraught with difficulties in the interpretation of results. In cross-sectional analysis it is hypothesized the functional status of the patient will be found to explain a relatively high proportion of the variance of TCM, especially when compared with medical status.

2. Total Cost per Hour of Service Received by the Client (CH). The hourly cost of providing services is presumed to be more dependent on the relationship of the provider to the client and on whether or not the client lives with the provider, than was the case with TCM. It is hypothesized the SES

categories will explain a higher proportion of the variance of TCH than was the case with the endogenous variable TCM.

3. Total Hours of Care Provided to the Client Each Month (THM). This variable measures the total amount of service time employed by the patient in any month of care. It is included in the analysis to confirm the hypothesis that relatives are more likely to work for longer hours than non-relatives for a patient of given functional status.

B. Exogenous Variables. With few exceptions, the exogenous variables and their categories described here are either dichotomous or polytomous. Non-binary variables are transformed into dichotomous variables. The three exogenous variables, functional status, socio-economic status and medical status, are disaggregated into the categories described below.

1. Functional Status Categories. Since chore type services are provided to compensate largely for functional disability, it is anticipated that a relatively large proportion of the variance in total cost per month will be accounted for under this category of variable.

Functional status variable categories will be specified in three different ways. Broadly these are: first, an inputed "cardinal" index of functional status, based on the five functional status variables, mobility, dexterity, sensory faculties, comprehension and home management; second, a binary index where each of the above mentioned functional status categories will be disaggregated into five binary categories or levels of disability; third the functional status of the patient will be classified according to the type of service they are receiving. This is justified in the belief that it is feasible to classify clients according to the everyday tasks they cannot perform regardless of the specific impairments or circumstances involved. Where appropriate, two of these three types of variable specifications may be combined in some form.

a) "Cardinal" and Binary Forms of the Functional Status Variable.

As they appear in the form DSS-3492, the functional status variables are recorded at three levels of incapacity, 'None,' 'Some,' and 'Much.' Yet in the data provided, they were recorded at five levels. For example, the category mobility in the data is an average of two indexes recorded in the DSS form, namely walking and climbing. The number scale used to code the range of three levels of a category in the DSS form was from one to five; one being the score given to an individual with no impairment in, say, walking, a score of three for some impairment, and a score of five for much impairment. The five levels for the constructed mobility category were determined by averaging the scores for walking and climbing. Thus, if walking received a score of three and climbing received a score of five, then the score for the constructed category mobility would be four, the average of these.

This same procedure was used to determine the other four constructed categories, the nearest whole number being recorded where fractions appeared. Original categories appearing on form DSS-3492 and making up the five functional status categories are referred to below in the course of discussing the constructed categories.

(i) Mobility (MOB or alternatively M1, M2, M3, M4, M5).

As mentioned above, mobility represents the aggregation of levels of categories, walking and climbing, from the form DSS-3492. Two approaches were taken when incorporating this category into the mathematical model presented above. The first way was to treat each level hierarchically by assigning the number designating the lowest level of impairment to represent no impairment (i.e., number 1) and the number 5 to represent the highest level of impairment. This approach represents mobility by an index between 1 and 5 called MOB. In this implicitly "cardinal" representation of functional status, the highest level of impairment is viewed as five times more disabling than no impairment. For this measure of mobility (MOB) it is hypothesized that the variable in the equation will be significant, indicating that the variation in total costs per month is significantly influenced by this category of the functional status variable.

The alternative specification of the category is to treat each of the five levels of disability as a dummy variable: M1, M2, M3, M4, M5. Here there is no implicit constrained relationship relating the magnitude of one level of impairment to another. It is hypothesized that this specification of the mobility category will explain more of the variance in the total costs than that explained by the mobility index category (MOB) which constrains the relationship between one level of disability and another. As impairment in mobility increases from M1 through M5, it is hypothesized that the coefficients will increase respectively across the sub-categories M2 to M5. If as a result of testing, such hypotheses prove plausible, then greater costs for higher levels of impairment would be indicated. It is also hypothesized that the mobility category will be significantly related to the endogenous variable Total Hours per Month in a positive manner.

(ii) Dexterity (DEX or alternatively D1, D2, D3, D4, D5).

The dexterity index category (DEX) is constructed similarly to the mobility index category (MOB) utilizing the clients degree of impairment with respect to grasping, bending, and lifting as taken from the form DSS-3492.

As with the mobility category this category also has the alternative binary specification utilizing the five levels of impairment as dummy variables: D1, D2, D3, D4 and D5).

The hypotheses stated with respect to the alternative specifications of the mobility category also apply to the dexterity category and to the sensory, comprehension and home management categories listed below.

(iii) Sensory (perception) (SEN and S1, S2, S3, S4, S5).

The alternative specification of this category is constructed from the client's sensory impairments in hearing and seeing.

(iv) Comprehension (COMP of C1, C2, C3, C4, C5).

The ability to manage finances and understand instructions, in other words the client's mental capacities, are incorporated in this category.

(v) Home Management (HOM or K1, K2, K3, K4, K5).

The clients level of impairment in light cleaning and heavy cleaning are used to construct this category.

There is some potential confusion in the use of this category since home management also appears to bear a strong relationship to mobility, dexterity and sensory impairments. It could basically be employed as a proxy category for those three categories enabling us to propose the hypothesis that the amount of TCM variance explained by the four categories mobility, dexterity, sensory and comprehension will not significantly differ from that explained by comprehension and home management. The implications of these observations will be further discussed in the section describing the econometric specification of the model.

It is important to note that the component parts of the attribute category home management (light cleaning and heavy cleaning) are identical designations for two of the services offered to the client, namely light cleaning and heavy cleaning. This raises an important issue which is discussed in the following section.

b) Functional status classified according to services received.

Since the major decision makers with respect to the appropriate care of the disabled and chronically ill are usually physicians, the medical model of diagnosis and treatment is the principal guide in determining the kinds of services patients require. In employing this model the physician, prior to designating the appropriate services, will seek to accurately categorize the problem. Thus a client might be classified as to the degree of such impairments as inability to walk, inability to bend down and tie shoe laces, inability to put a spoon to the mouth, difficulty with bathing, etc. The problem is that this process of deciding on patient classification has not been readily agreed upon by the various researchers or agencies (KATZ ET AL 1963, TOWNSEND 1963, MICHIGAN DEPARTMENT OF SOCIAL SERVICES, etc.).

Without such an agreement, it is not easy to communicate with others about the type of client being classified nor about the type of services that should be made available to the client. Another problem faced by the person designing categories is the degree to which the classifications employed will overlap. For example, the category "Difficulty with Washing" is presumed to be related to the category "Difficulty with Dressing" and the category "Meal Preparation" will overlap or be correlated to some extent with the category "Light Cleaning."

Consideration of these issues leads to another possible means of classifying functional status; that is by means of designating the services required by clients as the proxy for functional status. Thus instead of classifying clients according to their degree of impairment in mobility or dexterity, etc., the classification might be in terms of the service needs of the patient. Using this approach, a client might be classified as needing services such as light cleaning, meal preparation, transportation and yardwork. Clients with different attributes according to the previous definitions of functional status might require identical or different mixes of services. A possible limitation on the use of needed service mixes as indicators of functional status lies in their strictly qualitative form. The amount of a particular service required is unquantified in the data though the hours required to perform the total mix of services has been recorded.

Theoretically, there is a relationship between the functional status as measured by mobility, dexterity, etc., and the types of service provided to a client. This relationship, taking into consideration socio-economic status is analyzed in detail utilizing the entropy minimax pattern detection algorithm as described in Chapters 6 and 7. In that analysis, functional and socio-economic status are both specified as inputs or attributes. Services provided to the client are specified as outputs. In the analysis of costs under consideration in this chapter, levels of impairment and service mixes are both considered as alternative forms of specifying functional status. In some forms of the model, equations will be presented that encompass aspects of both forms of functional status.

Services received by patients are listed and described below:

(i) Light Cleaning (LK). This service includes such tasks as washing dishes, tidying rooms, making beds, and other light tasks performed on a routine basis.

(ii) Heavy Cleaning (HK). Under this designation appears such one-time tasks as washing and waxing floors, polishing furniture, cleaning bathrooms, appliances and windows.

(iii) Home Maintenance and Repair (HMR). Services performed under this heading include small paint jobs, repairs to furniture, attendance to storm windows and similar tasks.

(iv) Meal Preparation (MP). Most commonly this service involves the daily preparation of 2 meals for the client. Both meals may be prepared during the same visit, one of them being stored for later use by the client.

(v) Laundry (L). Regular washing of clothes, bedsheets and linens as required.

(vi) Shopping and Errands (S). With this service the client expects food and other goods to be purchased and then delivered to his/her home.

(vii) Financial Management (FM). Attending to income and expenses is sometimes beyond the client's ability especially if mental impairment is apparent. For this reason, these services are offered and cover such tasks as making out tax forms, applications for benefits, and attending to other money related problems.

(viii) Transportation (TT). Included here are: transport of clients to physicians, to recreational and cultural activities, to stores, churches and friends.

(ix) Attendant (A). This service involves a person being able to accompany the client when being transported to a hospital or to a doctor's office.

(x) Yardwork (Y). Yardwork largely entails cleaning of yard, lawn maintenance and snow removal.

(xi) Non-Nursing Personal Care (NNP). Personal care involves assisting the client with dressing, toileting, transferring from bed to chair or to another room, bathing and other personal care tasks.

The above listed clients services are entered into the cost models as binary exogenous variables, a procedure described later in this chapter under the discussion on the economic model. They are treated differently, in Chapters 7 and 8 where they are considered to be outcomes which are somewhat equivalent to endogenous variables.

2. The Socio-Economic Status Variable and its Categories.

The socio-economic status variable is divided into 6 categories: age, sex, location, income, client contribution and relationship/domicile. They are described below.

a) AGE (A_1, A_2, A_3, A_4).

The ranges in age corresponding to the 4 part grouping are A_1: <60 years of age; A_2: 60 through 69; A_3: 70 through 79; and A_4: >80.

It is hypothesized that age is positively related to endogenous variables TCM and TH. The reason for this is that among the aged chronic conditions tend to be more numerous and are likely to interact with the level of disability resulting in higher costs and hours of service. We should find coefficients of categories in the higher age ranges to be relatively larger than those in the lower age ranges.

b) SEX (SF).

Traditionally men, compared with women, have done few of the domestic maintenance chores associated with running a household. On the basis of this it is hypothesized that all other things being equal, men will require more services than women. Since the binary nature of this category is such that for women a unit figure is indicated and for men a zero, a negative sign on this coefficient is hypothesized with respect to all 3 endogenous variables. Females are represented by 1, males by 0.

c) LOCATION (LOC).

Clients living in Standard Metropolitan Areas (SMSA's) are classified as urban dwellers. Others are classified as rural. The postulated hypothesis is that monthly and hourly costs of care will be higher for urban dwellers. This is expected because opportunity costs are higher for providers who have both more employment opportunities and higher wages in urban areas. An urban client is represented by unity, the rural client by 0.

d) INCOME (INC1, INC2, INC3, INC4, INC5).

Clients income largely derives from Supplemental Security Income (SSI), Retirement Disability Survivors Insurance (RDSI) or a combination of the two. The few clients receiving less than the minimum SSI amount are probably in the process of applying for it. Income is stratified into 5 binary categories: INC1, >$250.01 per month; INC2, $200.01 to $250.00; INC3, $150.01 to $200.00 per month, INC4, $100.01 to $150.00; and INC5, <$100.00.

e) CLIENT CONTRIBUTION (CLT1, CLT2, CLT3, CLT4).

When assessing a client's eligibility for chore services, client income is reviewed. The cost of the service required by the client is also determined. If the client's income or assets lie above the SSI standard, the full amount by which it exceeds the SSI standard is deducted from the cost of the services.

The chore service payments, in effect, make up the difference between that which a client can afford and the total cost of the services. At least 80% of clients have neither sufficient incomes nor sufficient assets to enable them to pay part of the cost of chore services required.

Client contribution is broken down into 4 binary categories: CLT1, >$50.01; CLT2, $25.01 to $50.00; CLT3, $0.01 to $25.00; and CLT4 = 0.

It is hypothesized that as client contribution increases, the cost of care increases. The basis for this hypothesis lies in the assumption that for low cost chore services the client with income in excess of the (SSI) standard is likely to pay for his/her own services and not apply for chore service subsidization.

f) RELATIONSHIP/DOMICILE (RL1, RL2, RL3, RL4).

This category refers to the nature of the relationship between the provider and the client as well as the living arrangements of the client vis-a-vis the provider. This category is subdivided into 4 sub-categories.

(i) The client's relative providing services in the home which both share (RL1).

This subcategory of the SES variables is characterized by a situation in which both the client and the provider live under the same roof. Care is provided to the client by a son, daughter, parent or some other relative. It is believed that relatives will work for lower incomes than non-relatives since part of their income will be psychic to the extent they feel an obligation, or like, to care for their relative. It is also believed they will put in longer hours of service than the non-relative in an equivalent situation. For these reasons we hypothesize that the subcategory will be negatively related to TCM and CH but positively related to THM.

(ii) The client's relative providing services in the client's home but living elsewhere (RL3).

The relatives providing care for clients in other households are presumed to require greater benefits to compensate for the greater opportunity costs incurred in leaving their own home and providing services in that of the client. With this in mind it is hypothesized the coefficient of this subcategory will be relatively larger than that of RL1 for TCM and CH but smaller for THM.

(iii) Non-relative providing services in the home where both the provider and the client are domiciled (RL3).

Since the non-relative is assumed to enjoy less psychic income from the provision of care for a non-relative than would be the case with care of a relative, the hypothesis with respect to this subcategory would lead us to expect the coefficient of RL3 to be larger in a positive direction than that of RL1 for TCM and CH but smaller for THM.

(iv) A non-relative providing services in a household of a client who lives separately from the provider (RL4).

Based on theorizing with respect to RL1, RL2 and RL3 above, it is hypothesized here that the coefficient of RL4 will be greater in a positive sense than all other coefficients of the RL subcategories for TCM and CH but small for THM.

g) OMITTED CATEGORIES OF THE SERVICE VARIABLE.

Variable categories omitted from the equations in which categories of the service variable appear were GUIDE DOG, INTERPRETER and OTHER. These were left out because they were supplied to only a small number of clients.

3. The Medical Status Variable.

Medical Status was included in the data as transcribed from the form DSS-3492 and subdivided into 9 diagnostic conditions: Mental retardation (HS1), mental illness (HS2), heart-related problems such as cardiac condition, coronary problems, hypertension and strokes (HS3), respiratory and related diseases such as paralysis, arthritis and paraplegic conditions (HS5), diabetes (HS6), cancer (HS7), recuperation disabilities resulting from a recent operation, accident or bone fracture (HS8), and other (HS9).

Medical status was also categorized according to whether the medical condition was the first (or primary) health problem, second health problem, third or fourth health related problem. Since the response rate on the client encounter data sheet (DSS-3492) was incomplete, 20% of the forms ommiting any reference to health related problems, only the first health related problem was employed in this analysis.

In the Greenberg study (ANDERSON ET AL 1974) medical status was found to have no significant influence on costs of services provided in the home even where costs of medical professionals were included. Medical status is included in this research to confirm or question those results. Costs of medical professionals have not been included in this data so comparisons will relate only to the chore services.

The hypothesis with respect to the medical status variable is that its inclusion or exclusion from the equations being estimated will make no significant difference to the variation in any of the 3 endogenous variables.

C. Economic Model.

Each of the mathematical models posited at the beginning of this chapter has so far been characterized by 3 exogenous variables, functional status (FS), socio-economic status (SES), and medical status (HS). Model A is characterized by the endogenous variable total cost of chore services per month (TCM); Model B by total hours of service per month (THM); and Model C by cost per hour (CH).

In economic terms the models embody behavioral relationships between the endogenous variables and the exogenous variables. Each exogenous variable is subclassified according to specified categories which are in turn subdivided into between 2 and 5 subcategories. These were all listed and described above.

Within the relevant range in each model, the relationship between each category of the 3 exogenous variables and the endogenous variable is assumed to be linear.

The structural equation established in each model has a set of structural parameters:

β_1 to β_ℓ for categories of functional status

γ_1 to γ_m for categories of socio-economic status

δ_1 to δ_n for categories of medical status

The parameters are identically designated for each model. Thus in Model A the full equation would appear as follows:

$$TCM = \beta FS' + \gamma SES' + \delta HS' + U' \qquad \text{EQUATION (5.1)}$$

where β is a vector of functional status parameters

γ is a vector of socio-economic status parameters

δ is a vector of medical status parameters

FS is a vector of functional status categories

SES is a vector of socio-economic categories

HS is a vector of medical status categories

U is an error term.

In each of the models, Functional Status has 3 alternative forms. Taking as our example the Functional Status portion of Model A, the alternative Functional Status categories would appear as designated below:

1. Functional Status Specification in "cardinal" form.

$$TCM = \beta_{MOB}\ MOB + \beta_{DEX}\ DEX + \beta_{SEN}\ SEN + \beta_{COMP}\ COMP + \beta_{HOM}\ HOM \qquad \text{EQUATION (5.2)}$$

where MOB = mobility on rising impairment scale from 1 to 5

DEX = dexterity 1 to 5

SEN = sensory perception 1 to 5

COMP = comprehension 1 to 5

HOM = home management 1 to 5

2. Functional Status Specification in Binary Form

$$TCM = \beta_{M1}M_1 + \beta_{M2}M_2 + \dots + \beta_{M5}M_5 + \beta_{D1}D_1 + \dots + \beta_{D5}D_5 + \beta_{S1}S_1 + \dots + \beta_{S5}S_5 + \beta_{C1}C_1 + \dots + \beta_{C5}C_5 + \beta_{K1}K_1 + \dots + \beta_{K5}K_5 \qquad \text{EQUATION (5.3)}$$

where (all are specified as 1 if present, 0 if not):

M1 = mobility with no impairment

M2 = is between M1 and M3

M3 = mobility with some impairment

M4 = is between M3 and M5

M5 = much impairment in mobility.

D_1 = dexterity with no impairment

D_2 = is between D1 and D3

D_3 = dexterity with some impairment

D_4 = is between D3 and D5

D_5 = much impairment in dexterity

S_1 = sensory with no impairment

S_2 = is between S1 and S3

S_3 = sensory with some impairment

S_4 = is between S3 and S5

S_5 = much impairment in sensory

C_1 = comprehension with no impairment

C_2 = is between C1 and C3

C_3 = comprehension with some impairment

C_4 = is between C3 and C5

C_5 = much impairment in comprehension

K_1 = home management with no impairment

K_2 = is between K1 and K3

K_3 = home management with some impairment

K_4 = is between K3 and K5

K_5 = much impairment in home management

3. Functional Status Specified in the Form of Services Required:

$$\begin{aligned} TCM = {} & \beta_{LK}LK + \beta_{HK}HK + \beta_{HMR}HMR + \\ & \beta_{MP}MP + \beta_{L}L + \beta_{S}S + \\ & \beta_{FM}FM + \beta_{TT}TT + \beta_{A}A + \\ & \beta_{Y}Y + \beta_{NNP}NNP \end{aligned} \qquad \text{EQUATION (5.4)}$$

where (all are specified as 1 if utilized and 0 if not):

LK = light cleaning services

HK = heavy cleaning services

HRM = home maintenance and repair

MP = meal preparation

L = location

S = shopping and/or errands

FM = financial management

TT = transportation

A = attendant

Y = yardwork

NNP = non-nursing personal services

The full specification of the equation of Model A appears below with the provision that any of the alternative forms of Functional Status could be signified by the term βFS in the following equation.

$$\begin{aligned} TCM = {} & \beta FS' + \gamma_{A1}A_1 + \gamma_{A2}A2 + \\ & \gamma_{A3}A3 + \gamma_{A4}A4 + \gamma_{SF}SF + \\ & \gamma_L LOC + \gamma_{INC1}INC1 + \\ & \gamma_{INC2}INC2 + \gamma_{INC3}INC3 + \\ & \gamma_{INC4}INC4 + \gamma_{INC5}INC5 + \\ & \gamma_{CLT1}CLT1 + \gamma_{CLT2}CLT2 + \\ & \gamma_{CLT3}CLT3 + \gamma_{CLT4}CLT4 + \\ & \gamma_{RL1}RL1 + \gamma_{RL2}RL2 + \gamma_{RL3}RL3 + \\ & \gamma_{RL4}RL4 + \delta_1 HS1 + \delta_2 HS2 + \\ & \delta_3 HS_3 + \delta_4 HS_4 + \delta_5 HS5 + \\ & \delta_6 HS6 + \delta_7 HS7 + \delta_8 HS8 + \delta_9 HS9 \end{aligned} \qquad \text{EQUATION (5.5)}$$

where (all are specified as if 1 indicates present, 0 indicates not present):

FS = vector of functional status categories

A_1 = age, <60 years old

A_2 = age, 60-69 years old

A_3 = age, 70-79 years old

A_4 = age, >80 years old

SF = sex, when SF = 1 client is female

LOC = location, when LOC = 1 location is urban

INC1 = income >$250.01 per month

INC2 = income $201.00 to $250.00 per month

INC3 = income $150.01 to $200.00 per month

INC4 = income $100.01 to $150.00 per month

INC5 = income <$100.00 per month

CLT1 = clients contribution >$50.01 per month

CLT2 = clients contribution $25.01 to $50.00 per month

CLT3 = clients contribution $0.01 to $25.00 per month

CLT4 = clients contribution of zero dollars per month

RL1 = client is related to the provider and they are living together

RL2 = client is related to the provider and they are living apart

RL3 = client is not related to the provider and living together

RL4 = client is not related to the provider and living apart

HS1 = mental retardation

HS2 = mental illness

HS3 = heart related problems

HS4 = respiratory related problems

HS5 = mobility related problems

HS6 = diabetes related problems

HS7 = cancer related problems

HS8 = recuperation related problems

HS9 = other problems

One result we seek is the proportion of the inter-client variation in each of the 3 endogenous variables which can be explained by each of the 3 exogenous variables. It should be realized that among the 3 exogenous variables, functional status is a special case since there are several alternative forms and combinations of those forms by which it may be entered into the equations. With the proviso that multicollinearity will be taken into consideration, we can by dropping each variable in turn, but not simultaneously, from the full equation, determine the amount of inter-client variation in the endogenous variables associated with the variable in question. In so doing we should be able to determine the amount of inter-client variation attributable to each of the 3 exogenous variables in each of the 3 models.

In the case of the socio-economic status variable we can drop simultaneously all levels of each category from each of the 3 equations one category at a time to determine which categories account for significant changes in inter-client variation in the 3 endogenous variables.

Finally, we can observe the signs and significance of coefficients at each category level and determine within categories, the relationship between one level of a category and another level. Such levels within categories are sometimes cardinal, occasionally ordinal, but usually qualitative. Problems of interpretation of coefficients is anticipated in the event that the coefficients are unstable, that is exhibiting both positive and negative signs or manifesting large standard errors. Such problems may be exacerbated by evidence of multi-collinearity.

The client attributes included in the models are hypothesized to be important categories in explaining the variation in cost or hours worked by the provider. Client attributes such as those associated with socio-economic status (e.g., location, age, etc., as distinct from the variable SES) are also included to correct (adjust) for differences among clients who have the same functional status and medical status but different costs due to the SES attributes; for example, urban-rural cost differentials. We will examine the significance of these categories to determine which ones to include in future studies. Inclusion of significant attributes reduces the disturbance term "U" and increases the accuracy of our estimates of other coefficients. Also, if attributes are correlated with included variables, their exclusion leads to bias.

Though bias is incurred, we can also get some indication of category influence on variation in the endogenous variable by deleting it from the equation and observing the extent to which $\bar{R}^2$ changes.

D. Econometric Specifications and Assumptions.

Multiple regression will be used to analyze statistically the models using the standard ordinary least squares algorithm. The estimated coefficients are assumed to be unbiased, efficient and consistent if the disturbance term "U," appended to the end of each equation, meets the standard assumption that its distribution is normal and its mean, zero. There is no reason to doubt the assumption.

The likelihood that the other standard assumptions (KMENTA 1973) are met is also high. Homoskedasticity can be assumed because the state puts an upper limit on costs. The assumption of linearly independent exogenous categories can also be assumed because the study is cross-sectional.

