University of Cape Town

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

The Demand for Health Care Services in Nigeria

- A Nested Logit Model

Hyacinth Ementa Ichoku

University of Cape Town

A research report presented to the School of Economics/ Health Economics Unit – University of Cape Town, in partial fulfilment of the requirements for Master of Social Science degree in Health Economics.

July 2000

To

Tony & Imma

University of Cape Town

ACKNOWLEDGEMENTS

I wish to express my thanks to all those who assisted me in anyway in the course of this research. My deep appreciation goes to the Swedish International Development Agency (SIDA) who provided me the scholarship for this program. I thank the staff of the School of Economics and Health Economics Unit of the University of Cape Town, the staff of the Dept of Economics – University of Nigeria, Nsukka - for their support. I am deeply indebted to the staff and students of St Francis Xavier Seminary, Athlone, for their love and care.

I appreciate the kind assistance of the following individuals for making their time and resources available to me:

Professor Murray Leibbrandt, School of Economics, UCT,

who supervised this research and guided me at the most difficult periods

Charlotte Muheki, Health Economics Unit, UCT,

who co-supervised this research and made very useful suggestions.

Eyob Zere, Health Economics Unit, UCT.

Dr Reginald Chima, Dept of Economics, University of Nigeria, Nsukka.

Fr Michael van Heerden, St Francis Xavier Seminary –Athlone, Cape Town.

Okey Esimone, Faculty of Pharmacy – University of Nigeria - Nsukka.

To the Almighty, Invisible and Eternal God be all glory forever.

University of Cape Town

TERMS OF REFERENCE

This research report was prepared for the School of Economics and Health Economics Unit of the University of Cape Town in partial fulfilment of the Master of Arts degree in Economics.

The objective of the research is to investigate the factors that influence health care demand in Nigeria.

University of Cape Town

I certify that this research report has been prepared by me

Signed ....

H.E. Ichoku

Abstract

The main aim of this study is to understand better the factors that influence the health care demand decisions of Nigerian households. The achievement of this objective involves the estimation of the parameters of the demand for health care services in order to understand the nature of health care choices that Nigerian households make under the present depressed economy. These demand parameter estimates are considered valuable inputs into health care policy. Yet, to date, there is neither sufficient information on the factors that shape households' utilization of health care services nor is there sufficient information on the relative importance of health care alternatives available to them. This is the knowledge gap this study helps to fill. Due to the dearth of data on the subject and the costs of collecting such data for the whole country, this study has been limited to a particular geographical area – Nsukka local government.

The study is organized in seven chapters. Chapter 1 provides a brief characterization of the health situation of the Nigerian population and specifies the research objectives. Chapter 2 reviews the existing international literature on the demand for health care services focusing mainly on empirical studies done in developing countries. One of the major problems that arise out of the literature is the problem of specifying the appropriate model for estimating the demand for health care services. This problem is taken up in chapter three.

Using the Grossman-Wagstaff model, Chapter 3 shows that the demand for health care is a utility-maximizing behavior. It then reviews the models used in previous studies. It shows that while the Luce family of models as a whole conforms to the axioms of random utility maximization (RUM), the multinomial logit model suffers from the independence of irrelevant alternative (IIA) assumption. It then specifies the nested logit model of the discrete choice variant as one that helps us to overcome the limitations of the other models.

University of Cape Town

Chapters 4 and 5 are concerned with the data. Chapter 4 describes the method used in generating the data while chapter 5 provides the descriptive statistics of the data collected. These descriptive statistics are intended to serve as a preview of the output from the estimation of the model.

Chapter 6 analyzes the results of estimating a three-level nested model of the demand for health care. It shows that distance to a facility, household size, waiting time and treatment cost all have negative relationship with facility utilization while income, age, level of education and severity vary positively with utilization. The study suggests that severity is the most important variable in the choice of facility. It shows that many households usually sought health care from the patent medicine dealers within the first four days of illness but moved to other alternatives - public and private health facilities - if illness persists.

Chapter 7 examines the key policy issues arising from the results. Essentially, the estimated parameters indicate different forms of inequity in access to health care services in the area. The clear dominance of the private sector in the supply of health care services in the area and the well-acknowledged problem of information asymmetry in the medical market requires a policy that effectively regulates the sector. While the policy of using the public provision of health care services to check the effect of a private-sector dominated health care market is seen to be effective, the long run policy should aim at removing the barriers to access. The relative importance of patent medicine dealers and the traditional health care suppliers suggests a policy that integrates these groups of suppliers into the mainstream of the health care system.

University of Cape Town

Contents

University of Cape Town

University of Cape Town

List of Illustrations

Tables

Figures

University of Cape Town

List of Abbreviations used in this Report.

2SLS	*Two Stage Least Squares*
AIC	*Akaike Information Criterion*
ASCs	Alternative Specific Constants
NAVYRSED	Average Years of Education
BFGS	Broyder, Fletcher, Goldfarb and Shanno
BHHH	Berndt, Hall, Hall and Houseman
CIHI	Center for International Health Information
CLM	Conditional Logit Model
CBR	Crude Birth Rate
CDR	Crude Death Rate
DFP	Davidson, Fletcher and Powell
exp	exponential
FIML	Full Information Maximum Likelihood
GLS	Generalized Least Squares
H/Care	Home Care
HH_SIZE	Household Size
IIA	Independence of Irrelevant Alternatives
ILS	Indirect Least Squares
IMR	Infant Mortality Rate
IQR	Inter-quartile Range
LGA	Local Government Area
LIML	Limited Information Maximum Likelihood
Loc	Location
LR	Likelihood Ratio
MLE	Maximum Likelihood Estimates
MNLM	Multinomial Logit Model
Mo H	Ministry of Health
N/Care	No Care
NMNLM	Nested Multinomial Logit Model
NISER	Nigerian Institute of Social and Economic Research
OLS	Ordinary Least Squares
Pcfdx	Per capita food expenditure
P/Med	Pharmacy/Patent Medicine
PrvtH	Private Hospitals/Clinics
PubH	Public Hospitals/Clinics
RUM	Random Utility Maximization
S_age	Age of a sick person
SteDes	Steepest Descent
T/Med	Traditional Medicine
Tmtcost	Treatment Cost
U5MR	Under 5 Mortality Rate
W	Wald Statistic
W_time	Waiting time

University of Cape Town

Chapter 1

Background to the Study

1.1 Introduction

Until the decade of the 1980s, a primary preoccupation of governments and health policy makers was the supply of health care services to their populations. Public policies were directed towards health care provision that catered for all the segments of the society. This approach tallied with the prevailing ideology of the welfare state and the assumption that good health was the most important goal in the choices individuals make. The economic crises of the '80s especially among the developing countries brought to the fore the severity of resource constraints and the limits to which national resources can support unrestrained health resource consumption. With the active support of the World Bank and World Health Organization (WHO) many developing countries started a process of health sector reform with a view to diversifying their health sector resource base and towards making the health sector more market oriented. Increasingly more attention was directed towards the demand side of the health care market[1]. The purpose of this new focus was to provide a better understanding of the principles and factors that influence people's health care consumption behavior. The analyses of such factors would ultimately provide policy makers with the basic tools to reform and manage the health sector more effectively.

Among the several studies that have been undertaken to foster this new policy desire were those of Akin et. al (1981), (1985), (1986), (1994); Heller (1981), Mwabu (1984), (1986); Gertler, Locay and Sanderson, (1987); Dor et. al., (1987). Other major contributions to international literature on demand for health care literature include the works of Ellis and Mwabu, (1991), Schwartz et al, (1993); Ellis et al. (1994); Dow (1996). In spite of these outstanding contributions to our understanding of the health care demand behavior of households, it should be noted that econometric

[1] Demand for health care: this is the expression of desire for health care arising from the more fundamental desire for the basic commodity of good health. In this context, I will use demand for health care interchangeably with the demand for medical care since health care is, in the main, obtained through some form of medical attention.

University of Cape Town

estimation results are liable to a wide variety of data generation processes, model specification and econometric assumptions underlying the estimations process[2] Leibbrandt and Woolard (1999). This research attempts to apply the most recent refinements of the discrete choice methodology to the estimation of health care demand data generated from a demographically homogenous population of a local government area in Nigeria.

1.2 Some Nigerian Economic and Health Indicators

The decade of the 90s was a very difficult period for many Nigerian households. The ripple effects of the Structural Adjustment Program (SAP) that was introduced in 1986 combined with political instability in the country to send devastating economic shocks across the country Egwu (1998). There was rapid decline in most major macroeconomic indicators. The economy that had been growing at the rate of 7.5% in the '70s recorded on the average zero growth for the greater part of the decade Watts and Lubeck (1984). Per capita income fell from $860 in the early 1980s to below $300 in the early 1990s Pearce and Falola (1994). The national currency, the Naira, was devalued by more than 5000%. External debt grew to over $30 billion. Inflation sometimes stayed above 50% and industrial capacity utilization was as low as 30% in some years Central Bank of Nigeria/ NISER (1995), Ogunbekun (1991). Political instability brought insecurity and social welfare to intolerable levels. Social infrastructure and the physical environment deteriorated.

University of Cape Town

In the face of continued severe economic and social crises, many households could no longer afford the basic necessities of life for their members. The ability of households to cope with adverse economic conditions has been strained. Difficult trade-offs continue to be made in attempt to keep households afloat. Nutritional intake and other health-enhancing inputs into the household health production function such as leisure and sports have either been reduced or eliminated altogether from the household schedule. These social upheavals have led to breakdowns in the health of individuals, households and communities. The advances made against some of the communicable

[2] Household: This is defined according to context but it used here to mean all the people who live under the same roof at least 15 days in a month for a period of one year and share food from a common source. (See South African Project for Statistics on Living Standards and Development 1993)

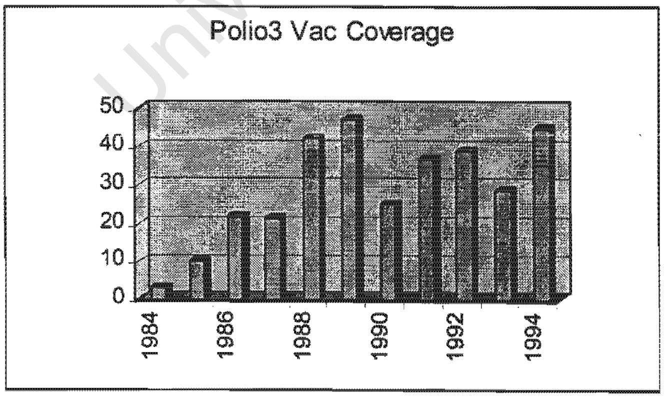
and preventable diseases during the period of economic growth of the '70s and early '80s have been eroded Pearce and Falola (1994). On account of economic barriers many households can hardly afford medical care. Reporting illnesses are delayed until they become severe since cost of medical care has to be weighed against other pressing household needs such as food and education.

Under conditions such as these, children are usually the most vulnerable group given their physical weakness Vogel (1988). The struggle to provide for the household especially among the poor means reallocation of time and resources. Parents are often away from home for the greater part of the day. Adults have less time to care for children. These children become victims of hunger and disease. Their physical and mental development is at risk of impairment and their survival and future have become imperiled.

The following charts and graphs illustrate some of the health indicators of Nigeria. While emphasis is laid on child survival indicators, the bad health situation affects almost all segments of the Nigerian population.

[The data source for the following charts and graphs is *Center for International Health Information*]

Fig 1 **Polio3 Vaccination Coverage in Nigeria 1984-94**



Source of data: CIHI

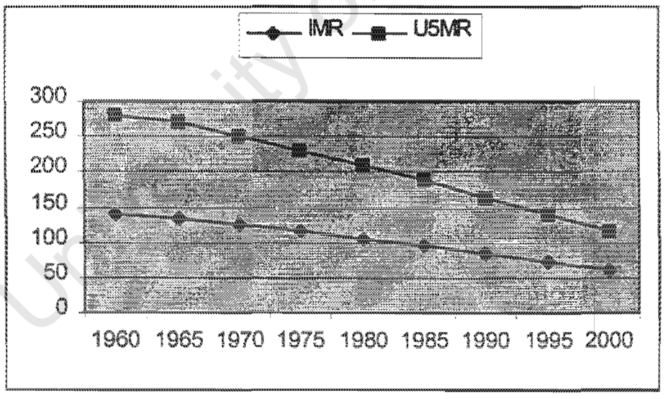
This chart reflects the poor implementation of the polio vaccination program. This program was supposed to immunize all children below 12 months of age against polio. As can be read from the chart, the program started on a low key with only 3%

University of Cape Town

of children covered in 1984. The highest coverage was in 1989 when 47% of children were immunized. Subsequently, owing to political instability and policy inconsistencies, the program started to decline. It, however, picked up again in 1994, two year after a major political crisis in 1992 that brought almost every social development to a halt.

Figure 2 below shows the decline in infant mortality rate (IMR) and under 5 mortality rate (U5MR) beginning from 1960. Although there has been a steady decline in both rates, it can be seen that the U5MR has declined a little faster than the IMR. As at 1995 the rate was about 73 deaths per 1000 infants under age 1. U5MR was about 120 deaths per 1000 children under age 5. With IMR of 73, Nigeria may have performed well relative to the rest of Sub-Saharan Africa or other Low Income Countries where IMR in 1995 were 97 and 98 respectively. But in relation to the rest of the developing countries with IMR of 62, the country still lags behind.

Fig 2 **Infant Mortality Rate (IMR) and Under 5 Mortality Rate (U5MR) in Nigeria 1960-2000**



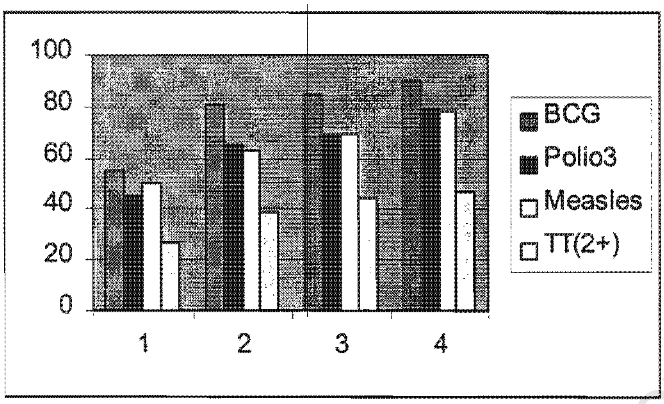
Source of data: CIHI

The country's performance on the Child Survival Indicators is compared with the performance of countries in Sub-Sahara Africa, Low Income Countries, and Developing Countries. The relative performance is shown in figure 3 below

University of Cape Town

Fig 3

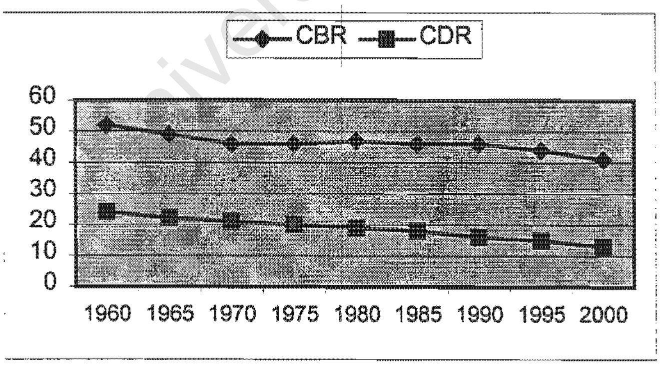
Relative Performance in Child Survival Indicators: Nigeria (1), Sub-Saharan Africa (2) Low Income Countries (3), All Developing Countries (4)



Source of data: CIHI

The emphasis has been on the child survival indicators because this group constitutes the most vulnerable group in any society. As such their well-being can be said to be an important indicator of the health of any society. The adult health indicators are not radically different from those of children as next figure shows

Fig 4 **Crude Birth Rate and Crude Death Rate in Nigeria 1960-2000**



Source of data: CIHI

As figure 4 shows the decline in both the crude birth and crude death rates have been very slow. This may reflect the slow advances made against the major prevailing diseases such as malaria. It also reflects the slow economic progress in the country within the period.

University of Cape Town

The introduction of cost recovery schemes in the health sector such as user charges even in the primary health care services means that many households may not afford basic health care needs. Even when such health care provisions are free, the time cost and distance may become barriers to access.

In summary, the health condition of the average Nigerian household in the decade of the 1990s has been precarious. Economic and political crises have frustrated many households. Under these circumstances, it is necessary to analyze how the households make decisions about the health of members and how trad-offs are made between health and other consumption goods and services.

The general characterization of the Nigerian health situation above is reflected in the health situation in Nsukka local government area of Enugu State. It is on this account that this area is being used as a study site for the general health care situation in Nigeria. A more detailed description of this area will be given in chapter 4 where the survey site is described.

1.3 Statement of the Problem

The brief characterization of the worrisome situation of the general health condition of the Nigerian population helps one to appreciate the magnitude of the task before the Nigerian health policy makers. Under severe shortage of resources, the provision of health care for the population has become a herculean task. A lot of attention has been paid to policies that improved the supply side of health care services such as building of new clinics and hospitals, deregulation of medical practices, provision of more drugs to hospitals. But not enough attention has been paid to the demand side of medical care services. For example, scant attention has been paid to factors such as the behavior of the consumer in response to changes in the price of health care goods and services. The vital parameters that underlie the demand for medical services have been often assumed without explicit attempt to measure them. But these parameters are the pillars upon which effective policies for the provision of health care services ought to be based. Policies aimed at improving the health situation of the Nigerian should, therefore, be made with a clear understanding of how the Nigerian households make decisions about their health care consumption.

University of Cape Town

More specifically, from the brief characterization of the socio-economic background of Nigeria especially in the period of 1990s one may raise the following research questions:

- How do Nigerian households make decisions about resource allocation in the face of competing needs and increased opportunity costs of such resources?
- What are the major factors that influence households' utilization of health care facilities?
- When a household reports illness, what are the factors that influence the type of health care provider from whom they seek treatment?
- What is the sequence of choices in this decision process and how can these be specified econometrically?

1.4 Objectives of the Study

The main objective of this study is to assist policy making by providing empirical estimates of some demand parameters.

More specifically, the research attempts to:

(a) Identify and estimate the most important factors that influence the demand for health care services in Nigeria

(b) Find out the relative importance of the different health care providers in the health care market in the study location.

(c) Analyze the health care-seeking behavior of Nigerian households and suggest ways that public policies may assist in improving their health care decisions.

1.5 Justification and Significance of the Study

A Nigerian health policy maker should be worried about how to effectively reduce pain and grief in the households arising from high incidence of disease and death. Policy makers and health care planners need information on the important barriers of access to medical care experienced by households. For example, policy makers may want to know the extent to which financial costs and distance to health facilities are barriers to access to medical care. Such information is, no doubt, of vital importance for health policy-making and health care investment planning. It is important, for

University of Cape Town

example, for policy makers to know the extent to which medical subsidies may be made to households through tax allowances or medical aid.

A better understanding of the health care consumption of the people would help the providers of health care services to make optimal supply decisions. For example, if a policy maker knows that the members of a given community are ready to go an extra distance to a hospital in the neighboring community to obtain quality treatment, then it might be more efficient to aim at centralizing rather than decentralizing the health care services.

For medical care providers, the study will highlight those factors that influence demand for health care and what factors lead households to prefer one type of provider against another. For example, a medical practitioner who understands that consumers in his or her area can afford to wait to obtain quality drugs may decide to invest more on drugs rather than hire more nurses. This is clearly important for investment purposes especially among private providers of care.

1.6 Scope and Limitations of the Study

Geographically, this study is limited to Nsukka local government area of Enugu state. In time frame, it is limited to households that made decisions about health care in the month preceding the interview. These two frames constitute essential limitations to the study. In the first place Nigeria is a vast country culturally and geographically. Nsukka cultural and political zone is only a unit among several others. This implies that the results of the study may not necessarily be generalized in all aspects to the rest of Nigeria. Secondly, cross-sectional data such as we shall be analyzing provide information about the short run response behavior to stimulus. It is less accurate in predicting inter-temporal or lung run response behavior, as would a time series data for instance. However, cross-sectional data remain very useful in predicting households' responses to changes in policy and other changes in the socioeconomic environment.