In order to test the models, the equations were run in a form different from that specified in the Economic Model discussed in the previous section. As mentioned, several of the variable categories were qualitative. In these cases, dummy (0-1) variables had to be employed. Where this was the case, one of the subcategories had to be dropped to avoid the problem of attempting the inversion of a singular matrix. The effects of dropping these subcategories was picked up in a new constant term "α" which became appended to the front of the right hand side of the equation. Employing the second alternative form of the functional status variable, in Model A (EQUATION 5.3) the modified equation appears below.

$$
\begin{aligned}
TCM = \alpha &+ \beta_{M2}M_2 + \ldots + \beta_{M5}M_5 + \\
&\beta_{D2}D_2 + \ldots + \beta_{D5}D_5 + \\
&\beta_{S2}S_2 + \ldots + \beta_{S5}S_5 + \beta_{C2}C_2 + \\
&\ldots + \beta_{C5}C_5 + \beta_{K2}HOM_2 + \ldots + \\
&\beta_{K5}HOM + \gamma_{A2}A_2 + \ldots + \qquad \text{EQUATION (5.6)} \\
&\gamma_{A4}A_4 + \gamma_{SF}SF + \gamma_{L}LOC + \gamma_{INC2}INC2 \\
&+ \ldots + \gamma_{INC5}INC5 + \gamma_{CLT2}CLT2 + \\
&\ldots + \gamma_{CLT4}CLT4 + \gamma_{RL2}RL2 + \\
&\ldots + \gamma_{RL4}RL4 + \delta_1 HS1 + \delta_2 HS2 + \\
&\delta_4 HS4 + \ldots + \delta_9 HS9 + U
\end{aligned}
$$

It can be seen that in each category, the first subcategory was dropped. The exception to this was in the medical status variable where the third category (heart problems) was dropped, this being the most mentioned physical manifestation of poor health.

As a check on the accuracy of the estimated equations, approximately one third of the 628 cases were randomly assigned to a "Test set." The models will be estimated and modified using

428 "training" cases and the final modified models will then be re-estimated using the 200 "test" cases. When comparing the results from the 2 sets of data we will be looking for consistency in: (i) the magnitudes and signs of coefficients; (ii) the significance of F-tests associated with each coefficient; and (iii) the magnitude of $\overline{R}^2$. The larger sample of data is hereafter referred to as the 'training set' and and the smaller as the 'test set.'

III. Conclusion

Since the description of variables is more readily understood in the context of a common form of analysis, such as regression analysis, the emphasis in the chapter was placed on the variables as they were entered into the cost analyses. A description of the economic models and the econometric specification of those models was presented in this chapter. The empirical analysis of data based on the models presented in this chapter is to be found in Chapter 8.

Chapters 6 and 7 draw on this chapter in so far as it describes the variables to be employed in analyzing the relationships between client attributes and services. Chapter 6, the following chapter, describes the theory associated with the entropy minimax algorithm and presents a model for estimating the relationships between attributes and services.

CHAPTER 6

THEORY UNDERLYING THE ENTROPY MINIMAX PATTERN DETECTION ALGORITHM AND ITS UTILITY IN DETERMINING THE RELATIONSHIPS BETWEEN CLIENT ATTRIBUTES AND SERVICES

I. Introduction

This chapter addresses itself to describing a method of discovering relationships between a client's profile of attributes and the services received by the client. If such relationships are detected and found to hold up under further testing, on different data, then it should be possible to determine the types of service received and estimate their costs on the basis of a client's profile. Since such relationships have not been statistically demonstrated and since there is no established theoretical basis on which to posit testable hypotheses, a method was sought that would detect such relationships.

II. Choice of Methodology to Detect Relationship Between Sets of Multiple Outcomes and Sets of Characteristic Attributes

The problem at hand is to find an algorithm which will enable us to estimate relationships between multiple outcomes and multiple attributes. For example, given a client with a known attribute profile, is it possible to determine the likelihood or probability of him/her receiving specific services? Since there is usually more than one service assigned to a client, the algorithm should determine the probability of one or more services assigned to a client, given that client's attribute profile.

Two regression techniques, 'Probit' and 'Logit' analysis offered possible avenues to approach the problem since, in both, the dependent variable can be dichotomous (THEIL 1971). In both, probability of an event or outcome is related to a set of attributes or independent variables.

In the case of probit analysis, the probability of an event P is related to the independent variables X as follows:

$$P = 1 - F\ (\alpha + \beta X)$$

subject to the constraint $0 < P < 1$.

There are several major drawbacks to using this approach:

1) The normal integral $F\ (\alpha + \beta X)$ implies a normal distribution. There are no grounds on which to maintain this hypothesis.
2) In the above form, only one outcome can be accommodated.
3) Though other forms may be developed that accommodate more than one outcome, the computational difficulties are enormous since it becomes difficult to specify a maximum likelihood estimator.

4) Given the number of outcomes and attributes involved in this analysis, the number of equations to be estimated would be unrealistically large since there is no theory to relate certain outcomes to given attributes.

In logit analysis the functional form is:

$$P = \frac{e^{\alpha + \beta X}}{1 + e^{\alpha + \beta X}}$$

which transforms to:

$$\ln \frac{P}{1 - P} = \alpha + \beta X$$

subject to the constraint $P > 0$.

The objections to employing this form are similar to those detailed under probit analysis. Both approaches assume the data conform to given specific distributions which may or may not be the case. We desire an algorithm that has no such distributional assumptions. One exists in the literature on information theory. It is called the entropy minimax algorithm and is used as a hypothesis discovery procedure. Once relationships between outcomes and attributes are discovered, using entropy minimax, they may be tested using the logit analysis, especially where single outcomes are involved. Such testing would seek to demonstrate that for a given outcome, coefficients of attributes were significant. To ensure the absence of contrived relationships, the testing of the discovered hypotheses (if any) is to be conducted on a second 'test' sample of data. Other clustering algorithms might have been chosen such as factor analysis. These were not adopted since, in a recent survey of tests on clustering algorithms, the entropy minimax algorithm was found to converge more quickly than the others (JOHNSTON 1976, GIFT 1978a). It had a smaller error rate and was not constrained by the assumtion of an apriori distribution.

III. Background on the Entropy Minimax Algorithm

Since little has been published on the entropy minimax pattern-detection algorithm, the description of the background and procedure employed here is somewhat more detailed than would be the case with a more commonly used form of analysis. Much of the following discussion is based on papers by Schonbein and Gift (SCHONBEIN 1978, GIFT 1978b). The discussion here diverges from theirs insofar as the application is different in both subject and interpretation procedures.

Entropy minimax is employed as an application of statistics in the field of communication. It was developed by Christensen and enables a researcher to simplify the structure of a data source and reduce uncertainty about the data (CHRISTENSEN 1972, 1973). Its major utility has been realized in seeking patterns or knowledge in seemingly random and complex data. To the extent that knowledge can be gleaned from apparently random data, the degree of randomness (or entropy as it is referred to in information theory) will be reduced. The specific amount by which the randomness or entropy is reduced can be measured. This measure is an indirect indicator of the quantity of knowledge extracted from the data after applying to it the entropy minimax pattern detection procedure.

Entropy, in the informational theoretic setting, is similar in concept to that of thermal energy in physics. In physics, entropy is a measure of the unavailability of a system's thermal energy for conversion into mechanical work. Entropy, as defined in information theory, is a measure of the unavailability of a system's random data for conversion into knowledge.

For any system of data, there are two components: (i) entropy or uncertainty, and (ii) knowledge. Seemingly random data would appear to yield no knowledge, that is, to consist entirely of entropy (uncertainty). After applying a pattern recognition algorithm to the seemingly random data, it may be possible to detect patterns or gain knowledge or, equivalently, eliminate some uncertainty. To the extent knowledge is gained, entropy is reduced. In this approach, knowledge *per se* cannot be measured; only the reduction in uncertainty can be quantified. The reduction in uncertainty is measurable and is called 'information.' A more general definition of information is, "the change in the random state of knowledge."

Uncertainty is at a maximum when complete randomness is evident, that is, when all events are equally likely.

Uncertainty is zero when an event is certain, either to occur or not occur.

For analytic convenience, it is helpful if measures of uncertainty are additive across the number of possible events in the data. A commonly used measured of entropy, and one meeting the additive criterion, was developed by Shannon (SHANNON 1948). It is represented by the expected or average amount of information one would receive upon observation of an event occurrence from the system of data. For example, in the case of chore services we might, by examining a client profile of attributes (an event occurrence), gain information about probable services appropriate for that client.

Data to which the entropy minimax procedure is applied is broken down into two categories; an attribute list or inputs, and an outcome or class membership list. Attributes are commonly measurements or observations on sample cases such as sex, age, income, functional status, etc. The outcome signifies the classification of a system of events; in this research the full range of the types of services received by clients such as light cleaning, laundry, meal preparation, etc. The data source is a sample from the population of clients.

Assumptions associated with the employment of the entropy minimax procedure in this research are the following:

A. Individual samples drawn from the larger population of events are independently and identically distributed.

B. Statistical regularity in the population with respect to client management decisions is evident and the conditional probabilities for various outcomes exist.

C. Clients with similar attribute profiles are assigned similar services signifying consistency in decision making. This assumption implies that in approving certain services for clients, the social worker making the assignments, makes his/her decision on the basis of the client's recorded functional, socio-economic and medical status. This is a key assumption since, should no patterns be detected, the implication could be that this assumption is violated. It might then be hypothesized that the social worker's assignment of services is based on other than the recorded functional, socio-economic and medical status of the client.

D. Attributes characterizing the client profile are significant in relation to outcomes and are important or controlled - to obviate affecting the results. Attribute vectors are assumed to contain a sampling of factors affecting the decision-maker's choice of outcome.

Unlike Assumption C above, this assumption implies there is a relationship between a client's attributes and the services assigned to him/her. An underlying assumption is that a social worker, if consistently employing the attribute profile of the client, would assign services on a systematic and statistically detectable basis. The absence of detectable patterns would signify a violation of this assumption. Such findings could flow from a consistent social worker seemingly assigning services on the basis of attribute characteristics but there being no systematic relationship between the two. Such a violation would suggest the specific attribute profile analyzed to be of no value in assigning services.

An inherent limitation of the procedure lies in the necessity for judgement and intuition in the specification of components of the attribute vector. A related problem is the identification of a client's characteristics and the accurate transformation of them into the attribute profile format necessary for manipulation in the entropy minimax procedure.

IV. The Entropy Minimax Procedure

This section of the chapter describes how the entropy minimax procedure operates conceptually. The data set of interest is classified according to N clients, M individual attributes, and K mutually exclusive outcomes. A schematic representation of the decision process is presented below:

$$\begin{array}{l} \text{Client Attributes } (\bar{A} = a_1, a_2 \ldots a_m) \rightarrow \\ \qquad \text{Decision Process} \rightarrow \text{Outcomes } (Q_1, Q_2 \ldots Q_K) \end{array}$$

Data processed by the entropy minimax procedure are generally in binary form; therefore, continuous variables are transformed into two or more parts, each part becoming, in effect, a newly specified attribute. The outcomes should be mutually exclusive but there are means of running the procedure when this restriction is not met.

The set of observable data would appear as illustrated in Table 6.1. We are interested in testing the hypothesis that a patterned relationship exists between attributes and outcomes. We are also interested in generating a hypothesis to the effect that a given set of attributes is probabilistically related to a given outcome or set of outcomes. This hypothesis could be then tested on different data using classical statistical methods such as regression analysis, in particular, logit analysis.

Since the entropy minimax procedure measures the extent of any statistical relationship between attributes and services, it would appear to be a useful tool for comparing one attribute profile format with another format. Each format would presumably have different categories or subcategories

TABLE 6.1.--Data format for individual client attributes and their outcome states

Client	ATTRIBUTE VECTOR					OUTCOME STATE				
	a_1	a_2	a_3	......	a_M	Q_1	Q_2	Q_3	......	Q_K
1	0	0	1	......	0	0	1	0	......	0
2	1	0	1	......	1	0	0	1	......	0
3	0	0	1	......	1	1	0	0	......	0
4	1	1	0	......	0	0	1	0	......	1

1 indicates the attribute or outcome is present; 0 it is not.

and perhaps different specifications of variables. The various profiles of activities of daily living could be compared with each other on the basis of how accurately they predict outcomes, i.e., services.

To simplify the explanation of the entropy minimax procedure we will take as an example, a sample of clients, each characterized by an attribute profile of three identical categories. A point in three dimensional space would thus encompass all three attribute categories. The purpose of employing the entropy minimax pattern detection procedure is to divide up the space or volume (a process known as partitioning, screening or analogical screening) into blocks such that each group or individual classification of outcomes is uniquely represented by one of the blocks. In terms of the chore service program, after such a screening program has been run, we should be in a position to take an additional client's attribute profile and predict: (i) in which block it should be placed, i.e., which services would be utilized by that client, and (ii) the probability of the client falling into that block, i.e., actually being assigned those services represented by that block.

For example, in Figure 6.1 we have a three dimensional attribute space: $\bar{A} = (a_1, a_2, a_3)$. Let us assume the dimension a_1 represents the sex of the client and equals "f" if the client is female, "M" if the client is male. Let the second dimension a_2 represent the relationship of the client to the provider, where a_2 equals "r" if the provider of

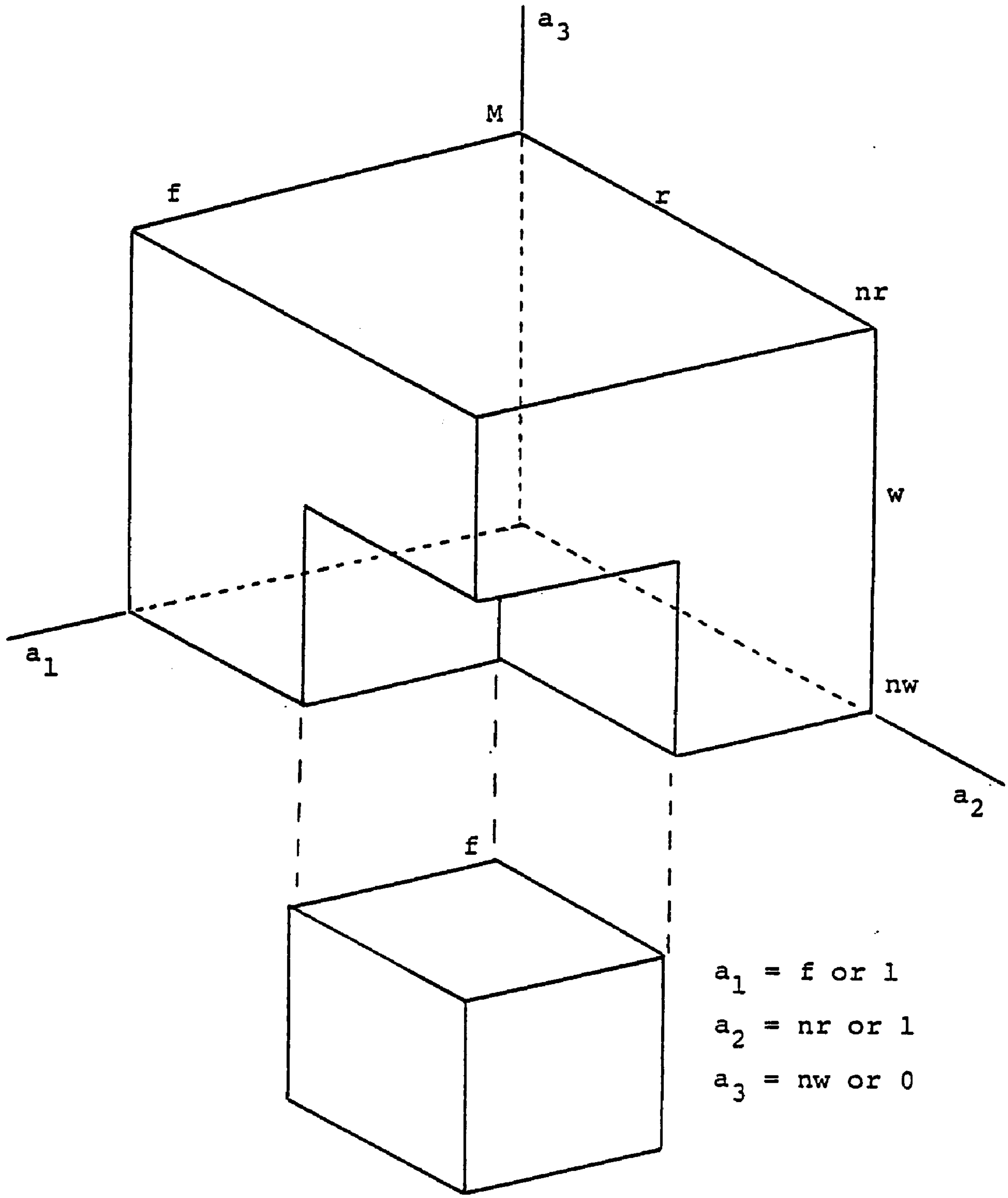


FIGURE 6.1.--Screening of a three dimensional space into two outcomes

services is a relative and equals "nr" if the provider is a non-relative. The third dimension a_3 represents mobility where a_3 equals "w" if the client is mobile and equals "nw" if the client is immobile.

There are potentially eight elementary cells for which there is only one attribute value for each cell. If we assume two outcome states, say meal service represented by the small detached block in Figure 6.1 and other services (OS) represented by the remaining seven-eights of the block, then the tabular representation of the results would appear as in Table 6.2. Let the probability of employing meal services be P(M). The conditional probability $P(Q_k/\bar{A}_n)$ represents the probability of employing meal services given that the profile $\bar{A}_n$ is observed.

In this case, cell number seven is partitioned from the rest of the attribute space and meal service is provided to clients falling into that cell according to a quantified conditional probability. Other services are provided to other clients according to a different conditional probability.

To take a more complicated example, a different screening has been performed on the three dimensional attribute space partitioned in Figure 6.2. This latter screening has partitioned the attribute space into four outcomes; Q_1, Q_2, Q_3, and Q_4. These are shown in Figure 6.2.

From Figure 6.2 we could draw the conclusion that within the class of clients for whom the value of $a_2 = 0$

TABLE 6.2.--Conditional probabilities of specific outcomes given characteristic attribute profiles

Attribute Vector				Outcome State	
				Q_K (observed) M	Q_K (not observed) Q
	a_1	a_2	a_3		
1)	0	0	0	$P[M/(a_1=0,a_2=0,a_3=0)]$	$P[Q/(a_1=0,a_2=0,a_3=0)]$
2)	0	0	1	$P[M/(a_1=0,a_2=0,a_3=1)]$	$P[Q/(a_1=0,a_2=0,a_3=1)]$
3)	0	1	0	$P[M/(a_1=0,a_2=1,a_3=0)]$	$P[Q/(a_1=0,a_2=1,a_3=0)]$
4)	0	1	1	$P[M/(a_1=0,a_2=1,a_3=1)]$	$P[Q/(a_1=0,a_2=1,a_3=1)]$
5)	1	0	0	$P[M/(a_1=1,a_2=0,a_3=0)]$	$P[Q/(a_1=1,a_a=0,a_3=0)]$
6)	1	0	1	$P[M/(a_1=1,a_2=0,a_3=1)]$	$P[Q/(a_1=1,a_2=0,a_3=1)]$
7)	1	1	0	$P[M/(a_1=1,a_2=1,a_3=0)]$	$P[Q/(a_1=1,a_2=1,a_3=0)]$
8)	1	1	1	$P[M/(a_1=1,a_2=1,a_3=1)]$	$P[Q/(a_1=1,a_2=1,a_3=1)]$

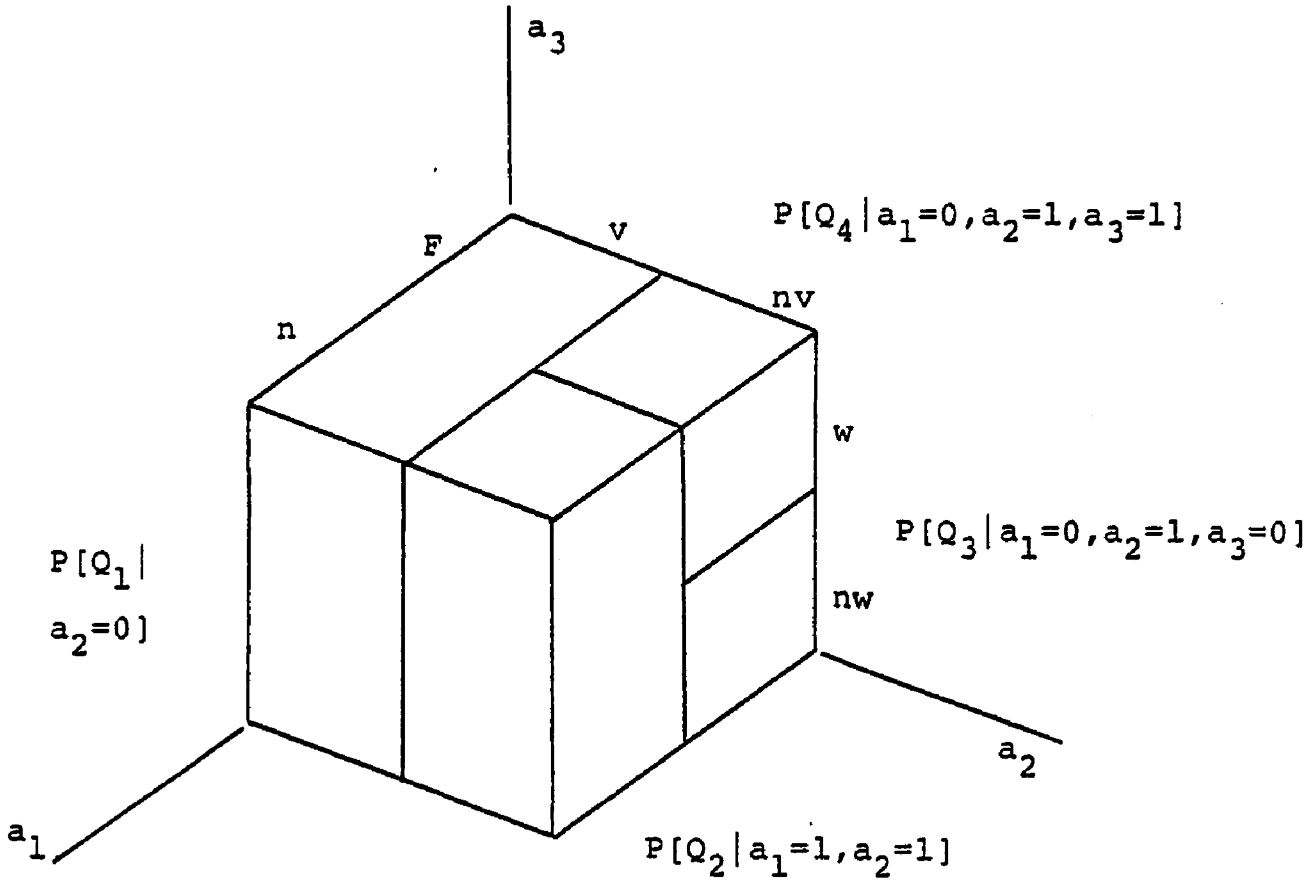


FIGURE 6.2.--Screening three dimensional space into four outcomes.

(those with relatives as providers), the values of the other two attributes, a_1 (sex) and a_2 (mobility) are immaterial in predicting the outcome Q_1. In the case of outcome Q_2, two attributes a_1 and a_2 are pertinent in the determination of that outcome; the mobility vector a_3 is immaterial. Outcomes Q_3 and Q_4 are both dependent on the level of all three attributes.

The entropy minimax procedure consists of a stepwise algorithm to partition the attribute space. The first step is to find the most populous sub-volume with the lowest entropy; in Figure 6.2 this volume would be represented by $P(Q_1/a_2=0)$. This sub-volume is then removed together with all the cases (clients) in it. The next step searches for the next most populous sub-volume, e.g., that associated with outcome Q_2 which turns out to have the lowest entropy for its size. These steps continue until the pre-set limit of steps is reached or until no more volume of attribute space remains to be searched. The number of partitions performed on the space can also be limited by a minimum degree of significance attributable to the last sub-space removed.