University of Cape Town

Chapter 2

Review of Relevant Literature

Recent research on demand for health care services has been inspired, in the main, by Grossman's seminal work: *On the Concept of Health Capital and the Demand for Health*. The major insight of this work is that it showed that the demand for medical care is a derived demand. It arises from the demand for a basic commodity, health. The conclusions from this work provided the platform for economic analysis of the demand for medical care. Following the publication of this work in 1972, a number of empirical studies have been conducted in the area of demand for medical care Heller (1981); Akin et al (1981), (1986); Mwabu (1984), (1986); Gertler, Locay and Sanderson (1987). More recent research in this area has drawn further impetus from new developments in health care reform Gertler and van der Gaag (1990); Ellis and Mwabu (1991); Bartholome and Vosti (1995); Dow (1996). Of particular interest to many of these researchers is the impact of user charges on the demand for health care in developing countries. In the developed countries the research question assumes a slightly different perspective, namely, the impact of mode of payment on the utilization of health care services.

University of Cape Town

Among the first empirical works in the context of developing countries is Heller (1981). This work estimated the demand for medical care services using data obtained from a 1975 household survey in Peninsular Malaysia. His findings showed that total medical demand proxied by absolute volume of inpatients is inelastic to cash price and time costs of utilization. Furthermore, Heller finds that income is of little significance in the demand for medical care but sufficiently significant in the choice of medical care provider. For price and income to make little impact on the demand for medical care seems contrary to economic expectation. For example, one implication would be that the introduction of user charges is not a barrier to health care utilization among the poor.

J.S Akin (1981) investigated the determinants of demand for child health services in the Philippines using the logit model. Like Heller, his model assumed a multi-categorical dependent variable. Patients were assumed to face provider options that included government facilities, private clinics, traditional practitioners and self- care. Although the study found variables such as waiting costs of the mother, and education of the mother to be significant, surprisingly, like Heller's work, the result shows that income is an insignificant variable.

There are two basic problems with this study. In the first place, the combination of two sets of data from two different surveys, one obtained in 1978 and the second from a supplementary survey conducted three years latter casts doubt on the validity of the results. Gertler and van der Gaag (1990) have pointed out that the functional form of the model used in obtaining the results was miss-specified. The miss-specification consists in including income as a separate explanatory variable among other explanatory variables rather than allowing income to have an indirect effect on demand for health care through its influence on consumption. However, this later position is not altogether conclusive as some other researchers W.Dow (1996) think that it is not necessary for income to enter the utility function indirectly through its effect on consumption.

The same criticism made against Akin (1981) can be made against Akin et al (1986). Akin et al. investigated the determinants of demand for primary health care services in the Bricol region of the Philippines. The authors found the multinomial logit model adequate for the study. Like the previous studies, this study found that prices and distance are unimportant explanatory variables in the demand for outpatient health care services in the region. This is contrary to basic microeconomic theory of demand for normal goods and services unless it is assumed that due to externalities in their consumption, health care services are public goods. It is therefore important to inquire into this seeming contradiction between neoclassical demand theory and the empirical results from these studies.

University of Cape Town

Unlike the work of researchers reviewed earlier, Gertler, Locay and Sanderson (1987) which employed data from urban Peru found price to be an important determinant of health care demand. Their finding also differed from the findings of some other researchers not reviewed here including Musgrove (1983) and Mwabu (1986). The finding that price is an important factor in the demand for health care is in consonance with basic microeconomic theory. Ellis and Mwabu (1991) have, howerver, pointed out a major weakness of this study by Gertler et al. It used the price of a single episode of illness of an individual rather than the entire monthly health care bill of the household to measure the influence of price on the demand for health care. While limiting price of health care to that of an individual in the household, the study at the same time employed household monthly income to measure the impact of income on demand for health care. The problem here is that of confusion about the unit of analysis. This same weakness may also be noted against the work of Dor et al (1987) in rural Cote D'Ivoire.

Ellis and Mwabu (1991) advanced the study of the demand for medical care in developing countries by introducing MacFadden's conditional choice model into the estimation of the demand for medical care. Their study, based on outpatient medical care in rural Kenya, first of all examined the issue of whether consumers are willing to pay higher prices for higher quality facilities. They also examined the relationship between utilization of health care services and demographic variables such age and education and whether differences in utilization arise from differences in the probability of becoming ill or from differences in health care-seeking behavior. The authors used a four-level nested model of decision making to capture the process of household decision at four different levels leading to health care utilization. The first level of decision was whether or not to report an illness. At the second stage, if an illness was reported did the patient seek formal treatment or not. At the third stage, if the patient sought formal treatment what choice of provider did he/she make. The fourth and final stage concerned the means of transport to the facility. It is surprising that none of the demographic characteristics had a significant influence on the choice of health care provider.

University of Cape Town

Perhaps the major strength of this study is that it avoided a major weakness that plagued some of the studies that were based on the multinomial logit models, namely, the correlation among the error terms. This correlation in error terms which tends to characterize multinomial models will be considered in fuller details when we consider the assumption of independence from irrelevant alternatives (IIA) in chapter 3.

Schwartz, Akin and Popkin (1993) examined the economic determinants of demand for modern infant delivery in the Cebu region of the Philippines. The study used the multinomial logit technique with information on about 3000 mother-infant pairs. The results suggest that choice of delivery-service type is relatively insensitive to changes in money prices. This, again, implies that price elasticity of demand is low.

Bartholome and Vosti (1995) investigated the choice between public and private health care providers. This study analyzed the consumer choice of providers in the treatment of malaria in Brazil. This study concentrated on a homogenous demand for the treatment of malaria; that is to say, the demand for the treatment of a specific type of illness rather than illnesses. This is in line with the suggestion made by McGuire, Henderson and Mooney (1987) that the demand for health care is heterogeneous and researchers should seek to concentrate on homogenous demand for health care. In other words, a focus on specific types rather than on an aggregation of illnesses may tend to produce more accurate results. Using a binary dependent variable, and a vector of explanatory variables that includes economic, demographic and other household characteristics, they found that distance and wealth proxied by monthly expenditure are significant in the choice of treatment between public and private facilities. While this study appears more focused, which may contribute to greater estimation precision, it nevertheless suffers from some of the weaknesses of the other previous studies already reviewed. Constraining the decision to only public and private facilities seems unrealistic as patients may decide not to seek a cure at all.

Efforts have also been made by some researchers to introduce a time dimension into the estimation of demand for medical care. One such effort is W.H Dow (1996). This work shows that the introduction of the long-run horizon and the inclusion of non-sick into the estimation reduce the marginal effects of the explanatory variables on the

University of Cape Town

demand for health care. In other words, the long run demand response to changes in exogenous variables such as introduction of user charges may differ in magnitude from demand response to changes in the short run. This clearly needs further research.

This brief review of international literature on demand for medical care has shown that there are two streams of research findings. First, there are those that found financial and economic cost insignificant as determinants of the demand for medical care and so justifying the policy of user charges in public health care systems. Many of the earlier studies belong to this stream. The second stream consists of those that found financial costs and economic factors to be strong determinants of demand for health care services. They, therefore consider the introduction of or increases to user charges in health facilities to be significant barriers to utilization of medical services especially among the poor.

University of Cape Town

Perhaps, it is important to compare the findings of these studies based on the developing countries with those based on studies in the industrialized countries. About the most comprehensive of such studies based on the developed countries' experience is study by Manning et al. (1987). This study was conducted for the Rand Corporation in the US. The team used a controlled, randomized experiment to study how the proportion of the bill paid by the patient in a cost sharing arrangement affects the patient's demand for health care. Their study found that price was a statistically significant determinant of health care demand. This conclusion apparently raises a problem. How does one explain the fact that financial prices were significant for the rich countries and insignificant for the poor countries assuming that both operate under the similar strengths and weakness of the medical care market? These results do not seem to conform to the theoretical predictions of the Engel's function that shows that as income grows demand for basic needs such as health care become inelastic. The contrary would be the case only if one assumes that health care is not a normal good.

Furthermore, many of the studies assumed that the multinomial logit model is the correct functional specification for health care demand. This specification assumes that a household's utility from seeking health care from facility A, for instance, is independent of its decision to seek no care all or to treat illness at home. More technically, the studies assumed that the error terms in such decisions are un-correlated. This fact needs to be proved in each case not assumed.

For this study I intend to use the nested logit model (also referred to as the Conditional Logit Model) to investigate the pattern of demand for health care services in Nsukka local government area. This model is due to D. McFadden (1981) who originated it in the context of transportation studies. It was applied in health care demand studies by Ellis and Mwabu (1991) who used it in a four-level nested model of demand for health care in Kenya. As was stated earlier on, this model, unlike many other models used in previous studies, avoids the IIA assumption.

University of Cape Town

Chapter 3

The Conceptual Framework

3.1 Introduction

In the preceding chapter, we critically reviewed some of the studies that have been done in the area of demand for medical services. Their strengths and weaknesses were pointed out. In particular it was shown that model specification was a major weakness in most of the studies. It was also noted that the nested logit model represents an appropriate functional form of the demand for health care services. The focus of this chapter is the functional specification of the nested logit model. It will be shown how this model overcomes the limitations of some of the other models. Most importantly, it will be shown that this model derives from the basic economic principles of consumer behavior, namely, maximization of utility under resource constraints. To lay the foundation for the functional specification, I shall first describe the basic theoretical assumptions behind the demand for health care services.

3.2 Theoretical Framework

As was noted earlier, much of the recent progress in health care demand research is due to the theoretical insights of Grossman (1972). Essentially, the Grossman's proposition is that health, like knowledge, is a durable capital stock. But while the stock of knowledge affects a person's market and non-market productivity, the stock of health affects the output of healthy time. A large stock of health increases the time available for productive activities. Therefore, health is demanded as an argument in the direct utility function of the individual. That is, by itself health yields utility to the possessor and sickness is a source of dis-utility. It is also demanded as a capital or investment good because it determines the total amount of time the economic agent can spend in money earning production and the production of other goods and services. As a capital stock, health is liable to depreciation with age. Death occurs when, as a result of depreciation, the health stock falls below a certain level.

University of Cape Town

As a basic commodity, households produce health with inputs of market goods and their own time. "Since goods and services are inputs into the production of commodities, the demand for these goods and services is a derived demand" Grossman (1972:224).

The demand for health care is therefore a derived demand. Health care, unlike health, is not demanded for its own sake but arises from the desire to replenish the stock of health that is a fundamental object of choice. The optimal stock of health a household can produce depends on a number of variables. These variables include the environment, the total amount of resources available to the household, its level of education, the quantity and prices of other consumption goods and services.

Under these conditions, Wagstaff (1986) shows that the Grossman's model builds on the fundamental assumptions of neoclassical microeconomics, namely, a downwards-sloping demand curve, a set of convex indifference curves, a budget constraint and a production function that is subject to diminishing marginal returns. A downward–sloping demand curve implies that the shadow price of health expenditures is not only the price the household pays the health care provider but also the index of all other consumption goods that are forgone. A convex utility function implies that trade-offs have to be made between the quantity of health producing goods and the quantity of other consumption goods that are available to the household. As Grossman himself puts it "shifts in these variables alter the optimal amount of health and also alter the derived demand for gross investment, measured, say, by medical expenditures". (Op.cit p.225)

Since the household budget is necessarily limited, choices must be made with a view to maximizing the total utility of the household subject to budget constraints. If a household member falls ill, the utility that the household would derive from giving the individual medical treatment which produces improvement in health status has an opportunity cost in terms of the reduction of the quantity of other consumption goods available to the household. Therefore, given its resources, the household has to make a decision whether to seek medical care or not for this

University of Cape Town

member depending on which decision is consistent with its utility maximization objectives. Once a decision is made to seek medical care, a further decision has to be made regarding from which of the available health care providers this treatment is to be obtained. A feasible option is to treat the person at home; that is, home care. Since the household is assumed to be rational and inclined to maximizing its utility, it has to weigh the quality of health gained from a given provider against the cost of seeking care from that provider or category of providers. The home care option may then be considered as the normalizing alternative.

In the conventional form, let us consider the utility maximizing behavior of a household who faces a choice between two health providers. Subject to resource constraints, this household will have an utility function that measures the desirability of an option. This utility function may be specified as:

$$U_i = U(h_{ik}, x_i, \varepsilon_i) \quad \text{..} 3.1$$

Where h may be considered as the attributes of kth provider or facility which define its health production function while i indexes the household. The x refers to a vector of the household's socio-economic or demographic variables (eg level of education) which determine its own health-enhancing production. The error term ε.
captures the unobservables, measurement errors about the characteristics of the household and mis-judgements about the attributes of the facilities. If the households were sampled randomly from a population with common socio-economic characteristics x and assuming that the same providers or facilities are available then, following Domencich and McFadden (1975:50-52) the utility function can be said to be stochastic. In other words, the error term is randomly distributed. This is clearly so because the value of the utility function will depend on the household drawn at random from the population and the particular provider alternative chosen.

To simplify our notation let us redefine the utility function in (3.1) in index form so that we obtain

$$U = \mu + \varepsilon \quad \text{...3.2}$$

Where μ is a vector of covariates and associated parameters. A respondent household will have the following utility functions

$$\begin{aligned} U_0 &= \mu_0 + \varepsilon_0 \quad \textit{from} \quad \textit{alternative} \quad 0 \\ U_1 &= \mu_1 + \varepsilon_1 \quad \textit{from} \quad \textit{alternative} \quad 1 \end{aligned} \quad \text{............................3.3}$$

The household will choose provider 1 if this is the alternative that maximizes its utility, subject to household constraints. Formally then,

$$U_{i1} \succ U_{i0} \quad \text{......................................3.4}$$

We may define y = 1 to be the probability (P) that an event occurs; and y = 0 to be the probability of its non-occurrence. The event that a household drawn randomly from the population with the same household socio-economic characteristics chooses U_{i1} over U_{i0} then gives the probability function

$$P(y_i = 1 \mid \mu) = \Pr ob(U_{i1} \succ U_{i0}) \quad \text{...................................3.5}$$

Equation (3.5) may then be expressed more fully as

$$P(y_i = 1 \mid \mu) = \Pr ob(\varepsilon_{io} - \varepsilon_{i1} \leq \mu_{i1} - \mu_{i0}) \text{.....................................3.6}$$

Or

$$P(y_i = 1 \mid \mu) = \Pr ob(\varepsilon_{i0} - \varepsilon_{i1} \leq \mu_{i0} - \mu_{i1}) \quad \text{..............................3.7}$$

The probability can easily be seen as a function which depends on the difference between the two stochastic utility functions defined in equation (3.3). That is the difference between $(\mu_{i0} + \varepsilon_{i0})$ *and* $(\mu_{i1} + \varepsilon_{i1})$ This difference becomes the argument in the function that may be designated as H.

We can therefore express the probability (P) that $U_{i1} > U_{i0}$ as

University of Cape Town

$$P(y_i = 1 \mid \mu) = H[U(\mu_{i0}) - U(\mu_{i1})] \qquad 3.8$$

(Note that for simplicity the error terms have been subsumed into μ)
Assumptions made about the variance of H give rise to several forms of probabilistic models. For example, in the probit model it is assumed that mean of $(\varepsilon_0 \ - \ \varepsilon_1)$ is 0 and the variance is 1 which gives a normal distribution. In the logistic model the mean is 0 , the variance is assumed to be $\pi^2/3$. Although both the probit and logistic distributions resemble the normal distribution the probit distribution cannot be evaluated without numerical multivariate integration. The logistic distribution, on the other hand, is computationally more tractable and is therefore preferred by researchers. For the logistic distribution

$$H(U) = \frac{1}{1+\exp[(\mu_{i0})-(\mu_{i1})]} \qquad 3.9$$

Where the second expression in the denominator is the natural logarithm raised to the power $(U_{i0} - U_{i1})$. This distribution is often expressed as log of odd

$$\ln\left[\frac{\Pr(y_i = 1 \mid \mu)}{1-\Pr(y_i = 1 \mid \mu)}\right] \qquad 3.10$$

This is the binary logit model which is the simplest case of what is known as the Luce family of models. Other members of this family include the multinomial and nested logit models D.McFadden (1981). The nested logit model that will be used in this study is simply a development of this basic binary logistic model.

The effort so far has been to show that the logit family of models that will be used in this study is consistent with random utility maximization (RUM). Consistency with this property is crucial if a model is to be used for welfare analysis. In the logit model this important property follows from the assumption about the distribution of the difference of the error terms in equation (3.7). It is assumed that the difference in the error terms is independently and identically distributed which implies that there is no correlation between the error term in one utility function and the error term in another utility function.

University of Cape Town

The stochastic term

$$\varepsilon \quad = \quad (\varepsilon_{i0} - \varepsilon_{i1}) \qquad \text{.....................................3.11}$$

is, therefore, assumed to be independently distributed. This distribution is said to follow Type I extreme value distribution or log Weibull distribution. It is specified as

$$\Pr ob\,(\varepsilon_{i0} \leq \varepsilon_{i1}) \quad = \quad \exp[-\exp(-\varepsilon)] \qquad \text{..............................3.12}$$

This distribution resembles the normal distribution. It is bell-shaped. However, it has a thicker tail on the right and a thinner left tail. It has a mode at 0, mean at 0.575 and variance of 1.622 in contrast to normal distribution with variance = 1. Two properties of a Type I extreme value distribution or the Weibull distribution which are important for our purpose here may be summarized as follows:

(a) The Weibull distribution like the normal distribution is stable under maximization. That is to say, the maximum of two independent Weibull distributions is also a Weibull distribution (Just as the sum of two normal distributions is a normal distribution). In other words, as the normal distribution is stable under addition, the Weibull distribution is stable under maximization.

(b) The difference of independently distributed Weibull distributions has a binary logistic distribution.

These properties are stated and proved in Domencich and McFadden (1975:63-65). The importance of these properties in our study is that they guarantee that our model is consistent with the theory of utility maximization and therefore useful for analyzing household health care demand behavior.

3.3 Empirical Specification of the Functional Form

The preceding section of this chapter established a very important foundation for our model. It was shown that the logit family of probablistic choice model shares the important property of RUM. This was illustrated in the case of the binary logit model. We can now extend the same principle beyond the binary to the multi-

University of Cape Town

response level that will enable us to specify the model of nested logit that is proposed for this study.

For the case of multiple choice indexed over k, for k = 0,1,2,...K, we may define $y_{ik} = 1$ if the ith household chooses the kth health care provider and 0 if not.

We may further define

$$V(\mu)$$

to be the function of the explanatory variables and associated but unknown parameters. The error term has been subsumed μ into μ Then, analogous to equation (3.9), the case of a choice among three alternatives by the ith household may be specified as

$$P(y_i = 0 \mid \mu) = \frac{1}{1+\exp(\mu_1)+\exp(\mu_2)} \quad \text{..................................3.14}$$

(where the chosen alternative k = 0)

For alternative (1) the probability of the ith household choosing this option is given by

$$P(y_i = 1 \mid \mu) = \frac{\exp(\mu_1)}{1+\exp(\mu_1)+\exp(\mu_2)} \quad \text{..................................3.15}$$

For the choice of the third of the alternatives, the probability becomes

$$P(y_i = 2 \mid \mu) = \frac{\exp(\mu_2)}{1+\exp(\mu_1)+\exp(\mu_2)} \quad \text{..................................3.16}$$

Equations (3.14 – 3.16) specify the multinomial logit model

Domencich and McFadden (1975:69) called the function $V[\exp(\mu)]$ the "strict utility function". It is easily seen that the probability of an alternative being chosen is given by its strict utility divided by 1 plus the sum of all the strict utilities including that of the new addition. The sum of all the probabilities must then sum to 1.