Each volume or space identified is characterized by: (i) the attributes within that space; (ii) the outcome associated with that space; (iii) the conditional probability relating a specific outcome to a specific vector of attributes ($\bar{A}_i$), i.e., $P(Q_k/A_i)$; (iv) the amount of entropy or uncertainty removed from the total entropy as a result of removing the sub-space volume. It is also possible to estimate to what degree the entropy removed is significant. Possible extremes that can be achieved in the partitioning or screening process are: (i) for the space to be divided into the number of partitions equal to the number of events (clients), in which case the amount of data in any one cell is so small that conclusions as to probabilities cannot be determined; (ii) to take all the data and treat it as one cell or one volume thus ignoring potential information associated with independent variables.

Probabilities are based on directly observable frequencies, there being many logically consistent sets of probabilities. Probability estimates assigned to conditional relationships between attributes and outcomes depend on the specific screening used to partition the sample data; that is they depend on the way in which the space is "chopped up."

The entropy of the total sample space, when that space is treated as an undivided cell, is a measurable entity constituting the maximum entropy associated with that space. When the entropy of partitioned space is lower than that of unpartitioned space the existence of a patterned structure within the seemingly random data is evident. The measure of the degree to which entropy is reduced by the partitioning is an indicator of the unquantifiable amount of knowledge gained by partitioning the space. Entropy is thus both a measure of the reduction in uncertainty resulting from the screening, and a useful criterion by which to measure the relative efficacy or "goodness" of each screening. The "best" partitioning is signified by that screening yielding the maximum amount of entropy removed. The entropy minimax algorithm selects a screening which maximizes the information extracted from the sample and minimizes the amount of "bias"* in the result. When classical statistical procedures are used to analyze seemingly random data, it is common practice to ignore the screening problem. That is, the data is divided into variable categories on theoretical grounds or intuitively before classical techniques are applied. Where no theoretical grouns exist for such divisions, the entropy minimax procedure would likely yield results superior to the intuitive approach. Furthermore, it is much less likely to result in spurious relationships.

V. The Mathematics of Entropy Minimax

Entropy, otherwise known as informational entropy or the expected value of information is commonly notated as $H(Q\,;\bar{A})$ and is defined by the following equation:

$$H(Q;\bar{A}) = -\sum_{k=1}\sum_{i=1} P(Q_k,\bar{A}_i)\ I(Q_k,\bar{A}_i) \qquad \text{EQUATION (6.1)}$$

---

*In the context of entropy minimax, bias refers to the possible selection of a screening in which the maximum entropy has not been determined or arrived at before the entropy of that screening is compared with the entropy of other screenings. Unbiasedness indicates that for any given screening the maximum entropy of that screening is determined and compared with the maximum entropies of other screenings. It is the minimum entropy among these maximum entropies that indicates the best screening...thus the term Entropy 'Minimax.'

where $I(Q_k, \bar{A}_i)$, commonly referred to as 'mutual information' represents the amount of information with respect to Q_k, and is defined as:

$$I(Q_k, \bar{A}_i) = \log_e \frac{P(Q_k | A_i)}{P(Q_k)} \qquad \text{EQUATION (6.2)}$$

Manipulation of Equation 6.1 (SCHONBEIN 1978) leads to the following equation:

$$H(Q;\bar{A}) = H(Q) - H(Q|\bar{A}) \qquad \text{EQUATION (6.3)}$$

where $H(Q)$ represents the self information of the outcome states or the maximum value of entropy in the data. $H(Q\ \bar{A})$ is the conditional entropy of uncertainty remaining after the attribute vector has been observed. We seek to minimize this measure of entropy. This measure might be considered to be the information-theoretical counterpart of $1-R^2$ in regression analysis.

The residual uncertainty $H(Q|\bar{A})$ is that entropy remaining after observing an attribute vector $\bar{A}$ when that vector fails to completely predict an outcome. Certain prediction of an outcome Q that is based on observation of attribute vector $\bar{A}$ would result in the entropy $H(Q|\bar{A})$ becoming equivalent to $H(Q)$ thus rendering total entropy $H(Q;\bar{A}) = 0$.

For any given partition the algorithm seeks to maximize the entropy; equivalently, this entails the minimization of $H(Q|\bar{A})$. After recording the entropy for the first partition or the first cell of the algorithm, the algorithm will continue scanning a variety of candidates for first cell until it encounters one of a lower maximum entropy which it records. The scanning procedure continues and terminates when it encounters no cell of lower maximum entropy than that previously recorded. The program records the probability of the outcome identified with that cell given the attributes associated with the cell. It also records the amount of entropy removed from the total entropy as a result of removing that cell. The data associated with that cell is then dropped from the complete set of original data and the search procedure starts again on the remaining data to identify a second cell. The second cell will usually be found to remove less entropy from the remaining data than did the first cell from the original data. After recording the relevant statistics associated with the second cell, the data embodied by that cell are discarded and the search procedure continues until a third cell is identified. The process can continue until some predetermined number of cells has been identified or until there is no entropy left in the data.

The entropy minimax procedure is different from classical decision making in that in the latter, the significant or relevant attributes are assumed to be known. In addition, in classical procedures, *a priori* information with respect to the distributions or the likelihood of observing various outputs as a function of certain inputs is also assumed. In the entropy minimax procedure, no *a priori* information, or theories as to which attributes are strongly related to which outcomes, is assumed. Nor are the distributions of outcomes assumed. The procedure merely discovers aggregations of attributes that are strongly associated with different outcomes of which there may be up to six or seven in number. The decision rules associating given attribute vectors with given outcomes is assumed to be implicit in the data and will be discovered. Such a discovery, or the detection of a pattern would lead to a hypothesis pertinent to the nature of that decision rule. The hypothesis can then be tested by classical methods.

Schonbein has shown that strategies based on the maximization of likelihood ratios such as minimizing the probability of error, also have the property of maximizing mutual information and visa versa (SCHONBEIN 1978). Since discriminant analysis techniques are essentially likelihood ratio-maximizing techniques, it is clear that a technique which is optimal in information theory must simultaneously be an optimal discriminant analysis technique.

Schonbein has also shown that outcomes may be predicted by fewer attributes than those characterizing the complete attribute vector ($\bar{A}$). Such a truncated attribute vector, let it be labelled $g(\bar{A})$, is known as a sufficient statistic in that it conveys all the information that would be conveyed by observation of the complete attribute vector. Any vector of attributes $g(\bar{A})$ conveying no less than the full attribute vector $G(\bar{A})$, is a sufficient statistic and could be any vector of Boolian logical functions, e.g.,

$$G(\bar{A}) = [g_1(X_2=1),\ g_2(X_1=0,X_3=1),\ g_3(X_1=0,X_2=0,X_3=1)] \qquad \text{EQUATION (6.4)}$$

Cells are defined in attribute space by each of these individual Boolian expressions for which there exists a unique set of outcome probabilities; each outcome probability being a function of only those attributes in the cell.

From Equation 6.3, the following expression for entropy can be derived:

$$H[Q;G_j(\bar{A})] = -\sum_{k=1} P[Q_k]\log_e P[Q_k] + \sum_{k=1} P[g_{ij}(\bar{A})] \sum_{i=1} P[Q_k;j_{ij}(\bar{A})] \log_e P[Q_k;g_{ij}(\bar{A})] \qquad \text{EQUATION (6.5)}$$

where G_j $(\bar{A})$ specifies the j^{th} set of Boolian expressions (g_{ij}) which are mutually exclusive and span attribute space. $P[g_{ij}(\bar{A})]$ is the probability of choosing at random an attribute vector that is identical to that vector of the attribute of the i^{th} cell or Boolian expression in the j^{th} screening.

For each partitioning, indeed before any partitioning, the self information of the outcome set $H(Q)$ is the same; thus, it can be ignored. The goal of the entropy minimax algorithm is to minimize $H[Q;G_j(\bar{A})]$ and thereby to determine the sufficient statistic with the minimum number of attributes.

To determine that the sufficient statistic, conditional entropy, is minimized:

$$H_j[Q|G_j(\bar{A})] = \sum_{i=1} P[g_{ij}(\bar{A})] \sum_{k=1} P[Q_k| g_{ij}(\bar{A})] \log P[Q_k|g_{ij}(\bar{A})] \quad \text{EQUATION (6.6)}$$

$$= H_j[Q;G_j(\bar{A})] - H[Q] \quad \text{EQUATION (6.7)}$$

Since a priori probability distribution of outcomes (Q_k) need not be known in order to achieve a solution, the entropy minimax procedure is superior to classical procedures in this analysis.

In order to minimize conditional entropy, it is necessary to choose an estimator P_{ij} for the probability $P[Q_k| g_{ij}(\bar{A})]$ where P_{ij} represents the probability that in the j^{th} partitioning, attribute vector i falls in the same cell as output k. It must do this according to the principle of maximum uncertainty which states, that information from one additional sample with an identical vector of attributes to that in the cell, would constitute the maximum amount of information necessary to make a decision as to the determination of the outcome associated with that cell.

From another perspective ($P^k{}_{ij}$) is that set of probabilities which, in the j^{th} screening, maximizes the remaining entropy after all available data has been observed.

Through a process known as the entropy of variation, an estimator is found which maximizes the entropy $H[Q|G_j(\bar{A})]$. This estimator is

$$\hat{P}^k{}_{ij} = \frac{n^k{}_{ij} + t}{n_{ij} + 1} \qquad \text{EQUATION (6.8)}$$

where: $n^k{}_{ij}$ is the number of observed outcomes of type k in the i^{th} cell of the j^{th} screening

n_{ij} is the number of outcomes of all types in the i^{th} cell of the j^{th} screening

t is the a priori probability of an outcome state k.

The probability $P[g_{ij}(\bar{n})]$ is

$$P_{ij} = \frac{n_{ij} + 1}{n + J} \qquad \text{EQUATION (6.9)}$$

where: J is the number of cells in the j^{th} screening

n is the total number of cases in the sample or the number of events.

It is known that $\hat{P}^k{}_{ij}$ converges to a maximum conditional entropy for each partition and then identifies the minimum entropy among all possible maximum entropies. The estimator also converges in probability and is, therefore, a consistent estimator in the statistical sense, thus assuring us that the estimator of conditional entropy will also converge to the minimum entropy partitioning.

We seek to minimize

$$\{H_j[Q|G_j(\bar{A})]\} \text{ by minimizing } \{\sum_i \hat{P}_{ij} \sum_k \hat{P}^k{}_{ij} \log \hat{P}^k{}_{ij}\}$$

In so doing, we will determine if a patterned relationship exists between outcomes (services) and vectors of client attributes. We will learn which attributes are associated with specific service outcomes. Finally, we will learn the extent of the knowledge gained by determining, for each successive cell in the chosen partition, the quantity of entropy removed relative to the total entropy found in the unpartitioned data. Knowledge of such relationships should be of value in analyzing cost of chore services.

VI. The Application of the Entropy Minimax Algorithm in Pattern Detection

The conceptual and theoretical utility of the entropy minimax pattern detection algorithm was explained in previous sections. This section addresses the manipulation of the chore service data into forms appropriate to the detection of patterns and the generation of hypotheses.

Basically, the aim of using the procedure is to attempt to detect systematic relationships between a client's profile (encompassing categories of functional status and socio-economic status) and service categories. Description and justification of the alternative specifications of outcome classification (service categories and their combinations) and client attributes are found in the following subsections of this chapter.

A. Selection of Outcome Classes.

The potential number of outcome classifications in this analysis totaled 14. Three of the service categories (guide dog, interpreter, and other) were dropped from the analysis because very few clients employed them. Ideally, each outcome class would be represented by each of the remaining 11 service categories. In practice, the accuracy and reliability of the entropy minimax computer program markedly decreases as the number of outcome classifications rises above five or six. Means were sought to aggregate the 11 service categories into between three and six outcome classifications. Five procedures were followed:

1. Those service categories were aggregated which seemed to occur together in the data.
2. Individual service categories were employed as proxies for other service categories.
3. The choice of outcome classification was based on service categories that did not appear to be associated with each other.
4. Choice here was based on those service categories occurring most frequently in the data.
5. Finally, the choice was similar to that mentioned in number 2, above, but a highly frequent service category was also added.

Possibly, other service category combinations could be justified and tested. The limitations on computer funds and the largely exploratory nature of this application of the entropy minimax algorithm suggested further research should proceed only after thorough evaluation of the present results.

The careful reader will have noted that the assumption concerning mutual exclusivity of outcomes has been violated. This problem was overcome by designating each outcome as an individual case. Thus, in the case where a client was receiving five services, the client was 'split' into what amounts to five clients with the same profile, each with a different outcome.*

1. Aggregation of Service Categories.

   It was recognized in the early stages of this research that certain outcome classes would probably be associated with each other and it was on this basis the first classification of outcomes was made.

   a) In the first outcome group, if clients could not perform heavy cleaning, it was unlikely they would do their own shopping or laundry. Thus, such services

---

*In a private communication, I was assured by Mr. Schonbein, of the Department of Radiology, College of Human Medicine, Michigan State University, that the pattern detection procedure would still function appropriately and detect patterns if any existed.

would probably be associated with each other. Table 6.3 compiled from the full data set of 628 cases, lays out the frequency of services employed by clients according to the number of services employed by clients according to the number of services they received. A visual scan of the table indicates that the first five services are associated with each other, especially among those clients receiving more than four services. A further confirmation of the common association of the first five most frequently utilized services (shopping, laundry, heavy cleaning, light cleaning and meal preparation) was observed in a frequency count in the training sample in which all five were observed among 234 clients out of a sample total of 428. Only six clients failed to receive any of these services. Together, laundry, heavy cleaning, meal preparation and shopping were received by 256 clients; laundry, heavy cleaning and shopping by 302 clients.

b) Possible associations between other classifications of outcomes were sought.

TABLE 6.3.--Frequency of services received by clients according to total number of services provided

Total No. of Services	TYPE OF SERVICE										
	Laundry (L)	Shopping (S)	Heavy Clng. (HK)	Light Clng. (LK)	Meals (M)	Non-Nursing Personal (NNP)	Trans. (TT)	Fin. Mgt. (FM)	Attendant (A)	House Maint.Rep. (HMR)	Yard Work (Y)
1	0	1	2	0	1	6	0	0	0	0	2
2	10	4	13	4	5	6	2	0	0	0	2
3	31	20	26	15	8	11	8	0	2	3	2
4	55	52	47	34	23	15	18	3	1	7	5
5	96	84	72	74	75	36	24	13	6	11	4
6	135	134	116	116	119	76	55	23	20	20	8
7	112	113	102	105	105	85	68	43	26	19	13
8	83	84	83	82	83	76	65	42	37	22	15
9	39	39	39	39	39	37	36	28	25	18	12
10	12	12	12	12	11	12	12	6	11	12	8
11	7	7	7	7	7	7	7	7	7	7	7
Col. Total	580	550	519	488	476	367	295	165	135	119	78
Percentage of Clients	92.4	87.6	82.6	77.7	75.8	58.4	47.0	26.3	21.5	18.9	12.4

There appeared to be an association between the use of financial management services and non-nursing personal services since together they were utilized by 88 clients.

c) The classification of the <u>third outcome</u> was made by associating home repair and maintenance services with yardwork, these being less frequently required than other services and being associated in 30 cases. Neither service was used by 314 clients. The classification of the (d) fourth and (e) fifth outcomes were attendant services and transportation, respectively.

2. The Use of Proxy Service Categories to Classify Outcomes.

Since the computer analysis associated with the entropy minimax computer runs was relatively complicated compared to the cost analysis which utilized regression techniques, the two procedures were developed and run simultaneously. This approach facilitated some interaction between the two procedures. The choice of outcomes or service category proxies was thus based on the regression of services on total monthly costs (TCM). These regressions, explained in detail in Chapter 8, yielded significant coefficients in the categories: light cleaning (LK), meal preparation (MP), attendant (A), and non-nursing personnal services (NNP). These four outcome categories were used as four individual outcome classification in the entropy minimax algorithm. The remaining five outcome categories were not designated as outcome classifications and were left out of the analysis. It was decided that if the four included categories resulted in the discovery of significant patterned relationships between client attribute profiles and outcome classifications, then further work would proceed on relating these service categories to other service categories.

3. Outcome Classifications Represented by Service Categories That Were Unassociated with Each Other

In this classification, service categories chosen were those that appeared to be clearly not associated with each other in the sense that they seldom occurred together for any one client. It was hypothesized that such a lack of association may lead to easier separation of data in pattern detection.

Three service categories were chosen: transportation, attendant, and non-nursing personal services. The remaining six service categories were left out of this computer run.

4. Outcome Classifications Represented by the Categories of the Four Most Frequently Utilized Services

According to Table 6.3, the four most frequently used service categories were laundry (92.4 percent of clients), shopping/errands (87.6 percent), heavy cleaning (82.6 percent) and light cleaning (77.7 percent), respectively. These were designated outcome classifications for this run. This does not imply that these services, relative to other services, occupied the most time of the provider. It was hypothesized here that by including only the most frequently used services in the computer run, large cells would be constructed by the pattern detection algorithm.

5. Outcome Classification Based on a Modification of the Proxy Approach Used in Method (2) of Outcome Classification

The only modification made to Approach (2), described above, is the addition of the shopping/errands service. This was done to see if the addition of service categories made any difference to the amount of entropy removed from the total data entropy.

B. Choice of Attributes

Attributes were limited to categories of SES and functional status. Medical status was not included because (as mentioned in the literature review) other studies found no relationship between medical status and use of non-medical chore-type services. Furthermore, results from the regression analysis of costs in this research showed no relationship between health services and costs of services (see Chapter 8).

Each attribute category was transformed into a binary variable. This approach is somewhat different from that used in the regression analysis where each category was sometimes divided into five subcategories, each being represented by a binary variable. An effort was made within each category to subdivide the category between those subcategories where the difference between one subcategory and an adjacent subcategory was at a maximum. In the case of the impairment indicators, mobility, dexterity, sensory, and comprehension, these judgements were based on the array of data in Table 6.4. Home management, which does not appear in the array, was divided on the same basis.

Considering the socio-economic variable first, age was divided between those over 80 and those under 80 because it was thought the older person would use more services due to circumstances purely related to age.

Location was divided between urban and rural since it was felt those clients in urban areas would be more likely to request services for any given functional status than those in rural areas. This was based on the proposition that with respect to transaction costs for clients, obtaining services in the city was reckoned to be less costly than for those clients in rural areas.

A client's living arrangement with respect to the provider of services was divided between those who lived with the provider and those living separately from the provider.

The client's relationship to the provider was divided between clients who were relatives

TABLE 6.4.--The number of clients at different impairment levels and the service categories employed by them

	TYPES AND LEVELS OF IMPAIRMENT (INCREASING FROM 1 TO 5)																			
Service Category	Mobility					Dexterity					Sensory					Comprehension				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Laundry	51	17	152	121	239	44	57	111	243	125	238	96	178	48	20	245	96	82	79	78
Shopping/E.	49	19	141	111	230	42	54	105	233	116	212	96	171	46	18	229	91	78	77	75
Heavy Cleaning	36	16	137	111	219	29	49	103	228	110	200	94	162	45	18	243	89	72	54	61
Light Cleaning	34	10	122	103	214	29	38	96	211	114	183	89	157	41	18	208	86	71	56	67
Meal Prep.	41	9	116	97	213	37	37	86	204	112	175	88	151	46	16	182	77	74	65	78
Non-Nursing Personal	33	6	67	69	193	25	27	60	146	109	142	66	104	40	15	110	52	63	70	72
Transport	31	11	71	58	124	28	28	47	130	62	108	45	100	33	9	102	48	45	51	49
Financial Mgt.	32	3	24	26	70	29	11	29	55	41	54	22	57	26	6	7	22	30	49	57
Attendant	16	4	26	24	65	15	10	24	49	37	51	20	46	11	7	45	19	23	20	28
Home Mgt. & Rpr.	2	1	35	24	57	1	12	19	55	32	39	23	37	13	7	55	22	18	9	15
Yardwork	4	1	14	19	40	1	4	14	38	21	20	25	22	7	4	37	14	16	3	8

of the provider and those who were not relatives.

The dichotomization of categories of the functional status variable was less clear cut than was the case with the SES variable. In the raw data, each functional status indicator had five levels of impairment. Frequency counts were run on each level of each category, see Table 6.4. Judgements were made as to where the cut-off between relatively non-impaired and the relatively impaired should be. Mobility was divided between the second and third levels of impairment. Dexterity between third and fourth; sensory between third and fourth; comprehension between first and second; home management was divided between the second and third levels.

The nine actual computer runs utilizing the entropy minimax algorithm are described in the next Chapter where the computer printouts are presented.

CHAPTER 7

RESULTS FROM APPLICATION OF THE ENTROPY MINIMAX ALGORITHM

I. Introduction

In the last chapter the theory underlying the entropy minimax principle was discussed. This chapter examines the empirical results. The first section lays out the specific attributes and outcomes as they were entered into each computer run. This is followed by a description of the *a priori* probabilities of attributes and outcome classes in the context of the computer printouts. Input and cell entropies are then analyzed after which the outcome probabilities are examined. Finally, the results are discussed in light of the methodology employed and the data base.

II. Attributes and Outcome Classes

Nine pattern detection runs were conducted on various formats of the data. These are listed in Table 7.1. The first computer run was conducted using aggregations of the outcome classes in clusters. In addition to the first run, 2 series of four runs were conducted; one series used a format of functional status attributes including mobility, dexterity, sensory and comprehension. The other series dropped mobility, dexterity and sensory attributes, replacing

TABLE 7.1.--Forms of outcomes (service mixes) and client attributes employed in Entropy Minimax Computer Runs

Run No.	1	2	3	4	5	6	7	8	9
	(LK,HK,MP, L,S) (FM, NNP) (HMR,Y) A, TT	LK MP A NNP	TT A NNP	LK HK L S	LK MP S A NNP	LK MP A NNP	TT A NNP	LK HK L S	LK MP S A NNP
Impairment Categories									
Mobility	P	P	P	P	P	O	O	O	O
Dexterity	P	P	P	P	P	O	O	O	O
Sensory	P	P	P	P	P	O	O	O	O
Comprehension	P	P	P	P	P	P	P	P	P
Home Mgt.	P	P	P	P	P	P	P	P	P
Socio-Economic Status									
Age	P	P	P	P	P	P	P	P	P
Relationship	P	P	P	P	P	P	P	P	P
Living Arr.	P	P	P	P	P	P	P	P	P
Location	P	P	P	P	P	P	P	P	P

Client Attribute --O indicates the attribute was not included in the run
P indicates the attribute was included in the run

them with the home management attribute. The socio-economic attributes, age, relationship, living arrangement, and location had the same specification throughout the 9 computer runs.

Before examining the results of the runs, the computer printouts will be explained. Four printouts have been reproduced in Figures 7.1, 7.2, 7.3 and 7.4. The important features of Figure 7.1 are the total number of events or clients in the sample (NTOT 428) and the number of outcome class members (NOUT). In an actual printout of Figure 7.1, taken from run number 2, the outcome class members are four in number: light cleaning (LK), meal preparation (MP), attendant (A), and non-nursing personal services (NNP). The key process control parameter in Figure 7.1 is LSTEP, indicating the maximum number of processing steps through which the algorithm cycles before terminating. In all of the computer runs, with the exception of the first, processing steps were limited to 10, each step constituting the creation of a cell. The cells were considered sufficient since any useable information usually appears in the first 5 or 6 cells created.

III. A Priori Probability of Attributes and Outcome Classes

The utilization of the equal A PRIORI WEIGHTING FACTORS (see Figure 7.1) and their magnitudes (see Figure 7.2) is based on the assumption of no known a priori reduction in uncertainty characteristic of the data. Use of

```
****** Entropy Minimax Pattern Detection.  January 1976 Version.

RUN TITLE.  WARD TRAINING SAMPLES --  run number 2.

       Date.  22 August 1978

Parameters From Tape 9.

       Nentry      428   Number of Row Entries in D Array.
       NTOT        428   Total Number of Events.
       NV1           8   Number of Binary Variables.
       NV2           0   Number of Scaler Variables.
       NCO           1   Number of Outcome Classes.
       NOUT          4   Number of Outcome Class Members.

Process Control Parameters.

       NORD          8   Processing Order.
       LTEST         3   Preprocessing Order.
       LSTEP        10   Number of Processing Steps.
       KCODE         0   Scaler Variable Code 1=A/D 2=Midpoint.
       PERR       1.00   Error Varience Limit.

Apriori Weighting Factors.
WV(1)  =           .25
WV(2)  =           .25
WV(3)  =           .25
WV(4)  =           .25
Total Weighting Factor WW = 1.00
MAX. ARRAY DIMINSIONS, FPVCTR(11734) INVCTR(879)

Apriori Probability Array.
APP( 1)              =     1.767
APP( 2)              =     1.690
APP( 3)              =      .233
APP( 4)              =      .837
APP( 5)              =      .472
APP( 6)              =      .984
APP( 7)              =      .864
APP( 8)              =     1.721
APP( 9)              =     1.685
APP(10)              =     1.649
APP(11)              =      .444
APP(12)              =     1.251
```

Figure 7.1. Part 1 of computer printout from run number 2.

STEP-WISE APPROXIMATION PATTERN DISCOVERY PROGRAM

BILL SCHONBEIN
DEPARTMENT OF RADIOLOGY
MICHIGAN STATE UNIVERSITY

INPUT VARIABLES

INDICATED NUMBER OF DATA SETS IN INPUT SAMPLE 387

COMPUTED TOTAL NUMBER OF DATA EVENTS IN INPUT SAMPLE 973

INDICATED NUMBER OF QUALITATIVE(BINARY) ATTRIBUTES 8

INDICATED NUMBER OF QUALITATIVE(SCALAR) ATTRIBUTES , 0

COMPUTED TOTAL NUMBER OF ATTRIBUTES 8

INDICATED PROCESSING ORDER (I.E. MAX NUMBER OF VARIABLES USED TO DEFINE A ATTRIBUTE) 8

INDICATED PREPROCESSING ORDER (FOR ATTRIBUTE SUBSET SELECT ION BEFORE EACH PROCESSING STEP) 3

INDICATED NUMBER OF OUTCOME CLASSES UNDER CONSIDERATION 4

INDICATED A PRIORI CLASS PROBABILITIES.

.250 .250 .250 .250

INDICATED NUMBER OF DATA EVENTS A CELL MUST CONTAIN IN ORDER FOR THE CLASS FREQUENCIES FOR THESE EVENTS TO CONTAIN THE SAME AMOUNT OF OUTCOME INFORMATION FOR THE CELL AS THE A PRIORI CLASS PROBABILITIES 1.

BINARY VARIABLES LIST.
VARIABLE N., 1. MOBILITY
VARIABLE N., 2. DEXTERITY
VARIABLE N., 3. SENSORY
VARIABLE N., 4. COMPREHENSION
VARIABLE N., 5. AGE (80 YEARS OR OLDER≡)
VARIABLE N., 6. RELATIONSHIP OF PROVIDER (RELATIVE≡)
VARIABLE N., 7. LIVING ARRANGEMENT (SHARE HOUSE≡)
VARIABLE N., 8. LOCATION (URBAN≡)

OUTCOME CLASS NAME. OUPUTS (BASED ON FIRST MINIMAX RUN)

Figure 7.2. Part 2 of computer printout for run number 2.

```
***** ISTEP =  1   IN=387   INR=428   IW=  973

OUTCOME TOTALS BY CLASS MEMBER. T( 1)=  326,
                                T( 2)=  319,
                                T( 3)=   86,
                                T( 4)=  242,
D ARRAY AFTER REDUN. JN=  8   JNR=  8

PREPROCESSOR RESULTS. LTEST= 3

RESULTS FROM ENTRPY. CELL DEFINITION ICELL=  15
                NO.OF OUTCOMES IN CELL XM=    2.0
                       ENTROPY OF CELL BB=  .836988E+00

 EVENTS SATISFYING ALL THE FOLLOWING

          BINARY VARIABLE  3 =,  1    SENSORY
          BINARY VARIABLE  2 =,  0    DEXTERITY
          BINARY VARIABLE  1 =,  0    MOBILITY
          BINARY VARIABLE  5 =,  0    AGE  (80 YEARS OR OLDER=)
               THE FOLLOWING ATTRIBUTE(S) ARE CONSTANT ON THIS CELL BUT WERE NOT USED IN THE CELL DEFINITION
          BINARY VARIABLE  8 =,  1    LOCATION  (URBAN=)
               THE FOLLOWING ATTRIBUTE(S) WERE SELECTED BY CHOOSE AS SIGNIFICANT BUT WERE NOT USED IN THE CELL DEFINITION.
          BINARY VARIABLE  4(COMPREHENSION                        ).NE.0 IN THIS CELL WITH PROB.= .459  APRIORI PROB.= .419
          BINARY VARIABLE  6(RELATIONSHIP OF PROVIDER (RELATIVE=)  ).NE.0 IN THIS CELL WITH PROB.= .496  APRIORI PROB.= .492
          BINARY VARIABLE  7(LIVING ARRANGEMENT (SHARE HOUSE=)     ).NE.0 IN THIS CELL WITH PROB.= .466  APRIORI PROB.= .432

HAVE ESTIMATED FREQUENCIES


    + .217
 .083       LK
      .019

    + .088
 .750       MP
      .265

    + .207
 .083       AT
      .019

    + .207  NNP
 .083
      .019


ESTIMATED PROBABILITY OF EVENT FALLING IN THIS CELL = .002
 RESULTS FOLLOWING DECODE, IEND=  0
```

Figure 7.3. Part 3 of computer printout for run number 2.

```
TERMINATED BECAUSE LIMIT OF 10 STEPS WAS REACHED.

STEP, 11

REMAINING EVENTS AGGREGATED INTO SINGLE CELL GIVE FOLLOWING RESULTS

ANY REMAINING EVENTS

HAVE ESTIMATED FREQUENCIES

   + .015
.332      LK
     .017

   + .015
.323      MP
     .017

   + .012
.092      AT
     .001

   + .015
.253      NPN
     .005

ESTIMATED PROBABILITY OF EVENT FALLING IN THIS CELL = .954
 RESULTS FOLLOWING DECODE, IEND =  5

 D ARRAY  K1=  429   K2=  13   IN=  360   JN=  8
```

Figure 7.4.—Part 4 of computer printout for run number 2.

other than equal weighting factors might have been specified if previous research had indicated such weights were justifiable.

The A PRIORI PROBABILITY ARRAY in Figure 7.2 lists the client attributes and outcome class members (or services) by code. Listed first are the client attributes coded APP (1) to APP (8). They are respectively: mobility, dexterity, sensory, comprehension, age, relationship, living arrangement and location. Lower down the list from APP (9) through APP (12), the service outcomes are listed. They are respectively: light cleaning (LK), meal preparation (MP), attendant services (A), and non-nursing personal services (NNP). The number of the right hand side of the coded list represent the _a priori_ probability of the attributes or outcomes being observed. The actual _a priori_ probabilities of the attributes or outcomes can be calculated by dividing the numbers associated with the list by two. For example, the probability of a client with a significant impairment in mobility appearing in the list is .05 x 1.767 = 0.888.

IV. Entropy

It is from the probability data in the A PRIORI PROBABILITY ARRAY that the total entropy of the input data set is calculated before partitioning is initiated. For example, in run number 2, Figure 7.1, the entropy of the input data (H_{input}) is calculated using only probabilities associated with the outcome classifications APP (9) to APP (12).

Calculation of the input data entropy H_{input} is conducted according to the following equation which was derived from Equations 6.1 and 6.2:

$$H_{input} = \sum_i P_{outcome_i} \log_e P_{outcome_i} \qquad \text{EQUATION (7.1)}$$

Restating the right hand side of the equation in terms drawn directly from the A PRIORI PROBABILITY ARRAY we get:

$$H_{input} = -\sum_i \frac{\frac{APP_i}{2}}{\sum_i \frac{APP_i}{2}} \log_e \left[\frac{\frac{APP_i}{2}}{\sum_i \frac{APP_i}{2}}\right] \qquad \text{EQUATION (7.2)}$$

$$= -\sum_i \frac{APP_i}{\sum_i APP_i} \log_e \left[\frac{APP_i}{\sum_i APP_i}\right] \qquad \text{EQUATION (7.3)}$$

$$= -\left[\frac{1.685}{5.029} \log_e \left(\frac{1.685}{5.029}\right) + \frac{1.649}{5.029} \log_e \left(\frac{1.649}{5.029}\right) + \ldots + \frac{1.25}{5.029} \log_e \left(\frac{1.25}{5.029}\right)\right]$$

$$= 1.292$$

which is the magnitude of the entropy in the input data.

The features of the computer printout illustrating an actual cell, (Cell No. 1 in run number 2) are illustrated in Figure 7.3. For our purposes, the pertinent characteristics are the following:

A. NUMBER OF OUTCOMES IN CELL XM = 2.0 indicating that only 2 events fall into the first cell. Indications of an emerging pattern would be suggested by a cell encompassing a much higher number of events.

B. ENTROPY OF CELL BB = .836988E+00. By itself, this figure signifies little. In combination with the next characteristic it is important.

C. ESTIMATED PROBABILITY OF THE EVENT FALLING IN THIS CELL = .002. Multiplying this figure by the ENTROPY OF THE CELL BB, we obtain a measure of the entropy remaining in the cell relative to that entropy in the total data. In the case of Cell Number 1, this weighted entropy is 0.836988 x .002 = .0017. This is a low figure implying little entropy is associated with this cell and little entropy has been removed from the total entropy of the input data.

After the data have been partitioned into the 10 cells, each of which embodies some entropy within itself, a final cell is constructed, Figure 7.4. It encompasses the data remaining after the data associated with the first 10 cells was removed. This last cell also embodies an entropy which is calculated in a similar manner to that of the input entropy H_{input}. Let the entropy in the last cell be designated $H_{last\ cell}$, then

$$H_{last\ cell} = -\sum_k \frac{P(Q_k)}{\sum_k P(Q_k)} \log_e \left(\frac{P(Q_k)}{\sum_k P(Q_k)}\right)$$

where Q_k = LK or MP or A or NNP

$\Sigma P(Q_k) = 1$

$$H_{\text{last cell}} = -[0.332 \log_e 0.332 + 0.323 \log_e 0.323 + 0.092 \log_e 0.092 + 0.253 \log_e 0.253] = 1.298$$

Multiplying this entropy by THE ESTIMATED PROBABILITY OF EVENT FALLING IN THIS CELL = 0.954 (see Figure 7.4), and adding this result to the weighted entropies remaining in the 10 partitioned cells, resulted in generating the total entropy remaining in the data. The entropy in each of the 10 cells and the probability of an event falling into those cells is recorded in Table 7.2. For example, the top left hand probability of .083 is the probability of the light cleaning service appearing in cell number 1. In run number 2 this entropy amounted to 1.285 (Table 7.3).

The entropy removed (H_{removed}) from the data as a result of the partitioning was calculated by subtracting the sum of the weighted entropies embodied in the cells (including the last cell) from the total input entropy (H_{input}).

$$H_{\text{removed}} = H_{\text{input}} - \sum_i H_i \qquad \text{EQUATION (7.4)}$$

where H_i = weighted entropy in cell i

$$H_{\text{removed}} = 1.292 - 1.285 = 0.007$$

This result indicated that using the entropy minimax algorithm on data characterized by the attributes: mobility, dexterity, sensory, comprehension, age, relationship, living arrangement and location; and also characterized

TABLE 7.2.--Array of outcome classifications, attributes and probability and other results from run number 2

	Cell Number										
	I	II	III	IV	V	VI	VII	VIII	IX	X	Last Cell*
(9) Light Cleaning	.083	.464	.375	.050	.313	.563	.529	.083	.417	.438	.332
(10) Meal Prep'n	.750	.464	.542	.450	.063	.313	.296	.417	.083	.438	.323
(11) Attendant	.083	.036	.042	.050	.063	.063	.029	.083	.417	.021	.092
(12) Non-Nursing Personnel	.083	.036	.042	.450	.563	.063	.146	.417	.083	.104	.253
No. In Cell	2	6	5	4	3	3	7	2	2	11	384
Prob. of Cell	.002	.006	.005	.004	.003	.003	.008	.002	.002	.011	.954

*Probability of event falling in last cell.

TABLE 7.2.--Continued

Entropy	.84	.95	.96	1.10	1.00	1.00	1.10	1.10	1.14	1.00	--
(1) Mobility	0	-	0	-	1	0	-	0	0	-	
(2) Dexterity	0	0	-	0	-	0	0	0	1	-	
(3) Sensory	1	-	0	0	-	0	-	-	1	1	
(4) Compre-hension	0	-	0	1	0	0	0	1	1	1	
(5) Age	0	-	-	0	0	-	0	0	0	0	
(6) Rel	0	-	1	1	0	0	1	0	0	-	
(7) Liv. Arr.	0	0	-	1	1	0	0	-	-	0	
(8) Location	0	0	-	0	-	1	1	1	1	-	

TABLE 7.3. Summary of results from nine computer runs of pattern detection algorithm.

	Run Number	1	2	3	4	5	6	7	8	9
Names of Variables in the Run		(LK,HK,MP,L,S) (FM,NPN) (HMR,Y) A TT	LK MP A NNP	TT A NNP	LK HK L S	K MP S A NNP	LK MP A NNP	TT A NNP	LK HK L S	LK MP S A NNP
	Number of Steps	20	10	10	10	10	10	10	10	10
Impairment Levels	Mobility	1	1	1	1	1	0	.0	0	0
	Dexterity	1	1	1	1	1	0	0	0	0
	Sensory	1	1	1	1	1	0	0	0	0
	Comprehension	1	1	1	1	1	1	1	1	1
	Home Mgt.	0	0	0	0	0	1	1	1	1
Socio-Economic Status	Age	1	1	1	1	1	1	1	1	1
	Relationship	1	1	1	1	1	1	1	1	1
	Living Arrt.	1	1	1	1	1	1	1	1	1
	Location	1	1	1	1	1	1	1	1	1
Probability of being member of last cell at end of run		.954	.954	.898	.975	.957	.932	.875	.923	.947
	Entropy of Data	1.418	1.292	1.021	1.384	1.525	1.292	1.021	1.384	1.525
	Entropy in Cells	1.338	1.285	.925	1.323	1.579	1.292	.805	1.377	1.518
	Entropy Removed	.080	.007	.095	.060	.006	.000	.116	.007	.007
Total % of Entropy Removed		5.600	0.500	9.400	4.400	0.400	0.00	11.400	0.500	0.500
	Entropy in Cells Other Than Last Cell	.080	.047	.076	.027	.055	.082	.101	.099	.068

by the outcome classifications: light cleaning, meal preparation, attendant, and non-nursing personal services; entropy removed was negligible, therefore, no relationships were detected on which to base and test hypotheses. In more concrete terms, the assignation of services LK, MP, A, and NNP to clients with the above profiles appeared to be almost completely random. Therefore, the client's need for light cleaning, meal preparation, attendant and non-nursing personal services could not be predicted on the basis of the client's attribute profile. Similar results were obtained using the other 4 combinations of outcome classification and the alternative attribute profile, (see Table 7.3).

The configuration of outcome classifications and attributes yielding the largest reduction in entropy was that of run number 7 where the entropy removed totaled 0.116 (Table 7.3).

Since entropy was not appreciably reduced (no patterns were discovered) it was judged to be of no value basing hypotheses on these results.

V. Probability of Outcome Classes

Had the entropy of the data been significantly reduced by the subtraction of a cell and the events it embodied, then we would have examined closely the printout labelled HAVE ESTIMATED FREQUENCIES (see Figure 7.3). For example, in CELL 1 (indicated by ISTEP = 1) we had the likelihood of Light Cleaning (LK), Meal Preparation (MP), Attendant (A), and Non-nursing Personal (NNP) services falling into the cell with the probabilities of 0.083, 0.750, 0.083 and 0.083, respectively. The upper and lower range figures to the right of each probability in the printout represent one positive and one negative standard deviation from the mean. Since some of these standard deviations are large relative to the magnitude of the mean probability of an event falling into the cell, there is still considerable uncertainty about the magnitude of the probability. This accounts for the large amount of entropy (0.836988) associated with the cell.

Another cause for uncertainty (but not high cell entropy) stemmed from the relative magnitudes of the cell probabilities. Taking an example from Table 7.2, the probabilities of outcomes falling into cell number 2 are displayed. Here the probability of a person with a given set of attributes receiving Meal Preparation (MP) services was 0.464, the same probability for receiving Light Cleaning (LK) services. Such a client was slightly less likely than not to have received the services LK and MP. Probabilities in the region of 0.50 are characterized by a high degree of uncertainty. If the probabilities had been close to zero or unity, there would have been more certainty about the relationship between the client's attributes and the outcomes associated with those attributes. For example, from Cell Number 2, it is reasonable to conclude that clients with attributes characterizing that cell, were highly unlikely to

TABLE 7.4.--Array of outcome classifications, attributes and results from run number 1.

	Cell Number									
	I	II	III	IV	V	VI	VII	VIII	IX	X
Service Categories in Outcome Classifications										
(9) Light Cleaning, Heavy Cleaning, Meal Preparation, Laundry, Shopping/Errands	.050	.733	.067	.408	.733	.440	.300	.480	.050	.300
(10) Financial Mgt. Non Nursing Personal	.050	.067	.067	.008	.067	.040	.050	.040	.300	.050
(11) Home Main. and Repair Yardwork	.050	.067	.067	.008	.067	.040	.050	.040	.050	.050
(12) Attendant	.050	.067	.067	.088	.067	.040	.050	.040	.050	.050
(13) Transportation	.800	.067	.733	.488	.067	.440	.550	.440	.550	.550
Number in Cell	3	2	2	24	2	4	3	4	3	3
Probability of Cell	.005	.003	.003	.037	.003	.006	.005	.006	.005	.005
Entropy	.777	.949	.949	1.000	.949	1.100	1.130	1.100	1.140	1.140
Attributes										
(1) Mobility	0	-	0	-	0	1	0	0	-	1
(2) Dexterity	0	-	-	0	-	0	-	-	0	1
(3) Sensory	-	-	-	-	-	0	0	-	0	-
(4) Comprehension Attribute	1	1	0	0	0	1	0	-	1	0
(5) Age	-	-	1	-	-	0	0	0	0	0
(6) Relationship	0	0	-	-	1	0	1	0	1	0
(7) Living Arrangement	0	0	-	0	1	0	-	0	1	1
(8) Location	-	0	1	-	0	1	-	-	0	0

H_{Input} = 1.418 Total Entropy of Data Set.

H_{Output} = 1.418 - .080 = 1.338

have received Attendant (A) or Non-nursing Personal (NNP) services because of their low probabilities, both of which were 0.036.

The outcome probabilities are the estimates of P_{ij}, namely $\hat{P}^k_{ij}$ which are calculated in the computer program but not laid out explicitly in the printout. The constituent parts of the equation used to estimate P_{ij} can be implicitly determined by observing some results. By examining Cell Number 1, in run number 1, Table 7.4, the probability of observing the composite outcome classification (LK, HK, MP, L, S) was 0.050; that is almost zero. Since there were only 3 client events in this cell, it was concluded that this outcome was highly unlikely to appear in the cell. The same result applied for 3 of the other 4 outcomes.

From Equation 6.8 we have the estimated probability $\hat{P}^k_{ij}$:

$$\hat{P}^k_{ij} = \frac{n^k_{ij} + t}{n_{ij} + 1}$$

where $n^k_{ij} = 0$, the number of observed outcomes in the i^{th} cell of the j^{th} partitioning; in this cell the k^{th} outcome type is characterized by the outcome set (LK, HK, MP, L, S).

$n_{ij} = 3$, the number of outcomes of all types in the i^{th} cell of the j^{th} partitioning.

$t = 0.2$, the a priori probability of outcome k appearing in the cell, there being 5 possible outcome states.

The estimated probability of the outcome set k, (i.e., LK, HK, MP, L, S) appearing in the i^{th} cell of the j^{th} partitioning is:

$$P_{ij} = \frac{0 + 0.2}{3 + 1} = 0.050$$

The same applies to the 3 of the 4 other outcomes.

The final outcome listed is Transportation (TT) with a probability of 0.800. From this, it can be concluded that all of the outcomes in the cell were Transportation services (TT). Thus $n^k_{ij} = 3$, the number of client events in the cell. Hence,

$$P_{ij} = \frac{3 + 0.2}{3 + 1} = 0.800$$

The algorithm searched out this cell first, considering it to be the cell removing the most entropy from the input data and a cell with minimum entropy for all cells screened. The algorithm then drops from the data set those client events in Cell Number 1 and goes on to search the remaining data in order to construct the next cell.

VI. Discussion

No pattern was found to systematically link client attributes with the range of services assigned to clients. No testable hypotheses were discovered. In the light of the results of the regression analysis discussed in Chapter 8, these results were not totally unexpected. In Chapter 8, it was shown that with few exceptions, client attributes, identical to those employed above, were poor predictors of monthly costs. It was, therefore, not surprising they would also be poor predictors of assigned services.

Explanations of the above results could emanate from several sources:

A. The specification of client attributes on the survey form DSS-3492 (Rev. 2-75) could be inappropriate with respect to guiding the social worker in assigning services. If this is the case, a different profile of attributes might yield more systematic relationships with assigned services. This explanation constitutes a violation of Assumption D above.

B. The attribute profile may be appropriate but not used by the social worker in the assignment of services, thus violating Assumption C above. Here, two forces may be operating:

   1. Social workers may utilize the client's attribute profile in assigning services but do so in different ways from each other resulting in the inconsistent matching of services to clients attributes. The training of the social workers may have been inconsistent.

   2. Social workers may choose to ignore much of the recorded client profile, preferring instead to assign services on the basis of

other criteria. Another hypothesis might be that they assign services on the basis of services required. Thus, instead of checking out each attribute and deducing from the observations the types of care required, the social worker may make direct observations about the type of appropriate service required. For example, the social worker may observe directly that the client cannot prepare his/her own meals in which case it may be appropriate to assign meal services. In such an event, the social worker may be content to record sufficient profile characteristics that seemingly justify the assignment of meal services. However, the justification of meal service may vary across clients supervised by one social worker or across social workers.

C. The specification of the service outcome mixes for purposes of entering into the algorithm, may have been inappropriate. The combinations of assigned services were selected on the basis of five types of criteria as described in Chapter 7. Other criteria and combinations might have produced more defined patterns.

D. The attribute characteristics transformed into binary categories for the purpose of entering them into the entropy minimum algorithm may have been inappropriate. The specific transformations used were based on judgements as to the nature of discontinuities within each category. The dichotomization of each category may have been inappropriate.

CHAPTER 8

RESULTS OF REGRESSION MODELS

I. Introduction

The regression models were specified in Chapter 5. Three basic models were estimated, Model A, Model B and Model C. Rather than discuss these models equation by equation we shall view the results in several steps as variables, variable categories, or subcategories, are simultaneously added or deleted from the models.

The results of the regression analysis are discussed at various levels of aggregation. The analysis is first focused on examining the effects of the inclusion or exclusion in the model of the 3 variables: functional status (in its 3 alternative forms), socio-economic status, and medical status. In this approach, the emphasis is placed on the changes in the adjusted coefficient of multiple determination $\bar{R}^2$. Here also the signs and stability of the coefficients are assessed with a view to making judgements as to the degree of multicollinearity present.

Second, the results are examined to determine the most significant categories within each variable. In the case of functional status, these results are examined in the same context as the analysis of $\bar{R}^2$ referred to above.

Third, the coefficients of sub-categories are examined to determine the relationship between various sub-categories.

Each of the endogenous variables, Total Cost per Month (TCM), Total Hours per Month (THM), and Cost per Hour (CH) will be examined within each of the above sections.

Finally, the results on the 'test data', as distinct from the 'training data', will be assessed. It may be recalled that 428 observations were selected randomly from a total of 628 cases and used to estimate the equations. The 200 remaining observations were saved to compare the "best" models from the 'training data' result with their equivalents in the 'test data.'

II. The Effects of Variables on the Variation in Total Cost per Month, Total Hours per Month and Cost per Hour

Of interest in this section, is the effect of the different forms of functional status on variation in the endogenous variables TCM, THM, and CH in Models A, B and C, respectively.

Due to the results of previously mentioned research (ANDERSON 1974), it was believed that the exogenous variable, Medical Status (HS), would have little or no effect on the variation in either costs or hours of service. For this reason and to reduce computer costs, Medical Status was not included in the early computer runs on the models. Later computer runs confirmed the efficacy of this approach.