University of Cape Town

The strength of the multinomial model is that the addition of a new alternative while decreasing the probability that an alternative is chosen does not alter the relative odds of the existing ones. In other words, the ratio of the odds of choosing facility 0 to the odds of choosing facility 1 does not change because of the addition of facility 2. However, the main weakness is that for the relative odds of alternative 0 to alternative 1 to remain the same, the new alternative must be perceived as completely distinct and independent. Furthermore, the model implies that the effect of the introduction of the new alternative on the existing ones is the same. That is, the model implies constant cross elasticity of demand in which substitution is difficult. (Domencich and McFadden 1975)

A major conclusion to be drawn from the above is that unless the alternatives are completely dissimilar, the multinomial logit model cannot be applied. To put it another way, unless the error terms in the choices are independently distributed, the multinomial logit model would be inappropriate. The assumption that the error terms are completely independently distributed is what Domencich and McFadden (1975) called independence of irrelevant alternative (IIA). Amemiya (1981) shows that in many studies the absence of IIA cannot be assumed. Many economic phenomena would display some form of dependence or the other. Apparently, the choice among alternative health care providers manifests this quality. Obviously, a household seeking health care for its member, may not perceive the difference between a public and a private hospitals, for instance, as very clear and distinct. The similarity between the two categories may be perceived as high. Put in another way, the elasticity of substitution between a private and a public hospital may, generally, be perceived to be very high

3.4 The Nested Logit Model

In order to ameliorate the failure of the multinomial logit model (MNLM) in the presence of correlated error terms, McFadden (1981) developed the nested logit model. Instead of assuming that the error terms are independently distributed, this model assumes that the error terms are correlated.

The motivating idea behind the nested logit model is a sequential choice structure in which the economic agent is assumed to choose between independent classes of

University of Cape Town

utilities at each level. While the correlation between error terms in the independent classes are assumed to be zero, the error terms within the classes are allowed to be nonzero. How does this improve the problem of correlation among the error terms? The assumption here is that mean utility of 'best' alternative in each class approximates the maximum utility in that class. For example, for a household that is faced with choice between formal and informal care. Assuming that under informal care it has only home care on one hand, under formal care it has two options: a public hospital or a private hospital. This household now weighs the utility from home care against the 'best' of the two alternatives in the class of formal providers. Although the error terms within the class of formal providers may tend to correlate, the error terms in the formal and informal have zero correlation. This model gives rise to a tree-like structure of decision process.

Using the tree analogy to illustrate, it is assumed that the following conditions obtain.

Trunk : A household reports sick

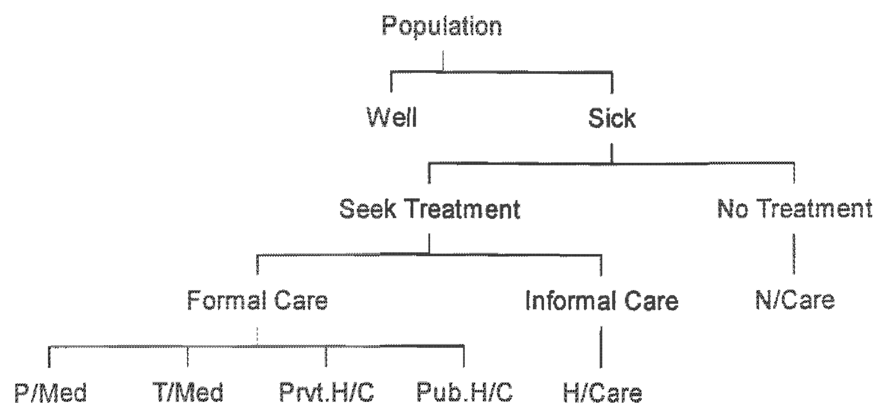
(a) Limb: Conditional upon a member reporting ill, the household either seeks treatment or does not seek treatment.

(b) Branch: Conditional upon the decision to seek health care service, it either seeks formal care from a professional health care provider/facility or informal care (i.e, the production of health care at home)

(c) Twig: Conditional upon the decision to seek professional care, it chooses from among the professional providers [government hospital/clinic (Pub.H/C), private hospital/clinic (Prvt.H/C), traditional medicine (T/Med) and pharmaceutical/patent medicine store (P/Med)]

University of Cape Town

Fig 3.1 **A Sequential Choice Structure of Demand for Health Care**



As the focus of this study is on the choices that households make when they are confronted with sickness, we ignore the trunk with the Well population and concentrate on the trunk with the Sick population.

There are three decision nodes on the trunk with the sick population representing the three-level nested model. At the lowest level is the choice among the available categories of health facilities (the twigs). This is nested on the decision to seek formal care or informal care. The decision to seek formal or informal care is nested on the third level of decision: seek treatment or not. The decision to seek treatment or not is nested on the fourth level: Well or Sick. But as was said earlier, this fourth level is outside our consideration here. The decision at each of these three levels can be modeled as a separate utility function. That is, the household may be seen as taking an independent decision at each decision node or point. The decision whether to treat this sickness or not may be modeled as a binary logit decision. Suppose the household takes the decision to treat this illness, then at the second stage the household takes another decision whether to treat it at home without consulting any medical expert (informal care) or to seek the help of a medical professional (formal care). This level, again, may be modeled as a logit model. Finally, suppose the household decides to consult a medical expert, it has to decide which of the available professional facilities to consult. Here the household is deemed as making a discrete choice among the facilities. Each of these decisions is taking with a view to maximizing the household's utility function. Domencich and McFadden (1975:78) show that this decision structure

University of Cape Town

provides a way out of the independence of irrelevant alternative (IIA) assumption because the decision at every level is deemed to be between distinct classes while within-class decision error terms are allowed to correlate. The model is also consistent with the strict utility model.

While it may be normal to consider the household as taking a separate decision at each step, it is also possible to consider the decisions as taking place simultaneously (Ellis and Mwabu, 1991) . That is, the household takes all the three decisions in one pass to maximize its utility function. In the later case, the unconditional utility function may be specified as

$$U_{ijk} = U(X_{ijk} + Y_{ij} + Z_i)\,\varepsilon_{ijk} + \varepsilon_{ij} + \varepsilon_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.17$$

Where X_{ijk} represents the set of observed attributes that vary with twig, branch, and limb. The number of twigs k is indexed as k=1,2,...K for all possible options at that level. (In this model K=6.) It is to be understood then that any attribute that varies among branches will lead to variation among the limbs and trunk. It is further to be understood that the set of variables represented by X_{ijk} influences the household's decision about which facility to seek health care. Y represents the set of attributes that varies according to branch. The number of branches may be indexed as j=1,2,...J (Here, J=2). Any attributes that vary among the branches will imply variation between the limbs. Z represents the set of attributes that characterize the tree at the limb level. The number of limbs may be indexed as i=1,2...S. (Here S=2). The residuals are captured as ε_{ijk}, ε_{ij} and ε_I respectively. They represent the influence of variables that may have been excluded from the model, variations in household perceptions, habits, imperfections in household judgements about choices that maximize its utility.

In the multinomial logit model the error terms in the stochastic utility function were assumed to be independently distributed (Type I extreme value distribution). In the nested logit model the error terms are assumed to follow a multivariate extreme value distribution or generalized extreme value (GEV)) distribution which allows for dependence among the choices in a class. McFadden (1981) defines this distribution as

$$F(\varepsilon_1, \varepsilon_2, ..., \varepsilon_m) \quad = \quad \exp\{-G[\exp(-\varepsilon_1), \exp(-\varepsilon_2), \; ...., \exp(-\varepsilon_m)]\} \quad ....3.18$$

Where

$$G(\exp(\mu_1), \exp(\mu_2), ..., \exp(\mu_m))$$

is a function of stochastic utility satisfying the condition

$$(\mu_1, \mu_2, ...\mu_m) \quad \geq \quad 0$$

and is homogenous of degree 1.

For the case of three alternatives in which alternative 1 is distinct and alternatives 2 and 3 are similar, the resulting probability, following Maddala (1983), will be

$$P(y_i = 1 \mid \mu) \quad = \quad \frac{\exp(\mu_1)}{\exp(\mu_1) + \exp(\mu_2) + \exp(\mu_3)} \quad ....3.19$$

The probability for alternative 2 will be

$$P(y_i = 2 \mid x) \quad = \quad \frac{\exp[\,\frac{\mu_1}{(1-\sigma)}\,]\left(\exp[\,\frac{\mu_2}{(1-\sigma)}\,] + \exp[\,\frac{\mu_3}{(1-\sigma)}\,]\right)}{\exp(\mu_1) + \exp(\mu_2) + \exp(\mu_3)} \quad ......3.20$$

The probability for alternative 3 will resemble that of 2. The correlation coefficient σ is a measure of the correlation between the alternatives. As a measure of correlation, its value lies between 0 and 1.

For a three-level nested model such as we are using here, the three levels may have different sets of variables specified at each level. Where it is feasible, a researcher may, for example, want to isolate a vector of variables (X_{ijk}) that influence decision at the level of choice among facilities. He or she may also want to isolate a different set of variables (Y_{ij}) that he/she suspects may have influence on the decision at the

University of Cape Town

branch level – to seek formal or informal care - and then specify them explicitly. Finally, the researcher may further want to isolate a set of variables (Z_i) that he/she suspects to exert influence at only the limb level. In these cases, he/she will have to specify different sets of variables at each of these levels of decision. On the other hand, the researcher may want to allow for interactions among the variables to occur at any of the levels of the decision process. Or he/she may want to allow for the possibility that variables specified at facility level may also influence the decision at the two higher levels (formal or informal care; seek treatment or not). In the latter cases, he/she would specify only the vector (X_{ijk}) at the utility function (facility level) Greene (1998:559-60). In this present model, there are no specified attributes at the branch and limb levels. Attributes are specified at the level of choice among health care alternatives.

The utility function may then be specified as

$$U_{ijk} = \mu_{ijk} = \beta' X_{ijk} + \alpha' Y_{ij} + \gamma' Z_i \quad \text{..........................3.21}$$

Note again that $\alpha ' Y$ and $\gamma ' Z$ are included in the specification because of the possible impact of the variables specified at the choice of facility level (X_{ijk}) on these higher levels.

The vector X_{ijk} represents attributes that vary with facility and so vary also with branch and limb. Y_{ij} is the vector of attributes that vary with the branch (formal/informal treatment) and so vary also with limb. Z_i represents the set of attributes that vary with limb (treatment/ no treatment). $\beta, \alpha \; and \; \gamma$ are vectors of unknown parameters to be estimated. The probability that a household chooses option k in branch j in limb i may be expressed as:

$$P_{ijk} = P_{k\backslash ij} \times P_{j\backslash i} \times P_i$$

or

$$P_{ijk} = \frac{\exp(\mu_{ijk})}{\sum_l \sum_m \sum_n \exp(\mu_{lmn})} \quad \text{...3.22}$$

Where the subscripts l, m and n tell us the number of choices available at the limb, branch and twigs respectively. For a given household the value P_{ijk} may be considered as its unconditional maximized utility. It is its contribution to the overall likelihood function of the entire sample.

The conditional probability $P_{k|ij}$ involves the impact of the set of attributes X_{ijk} at *only* the level of choice among the facilities. The set of parameters β estimates this impact at the branch level of choice among facilities. The conditional probability of k given j and i may be expressed as

$$P_{k|i,j} = \frac{\exp(\beta' x_{ijk})}{\sum_n \exp(\beta' X_{ijn})} \quad \text{...3.23}$$

Following Maddala (1983), we may define the inclusive value at the branch level to be the log of the sum of $\exp(\beta' X_{ijn})$ which is the denominator in equation (3.23). This has been interpreted to mean the expected contribution to overall utility of the choice made at the preceding level Ellis and Mwabu (1991).

To formalize, we may specify this value as

$$I_{ij} = \log\left(\sum_n \exp(\beta' X_{ijn}\right) \quad \text{..3.24}$$

Equation (3.24) specifies the branch inclusive value and has a coefficient that lies between 0 and 1. We can therefore express it as:

$$\log\left(\sum_n \exp(\beta' X_{ijn})\right) = (1-\sigma_k) I_{ij} \quad \text{.................................3.25}$$

University of Cape Town

The conditional probability that a household seeks formal care given that it seeks treatment may be expressed as

$$P_{j|i} = \frac{\exp(\alpha Y_{ij} + I_{ij})}{\sum_m \exp(\alpha' Y_{im} + I_{im})} \quad ....3.26$$

The inclusive value at the limb level is then the expression

$$\log\left(\sum_m \sum_n \exp(\beta' X_{imn} + \alpha' Y_{im})\right) = \log\left(\sum_m \exp(\alpha' Y_{im} + I_{im})\right) \quad ....3.27$$

The inclusive value parameter at this level also lies within the 0-1 range and thus we denote it as

$$(1-\sigma_j)J$$

Where J denotes the right hand expression in the inclusive value identity stated in equation (3.27)

Finally, the probability that a household reported sickness is given by

$$P_i = \frac{\exp(\gamma' Z_i) + \log\left(\sum_l \exp(\alpha' Y_{il} + I_{il})\right)}{\sum_l \exp(\gamma' Z_l) + \log\left(\sum_l \exp(\alpha' Y_l + I_l)\right)} \quad ....3.28$$

Replacing the second expression in both the numerator and denominator with J, this last equation can be expressed more compactly as

$$p_i = \frac{\exp(\gamma' Z_l) + J_l}{\sum_l \exp(\gamma' Z_l) + J_l} \quad ....(3.29)$$

This third level inclusive value parameter is expected to be unity because the probability of reporting ill for a household that reported ill is one since we are dealing with only households that reported illness.

University of Cape Town

The inclusive value parameters are not only important in themselves, they are even more important because they provide a specification test for the over all fit of the model. McFadden (1981) shows that a necessary and sufficient condition for this model to be consistent with rational utility maximization is that the inclusive value parameter must lie in the unit interval. If the estimated inclusive value parameter equals 1.0, then that is an indication that the appropriate functional form is one level model. However, if the estimated inclusive value parameter lies outside the range 0-1, Maddala (1983) suggests that that is and indication that the model does not suit the data generating process. In other words, the model does not fit the behavior pattern of households from which the data was generated. The condition requiring the estimated inclusive value parameters of the model to lie between 0-1 is apparently crucial if we are to ensure that the model approximates the data generating process.

3.5 Definition of Variables and Functional Relationships in the Model

We now wish to define the variables to be used in the estimation process and establish more clearly the functional relationships between them

The dependent variable at the utility function level of the estimation is FACILITY. This is an indicator variable with six categories that include NOCARE, HOMECARE, PUBLIC HOSPITALS/CLINICS, PRIVATE HOSPITALS/CLINICS, TRADITIONAL MEDICARE, and PHARMACEUTICAL/PATENT MEDICINE STORES. This implies that each observation will have five dummies. The indicator is 1 if a particular facility is chosen or 0 otherwise. Thus the probability that FACILITY = 1 is an index of choice made. The table below lists and defines the independent variables and also gives the expected signs of the coefficients to be obtained.

University of Cape Town

Table 3.1 Independent Variables

Variable	Definition	Expected Sign
LOC	This variable is the location dummy. It assumes the value 1 if a household is located within Nsukka urban area or 0 if the household is in any of the other rural communities that make up the local government area. This variable is chosen because it is generally believed that access to health facilities will differ considerably for those in urban and rural communities. Those in rural communities are presumed to have less access to facilities than those in urban communities	Since in this case the rural category becomes the reference group, the sign will be positive for the coefficient of the urban category
HH_SIZE	This is number of people who live under the same roof at least 15 days in a month for a period of one year and share food from a common source. This variable is expected to influence health care consumption through its negative effect on food consumption and leisure which are health-producing	The larger the size of the household the less accessible health care becomes
DISTANCE	This is the distance a household seeking health care has to travel to reach the chosen facility. It is measured in kilometers. This obviously a major determinant of access and is therefore included in the model	Distance constitutes a barrier to access and is therefore expected to be negatively signed
S_AGE	Age of the ill person. The inclusion of this variable is to indicate that consumption of health care will vary with a physiological state with higher consumption at the early and later life.	The predominance of children in population suggests that the sign will be negative
SEVERITY	This measures the seriousness of the illness. It is proxied by the number of days the ill person was unable to carry out his or her normal duties within the month preceding the interview. This variable is an obvious inclusion because severity will generally determine whether a person seeks medical help.	The sign is expected to be positive as the more severe the illness the more probable that help will be sought
W_TIME	The amount of time the patient had to wait in order to obtain treatment from the health	This is expected to be negatively

University of Cape Town

	in order to obtain treatment from the health care provider. Waiting time is a disutility that will influence health care consumption	to be negatively signed as waitin time is a disutility
TMTCOST	This is the financial cost of obtaining treatment from the provider. It excludes the transportation cost. This is also and obvious inclusion as cost will influence health care consumption	This should also be negatively signed
PCFDX	The per capita food expenditure of the household. This variable is included because health care consumption is part of basic household consumption	Positively signed because households with higher food expenditure are also likely to have greater medical budget
INCOME	The household's estimated total financial income in a month wills surely influence how much to spend on health care	The larger the income the more likely health care outlay will increase
AVYRSED	The average years of formal education of the members of the household. Education or information will certainly influence how a household handles its health care needs	It is expected that the more knowledgeable a household is, the more it will avail its members of health care opportunities.

It may be appropriate here to remark that there are two forms of nested logit models in research literature. One is the nested multinomial logit model (NMLM) and the other is the conditional logit model (CLM). Both are referred to as the nested logit model W.H Greene (1998:513); J.S Long (1997:178). The difference between them is that the variables in the multinomial model are usually the characteristics of the individual who makes the choice among the alternatives and so are invariant to the choices (eg income). In the conditional model, the variables are characteristics of the choices themselves (e.g waiting time in a given facility). However, Maddala (1983:42) shows that both the multinomial and conditional models are, algebraically, equivalent.

University of Cape Town

Furthermore, the two types of variables mentioned above may be combined in one model J.S Long (1997:180); T.F Liao (1994:60-61); Greene (1993:664 -665).

In the discrete choice probability model used here, without regard to the number of choices, there is only a single vector of parameters to be estimated Agresti (1990:321-24); T.F Liao (1994:58-61); Greene (1998:620)]. This also makes the model different from the multinomial logit model in which each alternative has its own set of explanatory variables. Thus in the case of MNLM if, for example, there are six alternatives to choose from there are also five set of parameters to estimate (the sixth set is reserved for identification purposes). This model proliferates parameters. The reason is that each in multinomial logit model, each variable has a different type of effect on each outcome of the dependent variable. In the nested logit or conditional logit model used here, the effect of each variable on each of the outcomes in the dependent variable is the same, it is only the amount of the effect that may differ from out come to outcome.

3.6 Full Information Maximum Likelihood Method

The preceding sections of this chapter have allowed us to derive the functional form of the model to be used in estimating the parameters of health care demand behavior. Now, we want to briefly consider the actual method of obtaining the parameter estimates. This method is the Maximum Likelihood Method (MLE)

The ML method works on the principle that a given sample generated by a random process could have come from a number of ~~parent populations~~. Each of these parent populations would have its own parameters, the mean and the variance. How do we know then, which of these parent populations has given rise to the observed sample? This question can be answered if we calculate the joint probability of obtaining the values in the given sample using the parameters of the parent populations. Those parameters, the mean and the variance, of that parent population that maximize the joint probability of obtaining the sample values are said to be the maximum likelihood

$$L(\beta_1, \beta_2, \ldots\ldots \beta_{10}; \sigma_k, \sigma_j, \sigma_i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.30$$

University of Cape Town

estimators. Thus, this method of estimation yields parameter estimates that maximize the probability of obtaining the observed data. Basically, the Maximum Likelihood estimator is an asymptotic sample estimator. Its efficiency increases as the sample size approaches infinity. It selects from all possible estimates of the populations those estimates that make the likelihood of observing the sample data as high as possible. The joint probability of observing all the elements of the sample is given by the maximum likelihood function of the variable. In this particular case the maximum likelihood function whose parameters are to be estimated is

The log likelihood function is given by

$$\ln L_{k|ij} = \sum_{k=0}^{K} D_{k|ij} \ln P_{k|ij} \qquad \text{.......................................}3.31$$

D_{ik} is a dummy that takes on 1 if a household chooses facility k and 0 if not

Limited Information Maximum Likelihood (LIML) and Full Information Maximum Likelihood (FIML) are two special cases of the Maximum Likelihood method. While the former is a single equation method of obtaining estimates from over-identified structural parameters, FIML is a system equation method that requires the complete specification of all the equations of the model. In both the assumption of normal distribution is crucial.

While the LIML estimates the nested logit model in two steps, the FIML does the estimation simultaneously. It has also been shown that FIML yields efficiency gains over LIML D.McFadden (1981); Amemiya (1981) Gertler and van der Gaag (1990:74)]. Both programs are available in *LIMDEP Version7.0* that will be used for the estimation. The FIML is used in the estimation.