The method of analysis used was to examine the component parts of the models with a view to learning which of the variables, functional status, socio-economic status or medical status, most fully explained the variation in each of the 3 endogenous variables. Then component categories of variables were deleted or added to determine those categories found to account for most of the variation resulting from the parent variable. These were then brought together into each of the full models.

This analysis is replete with considerable adjustments to equations in the sense that categories were dropped and added before the full model was estimated. There were several reasons for this approach. First, the early computer runs were some of the first for this analyst, and testing the model with a few variables left less room for error than if all the variables had been entered in the first few runs. Second, computer time was minimized using this approach. Third, the data had been divided up into two groups, a 'training set' and a 'test set.' Should the results from the 'training set' turn out to be contrived due to the excessive manipulation of the data, the 'test set' would probably fail to confirm the results based on the 'training set.'

For the sake of clarity and flow of exposition, the order in which the computer runs in each model are discussed is as listed in Tables 8.1, 8.2 and 8.3.

TABLE 8.1.--Results of Model A, total costs per month (TCM)

	Computer Run Number															
	AI	AII	AIII	AIV	AV	AVI	AVII	AVIII	AIX	AX	AXI	AXII	AXIII	AXIV	AXV	AXVI
R^2	.239	.08	.221	.273	.436	.149	.334	.308	.365	.495	.511	.536	.525	.512	.418	.492
$\bar{R}^2$	.229	.07	.217	.236	.420	.116	.299	.278	.304	.461	.474	.476	.481	.451	.412	.477
MOB	669 (.091)	1006 (.017)					627 (.102)									
DEX	400 (.360)	412 (.233)					652 (.127)									
SEN	131 (.686)	934 (.048)					374 (.259)									
COMP	1228 (.000)		1092 (.000)				974 (.001)	893 (.001)			277 (.307)			623 (.071)		352 (.119)
HOM	3492 (0)		4195 (.000)				3053 (.000)	3872 (0)			1450 (.001)			1620 (.001)		2595 (.000)
M2				-437 (.856)					139 (.953)			1615 (.432)				
M3				1020 (.611)					749 (.705)			96 (.955)				
M4				1179 (.579)					1041 (.618)			-57 (.977)				
M5				2174 (.209)					1779 (.384)			-67 (.970)				
D2				-2012 (.364)					-1218 (.582)			-180 (.927)				
D3				-2905 (.186)					-2192 (.310)			-1194 (.535)				
D4				-1400 (.566)					-262 (.906)			443 (.822)				
D5				-25 (.992)					1379 (.557)			1266 (.543)				
S2				-501 (.613)					-240 (.802)			-341 (.690)				
S3				233 (.796)					909 (.311)			404 (.615)				

TABLE 8.1 (continued)

	Computer Run Number															
	AI	AII	AIII	AIV	AV	AVI	AVII	AVIII	AIX	AX	AXI	AXII	AXIII	AXIV	AXV	AXVI
S4				1437 (.305)					1710 (.215)			707 (.560)				
S5				-351 (.876)					-385 (.863)			1034 (.596)				
C2				2997 (.004)					2594 (.011)			2020 (.026)	2123 (.015)			
C3				2655 (.013)					2428 (.024)			1608 (.099)	1744 (.061)			
C4				3265 (.005)					2562 (.033)			1282 (.260)	1451 (.179)			
C5				3904 (.003)					3611 (.007)			888 (.490)	952 (.435)			
K2				-4121 (.383)					-5387 (.238)			-5730 (.165)	-5200 (.191)			
K3				1736 (.691)					-711 (.866)			-1474 (.697)	-1380 (.707)			
K4				3464 (.410)					996 (.805)			-2059 (.579)	-1891 (.597)			
K5				8034 (.055)					5288 (.190)			182 (.960)	731 (.836)			
LK					-3238 (.000)					-2986 (.000)	-2161 (.014)	-2347 (.009)	-2153 (.014)	-1846 (.069)	-3226 (.000)	-2013 (.011)
HK					845 (.369)					1457 (.134)	1017 (.293)	762 (.451)	648 (.513)	378 (.730)		
HMR					-1560 (.058)					-1391 (.085)	-1189 (.137)	-1242 (.133)	-1390 (.082)	-1387 (.138)		
MP					-5298 (0)					-5403 (.000)	-4986 (.000)	-4732 (.000)	-4669 (.000)	-5183 (0)	-6044 (.000)	-5410 (.000)
L					-2344 (.051)					-1960 (.097)	-1665 (.154)	-1548 (.200)	-1783 (.129)	-1191 (.384)		
S					-412 (.678)					-434 (.656)	-403 (.677)	-117 (.906)	474 (.622)	-985 (.416)		
FM					-993 (.198					-947 (.238)	-518 (.554)	-406 (.647)	-384 (.661)	-240 (.806)		
TT					1128 (.080)					474 (.458)	485 (.443)	486 (.450)	460 (.465)	536 (.461)		

TABLE 8.1 (continued)

	Computer Run Number															
	AI	AII	AIII	AIV	AV	AVI	AVII	AVIII	AIX	AX	AXI	AXII	AXIII	AXIV	AXV	AXVI
A					-1836 (.021)					-1635 (.035)	-1508 (.050)	-1648 (.037)	-1649 (.032)	-1141 (.195)	-1917 (.013)	-1502 (.042)
Y					1008 (.290)					448 (.640)	499 (.598)	1019 (.313)	1153 (.238)	359 (.753)		
NNP					-5042 (0)					-5154 (.000)	-4661 (.000)	-4357 (.000)	-4548 (.000)	-4691 (0)	-5172 (.000)	-4592 (.000)
A2						-1685 (.134)	-1575 (.124)	-994 (.333)	-1411 (.175)	-1360 (.126)	-1154 (.192)	-1232 (.178)	-1043 (.238)	-535 (.508)		
A3						-128 (.907)	-873 (.391)	-47.2 (.962)	-862 (.403)	320 (.712)	328 (702)	-295 (.746)	164 (.849)	525 (.619)		
A4						512 (.625)	-1175 (.234)	-362 (.703)	-900 (.369)	-1016 (.223)	-1150 (.164)	-1515 (.086)	-1277 (.37)	-738 (.477)		-1107 (.112)
INC2						-3031 (.065)	-2709 (.065)	-2891 (.052)	-2374 (.112)	-2006 (.124)	-1997 (.121)	-1770 (.180)	-2045 (.110)	-3752 (.012)		
INC3						-2166 (.175	-2615 (.067)	-2631 (.069	-2319 (.118)	-2519 (.047)	-2591 (.039)	-2399 (.065)	-2735 (.028)	-3580 (.014)		-892 (.208)
INC4						-2426 (.187)	-3602 (.030)	-3697 (.028)	-3839 (.025)	-3348 (.023)	-3577 (.014)	-3684 (.015)	-3822 (.009)	-3937 (.018)		-1773 (.081)
INC5						-944 (.771)	-2775 (.343)	-2711 (.357)	-3772 (.207)	-2789 (.275)	-3153 (.212)	-4350 (.095)	-3760 (.137)	-3709 (.221)		
CLT2						1868 (.420)	-1644 (.427)	-1456 (.488)	-1512 (.476)	-620 (.738)	-720 (.695)	-533 (.778)	-490 (.789)	164 (.937)		
CLT3						-562 (.802)	-235 (.907)	-617 (.998)	262 (.897)	692 (.701)	673 (.705)	899 (.618)	750 (.673)	896 (.661)		
CLT4						-4636 (.009)	-3130 .048	-2891 (.072)	-2887 (.073)	-1124 (.436)	-1041 (.465)	-988 (.496)	898 (.528)	649 (.693)		-1305 (.115)
RL2						-4809 (.000)	-3611 (.002)	-3541 (.002)	-3309 (.004)	-1728 (.090)	-1684 (.095)	-1721 (.097)	-1709 (.091)	-830 (.481)		-1792 (.031)
RL3						241 (.880)	167 (.907)	334 (.818)	586 (.689)	165 (.897)	127 (.919)	10.9 (.993)	-172 (.891)	1566 (.284)		
RL4						-2994 (.002)	-1548 (.089)	-1516 (.100)	-994 (.289)	370 (.654)	334 (.689)	399 (.641)	448 (.589)	1088 (.248)		
LOC						4190 (.000)	3888 (.000)	3766 (.000)	3986 (.000)	3620 (.000)	3550 (.000)	3601 (.000)	3399 (.000)	3730 (.000)		3629 (.000)
SF						481 (.591)	250 (.754)	347 (.668)	366 (.655)	667 (.351)	578 (.412)	451 (.535)	554 (.434)	-144 (.857)		
HS1														-767 (.600)		
HS2														107 (.959)		

TABLE 8.1 (continued)

	Computer Run Number															
	AI	AII	AIII	AIV	AV	AVI	AVII	AVIII	AIX	AX	AXI	AXII	AXIII	AXIV	AXV	AXVI
HS4														-123 (.948)		
HS5														-89 (.922)		
HS6														992 (.471)		
HS7														4905 (.082)		
HS8														-1866 (.408)		
HS9														186 (.974)		

Omitted variables: A1, INC1, CLT1, RL1, M1, D1, S1, C1, K1, LOC (Rural), SF (Male), HS3 (Heart)

Parentheses indicate the level of significance of coefficients using an F-test (two-tailed).

TABLE 8.2.--Results of Model B, total hours per month (THM)

	Computer Run Number															
	BI	BIII	BIV	BV	BVI	BVII	BVIII	BIX	BX	BXI	BXII	BXIII	BXIV	BXV	BXVI	BXVII
R^2	.140	.133	.177	.185	.214	.274	.265	.303	.297	.312	.348	.321	.305	.163	.225	.303
$\bar{R}^2$	.129	.128	.134	.163	.184	.236	.232	.238	.249	.261	.264	.258	.218	.155	.201	.276
MOB	188 (.083)					222 (.031)										
DEX	-118 (.325)					-165 (.149										
SEN	- 64 (.472)					18 (.835)										
COMP	405 (0)	.377 (.000)				186 (.015)	177 (.016)			100 (.225)			81 (.460)		201 (.003)	103 (.152)
HOM	547 (.000)	604 (.000)				470 (.000)	524 (.000)			351 (.009)			410 (.01)		335 (.015)	352 (.007)
M2			-104 (.874)					146 (.816)			228 (.715)					
M3			591 (.282)					863 (.106)			882 (.096)					
M4			630 (.280)					949 (.092)			990 (.077)					
M5			773 (.178)					1022 (.064)			908 (.097)					
D2			-135 (.837)					355 (.552)			458 (.442)					
D3			-734 (.223					-654 (.261)			-619 (.292)					
D4			-709 (.253)					-552 (.359)			-496 (.409)					
D5			-450 (.491)					-407 (.519)			-424 (.504)					
S2			-113 (.677)					- 74 (.771)			-175 (.500)					
S3			-200 (.419)					- 26 (.912)			-211 (.390)					
S4			525 (.172)					511 (.169)			933 (.368)					
S5			-882 (.151)					-394 (.512)			-287 (.630)					
C2			308 (.276)					- 55 (.840)			- 75 (.785)					
C3			446 (.131)					34 (.906)			-110 (.710)	-87 (.760)				

TABLE 8.2 (continued)

	Computer Run Number															
	BI	BIII	BIV	BV	BVI	BVII	BVIII	BIX	BX	BXI	BXII	BXIII	BXIV	BXV	BXVI	BXVII
C4			1301 (.000)					688 (.033)			611 (.079)	562 (.091)				
C5			1456 (.000)					793 (.028)			428 (.275)	348 (.354)				
K2			-372 (.774)					-563 (.647)			-261 (.835)	17.2 (.989)				
K3			942 (.432)					737 (.516)			854 (.460)	907 (.423)				
K4			681 (.554)					681 (.532)			775 (.493)	731 (.507)				
K5			1589 (.165)					1319 (.225)			1208 (.283)	1274 (.202)				
LK				88 (.741)					123 (.629)	318 (.232)	266 (.330)	308 (.253)	509 (.117)	179 (.454)	172 (.491)	252 (.289)
HK				356 (.222)					92 (.755)	- 24 (.934)	-126 (.681)	- 55 (.855)	- 97 (.780)			
HMR				- 27 (.914)					93 (.702)	142 (.558)	232 (.357)	104 (.670)	210 (.481)			
MP				-944 (.000)					-910 (.000)	-799 (.002)	-828 (.002)	-811 (.002)	-851 (.009)	-1119 (.000)	-882 (.000)	-810 (.001)
L				-500 (.179)					-317 (.375)	-248 (.485)	- 56 (.878)	-250 (.488)	-367 (.401)			
S				21 (.945)					19 (.948)	13.9 (.962)	29 (.921)	6.5 (.983)	107 (.781)			
FM				-532 (.026)					-128 (.599)	21 (.937)	- 26 (.922)	40 (.880)	- 27 (.930)			
TT				174 (.381)					- 58 (.762)	- 53 (.780)	- 55 (.998)	- 40 (.836)	-141 (.544)			
A				-545 (.026)					-616 (.009)	-579 (.014)	-664 (.006)	-606 (.011)	-508 (.071)	- 653 (.006)	-579 (.014)	-577 (.010)
Y				- 76 (.797)					-294 (.313)	-278 (.335)	-388 (.208)	-213 (.478)	-186 (.609)			
NNP				-731 (.001)					-435 (.040)	-306 (.153)	-216 (.338)	-269 (.214)	-386 (.131)	- 909 (.000)	-518 (.016)	-297 (.150)
A2					50 (.857)	146 (.595)	171 (.528)	90 (.745)	104 (.698)	168 (.533)	115 (.678)	194 (.475)	369 (.245)			
A3					276 (.309)	271 (.321)	285 (.278)	228 (.412)	401 (.128)	399 (.126)	356 (.199)	424 (.111)	560 (.097)			
A4					378 (.146)	209 (.481)	246 (.331)	219 (.418)	261 (.301)	219 (.383)	222 (.409)	266 (.297)	447 (.178)		82 (.886)	76 (.718)

TABLE 8.2 (continued)

	Computer Run Number															
	BI	BIII	BIV	BV	BVI	BVII	BVIII	BIX	BX	BXI	BXII	BXIII	BXIV	BXV	BXVI	BXVII
INC2					48 (.906)	132 (.737)	61 (.875)	388 (.473)	92 (.816)	94 (.809)	284 (.480)	94 (.810)	50 (.915)			
INC3					156 (.692)	128 (.730)	82 (.829)	376 (.346)	1.67 (.997)	- 16 (.965)	269 (.497)	- 21 (.954)	- 84 (-855)		-157 (.486)	-139 (.526)
INC4					800 (.079)	668 (.136)	584 (-190)	882 (.056)	583 (.192)	516 (.245)	836 (.069)	514 (.249)	411 (.438)		551 (.088)	352 (.268)
INC5					389 (.628)	329 (.675)	125 (.872)	544 (.523)	67 (.930)	- 26 (.973)	349 (.660)	131 (.866)	216 (.823)			-222 (.748)
CLT2					126 (.827)	- 93 (.866)	- 51 (.926)	-114 (.842)	153 (.783)	135 (.809)	138 (.810)	89 (.874)	-159 (.810)			
CLT3					-620 (.266)	-568 (.294)	-522 (.334)	-483 (.377)	-239 (.662)	-240 (.658)	-167 (.761)	-230 (.674)	-460 (.480)			
CLT4					-334 (.442)	-108 (.798)	- 70 (.868)	-212 (.624)	211 (.630)	235 (.588)	129 (.770)	213 (.625)	- 21 (.968)		244 (.355)	237 (.348)
RL2					-1532 (.000)	-1350 (.000)	-1316 (.000)	-1286 (.000)	-1224 (.000)	-1201 (.000)	-1256 (.000)	-1128 (.000)	-1205 (.001)		-559 (.034)	-1109 (.000)
RL3					969 (.014)	1020 (.008)	1022 (.008)	1173 (.003)	1057 (.007)	1060 (.006)	1171 (.002)	1127 (-004)	804 (.083)			1066 (.005)
RL4					-1502 (0)	-1240 (.000)	-1233 (0)	-1165 (.000)	-1097 (0)	-1086 (.000)	-1090 (.000)	-1055 (.000)	-1109 (.000)			-1039 (.000)
LOC					204 (.427)	103 (.079)	134 (.590)	84 (.744)	159 (.536)	135 (.596)	37 (.889)	127 (.621)	22 (.943)		91.3 (.719)	118 (.625)
SF					- 80 (.715)	- 99 (.642)	-102 (.632)	-157 (.478)	-16.1 (.941)	- 41 (.847)	- 113 (.609)	- 48 (.825)	- 15 (.953)			
HS1													324 (.487)			
HS2													85 (.900)			
HS4													673 (.265)			
HS5													226 (.440)			
HS6													478 (.276)			
HS7													180 (.841)			
HS8													-583 (.418)			
HS9													145 (.699)			

TABLE 8.3.--Results of Model C, costs per hour (CH)

	Computer Run Number															
	CI	CIII	CIV	CV	CVI	CVII	CVIII	CIX	CX	CXI	CXII	CXIII	CXIV	CXV	CXVI	CXVII
R^2	.086	.080	.129	.099	.116	.160	.153	.202	.168	.188	.233	.207	.213	.073	.120	.167
$\bar{R}^2$	.075	.075	.083	.074	.082	.116	.116	.126	.111	.128	.135	.134	.116	.064	.093	.135
MOB	-6 (.196)					-7.3 (.117)										
DEX	4.6 (.372)					6.8 (.185)										
SEN	4.3 (.261)					2.2 (.583)										
COMP	-14.2 (.000)	-13.1 (.000)				8.7 (.011)	-8 (.015)			-6.9 (.067)			-6.6 (.165)		-9.1 (.005)	-5.8 (.081)
HOM	-16.6 (.006)	-17.9 (.001)				16 (.008)	-17 (.002)			-15 (.015)			-15.6 (.026)		-15.8 (.010)	-16.4 (.006)
M2			0.6 (.983)					-1.9 (.946)			-6.5 (.820)					
M3			-4.4 (.853)					-8.9 (.710)			-9 (.709)					
M4			-4.5 (.857)					-11.1 (.659)			-13.1 (.605)					
M5			-17.9 (.468)					-22 (.353)			-20.1 (.402)					
D2			-16.0 (.541)					-30.9 (.248)			-26.3 (.331)					
D3			9.3 (.721)					9.7 (.711)			14.1 (.598)					
D4			21.2 (.428)					18.2 (.501)			21.1 (.426)					
D5			-1.3 (.964)					-1 (.971)			7.1 (.806)					
S2			-4.3 (.789)					-3.6 (.758)			-1.8 (.881)					
S3			8.2 (.422)					5.1 (.637)			6.1 (.583)					
S4			-8.7 (.599)					-8.7 (.600)			-8.3 (.623)					
S5			15.6 (.555)					6.8 (.799)			7.7 (.775)					
C2			-5 (.968)					11 (.370)			10.5 (.402)	8.9 (.465)				

TABLE 8.3 (continued)

	Computer Run Number															
	CI	CIII	CIV	CV	CVI	CVII	CVIII	CIX	CX	CXI	CXII	CXIII	CXIV	CXV	CXVI	CXVII
C3			-4.9 (.700)					4.4 (.737)			9.4 (.486)	8.7 (.500)				
C4			-46.9 (.001)					32.4 (.026)			-30.3 (.055)	-30.5 (.043)				
C5			-48.6 (.002)					-30.2 (.061)			-23.3 (.192)	-24.8 (.146)				
K2			-52 (.351)					-53 (.328)			-74.2 (.194)	-80.5 (.846)				
K3			-81 (.117)					-87.9 (.065)			-110.1 (.037)	-105.7 (.040)				
K4			-77.2 (.120)					-89.6 (.067)			-111.7 (.030)	-99.5 (.047)				
K5			-102.6 (.038)					-107.6 (.028)			-125.1 (.014)	-117.8 (.017)				
LK				-7.8 (.509)					-9 (.441)	-16.9 (.164)	-13.7 (.271)	-15.8 (.195)	-22.2 (.114)	-16.5 (.117)	-17.5 (.117)	-20.1 (.067)
HK				-32.8 (.011)					-22.2 (.099)	-16.4 (.222)	-15.8 (.259)	-18.6 (.177)	-11.5 (.449)			
HMR				14 (.210)					10 (.366)	8 (.472)	5.4 (.635)	8.9 (.423)	6.3 (.624)			
MP				17.8 (.117)					18.8 (.102)	13.4 .249	17.6 (.138)	16.8 (.152)	11.1 (.428)	25.4 (.019)	15.4 (.161)	13.1 (.226)
L				26.8 (.102)					24.2 (.138)	21.6 (.184)	16.9 (.314)	19 (.222)	18.3 (.335)			
S				1 (.944)					- 1.9 (.890)	- 0.6 (.963)	- 3.5 (.801)	- 1.3 (.926)	- 6.4 (.705)			
FM				15.4 (.142)					3.5 (.751)	- 6.4 (.595)	- 3 (.807)	- 5.6 (.644)	- 1.6 (.906)			
TT				- 3.4 (.699)					- 9.8 (.266)	-10.3 (.241)	-11.6 (.194)	-11.8 (.194)	- 4.8 (.633)			
A				23 (.033)					27.2 (.011)	25.2 (.019)	25.1 (.021)	26 (.015)	18.9 (.121)	25.1 (.017)	22.3 .033	22.3 (.030)
Y				5.3 (.681)					9.2 (.490)	8.2 (.532)	14.8 (.290)	10 (.461)	12.5 (.430)			
NNP				21.7 (.023)					13.6 (.157)	7.4 (.448)	4.2 (.683)	6.4 (.516)	2.1 (.848)	28.5 (.002)	14.5 (.130)	7.2 (.440
A2					-2.2 (.859)	-8.0 (.517)	-7.2 (.551)	-5 (.690)	- 5.7 (.639)	- 9.5 (.438)	- 7.8 (.536)	-10.7 (.386)	- 8.7 (.528)			

TABLE 8.3 (continued)

	Computer Run Number															
	CI	CIII	CIV	CV	CVI	CVII	CVIII	CIX	CX	CXI	CXII	CXIII	CXIV	CXV	CXVI	CXVII
A3					- 7.8 (.516)	-10 (.398)	- 8 (.495)	- 8.9 (.478)	-13.8 (.250)	-13.5 (.255)	-14.9 (.237)	-15.8 (.189)	- 9.4 (.521)			
A4					1.2 (.916)	4.2 (.724)	6 (.596)	- 7 (.952)	2.7 (.817)	5.2 (.653)	- 0.3 (.983)	2.0 (.864)	- 3.7 (.795)		10.5 (.288)	9 (.351)
INC2					-18.7 (.298)	-21 (.233)	-19 (.283)	-26.2 (.147)	-20 (.261)	-20.5 (.252)	-27.2 (.137)	-21.3 (.232)	-17.1 (.407)			
INC3					-20.5 (.242)	-19 (.257)	-17.7 (.304)	-27.3 (.127)	-15 (.390)	-14.1 (.415)	-23 (.200)	-14.3 (.408)	-17.8 (.379)		- 2.5 (.803)	-2.9 (.771)
INC4					-33.4 (.099)	-26.7 (.183)	-24.5 (.220)	-33.1 (.109)	-23.7 (.244)	-20 (.321)	-29.1 (.163)	-21.4 (.290)	-21.8 (.344)		-14.6 (.311)	-7.7 (.598)
INC5					-16.6 (.642)	-15 (.671)	- 7 (.840)	-26.5 (.462)	-10.5 (.765)	- 6 (.863)	-23.9 (.507)	-17.8 (.614)	-6.7 (.874)			9.1 (.772)
CLT2					- 9.7 (.703)	-11 (.655)	12.8 (.608)	-14.7 (.567)	17.7 (.491)	-17.4 (.494)	-21.2 (.418)	-14.2 (.580)	-5 (.862)			
CLT3					- 5.9 (.811)	- 9.2 (.706)	-10 (.682)	-15.1 (.538)	-16 (.520)	-16.3 (.511)	-22.5 (.368)	-16.9 (.496)	.3 (.992)			
CLT4					-.9 (.961)	- 9.8 (.608)	-10.6 (.579)	- 8.9 (.648)	-18.4 (.358)	-19.8 (.318)	-20.4 (.311)	-18.9 (.341)	5.5 (.808)		- 3.2 (.782)	-3.1 (.786)
RL2					37.4 (.007)	29.3 (.033)	28.6 (.037)	24.1 (.085)	27.9 (.048)	26 (.004)	23.2 (.105)	21.8 (.122)	38.7 (.018)		5 (.670)	23.8 (.078)
RL3					-25.5 (.145)	-29 (.095)	-29 (.096)	-36.4 (.041)	-30.8 (.081)	-32 (.069)	-38.8 (.030)	-37 (.036)	-23.1 (.251)			-33.1 (.054)
RL4					51.9 (.000)	40 (.000)	40 (.283)	37.8 (.001)	39.3 (.001)	37.2 (.001)	36.5 (.002)	36.2 (.002)	45.8 (.001)			34.9 (.001)
LOC					13.4 (.241)	17.5 (.120)	16.2 (.151)	15.7 (.178)	14.6 (.214)	16 (.167)	16 (.186)	15.3 (.193)	15 (.264)		17.3 (.127)	16.3 (.144)
SF					- 1.7 (.862)	- 1 (.927)	- 1 (.931)	0.5 (.962)	- 1.4 (.884)	.14 (.999)	.7 (.942)	- .5 (.965)	- 5.9 (.599)			
HS1													- 3 (.907)			
HS2													- 3 (.918)			
HS4													-25.3 (.334)			
HS5													2.6 (.840)			
HS6													-13.8 (.469)			
HS7													38 (.329)			
HS8													38.7 (.215)			
HS9													7.7 (.636)			

A. <u>Functional Status</u>.

The first results to be examined in these tables were observed in the first five equations; equations which embody various alternative forms of the variable functional status, together with combinations of these forms. Here we were seeking to determine which specific functional status indicator was most appropriate for inclusion in each of the models.