This chapter has provided the basic framework for the empirical estimation and the interpretation of the output of the model. In the next two chapters, we shall discuss how the data to be used in the estimation was generated before we carry out the actual estimation exercise in chapter 6

University of Cape Town

Chapter 4

The Data Collection Method

4.1 Introduction

In this chapter I shall describe how the data to be used in this study were generated. The principles that guided the survey design and the data collection method will be described. Some of the experiences that were gained in the process of collecting the data will also be described.

4.2 The Survey Site

The data for this study were generated from a household survey I conducted in Nsukka local government area (LGA) between January and February 2000. The area was selected for this study because of the researcher's relative familiarity with the social and economic organization of the local government area. This relative knowledge of the area proved to be of much use during the data collection exercise.

Nsukka LGA is located in the northernmost part of Enugu State, South-East of Nigeria. It comprises 15 communities with populations: Anuka (776), Okutu (4,022), Ibeagwagu (1,304), Okpuje (9,259), Ibeagwani (9,443), Okpaligbo (2,500), Obukpa (20,056), Aloruno (6,530), Edem (16,661), Obimo (12,753), Lejja (15,325), Edeoballa 14,368), Opi (25,384), Ehalumona (36,129) and Nsukka (79,913). Total population as at 1996 was 254,422[1]. The LGA spans an area of about 1,200 km sq. Compared to other parts of South Eastern Nigeria, the LGA, is relatively poor. The majority of the population is engaged in farming, crafts and other rural occupations. Nsukka itself is a

[1] These population figures were obtained from the National Population Commission

University of Cape Town

semi-urban town with substantial trading and transportation activities that serve the population of the University of Nigeria located in this area.

There are a number of health facilities that serve the communities in this LGA. At the time of the interview, there were 2 general hospitals ran by Enugu State government, 20 Primary Health Care clinics ran by Nsukka Local Government. There were 20 private and mission hospitals, 25 private maternity centers, 11 private clinics and a medical center located inside the university[2] There were also a couple of pharmacy shops and several patent medicine stores at almost every nook and corner of the communities. Traditional medicine healers constitute a significant proportion of health care suppliers in the health market of this area.

From an epidemiological perspective, the most prevalent diseases as at the time of the survey, were malaria, typhoid fever, tuberculosis and children's diarrhoea. These were reported in many of the health facilities visited and confirmed by the Head of Department of the Local Government Department of Health.

4.3 Sample Size

In a household survey framework, sample sizes usually vary depending on the purpose and size of the population. In national household surveys it is usual to have sample to population ratio of 1:500 or even 1:2500 Deaton (1998:10). However, since ours is a much smaller population than national population the ratio had to be made as high as possible. With a population of about 250,000 living in about 44,000 households in Nsukka local government area, the sample to population ratio was set at 1:35. This implied that about 1250 households were to be sampled from the population.

4.4 Sampling Strategy

A two-stage sampling design was adopted for the sampling process. The first stage was the selection of the clusters to be sampled. The second stage was the selection of

---

[2] The figures were obtained from Enugu State Ministry of Health and Nsukka Local Government department of health.

University of Cape Town

the households to be interviewed. For the selection of the clusters, each of the 15 autonomous communities was assumed to constitute a natural cluster. Nsukka community with a population of about 80,000 and fairly developed social infrastructure made up the urban cluster while the rest of the other communities were rural.

Simple random sampling in which the respondent households are selected at random would have been ideal. That method of sampling would have required making a household roaster of the local government population and them choosing households to be interviewed randomly from the roaster. Although simple random sampling is generally regarded as an ideal, often, however, there are compelling reasons for choosing other sampling methods that may be more cost effective than simple random sampling. As Deaton (1998:13) notes, apart from cost considerations, the accuracy of an econometric estimate can sometimes be enhanced by the researcher's prior information about the population being surveyed. Such prior knowledge can assist the surveyor in designing an effective method that will strengthen the efficiency of the statistical inference about the population under study.

In designing this survey, the two-stage cluster sampling was considered optimal for the data collection. While cost consideration was important, there were other advantages that came to the fore in selecting the method. From the cost perspective, it was considered more cost effective to assign a group of field workers to a village community where households live in contiguous compounds rather than have them travel for long distances to interview households that live far apart. Deaton (1998:12-15) discusses other advantages of this method which were actually experienced in the field. In the process of data collection it was often necessary to revisit a respondent household to obtain missing information from the previous interview or to clarify some information that was provided in the first visit. This was particularly the case in some of the village communities where the heads of the households went to their distant farms during the day and came back only in the night. Information provided by other adults in the household during the first visit sometimes needed to be confirmed by the head of the household.

University of Cape Town

There is another advantage of the method adopted in this survey, also discussed by Deaton (1998:13), and which proved very useful during the survey exercise. Since households that were interviewed lived in clusters, it was easy to collect village level information. For example, it was easy to estimate the distance to the health facilities where a majority of the households living within the cluster went to obtain treatment. Furthermore, it was easy to collect market prices of consumption goods from the communities. (See also Ellis and Mwabu 1991)

Of the 15 communities that made up the local government, ten were picked at random for the survey. From the chosen ten communities households were selected in proportion to the population size of the communities and then interviewed. The 10 cluster communities that were picked in the random sample were Anuka, Ibeagwagu, Ibeagwani, Aloruno, Edem, Obimo, Lejja, Edeoballa, Opi and Nsukka.

4.5 The Survey Instrument and the Data Collection Process

A household questionnaire was the main instrument used for the collection of data. Facility questionnaires were also administered to some health facilities. The questions were generally structured questions in which multiple options were provided to the respondents. The questions were made as simple as possible. The respondents were required simply to select the appropriate responses from a list of possible responses. The first part of the interview revolved around household demographic structure, the sickness variables and problems associated with seeking care. The next part was concerned with economic variables. These included questions about household income and expenditure. The last part of the question was reserved for the operators of health facilities. The details of these questions are shown in Appendix 3.

The instrument was not translated into local language. This was deemed unnecessary since virtually every household in the area had some adults who were literate in English. The interpretation of the instrument was part of the four days of training given to the interviewers. An additional support was provided. The interviewers were graduate students of the university and spoke the same language (Igbo) as the people though some were not very familiar with Nsukka dialect of the Igbo language. They were accompanied with assistants from the local communities. The latter were versed

University of Cape Town

in the nuances of local dialects and did on the spot translations when the need arose. This strategy also proved effective in other ways. The presence of a familiar face though with its own drawbacks, gave a reasonable measure of security to the respondent households. The respondents were, therefore, more open and less economical with responses though there were still a number of cases where prospective respondents were unwilling to volunteer information because they were suspicious of the purpose of the interview.

To begin the interview a general permission was sought for and obtained from the secretary to the local government and the head of department of health care services in Nsukka local government. Apart from the permission from these local government authorities, before the interview took off in any community the team of interviewers visited the local chief in the area they were going to work to inform him of the purpose of the interview. They also had to ask for his blessing as the tradition of the people required.

As often as possible the interview was conducted with the head or acting head of the household. When this was not possible the next most senior member of the household was interviewed on behalf of the household. The interview opened with the team introducing themselves to the household and greeting them in the cultural manner of the people. A brief discussion on familiar topics established rapport with the household and paved the path for the interview. To allay fears about being identified with the information supplied, households were not required to include their names in the responses. Both households with a sick person and households without a sick person were interviewed.

Before the main interviews started, some trial interviews were conducted. The purpose was to test the respondents' understanding of the questions and experience what difficulties that might arise during the actual interviews. Although the responses were satisfactory, the trial interviews revealed the need for local assistants especially for those interviewers who were not very familiar with Nsukka local dialect. This arrangement was completed before the actual interview took off.

University of Cape Town

In terms of the period of coverage, most of the questions were focused on household health and economic decisions within the month preceding the interview. Indeed, except for questions about household expenditure on durables, the rest were limited to this period. The reason for the focus on the preceding one month is that a number of studies have shown that household expenditures and consumption tend to diminish the longer the recall period Scott and Amenuvegbe (1990); Deaton (1998:25). Even with this, many households still found it difficult to draw a time line between their expenditure and consumption in the immediate preceding month and the two preceding months. There were sometimes tendencies for respondents to include amounts spent or consumed in preceding months. However, as much as possible, care was taken in the field to check the responses. Interviewers had to be very vigilant and constantly reminded the respondents of the time frame of the expenditures – the preceding one month.

4.6 Measuring Household Consumption Expenditure

It is perhaps in estimating household consumption expenditure that the problem of recall shows up most clearly. Various economic theories of consumption such as the Life Cycle Hypothesis and Friedman's Permanent income Hypothesis show that consumption is a better estimator of household permanent income than short run income Stevenson et al (1988:201); Deaton (1998:29-32)]. The basic argument is that household consumption does not generally fluctuate with fluctuations in transitory income. Rather, households tend to smooth consumption to reflect their estimates of their permanent income (putting into consideration the level of household wealth). On this ground many empirical research base estimates of household welfare on household expenditure (or more accurately on household consumption) rather than on their actual income which is often subject to wide fluctuations.

However, the measurement of household consumption is also fraught with difficulties. Apart from the problem of recall mentioned earlier, there are also the problems of imputation and seasonality. Imputation concerns the difficulties households generally experience in translating their non-marketable income (e.g consumption from own production and public goods) into market prices. The problem of seasonality arises from the fact that, in many rural environments, consumption depends to a large

University of Cape Town

measure on the agricultural season. Rural households tend to consume more during harvest periods than during the planting season. These difficulties were experienced in the process of collecting the data used for this exercise.

In order to minimize these difficulties, some strategies were adopted. One, when it was not possible to obtain the amount of monetary expenditure, the actual quantities of items consumed by households were estimated and then translated into money values using local market prices. Second, annual expenditures on household durables were also estimated. Households tended to remember their expenditure on such durable items that they bought occasionally than they do for the consumables. These occasional expenditures included furniture, electric appliances, clothing, kitchen equipment etc. (Bartholome and de Vosti, 1995) used the expenditure on this category of household expenditure alone to proxy household income).

Furthermore, in measuring household per capita expenditure, the usual problem of the difference between adult and child consumption arose. Following the lead of Praise and Houthaker (1971) researchers have adopted the practice of estimating adult-equivalent in consumption. In this study, an adult (defined as a person greater than age 16) was rated a unit in the consumption scale. A child (defined as a person below age 16) was rated 0.8. This, hopefully, took account of the effects of differences in consumption between adults and children.

4.7 Data Management

Altogether about 1,268 households with about 7,800 individuals were interviewed. However, during the process of cleaning the data set, about 331 households involving about 2000 individuals were eliminated. There were remaining 937 households and 5877 individuals. About 60% of those eliminated were because they had one or more missing variables. The others were discarded on various other grounds. Some of these were deemed to be outliers. For example, about 22 households reported travelling distances in excess of 100 km where the average distance was less than 7 km. Some had extra-ordinary health care expending and some others were discarded for giving inconsistent information.

University of Cape Town

4.8 Weighting

A basic assumption of statistical inference is that the sample is representative of the parent population from which it was drawn. If this is not so, then, it becomes difficult to make predictions about the behavior of the population based on the observed characteristics of the sample. For the sample to be representative of the population, every element of the population must have equal probability of being included in the sample. While this is the ideal of statistical inference, in practice there are often reasons for deviating from this ideal of equal opportunity of inclusion. Very often the reason is on grounds of economic costs or because some households refuse to co-operate Deaton (1998:15-16). The practical effect of this is that some household types are over-represented and some are under- represented in the sample. In our case, weighting had to be used because of the initial difficulties in obtaining accurate population figures of the communities in the area.

In designing this survey, the first idea was to collect sample from each cluster community in proportion to its actual census population. In this way every household would have equal probability of being included in the overall sample. However, since the official population figures for the communities were not immediately accessible from the National Population Commission, the exercise started with estimated unofficial figures. Sample proportions were allotted based on these estimated figures. It was mid-way through the exercise that I was able to the official population figures for each of the 15 communities as at 1996. By this time some communities had been over-sampled while some others were under-sampled.

Further more, in the process of cleaning the data some clusters had more missing variables from the respondent households than others. For example, Nsukka urban with a population of nearly 80,000 thousand was found to be under-represented while Edeobala with a population of 15,000 was over-represented. When households with missing observations or observations with extra-ordinary leverages (and outliers) were eliminated, it became obvious that some weighting had to be done to make the sample representative of the entire LGA.

University of Cape Town

Appropriate weighting factors were calculated and attached to the clusters. (The population of each community, the sample size collected and the weights attached are shown in Appendix 1 at the end of the report). Clusters that were over-represented were, consequently, down-weighted while those clusters that were under-represented were weighted up. These adjustments were taking into consideration in the actual estimation process.

4.9 Econometric Soft-ware

The data were first keyed into Excel spreadsheet before they were imported into *Stata* and *Limdep* programs for actual estimation and analysis. *Stata* program would be used mainly for obtaining the descriptive statistics while the NLOGIT program of LIMDEP Version 7.0 would be used for the actual estimation of the three-level nested logit model.

University of Cape Town

Chapter 5

Descriptive Statistics

5.1 Introduction

The preceding chapter described the process of collecting the data. This chapter uses the data collected from that process to provide a snapshot of demand for health care services in Nsukka local government area with the aid of simple statistical tools. The variables that will enter into the functional model are also described with the use of simple descriptive statistics. This, it is hoped, will serve as preview of the regression results that will be the subject of next chapter. An understanding of the nature of the underlying data places the reader in better stead to judge the outcome of the estimation and analysis.

University of Cape Town

5.2 Household Structure and Morbidity

The average size of a household in the study location is 6.2. The smallest households had only one person while the largest households had 24. The standard deviation is 2.86 and the variance is 8.17. The skewness and Kurtosis show that the distribution of individuals in the households was non-normal. It is largely skewed towards the y-axis. This implies that more households were of less than the average size. Skewness is generally a characteristic feature of most socioeconomic data. They are very often skewed towards the y-axis. This feature reflects the inequalities in the society and economy Mukherjee et al. (1998:97)

The ratio of children to adult is about 0.80 which suggests that within the scope of our earlier definition of the child as person below age 16, there are slightly more adults than children population in the survey area.

The population of households in the data is made up of two major groups: those households that had sick persons and those that had no sick person. Out of the 937 household retained in the final data set, 617 or about 65.85 percent reported having had a sick member in the month preceding the interview. Table 5.1 presents a snapshot of the demographic structure of the household and prevalence of morbidity in the population.

Table 5.1 **Household Structure and Morbidity**

	Households		Individuals	
	No	Percent	No	Percent
Adult male			1464	25.1
Adult female			1787	30.64
Male child			1346	23.1
Female child			1235	21.17
Total			5,832	100
Sick	617	65.85	890	15.26
Non-sick	320	34.15	4,942	84.74
Total	937	100	5,832	100

University of Cape Town

5.3 Eduction

The survey showed that the average number of years of formal education in the sample was 7.16. For the purpose of this presentation the average number of years households attended formal education has been categorized into: (1) primary (years of education 1- 6); (2) secondary (7- 12) and (3) tertiary (> 12). Table 5.2 shows the distribution of the sick according to average number of years of formal education in the households and according to the health care alternatives they chose.

Table 5.2 **Education Distribution and Choice among Health Care Alternatives**

(a)

Average Level of HH educ	Frequency	Percent
1 – 6 years	294	47.65
7- 12 years	268	43.44
> 12	55	8.91
Total	617	100

(b)

	PRIMARY	SECONDARY	TERTIARY
NOCARE	3.40	3.36	1.82
HCARE	7.82	6.34	14.55
PUBH/C	24.49	31.34	27.27
PRVTH/C	25.17	29.85	32.73
T/MED	7.14	7.46	5.46
P/MED	31.29	21.64	18.18
TOTAL	100%	100%	100%

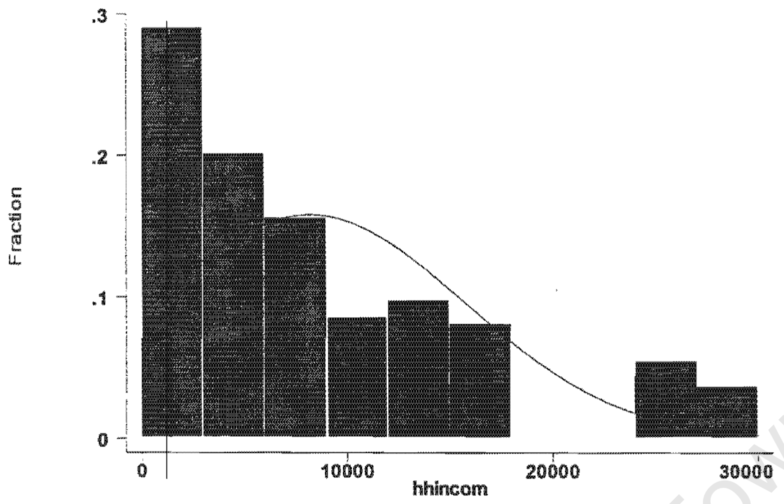
Table 5.2 (b) illustrates the fact that as the level of education of a household increases the preference for public and private hospitals and clinics increases. On the other hand, households with lower levels of education seem to utilize more of the other health care alternatives, particularly the patent medicine alternative. It should be noted, however, the that nearly 30% of the people who visited the traditional medicine practitioners was from the highly educated class. This may suggest that this group of providers were not irrelevant as is generally presumed.

5.4 Household Income and Per Capita Expenditure

The mean household income in this area is about N6,323.00 ($75.00) per month. But income is not evenly distributed. The standard deviation is about 6135. Household income tends to be skewed towards the y-axis implying that income is not normally distributed. Figure 5.1 shows that more households fall below the average income than households above it.

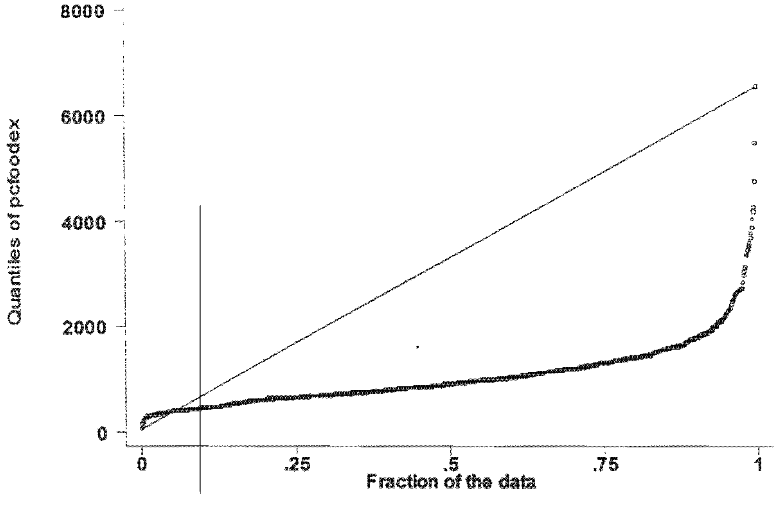
University of Cape Town

Fig 5.1 **Income Distribution Skewed towards the Vertical Axis**



Quantile plot is often useful for displaying a set of data in graphic form. It helps us to estimate some other useful statistics such the median and inter quartile range (IQR). This device is used here to illustrate the distribution of per capita food expenditure (Pcfoodex) in the population under study. For example, it can be seen from figure 5.2 that more than 90 per cent of the population has per capita food expenditure that lies within the first quartile. That means that 90 per cent of the sample spend less than N2000.00 on food. It shows the extent of economic inequality in the area.

Fig 5.2 **Quantile Plot of Per Capita Food Expenditure**



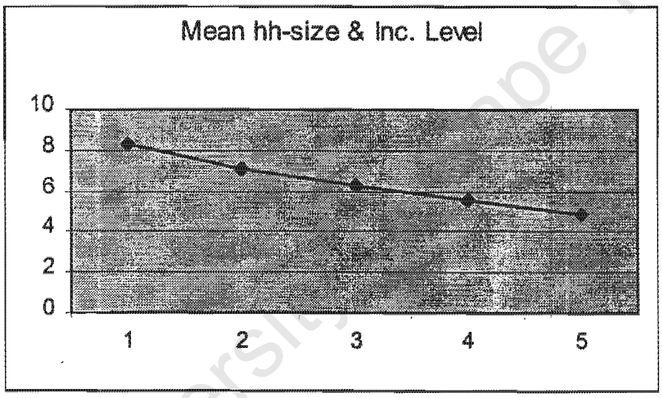
University of Cape Town

The quantile-normal plot in Fig 5.2 shows the extent of inequality in consumption expenditure. It could easily be inferred from the figure that income is very unevenly distributed.