Using the five functional status indicators, each of which was characterized by five levels of impairment: mobility (MOB), dexterity (DEX), sensory (SEN), comprehension (COMP) and home management (HOM) in Equation AI, Table 8.1, we observed that the adjusted coefficient of multiple determination ($\bar{R}^2$) in Model A was 0.229. Most significant were the sub-categories COMP and HOM.

Since multicollinearity was suspected between on the one hand HOM and COMP and on the other, the first three categories of functional status, MOB, DEX and SEN, the latter were entered into the equation alone in Model A, Equation AII, and yielded and $\bar{R}^2$ of 0.07. Equation AIII, in which only COMP and HOM were entered, yielded an $\bar{R}^2$ of 0.217. Multicollinearity appeared to be present and the categories MOB, DEX and SEN of the functional status variable explained little of the variation in cost. Another indicator of multicollinearity between the 2 groups of variables was indicated by large changes in both the coefficients and the significance among the MOB and SEN categories; in the first two equations in Model A, less change was evident among COMP and HOM categories which retained high degrees of significance in the first and third equations. Thus, COMP and HOM appeared to be the most appropriate indicators of the variable functional status in its implicit cardinal form.

Equation AIV presents the results of specifying functional status utilizing binary categories for each level of impairment. There was little difference between the $\bar{R}^2$ in Equation AI (0.229) and the 0.236 level of Equation AIV. This suggested equivalence between the two approaches. To simplify matters, the categories COMP and HOM still seemed to best explain variation in the endogenous variable.

Tables 8.2 and 8.3 provided comparisons among functional status specifications in Models B and C. The relationships between the various specifications of functional status categories were similar to those for Model A, but $\bar{R}^2$ was not as high in either model as in Model A. The

$\bar{R}^2$ fell in the range for Equations BI to BIV, of between 0.128 and 0.134 and the Equations CI to CIV of between 0.075 and 0.083.

The third alternative indicator of functional status, the services provided to the client, when entered into Model A alone, Equation V, resulted in an $\bar{R}^2$ of 0.42. Clearly, this was a superior indicator of variation in Total Cost per Month (TCM) to the impairment level indicators examined above.

However, when using the services indicator of functional status in Model B and Model C, there was only a marginal gain in the explanation of the variation in total hours worked per month (THM) (Equations BI to BV).* $\bar{R}^2$ rose from 0.134 to 0.163. There was no gain in the explanation of the endogenous variable cost per hour (CH) in Model C, (Equations CI to CV) where the $\bar{R}^2$ varied between 0.075 and 0.083.

B. Socio-Economic Status.

Before combining the socio-economic status variables (SES) with the above alternatives of

*Usually, with the exception of the endogenous variables, identical Equations were run for all 3 models. Several more equations than those listed in Table 8.4 were run but were dropped because they yielded no significant increase in information than that shown in the tables of equations, Tables 8.1, 8.2, and 8.3 and Table 8.4.

the functional status variables, the former was entered alone into the Models (Equations AVI, BVI, and CVI) where $\bar{R}^2$ was found to be 0.116, 0.184 and 0.082, respectively. From these results, it appears that SES was more important than functional status in explaining the variation in hours worked (THM) than it was in explaining month costs (TCM) or hourly costs (CH).

The effect of entering the alternative functional status indicators into the equations of all 3 models, together with the SES variable, is shown in Equations VII to XIII.

In Equation AVII, the 5 cardinal functional status indicators MOB, DEX, SEN, COMP and HOM were combined with all the socio-economic status categories. This raised the $\bar{R}^2$ from 0.229 to 0.299. In Models B and C, the $\bar{R}^2$ changed from 0.129 to 0.236 and from 0.075 to 0.116, respectively. When COMP and HOM alone were combined with the socio-economic status categories in Equations AVIII, BVIII, and CVIII, $\bar{R}^2$ rose from 0.217 to 0.278, from 0.128 to 0.232 and from 0.075 to 0.116, respectively. Clearly, in these equations, socio-economic status was an important variable in explaining all 3 endogenous variables, especially hours worked (TCH).

The binary form of the functional status categories were combined with socio-economic status in Equations AIX, BIX, and CIX. The magnitudes of $\bar{R}^2$ in each of these equations differed little from those in Equations AVIII, BVIII, and CVIII. However, the coefficients were much more unstable suggesting high degrees of multicollinearity in this form. It is worth noting that the coefficients of the various binary formulations of comprehension and home management (e.g. C_2, C_3, C_4, C_5, and K_2, K_5) were more stable than the equivalent coefficients of mobility, dexterity and sensory perception. These results coincided with the relatively stable coefficients for COMP and HOM in the cardinal form of the functional status variable.

The coefficient of multiple determination, $\bar{R}^2$, displayed a significant rise to 0.461 when the third alternative indicator of functional status namely, services, was combined with socio-economic status in Equation AX. In Model B, the same variable categories resulted in only a marginal increase in $\bar{R}^2$ to 0.249. In Model C, the same functional form resulted in a marginal decrease in $\bar{R}^2$ to 0.111.

In Chapter 7 no patterned relationship was found linking client attributes to services. This outcome indicates little correlation between attributes such as functional status ("cardinal" or binary indexes), or socio-economic status and services provided. Yet, intuitively, it would seem that in a regression frame work there is likely to be a high degree of multicollinearity between the cardinal or binary indicators of functional status and services rendered. In Equations XI and XII, in all models, this seems to be the case. In Equation XI, the COMP-HOM form of the cardinal functional status index has been entered into the equation together with the alternative, but not mutually exclusive, functional status indicator, services. There was a marginal gain of 0.013, 0.012 and 0.017 in the magnitude of $\overline{R}^2$ for Models A, B, and C, respectively. These marginal gains when viewed together with the large changes in coefficients, associated with COMP and HOM appeared to confirm the above mentioned intuition with respect to multicollinearity. The relatively large amounts of cost or time variances explained by the services categories compared with those of the impairment level categories suggested that the service index was the more powerful predictor.

This result was reinforced by the observation that the coefficients of the services indicator of functional status were relatively more stable than those of the alternative specifications of functional status. These results were also confirmed in Equations XII where the binary form of functional status was substituted for that of the cardinal form as in Equation XI. This substitution made little difference to the magnitude of $\bar{R}^2$. Neither was there much further change when only the C_2 to C_5 and K_2 to K_5 binary components of the functional status were entered into Equation XIII together with services categories and socio-economic status.

C. <u>Medical Status</u>.

As expected, the introduction of the variable medical status into the models (see Equation XIII), did not add explanatory power. In all models the decrease in degrees of freedom served to cause a decrease in the $\bar{R}^2$. For example, in Model A it dropped from 0.474 in Equation XI to 0.451.

D. <u>The Search for an Optimal Model</u>.

Two types of criteria were taken into consideration in defining an optimal model subject to constraints of the available data: (1) technical criteria and (2) policy criteria.

1. Technical criteria.

The number of degrees of freedom used up in the estimation of the models had an effect on the magnitude of the $\bar{R}^2$. This was apparent in the estimation of Equation XIII which included the Medical Status variable. It was apparent that by dropping what appeared to be insignificant categories, the explanatory power of the models might be increased.

2. Policy criteria.

The policy maker seeks to determine anticipated costs on the basis of a minimum amount of information. Cutting down the amounts and types of information required, reduces the propensity to commit errors and increases the time effectiveness of staff especially social workers who record the necessary information. Thus, an optimal model would preserve a high degree of predictive power and base it on a minimum of information. There is a direct trade-off to some extent between predictive power and information requirements.

Bearing in mind these criteria, two further forms of the model were estimated. Equation XV included only 4 categories of services, light cleaning, meal preparation, attendant and non-nursing personal services. These specifications reduced $\bar{R}^2$ to 0.412, 0.201 and 0.093 in Models A, B, and C, respectively. It is clear that though the reduction in $\bar{R}^2$ is evident, the four categories accounted for a large proportion of the variance in costs and hours worked.

By adding COMP and HOM together with one category of Age (80 years and above), 2 categories of income (INC3, INC4), one category of client contribution (CLT4), one category of the clients relationship to the provider (RL2), and location, the $\bar{R}^2$ was raised to virtually its highest level in Model A, that is 0.477. In Model B, the level of $\bar{R}^2$ rose even higher than that previously recorded to 0.276. Similar to Model A, the $\bar{R}^2$ in Model C remained at its previous highest recorded level of 0.135.

The results from Equations XV and XVI indicated the data requirements for cost predictions could be significantly reduced from current data recorded. The Equations explaining the largest amounts of variance in the endogenous variables were Equations XI, XII, XIII and XVI in all 3 models.

III. Effects of Variable Categories on the Variation in Total Cost Per Month, Total Hours Per Month and Cost Per Hour

The effect of including and excluding functional status categories was partially covered in the above section. It was found that the categories COMP and HOM in both the cardinal form and binary form, explained more of the variation in costs or hours than MOB, DEX, or SENS or their binary form equivalents.

The significant service variables, light cleaning (LK), meal preparation (MP), attendant (A), and non-nursing personal (NNP), when entered into the Model A along (Equation AXV) resulted in an $\bar{R}^2$ of 0.412. This compares with an $\bar{R}^2$ of 0.420 when all service categories are entered into the equation together (Equation AV). The coefficients of LK, MP, A, and NNP and their level of significance, remained relatively stable whether or not the other service variables were present suggesting little or no multicollinearity between the 2 groups of service categories.

Throughout all equations in Model B, the LK category had a significant coefficient. The effects of services alone in the equation accounted for an $\bar{R}^2$ in the range of 0.155 to 0.163. Otherwise, the results were similar to those found in Model A. In Model C, the most significant service category was Attendant (A) but the level of $\bar{R}^2$ varied between .064 and .074. Consequently, little importance can be attached to the different forms in which the service categories were entered into this model.

To estimate the effects of the different categories of socio-economic status, each category was dropped from Equation XI in each of the 3 models. Not more than one SES category was dropped from the equation for any one computer run. The results are presented in Table 8.4, run numbers XVII to XXI. In Model A, the location (LOC) category appeared to account for much of the variation attributed to SES, causing the $\bar{R}^2$ to drop from 0.471 to 0.450 when it was omitted. In Models B and C, the variation in hours and hourly cost can be largely attributed both to the relationship between the provider and client and their living arrangements relative to each other. These results are discussed more fully in the next section. The $\bar{R}^2$ dropped from 0.261 to 0.177 and from 0.128 to 0.086 when the relationship and living arrangement category (RL2, RL3 and $L4) were dropped from Models B and C, respectively.

IV. An Examination of Coefficients

Due to the occurrence of singularity in the estimation of coefficients in a model in which binary variables appear, certain subcategories of variables must be excluded. Estimation of coefficients would otherwise be impossible. Thus, the coefficients themselves cannot be interpreted with respect to their individual values. However, it is possible to gain some further knowledge of the inter-relationships among subcategories of variable categories by examining the

TABLE 8.4.--A comparison of $\bar{R}^2$ for all models

Training runs

		MODELS		
		A(TCM)	B (THM)	C(CH)
Run No.	Variables and Categories	$\bar{R}^2$	$\bar{R}^2$	$\bar{R}^2$
I	MOB, DEX, SEN, COMP, HOM	.229	.129	.075
II	MOB, DEX, SEN	.07		
III	COMP, HOM	.217	.128	.075
IV	M2-M5, D2-D5, S2-S5, C2-C5, K2-K5 (i.e., all FS, binary)	.236	.134	.083
V	LK, HK, HMR, MP, L, S, FM, TT, A, Y, NNP (i.e. all services)	.420	.163	.074
VI	A2-A4, SF, LOC, INC2-INC5, CLT2-CLT4, RL2-RL4, (i.e. all SES)	.116	.184	.082
VII	MOB, DEX, SEN, COMP, HOM, all SES	.299	.236	.116
VIII	COMP, HOM, all SES	.278	.232	.116
IX	All FS binary, all SES	.304	.238	.126
X	All services, all SES	.401	.249	.111
XI	COMP, HOM, all services, all SES	.474	.261	.128
XII	All FS binary, all services, all SES	.476	.264	.135
XIII	C2-C5, K2-K5, all services, all SES	.481	.334	.134
XIV	COMP, HOM, all services, all SES, all HS	.451	.218	.116
XV	LK, MP, A, NNP	.412	.155	.064
XVII	Same as run XI without A2-A4	.472	.262	.128

TABLE 8.4 (continued)

Training runs				
			MODELS	
		A(TCM)	B (THM)	C(CH)
Run No.	Variables and Categories	$\bar{R}^2$	$\bar{R}^2$	$\bar{R}^2$
XVIII	Same as run XI without CLT2-CLT4	.475	.263	.132
XIX	Same as run XI without INC2-INC5	.471	.261	
XX	Same as run XI without RL2-RL4	.471	.177	.086
XXI	Same as run XI without LOC	.440	.262	.126
XXII	COMP, HOM, LK, MP, A, NPN, and all SES minus A_2, A_3 INC2, CLT2, CLT3	.474	.276	.135

relative values of coefficients. These comparisons are useful to the degree that the coefficients are significant and remain relatively stable with respect to each other. A low level of confidence attributed to a coefficient would indicate the coefficient to be subject to large changes in magnitude. For example, in Models A and B, the consistently large positive magnitude of HOM relative to COMP indicates that the physical ability of the client to manage their home is more important in explaining variation in costs per month and hours per month, than is their mental ability.

The relative values of HOM and COMP coefficients were reversed in Model C when compared with Models A and B. Also, where the binary form of functional status was employed in Models A and B, the costs appeared to increase as the level of impairment increased. The opposite was the case in Model C where, as the level of impairment increased, the hourly cost of providing services decreased. This result could have been induced by the $270 monthly payment ceiling imposed by the Michigan State Government. If the ceiling were raised, it is conceivable hourly rates could also rise. Model C is of particular interest with respect to the binary representation of mental ability C2, C3, C4 and C5. The upper levels of impairment C4 and C5 showed both stability and relatively high level of significance. There was an inverse relationship between hourly costs and level of impairment.

It is interesting to note that when services were included in the Model A, for example Equation AXII, the relative magnitude within the Mobility subcategories (M2 to M5) and comprehension subcategories (C2 to C5) were reversed resulting in a decline in TCM as impairment within categories increases. This suggests a multicollinear effect between Mobility and Comprehension, on the one hand, and Services on the other. By the same token, some of the coefficients of service subcategories also changed in Model A when the binary form of functional status categories was added to the service and socio-economic status variables in the transformation from Equation AX to AXII. Heavy Cleaning (HK), Financial Management (FM) and Yardwork (Y) were especially affected. Though this observation also suggested the presence of multicollinearity, it should be viewed with caution since the levels of significance were relatively high. The same kind of effects were apparent when the truncated "cardinal" form of functional status (COMP and HOM) was added to the service and socio-economic status variables, Equations AX and AXI. The coefficients and significance of COMP, HOM, HK and FM were all affected. The probable explanation for the multicollinearity stems from the apparent relationship between the functional status category home management (HOM) or its binary equivalent K2 to K5 and service category heavy cleaning (HK); also between COMP or C2 to C5 and service category, financial management (FM).

Though there was apparent multicollinearity, the relative level of significance of COMP and HOM and the slight increase in explanatory power due to their presence, indicated it was useful to keep them in the models together with the service categories. Of the relatively significant categories among services (LK, HMR, MP, S, A and NNP) the attendant services, meal services and non-nursing personal services appeared to have the most effect on cost per month.

In Equations BXI, BXII, light cleaning (LK) appeared to have the most effect on hours worked by a provider. Laundry (L), meal preparation (MP), and attendant (A) services seemed to have the most impact on costs per hour in Model C. These statements should be viewed with caution in light of the low level of confidence associated with the coefficients.

Among the SES categories in Model A, location (LOC) and Relationship and Living Patterns between providers and clients (RL) were found to be often significant. Urban clients were $30 to $40 more costly on a monthly basis than rural clients, but were not significantly different from rural clients with respect to hours worked. Hourly costs were marginally greater for urban than for rural clients, on the order of 16 cents more per hour.

As evident in Equations AXI, AXII and AXIII, services provided to the client who was related to, but living separately from, the provider (RL2), cost less per month than the other 3 relationship-cum-living patterns. The coefficient for the related but separate provider was significant according to the F test at the 0.1 level. The other 3 categories of relationship-cum-living pattern were substantially different from each other. The monthly costs of an unrelated client living separately from the provider (RL4) was only marginally higher in cost than the living-in related provider (RL1 the omitted subcategory) or the non-relative in the same house (RL3). Since neither the coefficients of RL3 nor RL4 were significant, no firm conclusions can be drawn.

In Model B, all 3 relationship-cum-living pattern subcategories were significant at less than the 0.05 level for most computer runs.

Providers living separately from clients (RL2 and RL4) worked fewer hours than those living with the client. But non-relatives living with the client (RL3) supplied more hours of services than live-in relatives.

Cost per hour in Model C showed, as expected after observing the results in Models A and B, an inverse relationship to those described in Model B. Non-relatives providing services to clients living separately (RL4) resulted in the highest hourly cost of services. The next highest hourly costs were realized by provider's relatives living separately from the client, followed by the relative living with the client (RL2). The lowest hourly costs were incurred by the non-relative living with the client (RL3).

V. Results from the 'Test' Sample

Many equations were analyzed using data from the training set of data. The results were tested using a separate test set of data. The reason for this was to test whether the results based on the training sample were contrived.

The 200 case test set sample was analyzed by employing 3 equation specifications identical to 3 of the equations in each of the 3 models. The 'test' set equations and their training set equivalents, respectively, are as follows:

A_TXII (AXII); A_TXIII (AVII); A_TXVI (AXVI);
B_TXII (BXII); B_TXIII (BXIII); B_TXVII (BXVII);
C_TXII (CXII); C_TXIII (CXIII); C_TXVII (CXVII)

The 'test' set results are presented in Table 8.5 and designated by the subscript T.

In the case of Model A, $\bar{R}^2$ in the 'test' sample compares well with that of the 'training' sample results. Equations in Model A:

MODEL A EQUATIONS

	AXII	A_TXII	AXIII	A_TXIII	AXVI	A_TXVI
$\bar{R}^2$	.476	.494	.481	.453	.477	.435

In Model B the results were unusual because the test results are better than those from the training set.

MODEL B EQUATIONS

	BXII	B_TXII	BXIII	B_TXIII	BXVII	B_TXVII
$\bar{R}^2$	.264	.312	.258	.334	.276	.315

TABLE 8.5.--Test results - Models A, B, C

	Model A			Model B			Model C		
	$A^T{}_I$	$A^T{}_{II}$	$A^T{}_{III}$	$B^T{}_I$	$B^T{}_{II}$	$B^T{}_{III}$	$C^T{}_I$	$C^T{}_{II}$	$C^T{}_{III}$
R^2	0.617	0.551	0.471	0.479	0.458	0.319	0.514	0.460	0.350
$\bar{R}^2$	0.494	0.453	0.435	0.312	0.334	0.315	0.359	0.342	0.300
MOB									
DEX									
SEN									
COMP			508 (.091)			125 (.260)			- 7 (.061)
HOM			1016 (.101)			32 (.874)			0 (.998)
M2	3351 (.307)			10 (.993)			46 (.256)		
M3	5936 (.012)			-364 (.654)			12 (.674)		
M4	6749 (.010)			-367 (.683)			13 (.670)		
M5	5788 (.021)			- 61 (.943)			- 19 (.526)		
D2	-2197 (.323)			-603 (.434)			24 (.383)		
D3	-3131 (.167)			425 (.588)			0 (.986)		
D4	-2659 (.259)			77 (.925)			10 (.710)		
D5	270 (.914)			131 (.880)			29 (.347)		
S2	-1020 (.429)			-553 (.217)			15 (.345)		
S3	-1860 (.126)			-324 (.441)			1 (.962)		
S4	-1197 (.529)			-435 (.510)			45 (.054)		
S5	-9339 (.000)			111 (.899)			- 6 (.826)		
C2	2588 (.046)	1886 (.149)		- 85 (.848)	- 73 (.863)		16 (.305)	16 (.307)	
C3	-206 (.899)	15 (.992)		174 (.757)	- 40 (.937)		7 (.706)	14 (.457)	
C4	2076 (.213)	1146 (.471)		505 (.382)	343 (.510)		- 36 (.082)	- 29 (.124)	

TABLE 8.5 (continued)

	Model A			Model B			Model C		
	$A^T{}_I$	$A^T{}_{II}$	$A^T{}_{III}$	$B^T{}_I$	$B^T{}_{II}$	$B^T{}_{III}$	$C^T{}_I$	$C^T{}_{II}$	$C^T{}_{III}$
C5	3996	1994		310	353		- 34	- 29	
	(.016)	(.196)		(.586)	(.483)		(.093)	(.118)	
K2	30895	-18004		-1532	-1506		48	63	
	(.000)	(.008)		(.549)	(.497)		(.597)	(.436)	
K3	-32223	-18054		-3263	-2961		106	103	
	(.000)	(.004)		(.173)	(.150)		(.212)	(.173)	
K4	-30547	-15334		-2135	-2236		102	114	
	(.000)	(.012)		(.370)	(.258)		(.228)	(.117)	
K5	-29626	-13600		-2054	-2040		86	98	
	(.000)	(.023)		(.386)	(.293)		(.306)	(.172)	
LK	- 1156	- 674	- 437	- 374	- 313	- 570	30	27	36
	(.337)	(.578)	(.696)	(.371)	(.429)	(.127)	(.044)	(.066)	(.010)
HK	- 257	- 765		- 678	- 731		24	29	
	(.835)	(.535)		(.115)	(.071)		(.110)	(.050)	
HMR	1102	1248		104	223		- 24	- 26	
	(.402)	(.324)		(.819)	(.589)		(.136)	(.085)	
MP	- 5093	- 5510	-5305	- 963	- 848	- 885	19	19	22
	(.000)	(.000)	(.000)	(.016)	(.024)	(.012)	(.168)	(.158)	(.089)
L	- 368	369		1115	585		- 31	- 17	
	(.868)	(.854)		(.148)	(.374)		(.259)	(.461)	
S	-2254	- 1112		- 116	- 199		11	15	
	(.145)	(.469)		(.827)	(.691)		(.554)	(.393)	
FM	624	260		- 566	- 533		13	8	
	(.603)	(.827)		(.176)	(.157)		(.353)	(.568)	
TT	769	1881		634	521		- 11	- 7	
	(.447)	(.058)		(.072)	(.017)		(.352)	(.524)	
A	-2634	-2418	-1731	- 454	- 450	- 416	13	12	10
	(.014)	(.024)	(.078)	(.216)	(.197)	(.202	(.311)	(.327)	(.383)
Y	2419	2051		-1014	- 922		60	52	
	(.111)	(.171)		(.055)	(.060)		(.002)	(.604)	
NNP	-4519	-5208	-5208	- 324	- 451	- 483	19	21	22
	(.000)	(.000)	(.000)	(.377)	(.179)	(.130)	(.143)	(.085)	(.065)
A2	- 208	148		192	127		- 9	- 5	
	(.870	(.907)		(.663)	(.759)		(.551)	(.702)	
A3	- 257	- 295		- 24	- 165		- 1	1	
	(.847)	(.818)		(.958)	(.693)		(.923)	(.947)	
A4	2042	1591	1414	- 187	- 266	- 7	- 4	2	1
	(.144)	(.227)	(.154)	(.698)	(.535)	(.982)	(.815)	(.890)	(.938)
INC2	1782	2339		1437	1315		- 70	- 62	
	(.409)	(.288)		(.057)	(.069)		(.009)	(.019)	

TABLE 8.5 (continued)

	Model A			Model B			Model C		
	A^T_I	A^T_{II}	A^T_{III}	B^T_I	B^T_{II}	B^T_{III}	C^T_I	C^T_{II}	C^T_{III}
INC3	2424 (.220)	2332 (.245)	1087 (.377)	1002 (.145)	824 (.210)	- 268 (.525)	- 57 (.021)	44 (.066)	- 1 (.955)
INC4	2125 (.332)	2281 (.308)	777 (.616)	1047 (.169)	815 (.265)	- 445 (.408)	- 63 (.020)	- 51 (.056)	- 2 (.919)
INC5	422 (.927)	678 (.885)		- 717 (.656)	- 728 (.636)	-1731 (.208)	26 (.650)	32 (.563)	68 (.184)
CLT2	-2533 (.394)	-3427 (.245)		-1859 (.073)	-1900 (.050)		52 (.156)	53 (.135)	
CLT3	- 84 (.977)	-1241 (.667)		-2263 (.025)	-2464 (.010)		70 (.051)	72 (.038)	
CLT4	-3067 (.186)	-4462 (.058)	-2558 (.078)	-1159 (.150)	-1164 (.130)	489 (.320)	23 (.418)	16 (.558	- 22 (.224)
RL2	-1692 (.288)	-1755 (.277)	- 741 (.532)	-1706 (.002)	-1683 (.002)	-1757 (.000)	30 (.050)	38 (.050)	46 (.012)
RL3	-1483 (.398)	-1748 (.321)		162 (.789)	300 (.602)	238 (.671)	- 40 (.065)	- 45 (.036)	- 42 (.043)
RL4	-1246 (.324)	-1316 (.301)		-1978 (.000)	-1916 (.000)	-1798 (.000)	41 (.008)	45 (.003)	44 (.002)
LOC	2019 (.242)	864 (.619)	1924 (.214)	234 (.695)	- 191 (.731)	268 (.603)	7 (.716)	3 (.865)	1 (.927)
SF	- 527 (.627)	286 (.791)		597 (.114)	562 (.114)		- 19 (.145)	- 14 (.285)	

It is in Model C that the results have been found difficult to logically explain because the "test" results are clearly superior:

MODEL C EQUATIONS

	CXII	C_TXII	CXIII	C_TXIII	CXVII	C_TXVII
$\bar{R}^2$	.135	.359	.134	.348	.135	.300

In Model C, efforts were made to detect outliers causing the increase in $\bar{R}^2$, with no success. One feature that is noteworthy is the difference in means, and to a lesser extent, standard deviations of the endogenous variables. These are detailed in Table 8.6.