5.5 Household Size and Income Distribution

Figure 5.3 shows how consumption expenditure level varies with the average household size. Expenditure quintiles are plotted on the x-axis while average size of households in each expenditure quintile is plotted on the y-axis. The graph clearly shows that as the size of the household increases the level of consumption per head decreases

Fig 5.2 **Distribution of Expenditure Level and Household Size**

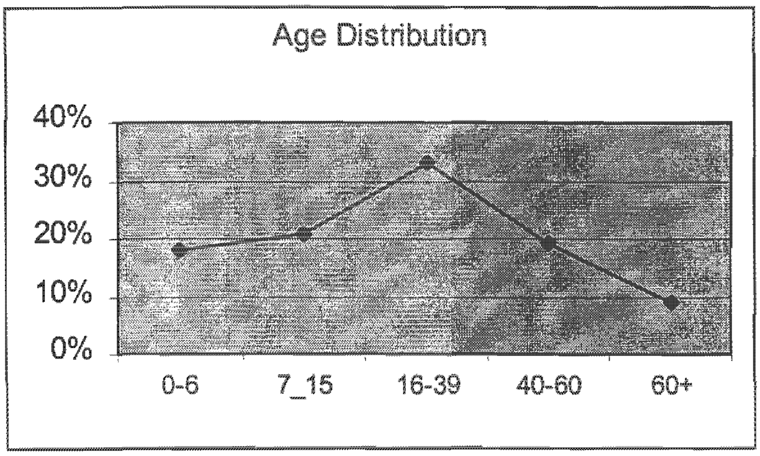


University of Cape Town

5.6 Age Distribution among the Sick

Among the households that reported sick, the age distribution is shown in figure 5.4 where agegrp1 = ages 0-6, sagegrp2 = 7-15, sagegrp3 = 16-39, sagegrp4 = 40-60, sagegrp5 = 60+

Fig 5.4 **Age Distribution of the Sick**

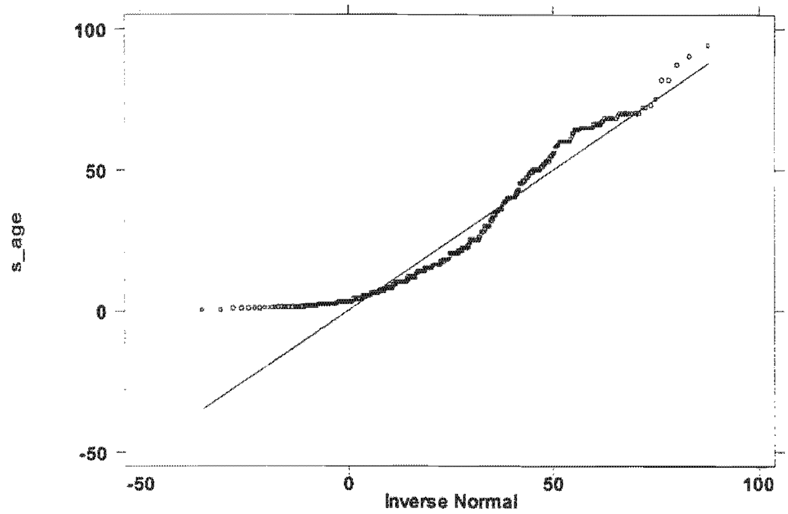


This figure shows that the greatest number of sick people was reported among the age group 16-39. The sick among this category make up the about 33 per cent of the total sick people. Age groups 0-6 and 7-15 constitute 18 and 21 per cent respectively. The age range 40 – 60 represents about 19 per cent while those above 60 years represent about 9 per cent of the population of the sick. The preponderance of sick people within age range 16-39 may not necessarily mean that this group is more sickly than others. It may rather be an indication that the population is skewed towards this age range.

Quantile-normal or normal probability plots are used to compare an empirical distribution with a theoretical normal distribution having the same mean and standard deviation. This help to show the extent the empirical distribution departs from normality. Figure 5.5 helps to illustrate the extent age distribution among the sick vary from the normal. If the age distribution were perfectly normal, the observation would like along the diagonal.

University of Cape Town

Fig 5.5 **Quantile-Normal Plot of Age Distribution of the Sick**



5.7 Severity of illness and Demand for Health Care

People's need for health care services often arise out of their perceived severity or level of seriousness of their health condition. Most people would not generally demand treatment for what they consider minor illness. This is especially so if treatment is not costless (for example if there is no universal social health insurance coverage) or if there is some barrier (for example financial or distance barrier) to consumption of health care services.

In the survey data, the level of severity of the sickness was proxied by the number of days the individual was unable to perform his or her normal daily activities on account of sickness. For the purpose of describing the data, Severity is here classified under five categories. A person is classified under Severity Level 1 if he or she did not abstain from normal duties at all. A person who missed one to seven days of duties is classified under Severity Level 2. One who missed between eight and fourteen days is classified under Severity Level 3. Similarly, if the person was unable to carry out normal duties for between fifteen and twenty-one days in the preceding month, he or she is classified under Severity Level 4. Finally, if the person had been unable to perform his or her usual activities through the preceding month, the person is

University of Cape Town

classified under Severity Level 5. The table below shows the number of people under each of the categories.

Table 5.3 **Severity Levels of Sick Persons**

Severity Level	Frequency	Percent	Cumulative
1	141	23.19	23.19
2	213	35.03	58.22
3	110	18.09	76.22
4	54	18.88	85.20
5	90	14.80	100

Of those who indicated sickness, 582 representing about 95% of the sick claimed to have sought treatment either in a health facility or at home. The others did not seek any form of health care.

5.8 Choice of Health Care Facility

Table 5.4 reflects the nature of decisions households in Nsukka LGA made in a month regarding choice of health care facilities. The first category in the choice list is the No Care category. This group reported sick but did not seek any form of treatment. In a different category were 48 households who decided to treat their sick at home.

The next category comprised those who were treated in government owned health facilities. These facilities include two state general hospitals and twenty local government Primary Health Care clinics in the area. Households that chose the facilities in this category from among those interviewed numbered 171 or about 29 percent of those who demanded health care services. About the same number of households demanded care in private health facilities. The private health facility category comprised of the 20 private and mission hospitals, all the 25 privately owned and operated clinics, and all the 10 privately owned maternity homes in the area.

University of Cape Town

Table 5.4 **Distribution of the Sickhouseholds according to Health Care Facility**

FACILITY	FREQUENCY	PERCENT
NO CARE	20	3.35
HOME CARE	48	7.06
PUB.HOSP/CLINICS	171	27.80
PRVT HOSP/CLINIC	172	27.97
TRAD. MEDICINE	44	7.15
PHARM/PAT MED	160	26.02

Some 44 or about 7 percent other households preferred to treat their sick in the traditional medicine homes. About 27 percent of all the households with sick persons demanded the services of pharmacy shops and patent medicine stores. This number indicates the relevance of this category of suppliers in the over all health care services in the area. Two major reasons could be adduced for their large influence. One, they provided quick and immediate services without long formalities. Second, they were relatively more accessible in terms of distance than other categories of health care providers in the area. The low number of seekers of traditional medicine may not necessarily indicate the waning influence of this group of health care providers in the health care market of the area. The fact that their charges are high relative to other facilities and that they are patronized by the rich and educated may rather suggest that they are preferred for certain types of care that other facilities do not provide. There were indications from the interviews that those who patronized the services of traditional healers were mainly those who suffered from illnesses that defied orthodox medicine such as poison and orthopaedic disorders.

5.9 Distance, Travel Time and Waiting Time.

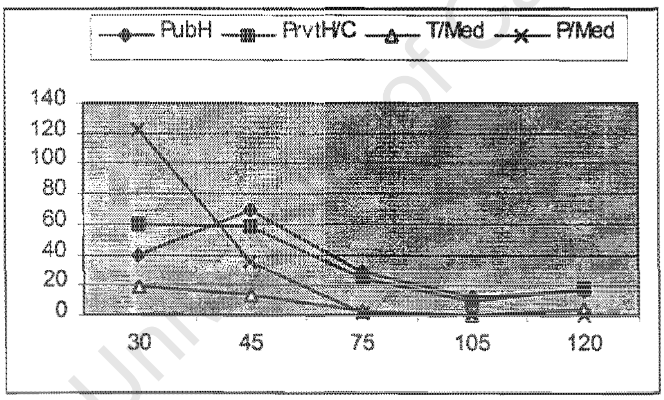
These three variables are generally considered important barriers to access and consumption of health care services. The distance a household has to travel to obtain treatment may discourage it from demanding health care services. It also impacts on the choice of health facility from which to buy care. In the area surveyed, the average distance that people usually traveled in order to get treated was 3km. The median was

University of Cape Town

1km. The standard deviation was 6km and variance. The distance was not normally distributed. Some households traveled more than 50 km in search of health care.

Travel time varied according to the location of the household, the distance of the household from the chosen health care facility, the severity of illness and means of transport. The most common means of transport were buses and cyclists. In many cases however, households walked to the facility. About 59% of all the households that demanded health care traveled less than 30 minutes by whatever means they used to obtain treatment. About 29% of the households traveled between 30 minutes to 1 hour to access health care. Some of the households traveled for between 1 hour and 1 hour 30 minutes to access the facilities of their choice. About 3% and 2% respectively had to travel in excess of 1 hour 30 minutes (but less than two hours) and 2 hours respectively in order to demand health care services.

Fig 5.6 **Distribution of the Sick According Waiting Time at Health Facility**



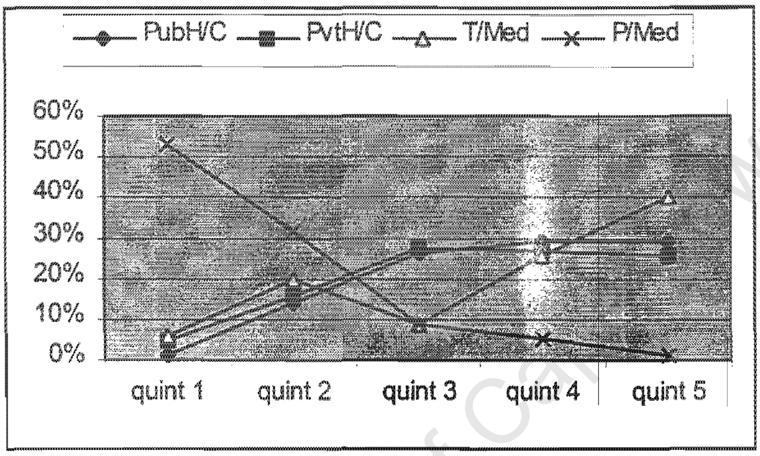
The mean waiting time in public and private health facilities averaged about 70 minutes and 65 minutes respectively. The mean waiting time at traditional medicine homes was about 50 minutes and the average waiting time at the pharmacists' shop was about 35 minutes. Figure 5.6 plots the number of sick people on the vertical axis and the waiting in facilities on the value axis. The time is measured in minutes. The figure shows that most sick people who sought care from patent medicine category spent less than 45 minutes on waiting time. The waiting time in public facilities is shown to be slightly above that in private facilities.

University of Cape Town

5.10 Treatment and Total Financial Costs

Treatment cost was estimated by the cost of registration in the facility, the cost of diagnosis and the cost of drugs. The treatment cost varied between N10.00 and N26200.00. The median value was N720.00 and the mean was N1530.00.

Fig 5.7 **Distribution of Treatment Cost According to Facilities.**



Treatment expenditure was partitioned into five treatment cost quintiles as follows: quint1: N0.00 – N132.00; quint2: N133.00 – N490.00; quint3: N491.00 – N1050.00; quint4: N1051.00 – N2200.00; quint5: N2201.00 – N26200.00. All the sick people that visited a given facility were divided according to treatment expenditure quintiles. The expenditure quintiles were then plotted against the percentage of people in that facility that fell within the given quintile. Fore example, figure 5.7 shows that more than 50% of the sick people that visited patent medicine stores fell within the first treatment expenditure quintile while almost 0% of the people in that facility fell within the 5th expenditure quintile. On the other hand, less than 10% of the people that visited either the public or private hospitals and clinics fell within the first treatment expenditure quintile. Similarly, about 40% of the people that visited traditional healers fell in the 5th expenditure quintile and about 25% fell within the 4th quintile. This again, indicates that either the cost of treatment in traditional medicine is very high relative to other providers or that sicknesses referred to this category of health care providers are specialized. It is further interesting to note the similarity between the choice structure of the public and private health care facilities. The choice between public health and private health facilities are very similar in all the quantiles.

Total financial cost includes treatment cost as well as any other financial expenditure the household may have incurred in the process of obtaining care for the sick member. For example, transport costs both for the patient and any other household member that accompanied the sick to treatment facility. This expenditure varied between N10.00 and N34000.00 The mean and median expenditures were N821.60 and N810.00 respectively. The standard deviation was N3028.61

Referring to the pie charts in figure 5.8, the first chart, (a) is plotted with the household income and food expenditure of households that did not demand health care services. Chart (b) is plotted with similar data from households that demanded health care services including the expenditure on such services. The purpose is to show the overall effect of health care demand expenditure seems to be a decrease in the nutritional intake of the household.

University of Cape Town

Fig 5.8 (a) **Food Expenditure as a Proportion of Household Income: No Sickness Reported**

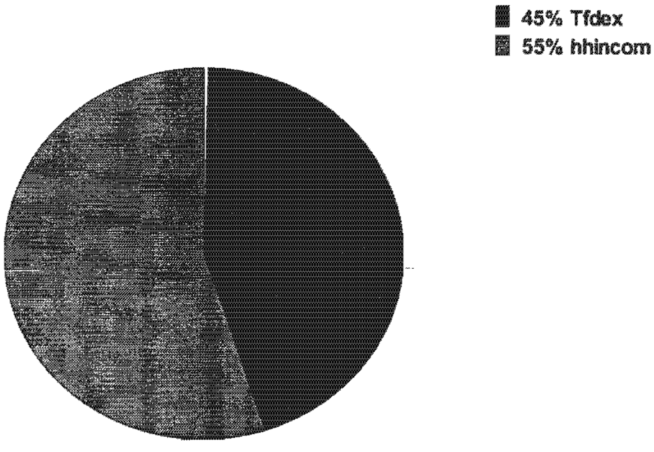
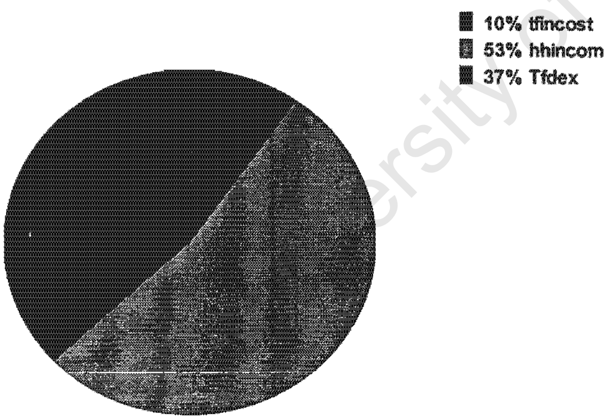


Fig.5.8(a) shows tha households that did no report sickness spent, o the average, about 45% o household income on food

5.8 (b) **Food and Financial Cost of Treatment as**



Proportions of Household Income: Sickness Reported

Fig 5.8 (b) shows that on th average, households tha sought treatment for sicknes spent about 10% o household's income o treatment. This led to shortfall of about 8% o consumption expenditure an only 2% shortfall on othe expenditures. Thus househol nutritional intake seemed t have been sacrificed in orde to finance cost of health care.

5.11 Quintile Distribution of Expenditure and Choice of Facility

The following tables group the households that chose each health care alternative according to expenditure quintiles. Each of the five quintiles consists of about 123 observations. The percentage of the quintiles are shown across each type of choice alternative for households that chose that alternative.

Table 5.5 **Choice of Health Care Alternatives According to Expenditure Quintiles**

	NOCARE	H/CARE	PUB.H/C	PRVTH/c	T/MED	P/MED
Quintile1	35.84	25.63	13.85	17.31	11.70	28.23
Quintile 2	5.00	22.88	11.74	25.04	31.98	21.13
Quintile 3	14.82	14.47	20.39	23.68	24.90	16.52
Quintile 4	24.91	16.68	35.10	18.05	9.14	15.37
Quintile 5	19.44	20.34	24.65	15.91	22.28	18.75
Total	100	100	100	100	100	100

From table 5.6 it can be seen that across quintiles, the highest percentage in the class of sick people who chose the NO CARE alternative came from the lowest quintile. Similarly, among those that treated their sickness at home, the highest percent came from the first quintile in the expenditure variable. Among those who were treated in public health facilities, the highest percent came from the 4th quintile. Surprisingly, among those who chose private facilities, the highest percent came from the 2nd quintile. The highest percentage of those who patronized the patent medicine health care providers came also from the lowest quintile.

University of Cape Town

Although these statistics do not lead to definitive conclusions, these statistics may well be indications that households from higher per capita expenditure brackets may have greater leverage in accessing better health care than those in the lower brackets.

This chapter has provided some preliminary insights into the nature of health care demand behavior of the people in Nsukka local government area. In chapter six that follows I shall present and analyze the outcome of the regression exercise. Chapter seven will consider the policy suggestions that will arise from the estimated model.

Chapter 6

The Results

6.1 Introduction

The preceding chapter provided relevant descriptive statistics of Nsukka local government area. In this chapter, the results from estimating a three-level nested logit model of demand for health care services in Nsukka local government area are presented. Statistical and econometric tests are used to analyze the robustness of the parameter estimates. Some of the problems associated with estimating nested logit models are also described.

Results

Table 6.1 below presents the output of the estimated three-level nested model of demand for health care services in Nsukka local government.

Table 6.1 **Parameter Estimates of a 3-Level Nested Logit Model of Demand for Health Care Services**

Variable	Parameter Est.	t-ratio	p-value
Set A : Utility Function Parameter Estimates			
DISTANCE	-0.108	-3.062	.0022
LOGINC	0.079	3.288	.0010
LOGPCFDX	0.123	6.330	.0000
HH_SIZE	-0.115	-4.644	.0000
SEVERITY	0.160	5.988	.0000
W_TIME	-0.069	-1.047	.2952
LTMTCOST	-0.051	-0.852	.3942
NAVYRSED	0.105	6.781	.0000
S_AGE	0.084	5.525	.0000
LOC	0.660	11.240	.0000
Pseudo R^2	0.27		
Set B : Inclusive-value Parameter Estimates			
Branch $(1-\sigma_k$	0.438	0.000	1.000
Limb $(1-\sigma_j)$	0.618	0.000	1.000
Trunk $(1-\sigma_i)$	1.000	0.000	FIXED

University of Cape Town

The results are divided into two sets of parameter estimates. The first, Set A, is the set of estimated utility function parameters with the Pseudo $R.^2$ Set B is the set of estimated inclusive-value parameters. The t-ratios and p-values are also included.

6.3 Overall Assessment of the Model

Perhaps the first question to ask in the assessment of the estimated model is whether the model that includes the variables gives us more information about the response variable than the model that contains only the constant term. In other words, does the model with variables included predict the underlying data generating process better than the model with only the constant? This comparison is achieved by using the likelihood ratio test (see note 1, Appendix 2).

In this model the log likelihood function for the unrestricted model, that is, the model with all the beta coefficients is -1121.35. The restricted model, that is the model with constant only, is –1534.64. The likelihood ratio (LR) test of validity of restriction, that is, a test that all the variables together are insignificant gives a Chi square of 826.58 with 10 degrees of freedom. The constraint was therefore rejected. This implies that all the variables as a group (but not necessarily individually) are significant at 95% confidence interval. That is, the unrestricted model gives better prediction of the data generating process than the restricted model.

With respect to individual coefficients, the Wald or t-statistic : $W_i = \dfrac{\hat{\beta}_i}{SE(\hat{\beta}_i)}$

is usually used to assess the significance of individual variables. This statistic simply divides the estimated parameter by its standard error. A two tail test of the null hypothesis at five percent level of significance either accepts or fails to accept the significance of the variables in question (see note 2, Appendix 2)

6.4 Interpretation of the Coefficients of the Estimated Model

The estimated coefficients in Set A of Table 6.1 reflect the impact of the variables on the utility that households derive from seeking care from health care facilities. As was indicated earlier, each household that decides to seek health care for a member is

University of Cape Town

assumed to compare the utility it will derive from one against the other options taking into account facility attributes and household variables. The household variables include its own health production function, income and demographic variables. The household then chooses, in the light of these variables, that option from which it hopes to maximize its utility function

There are several methods of interpreting estimates of probability models Long (1997:61-82). However, it has been noted that the easiest and most useful way is that of odds and odds ratios Liao (1994:14-15); Hosmer and Lemeshow (1989:41). The odds measure how likely or unlikely an event would occur (see note 3, Appendix 2)

There are ten variables in Set A. It is to be observed that there is no constant term shown in the table of coefficients. The reason is given in note 4, Appendix 2. The first variable in the set is DISTANCE measured in kilometers. This is the distance households have to travel to reach to the chosen health care facility. This variable was measured in kilometers. Obviously, this variable represents a dis-utility to the household that is seeking care. The coefficient on it is, therefore, negative as expected. Statistically, the estimated coefficient of the variable is significant at the usual level of 95 per cent confidence interval. (This same confidence interval is applied in the rest of our hypothesis tests). By taking the anti-log of the coefficient -0.108 we obtain 0.897. This is the decrease in the odds of seeking health care from health care facilities for every additional kilometer in distance to a health care facility given that a household decided to seek treatment and given that it sought formal care. This result indicates that distance to facilities is a major hindrance to access to health care utilization in the area under study.

Log of income [LOGINC] has a positive sign as expected. Its coefficient 0.079 is statistically significant as shown by the t-ratio or more precisely by the probability value [p-value = .0010]. The implication is that when we are comparing the effect of income on the choice between two health care alternatives, say k_1 and k_2, the odds of choosing one against the other will change by a factor exp(0.079) for every unit change in the log of income, holding every other variable constant. If, for example we consider the choice between home care and care in private hospital, this result

University of Cape Town

indicates that for every additional unit in the log of household income the odds that the household will seek care from a private hospital increases by a factor of exp(0.079). In other words, richer households have higher odds than poorer ones of seeking health care when they are sick. This implies also that low income is likely to be a hindrance to access to health care in the area.

Log of per capita food expenditure [LOGPCFDX]: This variable also has the expected sign and its parameter estimate of 0.123 is statistically significant. The anti-log is 1.13. The interpretation is essentially the same as in log of income but here it applied to food expenditure. Every 1% increase in log of food expenditure increases the odds seeking health care from a facility by 1.13%.

The demographic variable, household size [HH_SIZE] has a negative coefficient. The magnitude is statistically significant. The negative sign is also consistent with expectation. The coefficient is exp(-0.115) and the anti log is 0.891. Large households seem to decrease the odds of seeking health care. In other words the larger the household the less likely that a sick member of that household will be able to utilize a medical facility.

The severity of illness [SEVERITY] variable has a positive sign. This conforms to apriori expectation. The magnitude of the coefficient is exp(0.160) = 1.73. It is also statistically significant showing that severity is a major consideration for household when they seek health care services. An increase in the number of days one is unable to perform one's normal duties due to sickness will naturally prompt the household to seek medical attention for one.

Waiting time [W_TIME] has the expected negative sign but it appears to be an insignificant factor in health care decision in the area under study. Some other empirical studies have found a similar result [for example, Acton (1975); Akin et al. (1986)]. The non-significant influence of this variable on the demand for health care can be understood from the perspective that while a consumer may consider the time spent in obtaining treatment as important, he or she may place higher premium on other qualities of the facility. For example, a consumer may consider the time spent waiting for treatment as secondary to the quality of drugs, and attention obtained when he or she eventually gets treated. It may also imply that because of the number of facilities in the area patients do not have to wait long before they get treated. The latter may be more likely since from the descriptive statistics it was observed that most patients did not have to wait more than 45 minutes to obtain treatment.

The average years of formal education [NAVYRSED] of the household shows up as a positive and significant factor in the demand for health services from the sample included in this study. The average increase in the odds of seeking care from a facility increases by a factor of exp(0.105) = 1.11 as the average years of formal education of the household increases. However, some other studies have found education to be a non-significant factor in a household's health care utilization decisions. For example, Gertler and der Gaag (1990) using a nested multinomial logit model found this demographic variable insignificance in the decision of the individual patient to seek formal care in rural Cote d'Ivoire. Ellis and Mwabu (1991) found it to have a negative relationship with reporting illness while it has a positive relationship with the decision to seek formal treatment. It is arguable that the non-significance of the variable in the Ivorian study is due to the very low level of education in the area. (The average length of years of education for individuals in the sample was about one year, with little variation). This may be contrasted with household average of 7.16 years of formal education in the sample from Nsukka local government used in this study. Therefore, the contrast in the findings may not be altogether surprising.

The log of treatment cost variable [LTMTCOST] represents the financial cost to the household of seeking health care for the individual member of the household. This variable was expected to be negative and significant but it turns out to have the expected sign but statistically insignificant as noted in the review of literature. While it is economically reasonable to expect cost of obtaining treatment to be a hindrance to health care utilization quite a number of studies have found the this variable insignificant. One possible explanation for this aberrant behavior is that health care market suffers from information asymmetry between the buyer and the supplier. Consumers often rely completely on the prescriptions of the health care provider. In such an environment it is easy for consumers to use cost as the parameter for quality

University of Cape Town

of drugs. In such a situation prices my give contradictory signals to buyers. Higher prices may be seen as indicative of greater quality of care. In that case patients may attach higher value to drugs that cost more financially. This may attenuate the consideration for cost.

The age of the sick person [S_AGE] appears to influence the health care decisions of the household significantly. The model shows that there are greater odds by a factor of exp(0.084) = 1.09 in health care utilization as age increases. Although this may not be altogether expected, the result is consistent with some other studies that have found greater allocation of resources to working adults who must be in health condition in order to fend for the household and other dependants. The economic sustenance of the household depends on these working adults World Bank Report (1993); Sauerborn et al. (1996).

Location [LOC] is a dichotomous variable which assumes the value of 1 if the household was from Nsukka urban and zero otherwise. The coefficient of the dummy is positive and statistically significant. The calculated difference in the odds of a person from Nsukka urban area and his/her counterpart in the village in seeking health care from a facility is exp(0.660) = 1.93. This implies that a city dweller is about 2 times more likely to seek treatment from a facility than a village dweller judging from our sample. This is consistent with findings from other similar studies. For example, the World Bank Report 1993:69 shows that households living in urban areas seek and obtain health care services more often that their counter parts in rural areas.

Table 6.2 below summarizes the impact of the estimated variables on the odds of choosing a health care alternative.

University of Cape Town

Table 6.2 **Impact of Variables on the odds of utilizing Health Care Facilities**

Variable	Change in the odds/ unit change in Variable	Significant/Insignificant Y/N
DISTANCE	0.897	Y
LOGINC	1.082	Y
LOGPCFDX	1.131	Y
HH_SIZE	0.891	Y
SEVERITY	1.174	Y
W_TIME	0.933	N
LTMTCOST	0.950	N
NAVYRSED	1.111	Y
S_AGE	1.088	Y
LOC	1.935	Y

In general, it can be said that the estimated parameters of the variables included in the model are mostly consistent with economic theory. The findings are equally revealing. They indicate the variables that constitute dis-utilities to health care consumption. Higher income households are more likely to seek care from facilities than poor households, older people seem to get treated more than younger people,
big household size constitutes a barrier to health care consumption, more educated households are more likely to seek care from facilities when they are sick. It is in conformity with expectation that distance would be a hindrance to access to facilities but it is striking that waiting time and treatment costs do not seem to be serious barriers to health care consumption in the area. This is even more striking when it is considered that the area under study is not an economically wealthy area.

University of Cape Town

6.5 The Estimated Inclusive-Value Parameters

The estimates of the inclusive-value parameters are $(1-\sigma_k)$ and $(1-\sigma_j)$ for the branch and limb levels respectively. The trunk level inclusive-value parameter $(1-\sigma_i)$ should, obviously, be 1.0 since it is assumed that no other choice is available at this level.

The null hypothesis (Ho): $(1-\sigma_k) = 0$ was tested using Wald test. The estimated

$(1-\sigma_k)$

value was constrained to 0 but this constraint was obviously rejected. The value must therefore be assumed to be different from zero and different from 1.00. This is a rejection of the multinomial logit model in favor of the nested logit model. It also implies that the model is consistent with random utility maximization as explained in chapter 3. This implies further, that the decision about which health facility to seek care from is not independent of the prior decision to seek formal or informal treatment. Put in another way, an average household's decision to seek formal care from a health care professional or to resort to informal care is influenced by the availability of health care facilities. For example, a household may easily decide to consult a medical doctor if a hospital is next door. On the other hand if the hospital is very far away, or waiting time is very long, the household might decide to put up with the pain of illness or treat at home until it becomes very severe. Thus, the finding suggests that in the area under survey, access to health care facilities may influence people's decision to seek care from health care professionals as opposed to treating themselves at home. However, the fact that the inclusive-value parameter estimate at this level is different from 1.00 implies that there is no perfect correlation between any of the choice alternatives.

University of Cape Town

The hypothesis that the inclusive-value at the limb level $(1-\sigma_j)$ is equal to .00 or equal to 1.00 was similarly tested using the same Wald test. The constraints of equality were rejected. This implies that the average household's decision not to seek treatment or to seek treatment (whether from a professional health care facility or at home) given that it reported illness, is not an independent decision. The decision to seek treatment may be influenced by the decision either to treat at home or to go to a health care professional.

Furthermore, as Maddala (1983:73) shows, the values σ_k and σ_j on the inclusive value parameters indicate the correlation between the choices made within each subgroup – at the lowest level of choice (among facilities) and at the branch level (whether to seek or not to seek treatment). That is to say, these values on the inclusive value parameters provide estimates of the similarity or dissimilarity between the choices made within the subgroups at each level. Thus, for example, if $\sigma_k = 1$ implying that all the alternatives within the choice set (eg the facilities) are actually identical. In making a choice in this case the household will perceive all the alternatives within this subgroup as if it contains a single alternative. At the same time if the inclusive value $(1-\sigma_k) = 0$ it implies that the decision to seek formal care from professional health care providers or to seek informal care at home is not influenced by the availability of health care providers. It further implies that the error term arising from the utility function of seeking formal treatment does not correlate with the error term from the utility function of seeking informal care.

In more technical terms, if we assume *h* to represent informal and *f* to represent formal treatment, the implicit assumption when the inclusive value is 0 is that $\sigma_{hf} = 0$. That is, the correlation between the error term in the decision to seek formal treatment is un-correlated with the error in the decision to seek informal treatment. From this perspective, the test for the validity of the assumption of independence from irrelevant alternatives (IIA) is equivalent to testing that $\sigma = 1$. In the present model, using the Wald test, the values obtained for both σ_h and σ_f are significantly different from 1. (Hauseman and McFadden (1984) suggest the test of a reverse implication of IIA by eliminating one or more alternatives from the choice set to see if the underlying behavior would change. If the coefficients remain the same, then it is evidence of the validity of the IIA assumption; if not it is evidence that the assumption holds. In both cases the result obtained from this model suggests that:

(a) the facilities – Public Hospital/Clinic, Private Hospital/Clinic, Traditional medicine and Pharmaceutical/Patent Medicine stores are closer substitutes (but not identical) for one another more than any of these is with home treatment .

(b) Formal care (treatment at a facility) and informal care (home treatment) are closer substitutes than are the decision to seek care and not to seek care.

The inclusive value at the trunk level as noted earlier is obviously a unit and it is a fixed value. There is really no choice made at this level and the probability of reporting sick is 1.00 since. The model estimated the sample of only the sick. The inclusion of non-sick group in the model would result in extending the model to a fourth level.

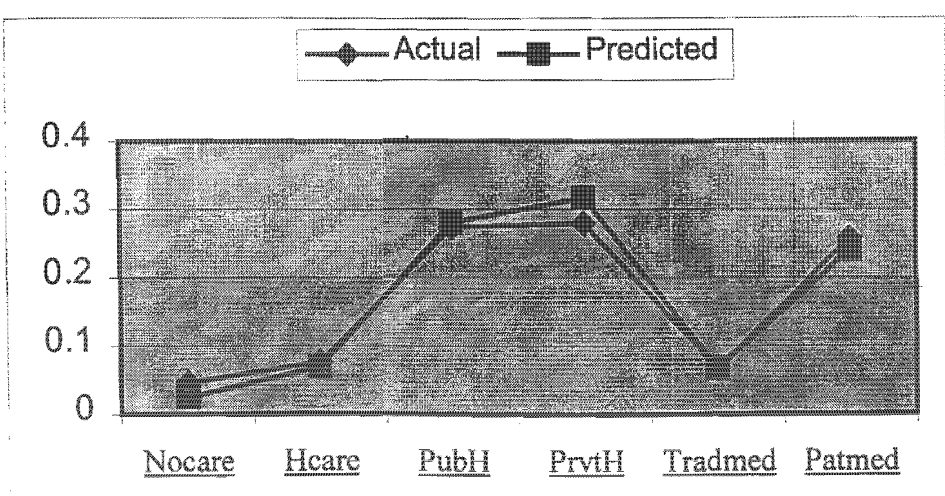
As was indicated earlier in chapter 3, a major test of the correct specification of the model is that the inclusive-value parameters should lie in the range of zero to one. Maddala (1983:73) states that if the estimated inclusive-value parameters lie outside the unit interval, we should view this as evidence of specification error and re-examine the model. Fortunately, the model that was estimated does not seem to have this flaw as the inclusive-values lie within the unit interval. One may therefore conclude that the model does not contain any major specification error. To confirm these results, the inclusive-value parameter estimates were constrained to equal one. Under the Wald test this constraint was rejected. Thus we conclude that the values are different from one. The pseudo R^2 is low (approx. 0.27) indicating perhaps that there are other important factors that influence the household health care decisions that were not captured by the model, although this statistics is not usually high in probability models unlike in linear models. The pseudo R^2 in logit models do not have the same meaning of explained-unexplained as in OLS regressions. Hamilton (1998:231) (Further details on the goodness-of-fit measure are given in note 6, Appendix 2).

University of Cape Town

6.6 Post Estimation Analysis of Predicted Probabilities

The results of the means of actual and predicted probabilities are plotted in the graph below. To obtain these results, LIMDEP was asked to list the predicted probabilities for all the alternatives for each observation. The mean predicted probability for each

Fig 6.1 Actual and Predicted Mean Probabilities of Choice of Facility

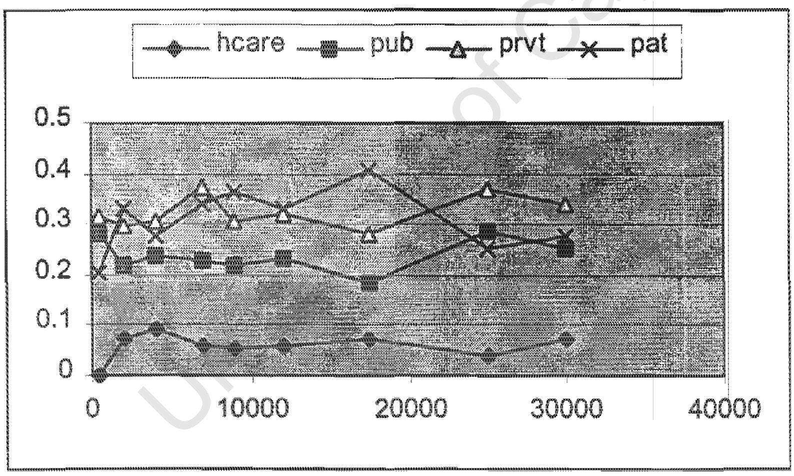


alternative was then calculated. For the actual, this was obtained by simply dividing the number of households that chose a particular facility by the total number of observations, that is 617. The graph suggests a close match between the means of actual choices made and the means of the predicted probabilities. Probability values lie on the vertical axis while the facilities are plotted on the horizontal axis.

The figure below shows the plot of income against predicted probabilities of choosing a given type of health care. In order not to clump the space the options No Care, and Traditional Medicine with relatively small numbers have not been included in the in the chart. On the vertical axis are the predicted probabilities while the X-axis is the different categories of income.

Fig 6.2

Household Income and Predicted Probabilities

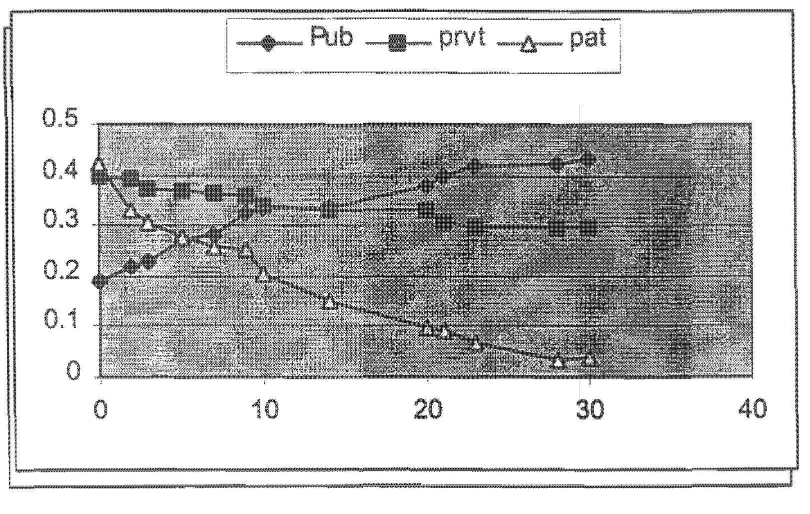


What seems obvious from the graph is that at almost all levels of income, the predicted probability of home care is much lower than that of any of the other three. Furthermore, controlling for every other variable, it appears that at almost all levels of income the probability of utilizing either a private hospital or clinic, and patent medicine is high than that of public health facilities.

Altogether, the graph seems to suggest that although statistically significant, the influence of income in the choice of medical facility in the area surveyed is not very decisive or seems to have been attenuated by the influence of some other factors. It was expected that the predicted probabilities especially for the private health care providers will rise steeply as income increases while the probability of choosing the home care option will fall as income rises. That this does not happen may be an indication that both the rich and poor delay treating their illness until it becomes sufficiently serious to prevent them from carrying out their normal duties. Alternatively, it may mean that both the rich and poor resort to first aid at home before consulting a medical facility. In the event of any of these, the impact of income on the demand for health care may be reduced though not altogether removed. In such a situation other variables may assume greater importance than income in the utilization of medical facilities. This suggestion is reinforced by the graph of severity against predicted probability of health care utilization shown below.

Figure 6.3 plots the predicted probabilities of utilizing the services of a public health facility, private health facility and patent medicine dealer on the vertical axis. The levels of severity of illness are shown on the category axis. The levels of severity range from zero to 30 representing the number of days the patient was unable to carry out normal duties.

Fig 6.3 **Severity and Predicted Probabilities**



University of Cape Town

As can be seen from the graph, the probability of consulting a public facility rises with relatively fast as the level of severity of the illness increases. At zero level of severity, the probability is about 0.20. As the level of severity increases, the probability of consulting a public health care provider increases steadily, tending towards 0.50 as the level of severity approaches the maximum.

On the contrary, the probability of patronizing a pharmacist or a patent medicine dealer starts off at a little above 0.40 when severity is zero. It declines steadily as the level of severity increases; reaching almost zero as the level of severity approaches its limit. This trend suggests the 'first-aid' importance of this group of health care suppliers in Nsukka local government area.

It is surprising however, that the probability of consulting a private health care provider shows a negative trend. It declines slowly from 0.40 to 0.30 as severity increases from zero to maximum. The expectation was that based on their often-claimed higher health care delivery efficiency, patients would resort to this group of providers for faster results on their health condition. That this is not the case suggests a closer investigation into the relative efficiencies of the different categories of health care providers operating in this area.

The major problems in the estimation of the model are discussed in note 6, Appendix 2. These relate mainly to the problems of identification of the model and numerical methods used in estimating maximum likelihood functions. Essentially, identification is a problem of appropriate formulation of the model that may lead to difficulties in estimation. In chapter 7, which is the last chapter of this report, the policy suggestions arising from the results will be discussed.