TABLE 8.6.--Means and standard deviations of endogenous variables

	Training Sample		Test Sample	
	Mean	Std. Dev.	Mean	Std. Dev.
MODEL A(1)	17939	7948	18465	7162
MODEL B(2)	1708	2047	1915	213[illegible]
MODEL C(3)	183	85	165	79

The standard deviations are relatively large, indicating the differences in means to be acceptable.

(1) In Model A, the mean and standard deviation are measured in cents; thus, 17939 indicates $179.39/month.

(2) In Model B, the mean and standard deviation are measured in tenths of an hour; thus, 1708 indicates 170.8 hrs/month.

(3) In Model C, the mean and standard deviation are measured in cents; thus, 183 indicates $1.83/hour.

In comparing variable categories and subcategories between the two samples, each model will be taken up in turn. Only the seemingly most revealing results will be presented, the criteria for their choice being the degree of significance and the extent of similarity between coefficients. The subcategories or categories have been divided into 3 groups. The first group (Group I) is comprised of those coefficients with similar magnitudes and high significance. The second group (Group II) embodies those coefficients which are less similar in magnitude or significance than those in the first group but for which interpretation of the results remains basically similar. The third group (Group III) shows contradictory results.

The comparison of results in Model A is presented in Table 8.7. Two equations in each sample are modified and used to demonstrate the similarities and differences between the two samples. The variable categories showing the highest degree of correspondence (designated Group I) were meal preparation (MP) and non-nursing personal services (NNP). Agreement between the 2 samples was not so clear in the categories LK, A and LOC and subcategory RL2 in Group II. Nevertheless, the interpretation of the results remain essentially similar. The magnitude of the location coefficient decreased by almost half and the significance dropped. The most common result of all the coefficients in Group II was the difference in significance of coefficients between the 'training' set and the 'test' set. Since the latter

TABLE 8.7.--A comparison of selected categories and subcategories between "training" and "test" samples from two equations in Model A

	Category	Equations (Incomplete)			
		AXII	A_TXII	AXVI	A_TXVI
Group I	MP	-4732 (.000)	-5093 (.000)	-5410 (.000)	-5305 (.000)
	NNP	-4357 (.000)	-4519 (.000)	-4592 (.000)	-5208 (.000)
Group II	LK	-2347 (.009)	-1156 (.337)	-2013 (.011)	- 437 (.696)
	A	-1648 (.037)	-2634 (.014)	-1502 (.042)	-1731 (.078)
	RL2	-1721 (.097)	-1692 (.288)	-1305 (.115)	- 741 (.532)
	LOC	3601 (.000)	2019 (.242)	3629 (.000)	1924 .241)
Group III	M_3	96 (.955)	5936 (.012)		
	M_4	- 57 (.977)	6749 (.010)		
	M_5	- 67 (.970)	5788 (.021)		

had fewer than one-half the cases of the 'training' set, the drop in significance should not be surprising.

The contradictory results characterizing Group III showed up only in the case of the binary variables representing mobility: M_3, M_4, and M_5. Significance increased to a significant level with the test results and the signs of two coefficients changed. These results, as with those of the other categories and subcategories not listed in Table 8.7 suggest these variables should be ignored in future research.

The comparison of training set and test set results in Model B is presented in Table 8.8 using the same format as that used for Model A. Again, meal preparation (MP) appeared to be the most significant predictor of hours of service provided per month to a client. Coefficients in the training set are Attendant services (A), all relationship-cum-living arrangements categories (RL2, RL3, and RL4), yardwork (Y), and non-nursing personal services (NNP). In each case, the significance and signs were such that interpretation of the results of the two sets were virtually the same.

The anomalies evident in the Group III occurred in the category HOM, and subcategories M_3, M_4, and M_5, the latter 3 being the same as those in Model A.

In Model C, the same number but not the same specific variable categories and subcategories were consistent across the training and test sets (Table 8.9).

TABLE 8.8.--A comparison of selected categories and subcategories between "training" and "test" samples from two equations in Model B

Category		Equations (Incomplete)			
		BXII	B_TXII	BXVI	B_TXVII
Group I	MP	- 828 (.002)	- 963 (.016)	- 810 (.001)	- 885 (.012)
Group II	A	- 664 (.006)	- 454 (.218)	- 577 (.010)	- 416 (.202)
	RL2	-1256 (.000)	-1706 (.002)	-1109 (.000)	-1757 (.000
	RL3	1171 (.003)	162 (.789)	1066 (.005)	238 (.671)
	RL4	-1090 (.000)	-1978 (.000)	-1039 (.000)	-1798 (.000)
	Y	- 338 (.208)	-1014 (.055)		
	NNP	- 216 (.336)	- 324 (.377)		
GROUP III	HOM			352 (.007)	32 (.872)
	M_3	882 (.096)	- 364 (.654)		
	M_4	990 (.077)	- 367 (.683)		
	M_5	908 (.097)	- 61 (.949)		

TABLE 8.9.--A comparison of selected categories and sub-categories between "training" and "test" samples from two equations in Model C

Category		Equations (Incomplete)			
		CXII	C_TXII	CXVII	C_TXVII
Group I	COMP			- 5.8 (.081)	- 7.0 (.061)
	C_4	- 30.3 (.055)	- 29 (.124)		
	RL2	23.2 (.105)	30 (.131)	23.8 (.078)	46 (.102)
	RL3	- 38.8 (.033)	- 40 (.065)	- 33.1 (.054)	- 42 (.043)
	RL3	36.5 (.002)	41 (.008)	34.9 (.001)	44 (.002)
Group II	A	25.1 (.021)	19 (.168)	22.3 (.030)	10 (.383)
Group III	LK	- 13.7 (.271)	30 (.044)	- 20.1 (.067)	36 (.010)
	K_3	-110 (0.037)	103 (.173)		
	K_4	-111 (.030)	114 (.117)		
	K_5	-125 (.014)	98 (.172)		

Due to the differences in $\bar{R}^2$ between the 'training' and 'test' sets, the results should be viewed with more caution. The significance of the relationship-cum-living subcategories RL2, RL3, and RL4 was clearly consistent across the two sets; hourly costs being predictable to some extent using these subcategories. Client comprehension in both one binary subcategory (C_4) and the 'cardinal' category (COMP) also appeared consistent suggesting that within the relevant range hourly costs were inversely related to mental impairment.

Only Attendant (A) service coefficients could be interpreted as similar (but not convincingly so) in Group II.

In Group 3, inconsistencies were found in the coefficients of light cleaning (LK) and 3 binary subcategories of Home Management (K_3, K_4, and K_5).

Further discussion of these results appear in the conclusion.

CHAPTER 9

CONCLUSIONS, FURTHER RESEARCH AND POLICY IMPLICATIONS

I. Introduction

This study has focused on the analysis of non-medical long-term care services assigned to the impaired under the Michigan Adult Chore Service Program. The services were supplied by providers who were hired by predominantly elderly clients. The client was financed by the Michigan Department of Social Services (MDSS) to acquire one or more of the following types of services: light and heavy cleaning, home repair and maintenance, shopping and errands, laundry, meal preparation, financial management, yardwork, attendant, transportation and non-nursing personal services.

The aim of the research was to examine two principal relationships: first, to determine whether or not there was a systematic relationship between the client's attributes and the services assigned to the client by the MDSS workers; second, the estimation of the relationship between the client's attributes and the cost of services. Implicitly, the study examines the decision making process by which MDSS workers assign services to clients.

Under the MDSS program, the assignment of services to clients by MDSS workers was purportedly based on the following attributes: functional status, socio-economic status and to a lesser extent, the medical status of the client. Functional status attributes included the degree of impairment in mobility, dexterity, sensory perception, comprehension and ability to manage the home. Socio-economic status covered income, age, sex, type of living arrangement, relationship of client to provider, and location (urban/rural). Medical status entailed diagnosis of a patient according to criteria such as respiratory problems, heart problems, cancer, rehabilitation problems, mental illness and retardation and others totaling 9 groups of diagnoses. If services are assigned to clients by MDSS workers on the basis of client attributes, it should be possible to demonstrate the types of services assigned to clients on the basis of an attribute profile. In addition, knowledge of the clients' attributes should facilitate estimation of service costs and hours of labor supplied by providers.

To determine whether systematic relationships existed between client attribute profiles and assigned services, a pattern discovery algorithm, called entropy minimax was employed. Stemming from developments in informational theory, entropy minimax was used because it could manipulate several assigned services and multiple attributes simultaneously and produce probabilistic relationships between assigned services and attributes. It also had the advantage of obviating the necessity of assuming a priori distributions in the data. The results appear in the next section.

The relationship between the client attributes as explanatory variables and service costs was estimated using regression analysis. The goal of the regression analysis was to identify, if possible, predictors of cost that would indicate where costs might be controlled. For example, we were seeking answers to such questions as: Do services provided to clients by relatives cost less than those provided by non-relatives? Is cost related to age of the client or to whether the provider lives with the client? Is cost related to whether the provider lives in a rural or urban area? In the hypotheses detailed in Chapter 5, theories on these and other questions have been postulated, e.g., assuming otherwise identical profiles, clients taken care of by relatives were less costly to care for than client's receiving services from non-relatives.

The two thrusts of the thesis are discussed below together with alternative explanations of the results. This is followed by a retrospective look at the results in the context of the aims of the thesis and future research. Finally, policy implications of this study are briefly discussed.

II. The Relationship Between Client Attributes and Assigned Services

The objective of this analysis was to seek to identify and utilize a methodology in order to determine if a systematic relationship could be detected between client attributes and assigned services. The entropy minimax procedure was the method used to search for patterns of service among a sample of 428 MDSS Chore Service clients or approximately 4% of the total client population in 1975-1976.

Using the available data on the Michigan Chore Service Program clients, together with the attribute profile associated with that program, no systematic relationship was discovered between client attributes and assigned services. Employment of the entropy minimax algorithm did not detect a patterned relationship between client attribute profiles and assigned services. Several features of the analysis conducted in this study might account for these results being inconclusive.

Methodological problems could have stemmed from misspecification of attribute categories. This implies the profile was inappropriate and that indicators other than, say, dexterity or mobility, might have better characterized the client. For example, the client's ability to dress himself might be a better indicator of impairment than dexterity.

Since some attributes were cardinal in nature and necessitated transforming to a binary form, a further methodological shortcoming of the analysis may have stemmed from inappropriately dichotomizing the category. The specific scale on which the transformation was performed may also have been inappropriate. Table 6.4 in Chapter 6 presents some of the specific categories transformed.

Another methodological problem may have resulted from the specification criteria employed to generate the five alternative outcome mixes. The reliability of the results of the entropy minimax computer program diminishes when more than 5 or 6 outcome categories are entered into the program. For that reason, it was necessary to either aggregate the 11 services into groups or employ some outcomes (services) as proxies for others. The criteria by which this specification of outcomes was accomplished is more fully described in Section VI-A of Chapter 6.

The regression analysis of Models A, B, and C tended to confirm the absence of a patterned relationship between client attributes and services rendered. This observation was based on thefinding that in regression Model A the service variable alone accounted for more variation in the costs per month ($\overline{R}^2 = 0.420$) than did either functional status alone ($\overline{R}^2 = 0.116$) or both of these variables combined ($\overline{R}^2 = .304$). When functional status, socio-economic status and services were combined, $\overline{R}^2 = 0.474$. Those categories of the functional status variable associated with almost negligible amounts of variation in monthly costs were mobility, dexterity and sensory perception. The implication of these findings is that the level of a client's dysfunction with respect to mobility, dexterity, and sensory perception were poor predictors of the client's assigned services.

The unexpected results of no relationships found by the entropy minimax analysis may stem from a lack of consistency among MDSS workers in their assignment of services. For example, MDSS workers may have no common working definitions of the variable categories, although the program had been running since 1972. Even if the definitions were common, the scaling of them may differ from worker to worker. Alternatively, the MDSS worker may assign services by implicitly taking into account more or fewer client attributes than those listed. Those key attributes used in each decision may vary from client to client and from worker to worker.

Another approach for the MDSS worker might be to inquire as to what services are perceived as required by the client, e.g., cleaning, non-nursing personal services, etc. In other words, to seek to determine what it is the <u>client cannot do that needs to be done</u> in order to maintain (produce) an independent life style. The inability to cope with some everyday tasks seems to be the primary motivation of the person requesting the MDSS services in the first place. Clearly it is not mandatory that a client need be mobile or dextrous to be independent. Neither does it follow that because the client is immobile or lacks dexterity that the MDSS will fund services such as cleaning or meal preparation. The specific assignment of services to a client by the MDSS worker would depend on both the client's inability to perform such services herself and on the MDSS worker's perception that without such services the client cannot continue to live appropriately in an independent setting. It is conjectured the MDSS worker reviews both the client's request for services and the consequences of doing without those services before assigning services. Implicit in this conjecture is the belief that the services are assigned to a client primarily on the basis of what the client is perceived to need by the MDSS worker rather than on the basis of the attributes of the client. Client attributes can subsequently be specified by the worker to justify the employment of services assigned by them to the client. Specific attributes need not be prerequisites for specific services. The one exception to this might have been the attribute comprehension which correlated with financial management services ($r = .55$).

As pointed out in the thesis, several approaches to determining long-term care are based on different client attribute profiles. In view of the foregoing, these approaches should be tested to determine if systematic linkages exist between attributes and services assigned. These methods of determining appropriate care are <u>indirect</u> approaches in that it is presupposed that the attribute profile of the client is first specified and then utilized in the assignment of services. A <u>direct</u> approach to assigning services would be to classify clients according to the services they require in order to maintain relatively independent lifestyles. Such an approach would only implicitly take into account functional, socio-economic and medical status. Thus, a person might be classified as requiring meal services, cleaning, laundry and shopping services. A more refined format might specify in the case of meal services, which meals are needed and the average length of time it takes to prepare them. This subjectivity of the decision making process in assigning non-medical services to clients is in apparent contrast to the more "objective" approach of assigning medical services used by the medical profession.

The approach to the assignment of subsidized services based directly on service needs might, due to its "subjectivity," jeopardize the program by rendering it more vulnerable to charges of abuse or fraud. The direct approach of stating needs for services might be a less "objective" justification for government assistance than a profile of readily observed attributes on which "need" or the assignment of services is supposed to be based. The formal specification and recording of the attribute profile together with the assigned services would appear to make abuse or fraud more difficult to perpetrate. Employed in this way, attribute profiles might serve less to justify the assignment of specific services than to justify services *per se*. The profile itself becomes an accountability check on the MDSS worker. If the worker views it exclusively as such, then his/her approach to assigning services is direct in that it is in direct *response to the service needs requested by the client*. The justification of services by the detailing of a client's attribute profile would appear to *respond to the MDSS requirement* that the MDSS worker, and in turn MDSS itself, be accountable to the legislature for responsible allocation of public funds. The primary purpose of utilizing an attribute profile thus becomes one of ensuring government accountability, not the justification of specific services. It provides an auditable record of why services were assigned. According to the results of pattern detection analysis, the attribute profile does not provide decision criteria as to what specific services ought to be assigned to the client. Another use to which the attribute profile may be put is the prediction of client outcome over time, given necessary support services to keep the client out of an institution.

III. The Determination of Categories Influencing Costs

The regression analysis was performed on three 'behavioral' as distinct from 'technical' cost functions. The distinction between these types of functions is explained in Chapter 5. Each of the cost functions had a different endogenous variable: Costs per month (Model A), hours per month (Model B), and costs per hour (Model C).

Model A best fitted the data in so far as it explained more variance (almost 50%) than the other models (33% and 13%, respectively). As noted above, the most significant variable in Model A was assigned services. Important categories of the service variable appeared to be meal preparation, light cleaning, attendant services and non-nursing personal services.

Among the categories of the socio-economic status variable, only location was significant in Model A, indicating higher costs in urban areas. This result was unconfirmed in the "test data," as distinct from the "training data."

The specific functional status variable categories mobility, dexterity and sensory perception were not as powerful explanatory categories as were the more general categories comprehension and ability to manage the home. Comprehension could be considered an aggregate measure of mental abilities. Home management could be an aggregate indicator of physical impairment. These findings question the value of creating more disaggregated indicators of physical and mental impairment with respect to estimating monthly costs, at least as understood by MDSS workers.

In Models B and C where the endogenous variables were hours of service employed and costs per hour, respectively, as hypothesized, the socio-economic status variable was more important. Surprisingly, the lowest hourly cost of services was incurred by non-relative providers living in the same residence as the client. Such providers also worked the longest hours. Fewer hours at higher cost were provided by relatives living with clients. Even fewer hours were supplied by relatives living separately from clients but at higher cost per hour than the previous two arrangements. Providers supplying the least hours at the highest hourly cost were non-relatives living separately from clients. This seeming indirectly proportional relationship between costs per hour and hours worked explained why monthly costs remained relatively constant. The provider related to the client was not less costly on a monthly basis than the non-related client. As hypothesized, hourly costs were higher for those clients living separately.

Variable categories that contributed significantly to the explanatory power of Models B and C were comprehension, home management and attendant services. Hours worked were higher for physical disabilities than for lack of comprehension but hourly costs of providing services to compensate for physical impairments were lower.

As expected, the medical status indicator appears to play no role in explaining monthly costs of chore services. This does not mean it plays no role in explaining costs of long-term care. Medical status was not addressed adequately in this analysis because the quality of the data was poor. The primary medical diagnosis underlying the client's impairment was not always recorded. Neither was it done systematically. This deficiency in data is a clear indication of the lack of systematic training among MDSS workers. The lack of adequately trained MDSS workers might also have affected the accuracy with which other client attributes, especially functional status, were recorded. This deficiency in turn might account for the lower explanatory power of functional status categories.

The functional status categories of mobility, dexterity and sensory perception were particularly poor predictors of cost. In this study, more comprehensive or general indicators of functional status such as level of comprehension or ability to manage the home were better categories with respect to explaining variation in costs.

From the regression results, it was apparent that services assigned to clients best explained variation in monthly costs. It was but a short intellectual leap to propose that classification of a client's functional status might better be based on those services required by the client to maintain an independent lifestyle. This approach to classifying clients more directly addresses the client's service requirements but might be subject to more abuse as discussed above.

IV. Other Conclusions and Implications for Research

This research started out with the assumption that a profile of client attributes including non-medical and medical attributes could be used to identify an appropriate package of care from among the whole spectrum of long-term care type services, medical and non-medical. This study did not address the provision of medical services which are assigned according to patient signs and symptoms and to a lesser extent, according to circumstances characteristic of the client. However, the procedure of observing the patient's signs and symptoms (or medical attributes) followed by the assignment of medical services based on those attributes is referred to in this study as the medical model of service assignment. The procedure by which MDSS purported to assign non-medical services to clients appeared to be based on a medical model. From the results arrived at in this study, it now seems that it is inappropriate to base the assignment of non-medical services on the basis of "signs and symptoms" of functional status. Future research, employing a pattern detection method such as entropy minimax should focus on the detectability of relationship between other forms of attribute profile and other specification of services. If future research also confirms the results of this study, the utility of current approaches to assigning services for long-term care clients on the basis of attribute profiles similar to that apparently used by MDSS is called into question.

One of the general aims of this dissertation was to suggest a methodology by which all long-term care clients or patients might be classified. From the foregoing discussion it is clear the client's attributes can be divided into at least 2 general categories: medical and non-medical classification of clients. The classification of the client's inability to perform necessary tasks associated with living independently appeared to be a superior method to that of classifying them according to physical and perhaps mental impairment attributes. The classification of clients according to deficiencies in service requirements rather than bodily or mental impairments or, with few exceptions, socio-economic status (e.g. location), seems to be the best approach to predicting service requirements and their costs.

Further research should be addressed to determining the best classification of clients according to their service requirements. Alternatively or concurrently the classification of services might also be clarified. One test of the efficacy of the classifications used is to examine the accuracy with which costs or hours worked can be estimated using those categories as explanatory variables in a regression analysis. From the research conducted here, it appears the more direct the relationship between client attribute classification and service rendered, the more powerful the explanatory power of the variables (categories) used. Thus, instead of basing the assignment of, say, dressing services to a client on a dexterity impairment index, it would be more appropriate to classify the client as being unable to dress himself. Dressing services in the context of the MDSS form would be subsumed under non-nursing personal services. But non-nursing personal services also covers bathing services, toileting services and others. The question is to what extent the "needs" of the client should be aggregated. Is it better to detail client "needs" such as dressing, bathing and toileting services or aggregate them under the heading non-nursing personal services? This also is a question for further research.

The first consideration in assigning services (either medical or non-medical) to any client is the medical status of the client. The medical status of the client is observed first. If the diagnosis indicates instability necessitating professional medical care, then the decision as to <u>where</u> it is best provided is considered. The diagnosis may necessitate hospitalization in which case non-medical attributes are of little importance. The medical condition of the client is always of great importance with respect to services required; therefore, it and implicitly medical services "needs," are considered first.