University of Cape Town

Chapter 7

Policy Considerations

7.1 Introduction

The preceding chapter highlighted the main findings from estimating a nested demand function for health care services in Nsukka local government area. Essentially, the results show the importance of such variables as distance, income, household size, food expenditure, level of household education among others. The results represent empirical estimates of health care demand behavior in the area. In this chapter I shall first, indicate some of the policy implications of these estimated parameters in the wider context of social welfare maximization before I go further to draw general policy conclusions from the model.

7.2 Public Policy as Social Welfare Maximization

One of the primary goals of public policy is to maximize social welfare. Just as firms behave in ways that maximize their profits and individuals in ways that maximize their utilities, the public policy maker aims at maximizing the social welfare function. However, like every other economic agent, the public policy maker is also constrained by scarcity of resources to satisfy competing needs. Rational decisions, therefore, have to be made in respect of allocation of resources to these competing needs in order to achieve maximum social benefit. Thus, while it is recognized that health is one of the greatest needs of the society, there is a limit to resources that can be allocated to it since society has other equally pressing needs. The guiding indicator in this allocation process is the relative prices of goods and services. How the public policy maker allocates social goods and services impact upon how the households allocate the resources available to them. It is within such a framework that we shall consider the policy implications of the estimated model.

7.3 Welfare Analysis of the Estimated Demand Function

University of Cape Town

It is intuitively obvious that in a developing society, distance to health care facilities is a barrier to access. The fact that estimated result shows that distance has a negative sign confirms this intuition in the case of Nsukka local government. The problem of distance may be compounded by other factors such as bad roads and financial cost of transportation. The fact that distance is an important factor in the estimated model suggests that there is unequal access to health facilities with those households that live furthest away from the facilities being the most disadvantaged. It also implies that those households that live furthest from the facilities pay more in terms of time cost of travel in order to access facilities while those who live nearer to the facilities and households that can afford transport stand greater chance of utilizing health facilities.

The significance of distance and the non-significance of treatment cost in the model suggest that non-monetary factors may, some times, pose greater barriers to health care services than monetary factors. This implies that policy makers must not rely solely on financial considerations when deciding health care charges at health facilities. The model further emphasizes the superior equity of income subsidies (such as tax reductions) to direct provision of goods and services as a redistribution policy instrument. In other words, while a policy of direct medical subsidies may be used to cushion the effect of cost of medical care on the poor, such a policy must be seen as a short term remedy. The long run policy must aim at improving the purchasing power of the poor.

University of Cape Town

The above argument has direct bearing with the significance of household income and per capita consumption expenditure in the estimated model. What the model suggests is that improving the income of the people of this area will shift the demand curve for health care services to the right. This will be brought about not only through the direct income effect on demand but also by minimizing the excessive barrier posed by distance since improved income will mean reducing the burden of the cost of transportation. The problems posed by low income and distance to utilization of facilities shows up clearly in the length of days that households allow sickness to develop (severity) before they begin to seek proper health care from facilities. This point will be developed further when we consider general policy issues in subsection (7.3).

Similar to the effect of income, the average years of education of a household has a positive impact on the utilization of health care facilities - a shift of the health care demand curve to the right. A household with a high level of literacy is likely to earn more income than one with low level of education and will make more enlightened health care decisions. While the effect of income is controlled for in the model, the effect of this better judgement about health care is not and is thus seen to impact on the out-come of health care choices. This is not an unexpected result. The level of enlightenment in a population affects its health care behavior in many other ways. It impacts on the environment, consumption and health-producing habits including physical exercise and recreation.

The positive influence of age on the utilization of health care seems rather surprising but parallels the findings of some research reports from other places. For instance Paul Gertler and van der Gaag (1990:92-93) reported a similar mode of health care seeking behavior among the Peruvian population. This raises a very crucial question about equity in intra-household resource allocation. Is the assumption of equality in intra-household resource allocation that is often made in cross sectional studies really justified? Do certain cultural and social norms place the weaker members of some households in disadvantage position in the allocation of household resources? These questions require further investigation.

Closely connected with the effects of level of education and income on health care demand decisions is the negative effect of large household size. Educated households generally tend to be smaller in size than households with low education. Household size negatively impacts on the utilization of health care facilities in the area. This is consistent with the findings of similar empirical studies (eg Ellis and Mwabu 1991). It is possible that this negative relationship arises from the fact that in a large household less attention may be given to the needs of individuals than in a small household. Expenditure per capita is likely to fall as the size of the household increases as was clearly demonstrated in figure 5.5. A large household is more likely to be poorer especially if children pre-dominate the population. However, a larger household is likely to be more efficient than a small-sized household in consumption. But this gain

University of Cape Town

in efficiency may not off-set the drop in average per capita for every marginal increase in size of the household.

The policy issue that arises here is how to encourage smaller household size. The problem of household size is linked to the whole gamut of social and cultural factors. Although in some countries, economic policies such as tax incentives and dis-incentives are often used to encourage or discourage large household sizes the success or failure of such policies would often depend on the political, cultural and social environment. Of equal importance is the policy of encouraging child spacing.

The location variable coefficient indicates a lopsided access to facilities between the urban and rural populations. The urban household is much more likely to consult a health facility than the rural counterpart. Reasons for this are obvious. There is often a greater concentration of social amenities in the urban area. These facilities become easily accessible to the urban dweller both in terms of nearness and variety. Health care providers themselves often want to live and raise their children in urban environment. Facilities sited at the urban areas tend to serve greater populations especially when the rural areas are sparsely populated.

The reasons sited above for the asymmetry in distribution of facilities between the rural and urban areas suggest that in the face of scarce resources more public resources should be directed towards provision of basic utilities to the rural populations. These rural people are likely to incur more non-financial costs in accessing facilities those who live in the urban areas. For example, a rural dweller may forego a whole day of farm work in order to obtain treatment from a distant facility. Private practitioners may be regulated to provide health care facilities in the urban areas while more of public resources are used to target rural populations. It is also possible to encourage the private practitioners to site their facilities in the rural area by given such practitioners tax relief incentives.

University of Cape Town

7.3 General Policy Issues

In general, the estimated parameters of an econometric regression model may individually have policy implications as was analyzed above. In practice it may be difficult to set each of such policy suggestions as a separate policy objective. This may not even be desirable as the multiple objectives will sometimes be in conflict. Furthermore, multiple policy objectives will require an even greater number of policy instruments that may involve greater cost in resources. It is sometimes more efficient, then, to have priority policy objectives and subsidiary objectives. In targeting the priority objectives, the secondary objectives may be achieved as well.

One thing that seems to constantly suggest itself in these brief policy considerations is the interconnectedness between health, economic and other social policy objectives. The policy environment often makes it difficult to target a single health care objective without paying attention to other related objectives. For example, a policy aimed solely at improving health care delivery in a community may achieve very little if other social and economic utilities are lacking. To provide a community with a hospital without addressing the problem of availability of potable water, literacy, and environmental cleanness will not give optimum social benefit.

One of the most important central issues that emerge from this study and which may, therefore, form a priority policy objective is the dominance of the private sector in the overall medical market in this local government. The private suppliers of health care services in this area consist of the private hospitals and clinics, the pharmacy and patent medicine stores and the traditional medicine practitioners. The public health care providers are made up of the Enugu State general hospitals and Nsukka local government health clinics and a Federal Government medical center. The average predicted probability that a household living in this location would seek treatment from a private hospital or clinic is about 0.31. The average predicted probability that the same household would seek treatment from a patent medicine or pharmacy store is about 0.26; and the probability of that household seeking care from a traditional medicine practitioner is about 0.075. If, for the sake of argument, we add up the chances a given household has of seeking care in a privately owned facility (be it private hospital or clinic, patent medicine and traditional medicine) then these chances

University of Cape Town

will add up to about 0.645. This may be contrasted with the chance of 0.28 that a household has of seeking care from a public health facility. The chance of home care and no care at all add up to only 0.075.

From these figures, it is seems clear that the influence of the private sector in this medical market is overwhelming. This has an advantage in the sense that it frees government resources for other equally pressing social services such as water, electricity and so on. However, it also implies that the price mechanism may have more significance in this medical market than it would with a greater presence of public suppliers since the profit motive is the dominant motive for a private supplier. To give free rein to the price system in such a crucial sector of the social welfare as health care will certainly be counter productive. This is more so when it is generally acknowledged that information asymmetry between supplier and consumer places the consumer at a disadvantaged position in the bargain process.

There are two policy instruments a public policy maker may use in addressing the unwanted outcome of an over privatized health care market such as the case in Nsukka local government. One such instrument is to strengthen the public sector facilities in such a way that the fees they charge impact significantly on the profit maximizing behavior of the private health care providers. This instrument depends on the high elasticity of substitution between a private and public health care provider. The government seems to have put this policy to good use. This may have also partly accounted for the non-significance of treatment cost in the model. Such a policy may have acted as a deterrent to private suppliers of health care from charging arbitrarily high fees. This instrument aims at an indirect control of the medical market.

The second and more direct instrument is the legal instrument. The policy maker may use his authority to register or de-register private health care practitioners to control their supply behavior. However, the effectiveness of this instrument depends on several other factors such as the power of private suppliers union, the ease of entry and exit into the market, the ability of the policy maker to supervise and implement laws and regulations in the market etc. The implementation of such regulations seems

University of Cape Town

crucial in Nsukka local government. As at the time the data for this exercise were being collected, these laws and regulations existed but were hardly implemented.

A corollary from the dominance of the private sector is relative importance of the pharmacy and patent medicine category in this medical market. From the analysis of the predicted probability and severity, it was shown that this group of health care providers serve mainly as a form of first-aid. More than 80% of those who used these facilities had illness severity levels less than 5 on the severity scale. It is when illness persisted that patients sought attention in hospitals and clinics.

From familiarity with this locality and the experience during the data collection, patent medicine stores were located in even the remotest corners of the local government area. Some of these operated without due limits as to the type of drugs they were suppose to sell to consumers. Most of the customers bought drugs on self-prescription or on the prescription of the dealer. This situation not only exposes the consumer to financial exploitation but still worse, it increases the risk of health complications.

While the importance of this category of suppliers in the medical market must be acknowledged, it would be worth while to put in place effective policy and regulations for this group. Issues about the medical qualification of patent dealers, the environment in which they operated, the type of drugs they should be permitted to sell and who should prescribe the drugs are serious health issues that need policy intervention.

From the estimated model, severity seems to assume a central consideration in the whole question of demand for health care and the choice of facility. Consequently, the facility people seek care from seems to depend to a great measure on the level of illness. As was noted in subsection 6.6, the emergence of this variable as a factor of central importance seems to have diminished the importance of income in the demand for health care services.

University of Cape Town

The dependence on this criterion for seeking treatment may have very undesirable social consequences. For one, some illnesses have mild symptoms and very short warning period. Such may result in sudden deaths. For another, the near-total dependence on patent medicine to treat illnesses that are considered not severe may lead to further health complications. It also suggests that Primary Health Care services should be given closer attention by the public policy maker than was the case at the time of data collection. Such a policy focus on Primary Health Care will help in early detection of serious illness symptoms and early referral of patients to curative health centers. It may also be useful to train and integrate the patent medicine dealers within the framework of Primary Health Care system in the area. Such integration would mandate the patent dealers to refer patients to hospitals and clinics after giving them the first aid treatment.

The category of people who buy health care from the traditional medicine practitioners raises a fundamental question about the continued neglect of this group of health care suppliers. Although they constitute a small part of the medical market, most of their clients came from the upper income quintiles and the well educated as was clearly demonstrated by figure 5.7 and table 5.5. The fact that the cost of treatment in these traditional health facilities seems to be the highest, and yet they continue to get clients indicates that they have a relevance in the health care system. Their periphery position in the health system need to be reviewed by health policy makers.

University of Cape Town

7.5 Conclusion

Among the main objectives of this study as set out in the objectives-of-study statement was to identify and estimate the parameters of the most important variables determining demand for health care services in Nsukka local government area of Enugu State – Nigeria. Also included among these objectives were to determine factors that influence households to choose one provider rather than the other; and analyze the policy implications of the health-seeking behavior of the area. The estimation of a three-level nested model showed that most of the variables included in the model collectively and individually provided substantial information on the nature of health care demand function of the area. The non-significance of waiting time and price was rather surprising and informative in itself. The model showed that the probability that a given facility will be chosen is sensitive to such variables as the distance of the facility to the household. Utilization also was sensitive to household size, income level of household education, location among, age of the patient, and severity of illness. The results also suggest the importance of treatment price and waiting-time but failed to demonstrate this through statistical significance. Most of these results were consistent with literature on demand for health care services in developing countries but also differed with some others. The policy implications of these were then discussed.

Some of the major policy conclusions suggest a continued policy of using public supply of health care to keep control of the liberalized private system in check. The importance of integrating patent medicine and traditional medicine practitioners into the mainstream health care system was highlighted. Furthermore, the model suggested the need for integrated policy approach to the health care system of the place. The improvement of the health care demand requires policies that also empower the people to buy health care.

It is interesting to know that this study showed, contrary to the assumption of many other studies, that a household's decisions to seek treatment or not; to seek formal or informal treatment; and choice of facility to treat do not seem to be independent household decisions. The first level of decision seems to influence the second and third; and the second seems to influence the third. This is the main conclusion drawn from the estimated inclusive-value parameters. This is a contribution of this study to methodological issues about modeling household decision process in developing countries.

These findings suggests that future research on demand for health care services should take into account the importance of these cross influences in specifying health care demand models. Future efforts to improve our understanding of the health care consumption behavior in the area under study should address more specific areas of the subject. It is possible that focus on specific health care demands such as primary health care, curative health care; or on specific types of illnesses may give a clear

University of Cape Town

understanding of the variation in demand for health services in Nsukka local government. Future research should also explore the impact of the quality of health care services such as level of training of the health care personnel, quality of drugs, reputation of the facility, acquaintance with staff etc, influence households demand behavior. Such studies may throw more light on how households make their health care decisions. This study has provided a basis for such future inquiries.

University of Cape Town

Appendix 1

Weights

The following table shows the estimated population of each of the sampled communities (column 2). Column 3 gives the population of households in the community as a percentage of all the households in the sampled communities. Column 4 gives the number of households interviewed in the community. Column 5 expresses the number of households interviewed in the community as a percentage of the households interviewed in all the sampled communities. Finally, column 6 gives the weight given to households from the community. The weights vary inversely with the probability that a household in a given community is selected.

	Community	Popn.	%	Sample Size (no of hh. interviewed	%	Weight
1	Lejja	15325	9.3	93	9.9	0.93
2	Edem	16661	10	93	10	1.00
3	Ibeagwagu	1304	0.72	15	1.6	0.45
4	Ibeagwani	9181	5.5	46	5.0	1.1
5	Opi	25384	15.2	113	12	1.27
6	Anuka	776	0.46	38	4	0.11
7	Nsukka	79913	59	190	20.2	2.92
8	Obimo	12753	7.6	84	9	0.84
9	Edeoballa	14368	8.6	200	21.3	0.4
10	Aluno	6530	4	64	6.8	0.6

University of Cape Town

Appendix 2

1. LR Test

This test is specified as

$$G = -2\ln\left[\frac{likelihood..with..constant..only}{likelihood..with\,variables..included}\right]$$

The multiplication by –2 is a means of obtaining a quantity whose distribution is known and can therefore be used for hypothesis testing (Hosmer and Lemeshow 1989:13-14) . In this case, the distribution has a Chi square distribution with degrees of freedom equal to the number equal to the number of estimated coefficients.

This ratio refers the ratio of the likelihood function of constant-only model to the full model or the saturated model which may be conceived as the model with as many variables as there are data points (see Hosmer and Lemeshow 1989). This ratio when multiplied by –2 is referred to as the deviance:

$$D = -2\ln\left[\frac{likelihood..of..estimated\,model}{likelihood..of..the..full..model}\right]$$

2. Test of Significance

In theory, this test is conducted (n-K) degrees of freedom. (Where n is the number of observations in the model and K is the number of parameters estimates). In practice, however, since the change in the value of t changes slowly between (n-K) = 8 and (n-K) = ∞, the degrees of freedom is often ignored in practice Koutsoyiannis (1979:90). If the value of the calculated t-statistic is greater than 2 or less than –2, the variable with the coefficient is said to be statistically significant. This implies that from the point of view of statistics, the variable has influence on the dependent variable. If, on the other hand, the calculated t-statistic is less than 2 or more than –2, the associated variable is considered to be statistically insignificant. Table 6.1 contains this ratio as t-stat and also the more accurate **p-value** (columns 3 and 4 respectively). The p-value gives the exact probability of committing a Type 1 error. That is, the probability that

University of Cape Town

one may reject the null hypothesis when in fact it is true. It is also referred to as the lowest significant level at which a null hypothesis may be rejected D.Gujaratti (1993:132).

3. Odds Ratio

The odds of an outcome is defined as

$$\frac{P(y=1\mid x)}{P(y=0\mid x)}=\frac{P(y=1\mid x)}{1-P(y=1\mid x)}$$

The odds ratio is may then be defined as

$$\psi=\frac{P(y=1\mid x)/1-P(y=1\mid x)}{P(y=0\mid x)/1-P(y=0\mid x)}$$

By taking the log of both sides we obtain the log of odds

There are four major desirable properties of the odds ratio that make it appealing:

(a) Odd ratios have clear interpretation. For example, if the odds ratio of A to B is 2 to 1, it is easily understood that event A is two times more likely to occur than event B.

(b) It is invariant to ordering of variables provided the signs are noted. Thus A/B = 1/(B/A)

(c) It does not change with sample size. An increase in sample size for instance will give approximately the same odds ratio.

(d) It can be applied in multivariate analysis F.T Liao (1994:14-15)

4. In the first place a constant terms captures the mean effect of omitted and other unobserved factors in the model. Since we are dealing with a system of equations, the expectation is to include alternative specific constants (ASCs). However, Amemiya (1981) points out that having ASCs is not consistent with discrete choice model since that would imply that we are recognizing the effect of a given facility beyond its attributes. Furthermore, T.F Liao (1994:61-62) shows that this is not necessary since some of the attributes are already choice specific, (eg treatment cost). The inclusion of alternative specific constants often complicates the problem of identification of the

University of Cape Town

model. The result of a model with alternative specific constants only is, however, shown in the appendix.

5. Measures of Goodness-of-fit

Econometric researchers have evolved –several summary statistics that can be used to assess the fit of an estimated model. Ordinarily, in the OLS models the summary statistics usually used is the R^2 and the adjusted R^2 This is defined as the proportion of the variation in the dependent variable that is accounted for by the variation in the explanatory variables. Formally the measure is given by:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

Probability models produce different values that therefore require the modification of this measure of fit. One of the measures is Likelihood Ratio Index suggested by McFadden. This measure is given the formula:

$$R^2 = 1 - \frac{\ln \hat{L}(M_\omega)}{\ln \hat{L}(M_\Omega)}$$

Where the log likelihood in the numerator is that of the model without restriction and the log likelihood in the numerator is that of the model with only the constant. This Index was the goodness-of-fit criterion used in this paper.

The Akaike Information criterion (AIC) is also popularly used in probability models especially where the model is non-nested. This criterion is specified as

$$AIC = \frac{-2\ln \hat{L}(M_\omega) + 2P}{N}$$

Where the log likelihood is as defined above, P is the number of parameters in the estimated model and N is the number of observations. By dividing by number of observations helps us to approximate the contribution of each observation to the likelihood function. Since the numerator is negative, the smaller the AIC value the better the fit. Many of these measures, however, have upper limits that are far below 1 Maddala (1983:40).