At the other end of the spectrum of care, after it has been observed the client is medically stable, the non-medical attributes assume importance in assigning services. A program like that of the MDSS studied here can meet such needs. It is for clients who fall between these extreme cases of hospitalization and employment of minimal non-medical service assignments that an accurate attribute profile becomes critical. A specific attribute profile may make a client eligible for a nursing home or eligible for chore type services and home health visits. As detailed in the review of literature, Chapter 3, it is claimed that 20-50% of patients in nursing homes are inappropriately placed. The implication is that many of these could have been cared for appropriately in a home setting if they had sufficient services to support them. Perhaps, also, there is another explanation for so called inappropriate institutionalization of people in nursing homes.

Perhaps considerable confusion arises from the fact that with age, the number of unstable episodes of illness increases and alternates with stable conditions. When this fluctuation between stable and unstable medical condition is frequent, the medical model takes precedence and professional health staff tend to take over. Since these staff usually work in institutions, the predisposition to institutionalize patients becomes logical. This may explain why, at any one time, some persons in institutions appear to be inappropriately placed. They may have been passing through a relatively shortlived stable episode of medical status when the assessments were made. Institutionalization of the medically unstable would not be inevitable if professional medical staff offered services in the homes of the chronically ill and disabled during periods of medical instability.

An alternative interpretation of the non-detectable relationship between client attributes and assigned services might be pertinent. Perhaps the specification of the profile outline and the assigned services is appropriate but the MDSS workers do not use it consistently. Implicit in this view is that if the MDSS workers consistently employed the attribute profile in the assignment of services, then patterns between attributes and services would become detectable.

Further research might focus on the actual means by which workers assigning services make their decisions. What criteria do they in fact use on which to base service assignment? This decision related research should encompass not only the workers making assignments of services but the clients themselves, their families and third party payors.

The objective functions of these groups were posited in Chapter 2. Just what specific criteria are considered by these groups and how are they weighted when decisions on assignment of services are necessary? When examining the basis on which decisions are made by the workers assigning services, it should be borne in mind that they must not only assign services to the client but maintain some form of accountability to the third party payor, in this case the state government.

V. Policy Implications

The present official MDSS approach to assigning services to clients on the basis of the current attribute profiles seems to be ineffective in that there appears to be no detectable relationship between client attributes and assigned services. The functional status variable and socio-economic status variable are less useful predictors of costs than the services themselves. It is probable that the actual utility of the functional and socio-economic status variables derives from their use for justification of service assignment *per se* rather than specific services. In this role, attribute records fulfill an accountability function and should be retained, at least in part, for that purpose. They meet the accountability concerns of the third party payor, in this case the state government.

Further emphasis might be given to specification of the client's service requirements within the MDSS Form DSS-3492.

With a view to more consistently documenting attributes and services assigned, the MDSS workers should be trained more systematically. To facilitate this, the MDSS form might be simplified. The parts of the form documenting types and levels of impairment might be removed or altered. More detail might be appropriate in the documentation of service assignment. For example, the estimated time required to perform each type of assigned service would be helpful in estimating costs and would also constitute a consistency check among MDSS service workers.

Third party payors such as the government, also seek to provide service packages for eligible clients at minimum cost. Within the context of the MDSS chore service program, the functional status categories that were significant in explaining costs were comprehension and ability to manage the home. These were broad indicators of mental status and physical impairments, respectively. The explanatory variable explaining the most variation in cost was the service variable. The most significant categories of that variable were light cleaning, meal preparation, attendant services and non-nursing personal services. That these categories were significant indicates they may be of value in establishing policies with respect to cost. For example, instead of the single monthly cost ceilings nowadopted by the MDSS for its chore service program, there might be several ceiling. Each ceiling might be associated with the assignment of one or more, but not all, of the above service categories. Clients for whom meal preparation is assigned might be subject to a higher cost ceiling than those for whom transportation or light cleaning is assigned. The higher cost ceilings for more labor intensive services such as meal preparation and non-nursing personal services might induce persons to remain in independent settings longer and thus delay possible entry into nursing homes. From the perspective of the third party payor (MDSS in this case) the appropriate cost ceiling for non-nursing personal services might be a percentage of the cost to the state of an average month of care in a nursing home. Thus, if the average cost of nursing care for Medicaid patients is $30/day and the state pays $16 of that, then a maximum ceiling cost for non-nursing personal services might be 16x30=$480 per month. Assuming medical costs constitute part of nursing home costs, the actual ceiling on a mix of services that include non-nursing personal services, may appropriately be reduced to between $300 and $450 per month. A lower ceiling might apply to a mix of services that does not include non-nursing personal services. The reason for suggesting the nursing home cost-related ceiling for non-nursing personal services was because these services most nearly approximate those provided in a nursing home, and therefore, might substitute for them. If the state offered this option for people who are clearly eligible for state subsidized nursing home services, and if the client opted not to enter a nursing home, then there would be a clear saving of federal dollars and a marginal saving of state resources.

Other lower ceilings than the current $270 might be imposed on service mixes which did not include non-nursing personal services. The overall cost of the Chore Service Program might rise or remain the same. As long as the rise was more than offset by demonstratable savings in nursing home costs, the application of multiple ceilings that depended on service mixes would be appropriate. On the other hand, the monetary savings to the state might justifiably be negligible if it can be demonstrated that the benefits associated with home care are greater than those derived from nursing homes.

ELIGIBILITY AND NECESSITY FOR ADULT PAID SOCIAL SERVICES
State of Michigan
Department of Social Services

INSTRUCTIONS:
1. *File original in service case file.*
2. Submit copy to S.O. attached to white DSS-13C

CLIENT'S NAME	BIRTHDATE	DSS CASE NUMBER	SOCIAL SECURITY NUMBER

I. ELIGIBILITY

A. STATUS OF ADULT (Check applicable boxes)

1 ☐ SSI Recipient

2 ☐ Potential SSI Recipient — Does recipient meet SSI Resources Criteria: Yes 1 ☐ No 2 ☐

If potential SSI recipient check one box.

1 ☐ Aged 2 ☐ Disabled 3 ☐ Blind

B. INCOME for all clients.

SSI ________

RSDI + ________

Other Pensions + ________

Earned Net + ________

Other + ________

TOTAL =

(INSERT IN COMPUTATION BOX ON LINE #1) → ________

C. COMPUTATION (for potential recipient's only)

TOTAL INCOME FORWARDED FROM INCOME BOX = ________

M.A. PATIENT PAY AMOUNT (Subtract) − ________

AVAILABLE INCOME (Subtotal) = ________

INCOME BASE (Subtract) − ________

TOTAL CLIENT PAY AMOUNT = ________

D. SERVICE COSTS

COST OF CHORE SERVICES (enter total from back) = ________

TOTAL CLIENT PAY AMOUNT (Enter from computation) (Subtract) − ________

TOTAL AUTHORIZED STATE PAYMENT = ________

II. NECESSITY

A. LIVING ARRANGEMENTS (Check applicable boxes)

1 ☐ Householder / Independent Living

☐ Owned ☐ Rented NUMBER OF ROOMS

SPECIAL CIRCUMSTANCES

2 ☐ Non-Householder / Living in household of another.

SPECIFY ARRANGEMENTS

B. HEALTH (if pertinent)

DOCTOR OR CLINIC'S NAME	ADDRESS NO. STREET CITY	PHONE NO.

MEDICAL PSYCHIATRIC DIAGNOSIS

Does client need health-related supervision Yes 1 ☐ No 2 ☐

COMMENTS

C. CLIENT LIMITATIONS
(The client has limitations as follows in the abilities:)
(Check Applicable Boxes)

DESCRIPTION	NONE	SOME	MUCH	DESCRIPTION	NONE	SOME	MUCH
Walk	1 ☐	2 ☐	3 ☐	Manage Finances	22 ☐	23 ☐	24 ☐
Climb Stairs	4 ☐	5 ☐	6 ☐	Remember Instructions	25 ☐	26 ☐	27 ☐
Grasp	7 ☐	8 ☐	9 ☐	Understand Instructions	28 ☐	29 ☐	30 ☐
Lift	10 ☐	11 ☐	12 ☐	Do Light Cleaning	31 ☐	32 ☐	33 ☐
Bend	13 ☐	14 ☐	15 ☐	Do Heavy Cleaning	34 ☐	35 ☐	36 ☐
See	16 ☐	17 ☐	18 ☐	Understands English	37 ☐	38 ☐	39 ☐
Hear	19 ☐	20 ☐	21 ☐	Other (Specify Below)	40 ☐	41 ☐	42 ☐

OTHER LIMITATIONS AND COMMENTS

II. NECESSITY (CONTINUED)

D. SERVICES NEEDED (Check applicable box)

1. Light Cleaning 1 ☐	6. Shopping and/or other errands 6 ☐	11. Yardwork (lawn, snow removal, etc.) 11 ☐
2. Heavy Cleaning 2 ☐	7. Guide Dog 7 ☐	12. Non-nursing personal care 12 ☐
3. Minor household repairs 3 ☐	8. Financial Management 8 ☐	13. Interpreter 13 ☐
4. Meal Preparation 4 ☐	9. Transportation 9 ☐	14. Other (Specify below) 14 ☐
5. Laundry 5 ☐	10. Attendant for transportation 10 ☐	

E. PROVIDER'S INFORMATION

DESCRIPTION	FIRST PROVIDER (21)	SECOND PROVIDER (22)	THIRD PROVIDER (23)	FOURTH PROVIDER
Name				
Relationship to adult	1-☐ Parent 2-☐ Daughter/Son 3-☐ Other Relative 4-☐ Non-Relative	1-☐ Parent 2-☐ Daughter/Son 3-☐ Other Relative 4-☐ Non-Relative	1-☐ Parent 2-☐ Daughter/Son 3-☐ Other Relative 4-☐ Non-Relative	1-☐ Parent 2-☐ Daughter/Son 3-☐ Other Rela[tive] 4-☐ Non-Relat[ive]
Age (approximate)				
Address				
Phone				
Service Provided (Insert codes as in D above)				
Hours per day				
Days per week				
Expected duration				
Rate of pay				
Monthly rate of pay				
Effective date				

F. APPROVED FOR: (Check applicable box)

1. Chore Services (Indiv. providers) 1 ☐	4. Family helper services 4 ☐	7. Guide dog 7 ☐
2. Chore Services (Agency provider) 2 ☐	5. Attendant services 5 ☐	
3. Homemaker services 3 ☐	6. Bilingual interpreter 6 ☐	

G. DENIED

Reason:

H. REVIEW NEEDED

I. CERTIFICATION

[I] certify that, in conjunction with the adult, I have determined that:

a) the adult possesses the above limitations and service needs.

b) each service provider has agreed to perform the required services.

c) each service provider is able to perform is performing the required services.

[illegible]

APPENDIX A

APPENDIX B

SAMPLE COUNTY RATE

CHORE SERVICE PAYMENT SCHEDULE

Type of Service	Rate
1. Light housekeeping - meal preparation, shopping, laundry, etc. - non-daily	$2.25/hr.
2. Heavy one-time house cleaning	2.75/hr.
3. Yard work - cleaning, snow shoveling, etc.	2.00/hr.
4. Daily household activities - include shopping, transportation, etc. at least 6 hrs. a day - 5 days a week	$200/mo.
5.* Live-in 24 hour personal care (ambulatory client) including all household activities	240/mo.
6.* Live-in 24 hour bedfast or chair to bed - total care including all household activities	240/mo.
7. A. Adult with parents or relative (1st degree), attending workshop or day activities - personal care needed (help dressing, eating, bathing, etc.)	150/mo.
B. Non-personal care, but cannot be alone extended periods of time	100/mo.
8. A. Adult with parents or relatives (1st degree) not attending workshop or day activities - personal care required	200/mo.
B. Non-personal care - cannot be left alone extended periods of time	150/mo.
9. Adults with parents or relatives severely retarded or handicapped needing total care	270/mo.

NOTE: These are only a suggested rate schedule and could be adjusted downward when possible.

*Own arrangements to be made with client for time off.

PAID SOCIAL SERVICES

The following information complements Services Manual Appendix E in the area of establishing county maximums under the Paid Social Services Program. Many considerations were taken into account in arriving at county maximums that are in line with the "going rate" for the cost of such services in Mason County.

Chore Services have been broken down into specific task functions. A particular client may be eligible for one or any combination of the specified task functions. Eligibility for specified Home Services shall be based on an evaluative interview between the social services worker and the client. Eligibility for specified Home Services, is based on the requirements described in Sections A and C of SM Appendix E. The narrative in the services record is to indicate the services provided, and must confirm the exact needs and duration of the specific services to be purchased.

	Task Function	County Maximum	Criteria
A.	MEAL PREPARATION	$5.00 per day	2½-3 hrs. daily for 2 meals - client has tools
B.	HOUSE CLEANING	$10.00 per week	3 hrs. incl. transportation client has tools
C.	LAUNDRY (in-home)	$3.00 per day	as needed: every day, twice weekly, once weekly, etc.
D.	SHOPPING FOR CLIENT	$4.00 per week	2 hrs. incl.transportation
E.	LAWN MAINTENANCE	$4.00 per week	2-3 hrs. - provider has tools
F.	GARBAGE DISPOSAL	$5.00 per month	provider transports
G.	LIVE-IN ATTENDANTS		
	a. 5 nights	$105.00 per month	______
	b. 6 nights	$130.00 per month	______
	c. 7 nights	$160.00 per month	______
	d. relative	$54.00 per month	personal needs plus $1.00 hh needs
H.	SNOW REMOVAL		
	a. Manual or hand plow	$2.00 per hour	Single payment-provider has tools
	b. Automotive	$3-6 per hour	Single payment-provider has tools

I. NON-NURSING PERSONAL CARE $2.00 per day _______

Any combination of these services may not exceed $270.00 per month.

The services worker should take care in determining the client's present needs and in projecting anticipated needs, such as a need for snow removal in the winter, so that the total maximum of $270.00 is not reached before adding the cost of snow removal.

11-17-73* *Not updated since then. Received by reviewer on 12-30-75

BIBLIOGRAPHY

BIBLIOGRAPHY

Anderson, N. N., A Planning Study of Services to Non-Institutionalized Older Persons in Minnesota, The Governor's Citizens Council on Aging, State of Minnesota, 1974.

Barney, J. L., Patients in Michigan's Nursing Homes: Who Are They? How Are They? Why Are They There? Institute of Gerontology, University of Michigan, Wayne State University, November 1973.

Barney, J. L., "The Well Being Report: A Look at Supportive Services for the Elderly." Institute of Gerontology, University of Michigan, Wayne State University, 1975.

Berger, E. J., "Study and Analysis of Utilization and Cost Data Concerning the Provision of Home Health Service and Extended Care Services." St. Louis Labor Health Institute, St. Louis, Missouri, February 23, 1970.

Berkowitz, M. and Johnson, W. G., "Towards the Economics of Disability: The Magnitude and Structure of Transfer and Medical Costs." J. Human Res., Summer 1970, 5(3), pp. 271-97.

Blenkner, M. et al, "Protective Services for Older People, Findings from the Benjamin Rose Inst. Study." in Social Casework, October 1971, Vol. 52, No. 8.

Boulding and , Redistribution to the Rich and the Poor, Wadsworth Publishing, Belmont, California, 1972.

Burton, R. M., et al, "Nursing Home Cost and Care: An Investigation of Alternatives." GSBA Paper No. 12 (Revised August 5, 1974), Center for the Study of Aging and Human Development, Duke University Medical Center, Durham, NC.

Busse, E. W. and Pfeiffer, E., (eds.), Behavior and Adaptation in Late Life, Boston, Little Brown and Co., 1969.

California Association of Home Health Agencies, Cost Survey of California Certified Home Health Agencies and the Related Effects of the California State Department of Health Care Services. Schedules of Maximum Allowances, March 1973.

Caro, F. G., Organizing and Financing Personal Care Services: An Alternative to Institutionalization for the Disabled. Levinson Gerontological Policy Institute, Brandeis University, June 1972.

Christensen, R. A., "Entropy minimax Method of Pattern Discovery and Probability Determination," Arthur D. Little, Inc., Cambridge, Mass., 7 March 1972.

Congressional Budget Office: "Long-Term Care for the Elderly and Disabled," Budget Issue Paper, February 1977.

Emling, Diane Carpenter, Adult Chore Services. A Profile of In-Home Assistance. Studies in Welfare Policy No. 10, Michigan Department of Social Services, November 1976.

Lave, J. R., Lave, B. L. and Silverman, L., "Hospital Cost Estimation Controlling for Case Mix," Applied Economics 1972, 4, pp. 165-180.

Evans, R., "Behavioral Cost Functions for Hospitals," Canadian Journal of Economics, Vol. IV, No. 2, May 1972a.

Fitzgerald, J., "Home Health Services for the Aged in South Eastern Michigan," The Comprehensive Health Planning Council of Southeastern Michigan, November 1972.

Florida Department of Health and Rehabilitation Services, "Community Care for the Elderly," 1971.

Genesee Region Health Planning Council, Survey of Need for Inpatient Beds and Related Home Health Care Services, Monroe County, New York, 1969-70.

Gift, D. A., Comparison of Classifying Power of New Neighbor Clusterer, Entropy-Minimax, Pattern Detector and Multiple Regression. Working Memorandum, Thyroid Project File, Department of Radiology, Michigan State University, 1978a.

Gift, D. A., "An Introduction to Entropy-Minimax Pattern Detection and Its Use in the Determination of Diagnostic Test Efficiency." Unpublished paper, Department of Radiology, Michigan State University, 1978b.

Gruenberg, L. and Thoener, J., A Comparative Analysis of Long-Term Care Patient Assessment Techniques Employed by the Department of Public Health. Working Paper, Division of Medical Care, Department of Public Health, Massachusetts, 1972.

Hundert, M. J., "The Nursing Home Simulation Model: A Policy Tool for Long-Term Care," Medicos Systems Corporation and Illinois Department of Public Health, December 1974.

Hurtado, A., Greenlick, M., McCabe, M., and Saward, E., "The Utilization and Cost of Home Care and Extended Care Facility in a Comprehensive, Prepaid Group Practice Program," Medical Care, Jan. Feb. 1972, pp. 8-16.

Johnston, B. C., A Comparison of Three Pattern Detection Methods on Binary Data. Working Memorandum, Thyroid Project File, Department of Radiology, Michigan State University, 1976.

Kastenbaum, R. and Candy, S., Reviewed in Institutionalized Aged: The Four Percent Fallacy, Geriatric Focus, March 1973.

Katz, S., and Akpom, C. A. "A Measure of Primary Socio-Biological Functions," International Journal of Health Services, 1976, 6, pp. 493-508.

Katz, S. Downs, T. C., Cash, H. R., and Grotz, R. C. "Progress in Development in the Index of ADL," Gerontologist, 1970, 10, pp. 20-30.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., and Jaffe, M. W., "The Index of ADL: A Standardized Measure of Biological and Psycho-Social Functions, JAMA 134: 914 September 21, 1963.

Kistin, H. and Morris, R., Alternatives to Institutional Care for the Elderly and Disabled, The Gerontologist, Vol. 12, No. 2, Part I, Summer 1972.

Kmenta, J., Elements of Econometrics, Macmillan Co., N.Y., N.Y. 1973.

Knox, K. C., et al, Criteria and Guidelines for Evaluation of Aged and Chronically Ill Persons for Maintenance at Home or for Placement in an Appropriate Long-Term Care Institution, The Evaluation Placement Unit, Rochester, N.Y. 1973.

Lee, M. L. and Wallace, R. L., Problems in Estimating Multi-produce Cost Functions: An Application to Hospitals, Western Econ. J. Sept. 1973, 49(4), pp. 344-56.

Medicus Systems Corporation. Protocol reference manual. Section 222, Demonstration and evaluation of long-term care alternatives. Unpublished manuscript, Chicago, ILL. Developed for National Center for Health Service Research, March 6, 1975.

The Michigan Comprehensive Plan on Aging, Lansing, Michigan, Policy and Program Analysis Division, Office of Services to the Aging, August, 1975.

Morris, R. and Harris, E., "Home Health Services in Massachusetts 1971: Their Role in Case of the Long-Term Sick," American Journal of Public Health, Vol. 62, No. 8, August 1972.

National Center for Health Statistics, "Limitation of Activity Due to Chronic Conditions, U.S., 1969 and 1970." Vital Health Statistics, Series 10, Publication No. (HSM) 73-1506, April 1973.

Nielson, Blenkner and Bloom, Home Aid Service and the Aged: A Controlled Study. A Final Report, Benjamin Rose Institute Research Department, Cleveland, Ohio, August 1970.

Pollack, W., Costs of Alternative Care Settings for the Elderly. Working Paper 963-11 (Draft), Urban Institute, Washington, D.C., March 12, 1973.

Robinson, N., Shinn, E., Adam, E., and Moore, E., "Costs of Homemaker-Home Health Aide and Alternative Forms of Service, A Survey of the Literature," National Council for Homemaker-Home Health Aide Services, Inc., N.Y., N.Y. 10003, 1974.

Sager, A., Estimating the Costs of Diverting Patients from Nursing Homes to Home Care," Levinson Policy Institute, Brandeis University, Waltham, Mass., 1977.

Shanas, E., et al, Old People in Three Industrial Societies, New York: Atherton Press, 1968.

Shanas, E., "Index of Functional Impairment Tested and Used in a Study of Older People," Journal of Gerontology, 1971, 26, pp. 37-40.

Schonbein, W. R., "Evaluation Studies in Nuclear Medicine Research," Unpublished paper, Department of Radiology, College of Human Medicine, Michigan State University, 1976.

Schonbein, W. R., "An Experimental Approach to the Determination of the Efficacy of Diagnostic Procedures," Unpublished paper, Department of Radiology, Michigan State University, 1978.

Skinner, D. E. and Yett, D. E., "Estimation of Cost Functions for Health Services: The Nursing Home Case," Paper presented at the 40th Annual Conference of Southern Economic Association, Atlanta, November 12, 1970.

Stroud, M. W. et al, A Summary Report Concerning Patients at Bryn Mawr Hospital Rehabilitation Center, Unpublished manuscript, Department of Community Health Science, Michigan State University, East Lansing, MI, Jan. 1978.

Theil, H., Principles of Econometrics, a Wiley/Hamilton publication, John Wiley and Sons, Inc., N.Y., N.Y., 1971.

Touche, Ross and Co., "Study of Cost Differential Between Skilled Nursing Facilities and Intermediate Care Facilities," USDHEW, Social Rehabilization Service, Medical Services Administration, Washington, D.C., 1974.

Townsend, P., "Measuring Incapacity for Self-Care," in Williams, R. H., Tibbits, C., and Donahue, W., (editors), "Process of Aging," Atherton Press, 1963.

Tracy, M. B., "Constant Attendance Allowances for Non-Work-Related Disability, Soc. Sec. Bull., Vol. 37, No. 1, November 1974, pp. 32-47.

U.S. Department of Commerce, Bureaus of Census, Population Division, 1970 Census of Population, Detailed Characteristics-Michigan Final Report PC (1)-D 24, Washington, D.C., U.S. Government Printing Office, November 1972, Tables 153 and 154.

U.S.D.H.E.W., "The Older Population Revisited," First results of 1970 Census, Administration of Aging, Office of Human Development, USDHEW.

Wager, R., "Care of the Elderly: An Exercise in Cost Benefit Analysis," Institute of Municipal Treasurers and Accountants,1 Buckingham Place, London, SWI, March 1972.

White, N. M. and Tenbrunsel, T. W., "Alternatives to Acute Care," Report of the Capitol Area Comprehensive Health Planning Association, Lansing, Michigan, October 1971.

Wilder, C. S., "Limitation of Activity and Mobility Due to Chronic Condition," U.S. 1972, Vital and Health Statistics Series 10, DHEW Publ (HRA) 75-1523, HE 20.6209:10/96.

Williams, T. F., Hill, J. G., Fairbank, M. E., and Know, K. G., Appropriate Placement of the Chronically Ill and Aged, JAMA Vol. 226, No. 11, December 10, 1973.

Worthington, N. L., National Health Expenditures, 1927-74, Soc. Sec. Bull., Vol. 38 #12, February 1975.