University of Cape Town

6. Problems of Identification and Estimation

Identification is one of the greatest problems in any simultaneous equation models. The problem of identification concerns whether the simultaneous equation model has unique statistical parameters such that it is possible to say that the estimated parameters belong to the model and not to some other model. In other words, for the model to be identified, the set of equations must have a unique solution. One basic condition for a model to have a unique solution is that the model must be complete, implying that there must be as many independent equations as there are endogenous variables in the system and each of these equations must have a unique solution. A system of equations is said to be identified if the system has a unique solution, implying that each of the component equations has a unique solution. A system of simultaneous equations is unidentified or under identified if the system has no unique statistical solution, implying that at least one of the component equations has no unique solution. On the other hand, if a system of simultaneous equations has a unique solution it is said to be identified. A system that is identified may be exactly identified or over-identified. If a system is unidentified or under-identified, it becomes impossible to estimate its parameters. A system that is over-identified would yield more than one solution. When a system is exactly identified, it parameter estimates can be obtained. In that case the appropriate method to use in the estimation is the Indirect Least Squares (ILS). This basically involves first obtaining the reduced form parameters of the structural model. (That is to say, express the endogenous variables as functions of the exogenous variables and the error terms). The reduced form coefficients can then be obtained with OLS and then inserted into the structural model and then estimated as ILS coefficients. Two essential conditions for identifiability may briefly summarized as follows:

(a) The Order Condition

$$(K-M) \geq (G-1)$$

Where G is the total number of equations in the system

K is the total number of variables in the model

M is the number of variables (both endogenous and exogenous) in a particular equation in the system. If the "greater than" sign holds, the equation is over-identified. It is exactly identified if the equality sign holds.

University of Cape Town

(b) the rank of the matrix, (that is the order of the largest nonzero determinant that can be formed from a given matrix) to be formed from the rest of the variables excluded from that particular equation must be of the order (G-1). Where G is as defined in (a) above Koutsoyiannis (1979:353).

These conditions essentially constitute restrictions on the component equations of the system. The method of maximum likelihood is generally used to estimate the parameters of such a system of equations, though the method of two-stage least squares (2SLS) can also be used to obtain unique parameters. As was indicated earlier, Limited Information Maximum Likelihood (LIML) and Full Information Maximum Likelihood (FIML) are special applications of the principles of the ML method. In the LIML which is a single equation method, each equation in the system is estimated singly taking into account only the restrictions on that equation though information on other equations would have to be consider for the purpose of identification. This estimation is carried out in sequence of steps. In FIML on the other hand, all the equations in the model are estimated in one pass taking into account all the restrictions on all the equations at the same time. This would imply, of course, much more computational difficulties than the single equation method. It has also been pointed out that because the system equation method takes on all the equations at the same time, any miss-specification in one equation spills problem to other equations in the system. In other words, FIML would be more sensitive to specification error than LIML.

University of Cape Town

This fact was clearly experienced in the process of estimating the present model. In many cases, a three level-nested specification was run by LIML method sequentially in LIMDEP was repeated with FIML in the same program but reported "Singular Hessian" indicating that the model is over-identified or "Model Unidentified". The main problem encountered with the use of LIML is that the name of any variable that appeared more than once in the system of equations becomes automatically fixed. In that case, LIMDEP reports the value of the coefficient without reporting any values for the standard errors, the t-ratio and the p-value. It has been shown, however, that FIML is more efficient in estimation of parameters than LIML. It is further noted that there are substantial differences in the coefficients obtained from both methods. McFadden (1981).

In order to achieve identification, even with the FIML, the advice of W.H. Greene was followed. This involves estimating the parameters by first providing start values for some of the parameters (For a nested model, for instance, one could specify the start values as the ordinary binary logit values of the parameters. LIMDEP would provide these values by itself. From the coefficients obtained from this 'trial' estimation, one or two parameters could be fixed in order to achieve identification in the required model. W.H Greene (1998:578). It would often be easier to obtain first the inclusive value parameters from the trial model. Even for a nested model, if the start value is specified as logit, LIMDEP obtains the parameter estimates of the ordinary logit model and goes on to estimate the nested model. Greene however warns that the parameter estimates obtained from this trial model are not consistent estimates. Having obtained the inclusive value parameter estimates, a value lower than what was actually obtained was then provided as start values to run the main model which in many cases achieved identification. At some other times, the parameters are obtained with the standard errors, t-ratios and p-values fixed. In this model, for instance, LIMDEP was asked to run the nested model with logit as start values. (That is to use the coefficients of the logit model as the start values for the nested) The figures that were obtained for the inclusive value parameters at the branch and limb levels were approximately 0.3 and 0.5 respectively. In running the main model, start values of 0.25 and 0.35 were provided to LIMDEP from which it obtained estimates of 0.4 and 0.6. At some other times, the parameters are obtained with the standard errors, t-ratios and p-values fixed

In linear equation models, parameter estimates are often obtained using algebraic methods. However, with complex non-linear models the level of computation increases enormously. It becomes difficult to apply the algebraic method. The estimations of such models are therefore carried by iterative numerical methods. In these methods, initial values are assumed for the parameters. (It was stated above that in the process of estimating this model, some 'start' values were provided). These initial values are improved by adding adjustment factors to these start values. This

University of Cape Town

process of improving on the initial value is iterative. It continues until convergence is achieved. Two basic elements in this process are the gradient or the first derivative of the log function. This indicates the direction of change in the log likelihood. The other is the direction matrix which indicates how fast curvature is changing (eg, could be very steep or nearly flat). If the turning point has a global maximum that is sufficiently steep implying narrow variance, then convergence is easily achieved. If the variance is very wide, the iteration may continue for long or may be discontinued if it is not possible to achieve convergence. Thus, it is possible to control the number of iterations by setting maximum number of iterations in the estimation of the model or even abort the process.

There are a number of algorithms used in this process. Some of these are named after their authors. They include DFP (Davidson, Fletcher, Powell), BFGS (Broyden, Fletcher,Goldfarb, Shanno), Newton's method, BHHH (Berndt, Hall, Hall, and Houseman) and SteDes – that is, the method of steepest descent. All these are available in LIMDEP and one can specify, depending on the model one is estimating, which of the algorithms to use. In this estimation, it was discovered that SteDes takes by far many more iterations to achieve convergence than the rest of the algorithms though it will often achieve convergence. If LIMDEP starts off estimation with Newton's Method or BHHH, if it does not achieve convergence within the first few iterations due to identification problem, it reports it is unable to invert the Hessian and then switches over to BFGS. However, DFP is the one that seems to achieve convergence more often than the rest. That was the one used in this model estimation.

University of Cape Town

Appendix 3

Demand for Health Care Services in Nigeria: -A Nested Logit Model

Survey Instrument

Name of interviewer [ ] Code [ ]

Form No [ ] Cluster No [ ]

Date of Interview [ ]

Interviewer : NB: Interview is to be conducted for both household that report sickness and those that do not.Ask : *Has anybody, child or adult experienced sickness or injury in this household in the last one month?*

What is the size of this household?
(NB the household size consists of all persons who reside under the same roof at least 4 days a week or 15 days a month and share food from a common source.)

1.1 Number, gender and age group of household members.

Age group	Number	
	males	females
Age 0-6		
6< age < 16		
16 $\leq$ age $\leq$ 64		
Age 65+		
Total		

Interviewer :If somebody is reported sick proceed sequentially and complete the rest of the questions. If nobody is sick in the household you proceed to section 2.0 and complete the rest of the questions.

Interviewer : *now ask the head/acting head of the household or sick/injured person :*
1.2 (a) How old is the sick person (in completed yrs) ----------

1.2(b) How long has he/she been sick (or how long were you sick, if the sick is household head): --------days

University of Cape Town

2(c) How many days in the last one month have you been unable to do your normal duties? ----------number of days

1.3 How would you describe the condition of your health now

v. good	good	fair	bad	v. bad	too bad
01	02	03	04	05	06

1.4 a Did you seek any treatment?

yes	
no	

1.4 b. If yes which of the following did the household choose for your treatment:

Public hospital/clinic	01
Private hospital/clinic	02
Mission hospital/clinic	03
Treated at home	04
Traditional healer	05
Pharmacy/chemist store	06
Other (specify)	07

1.5 If you did not seek treatment at all, why did you not?

01	Did not want to	
02	Too expensive	
03	No money for transport	
04	Would lose pay for work	
05	Distance to facilities too much	
06	Too many people waiting	
07	Others (specify)	

1.6 If you were treated in a facility what is the distance from your house to that facility? ---------km

1.7 By what means did you get there:

On foot	01
Bicycle/cyclist	02
Car/bus	03
Others (specify)	04

1.8 How long did it take you to get there ?

Less than 30 mins	01
30 mins to 1 hr	02
1hr – 1 hr 30mins	03
1 hr 30mins – 2 hrs	04

University of Cape Town

More than 2hrs	05

1.9 If you did not go there on foot how much did you pay for transport to get there: N : K

1.10 How long did it take you to see the doctor or whoever treated you?

Less than 30 mins	01	
30 mins to 1 hr	02	
1hr – 1 hr 30mins	03	
1 hr 30mins – 2 hrs	04	
More than 2hrs	05	

1.11 How many times have you been there for treatment? ---- number of times

1.12 How much were you charged ?

Card	diagnosis	drugs	total
N	N	N	N

1.13 Did anybody accompany you to receive treatment? Yes/No (tick the correct one)

1.14 Who accompanied you

Father	mother	brother	sister	relation
01	02	03	04	05

1.15 How long did the person stay with you?

<1 hr	1-3 hrs	3 -6hrs	1 day	2-3 days	> 3days

1.15 How much did this household pay to get you treated?

card	diagnosis	drugs	Transp for you	Trans. for c/taker	others
01	02	03	04	05	06

1.16 Interviewer : *show the respondent the amounts below and ask him/her*:
Suppose the person who accompanied you were to use that time for taking you for treatment to do something else, either in the household or paid job how much would he/she have earned taking into account the type of work he/she does?

< N99	01
N100 - N199	02

N200 – N499	03
N500 - N 999	04
N1000 – N1999	05
N2000 – N 2999	06
N3000 – N 5000	07
> N5000	08

1.17 If you treated the sickness at home how much did it cost to treat it?

< N99	01
N100 - N199	02
N200 – N499	03
N500 - N 999	04
N1000 – N1999	05
N2000 – N 2999	06
N3000 – N 5000	07
> N5000	08

2.0 Household Expenditure

Interviewer: *Try to impress on the respondent the need not to exaggerate or understate the facts because if they do it would defeat the purpose of the exercise. The period of expenditure is the* *<u>last one month</u>*

2.1 How much does this household consume or spend or spend on the following food items in a month

code	Item	Eaten/ Bought Yes/No	Amount	How much received	How much eaten from own prodn
01	Yam				
02	Garri/akpu				
03	Rice				
04	Beans/okpa				
05	corn				
06	meat				
07	Fish				
08	vegetable				
09	Oil				
10	bread				
11	Tea/sugar				
12	milk				
13	fruits				
14	groundnuts				
15	Soft drinks				
16	Meals given to guests				
17	others				

University of Cape Town

Non Food Spending

2.2 How much does this household or individual members spend on the following items in a month?

code	Item	Amount
01	cigarettes	N
02	Beer, wine, spirits	N
03	Personal care items: soap, shampo, hairdressing etc	N
04	Newspapers/stationeries	N
05	Telephone bills	N
06	Transport	N
07	Petrol, oil care services etc	N
08	Electricity bills	N
09	Washing powder/detergents	N
10	Child care	N
11	Religious/charities	N
12	Informal taxation & donations	N
13	Payments to house assistants	N
14	others	N

Occasional Nonfood Spending

2.3 How much did the household spend on the following items in the last one year

code	Item	Amount in N
01	Kitchen equipment/utensils	N
02	Building maintenance	N
03	Bedding – blankets, bedspread et	N
04	Furniture	N
05	Electric appliances	N
06	Clothing (adults & children)	N
07	Shoes "	N
08	Curtains/ door & window blinds	N

2.4 Health and Health Care

code	Item	Amount
01	Doctors, dentists, nurses, healers	N
02	Hospital fees	N
03	Drugs, bandages etc	N
04	others	N

2.5 Household Durables

Interviewer : *introduce by saying*: there are some items listed here which a household may or may not own. Which of these does the house hold own

code	Item
01	Vehicles – cars, lorries, trucks etc

University of Cape Town

02	Motorcycles
03	TV
04	Fridge
05	Telephone
06	Gas-cooker
07	Radio
08	others

2.6 Dwelling Place

Interviewer: *Look around and find out the type of dwelling the household lives in and tick accordingly.*

Shack	01
Part of a house	02
Hut	03
Maisonette	04
Flat	05
Hostel	06
Combination of buildings	07
Others	08

2.7 Household Income

Interviewer: *ask the respondent to look into the table below and pick the letter which best describes the total income of all in this household in a month. Include all sources of income – salaries, pensions, income from income from investment, petty trading etc(if he/she cannot read, read the amount allowed and interprete)*

A	N0.00 - N999	01
B	N1,000 - N 2,999	02
C	N3,000 - N4,999	03
D	N5,000 - N 7,999	04
E	N8,000 - N9,999	05
F	N10,000 - N14,999	06
G	N15,000 - N19,999	07
H	N20,000 - N29,999	08
I	N30,000 +	09

Interviewer: *address the respondent*: Finally, let us talk about education of this household

2.8 Can you give me the number of years each member of this household attended formal education beginning with the head or acting head of this household

Code	Position in the household	Age	Total no. of yearsof formal

University of Cape Town

			education
01	Head/acting head		
02			
03			
04			
05			
06			

Interviewer : *Thank the household for their cooperation and go to* another *one.*

3.0 Facility Characteristics

Interviewer: Introduce the purpose of your visit to the management or the officer in charge.

3.1 What type of health care facility:

Pub. Hosp.	Pub Clin	Priv. Hosp	Priv. Clin.	Misn Hosp	Misn Clini	Trad. Med

3.2 How many doctors, qualified nurses and trad. healers does the facility have both full time and part time?

F.Time Drs	Part time Drs	Full time nurses	Part time nurses	Tradit. Healers

3.3 What is the average waiting time (i.e how does it take a patient to be attended to

< 10 mins	11-30mins	31-60mins	61-120min	>2hrs

3.4 What is the average cost of treatment?

<N100	N100-249	N250-499	N500-1000	>1000

Signed ------------------------------
Interviewer

Checked and certified by me
signed ------------------------------
Supervisor

University of Cape Town

REFERENCES:

Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley

Akin, J. (1981) The demand for children health services in the Phillipines in *Social Science and Medicine* 15, 247-257

Akin J, et al. (1986) The demand for Primary Health Care services in the Bicol Region of the Phillipines in *Economic Development and Cultural Changes* 34(4), 755-782

Aldrich, J.H & Nelson, F.D. (1984). *Linear Probability, logit and Probit* Models. Beverly Hills, CA: Sage.

Amemiya, T.(1981) Qualitative Response Models: A Survey *Journal of Econometric Literature* Vol.XIX

Amemiya, T. (1985) *Advances Econometric* Oxford: Basil Blackwell

Bartholome, C. and Vosti, S. (1995) Choosing between Public and Private health care – a case study of malaria treatment in Brazil in *Journal of Health Economics* 14 191 – 205.

Cramer, J.S. (1991) *The Logit Model*, New York: E Arnold

Deaton, A. (1998) *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: The Johns Hopkins University Press.

Domencich, T.A. and McFadden, D. (1975) *Urban Travel Demand : A Behavioral Analysis* Amsterdam: North-Holland Publishing Company

Dor, A. et al. (1987) Non price rationing and the choice of medical care provided in rural Cote d'Ivoire in *Journal of Health Economics* Vol. 6, 291 – 304

Dow, W.H. (1996) *Unconditional Demand for Health Care in Cote d'Ivoire* Washington: The World Bank LSMS Working Paper No. 127

Ellis, R. and Mwabu, G. (1991). "The Demand for Outpatient Medical Care in Rural Kenya." Boston, MA: Dept of Economics, Boston University, mimeo.

Egwu, S. (1998) *Structural Adjustment, Agrarian Change and Rural Ethnicity in Nigeria* Uppsala: Nordiska Afrikainstitutet

Gertler, P. and van der Gaag, J. (1990) *The Willingness to Pay for Medical Care: Evidence from Two Developing Countries*. Baltimore: John Hopkins University Press.

Greene, W.H. (1993) *Econometric Analysis* (2nd ed.). New York: Macmillan.

University of Cape Town

Greene, W.H. (1998). *LIMDEP Version 7.0.* Plainview, New York: Econometric Software.

Grossman, M. (1972) On the concept of health capital and the demand for health, Journal of Political Economy,Vol. 80 pp. 223 – 255.

Gujaratti, D. (1993) *Basic Econometrics* (3rd ed.) New York: McGraw Hill International Ed.

Hausman, J. and McFadden, D. (1984) Specification Tests for the Multinomial Logit Model in *Econometrica* Vol.52, No.5 1219-1239

Hosmer, D.W., and Lemeshow, S. (1989). *Applied Logistic Regression*. New York: John Wiley

Heller, P.S (1982). A model of the demand for medical and health services in Peninsular Malaysia; in *Social Science and Medicine* Vol.16, 267 – 84.

Johnston, J. and Dinnardo, J. (1997) *Econometric Methods* 4th ed. NY McGraw Hill International Ed.

Koutsoyiannis, A. (1977). *Theory of Econometrics* (2nd ed.) Houndmills: The Macmillan Press

Liao, T.F. (1994). *Interpreting Probability Models: Logit, Probit and Other Generalized Linear Models*. Thousand Oaks, CA: Sage.

Long, S. (1997*) Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks : Sage Publications

Maddala, G.S. (1983) *Qualitative Variables in Econometrics* Cambridge: Cambridge University Press.

Mackinnon, J. (1995) *Health as Informational Good: The Determinants of Child Nutrition and Mortality during Political and Economic Recovery in Uganda.* WPAS/95-9 Centre for the Study of African Economies – Oxford University.

Manning,W.G, Newhouse J.P, Duan, et.al (1987) *Health Insurance and the demand for Medical care: evidence from Randomized experiment*, American Economic Review 77(3)' 251-277

McFadden,D. (1981) Econometric Models of Probabilistic Choice in C.F. Manski andD McFadden (Eds.), *Structural Analysis of Discrte Data eith Econometric Application (pp198-272).* Cambridge: MIT Press

McGuire, A. Henderson, J., Mooney G., (1987) *The Economics of Health Care.* London: Routledge and Kegan Paul.

University of Cape Town

Mills.A., and Lee K., (1993) *Health Economics Research in Developing Countries.* Mills and Lee (eds) Oxford: Oxford University Press.

Mukherjee, C., White, H. and Wuyts, M. (1998) *Econometrics and Data Analysis for Developing Countries* London: Routledge

Musgrove, P. (1983) Family health care spending in Latin America. *Journal of Health Economics* 2(3) 245 – 58

Mwabu, G. (1986). Health care decisions at household level: results of a rural health survey in Kenya , *Social Science and Medicine* 2 2(3).

Mwabu, G., Ainsworth, M. and Nyamete A., (1993) Quality of Medical care and choice of medical treatment in Kenya: an empirical analysis . *Journal of Human Resources* XXVIII,4, 838-862.

Ogunbekun, I. (1991) Which direction for health care in Nigeria., *Health Policy and Planning* 6(3): 254 – 261

Pearce, T. and Falola, T. (1994) *Child Health in Nigeria*, Aldershot: Avebury

Praise, S., and Houthakker, H.(1971). *The Analysis of Family Budgets*. Cambridge, England: Cambridge University Press.

Scott, C. and Amenuvegbe, B. (1990) Effect of recall duration on reporting of household expenditures: an experimental study in Ghana, qtd in Deaton (1998)

South African *Project for Statistics on Living Standards and Development* 1993

Stevenson, A., Muscatelli.V. and Gregory, M. (1988). *Macroeconomic Theory and Stabilization Policy.* New York: Philip Allan.

Sauerborn, R. Adams, A. and Hien, M. (1996). "Household strategies to cope with the economic costs of illness". *Social Science and Medicine* Vol. 43 No. 3 pp 291-301

Van Vliet R, and van Doorslaer, E. (1990) Disaggregation of demand for hospital care; in *Disaggregation in Econometric Modelling*. Barker and Pessaran (eds) London: Routledge (pp 188-205)

Vogel, R. (1988) *Cost Recovery in the Health Care Sector: selected country studies in W.Africa.* Washington DC: The world Bank

Wagstaff, A. (1986). The demand for health: theory and applications. *Journal of Epidemiology and Community Health*, 40, 1-11

Watts, M. and Lubeck, P. (1984). Popular Classes and the Oil Boom: A Political Economy of Rural and Urban Poverty. In Zartman, W. (Ed.) *The Political Economy of Nigeria*. Prager Scientific.

University of Cape Town

World Bank. 1993. World Development Report 1993. New York: Oxford University Press.

Web site:
Center for International Health Information: www.cihi.com
NIGERIA: Health Statistics

University of Cape Town