A Distributional Analysis of Healthcare Financing in a Developing Country: A Nigerian Case Study Applying a Decomposable Gini Index

Hyacinth Ementa Ichoku

Thesis Presented for the Degree of DOCTOR OF PHILOSOPHY in the Health Economics Unit, School of Public Health and Family Medicine, University of Cape Town

August 2005
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.
Declaration

A Distributional Analysis of Healthcare Financing in a Developing Country: A Nigerian Case Study Applying a Decomposable Gini Index

I, Hyacinth Ementa Ichoku, hereby

(a) grant the University of Cape Town free licence to reproduce the above thesis in whole or in part, for the purpose of research;
(b) declare that:

(i) the above thesis is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisors, I have received no assistance except as stated below:
(ii) except as stated below, neither the substance nor any part of the above thesis has been submitted in the past, or is to be submitted for a degree in the University or any other university.

The thesis has been presented by me for examination for the degree of PhD.

Signed ........................................
Hyacinth Ementa Ichoku
Acknowledgments

It's a pleasure to express my deep appreciation to all who have assisted me in my long academic journey. First, to Him who makes all things possible, may He forever be glorified. I thank Bishop S.A Okafor of Awka Diocese who has given me lots of encouragement in my academic pursuits, and Msgr J. K. Aniagwu who has been like a father and friend to me. My deepest appreciation goes to Dr Michael Thiede, of the Health Economics Unit, and Prof. Murray Leibbrandt, of the School of Economics, University of Cape Town. It has been immense pleasure for me working with you as supervisors of this project. Many thanks to Prof. Di McIntyre, the moving spirit in the Health Economics Unit.

I sincerely thank the Board of African Economic Research Consortium (AERC) for providing the financial support that enabled me carry out the field survey to generate the data for this study. My deep thanks to all my colleagues in the Department of Economics, University of Nigeria, Nsukka. Of particular mention are Prof. A. Okore for providing me academic guidance, and Dr Reginald Chima who set the ball rolling. Willy Fonta, who is about now preparing his PhD defense, John Ataguba (Technical Master), and Ezebilo Ukwueze, I can’t thank you enough for what you have been to me. To my friends, Jude Mgobukwu, Amaechi Awurum, Augustine Nwosu, all of who are currently writing their own doctoral theses, Rev. Dr Hill Achunike and Sylvester Dunu, I appreciate your love and friendship.

Words are hardly enough to express the depth of my gratitude to Rev Dr Michael van Heerden who has been a great friend to me, Fr. Chris, the inspiring Rector, and the rest of the staff of St Francis Xavier Seminary, Cape Town for their love and care. My tribute and appreciation to Mr Winston Jacobs and Virginia O’connor, the Ramdeens: Peter (bless his soul) and Barbara, the Wigmores: Les and Bennie, you have been great friends. I cherish you all.
Dedication

To

The eternal memory of two great brothers:

Peter Ichoku

Peter Ramsdeen
# Table of Contents

Declaration ........................................................................................................................... i
Acknowledgments .................................................................................................................. ii
Dedication ............................................................................................................................. iii
Table of Contents .................................................................................................................... iv
List of Tables .......................................................................................................................... viii
List of Tables .......................................................................................................................... viii
List of Figures ........................................................................................................................ ix
List of Acronyms .................................................................................................................... ix
Summary .................................................................................................................................... xii

## CHAPTER 1

INTRODUCTION .................................................................................................................. 1
1.1 BACKGROUND TO THE STUDY .............................................................................. 1
1.2 THE MOTIVATION OF THE RESEARCH ............................................................... 6
1.3 THE POLICY CONCERN OF THE STUDY ............................................................. 8
1.4 OBJECTIVES OF THE STUDY ............................................................................... 9
1.5 JUSTIFICATION OF THE STUDY ......................................................................... 11
1.6 SIGNIFICANCE OF THE STUDY ......................................................................... 12
1.7 LIMITATIONS AND ASSUMPTIONS OF THE STUDY ........................................... 13

## CHAPTER 2

THE SOCIOECONOMIC AND POLICY CONTEXTS OF THE STUDY ......................... 16
2.1 INTRODUCTION ..................................................................................................... 16
2.2 THE NIGERIAN SOCIO-POLITICAL CONTEXT .................................................... 17
2.3 MACROECONOMIC POLICIES AND THE HEALTH SECTOR .............................. 19
2.4 POVERTY AND INCOME INEQUALITIES IN NIGERIA ........................................ 22
2.5 THE HEALTH STATUS OF THE NIGERIAN POPULATION ................................ 24
2.6 FISCAL FEDERALISM IN NIGERIA ................................................................... 27
2.7 THE STRUCTURE OF THE HEALTHCARE SYSTEM .......................................... 28
2.8 THE STRUCTURE OF THE NIGERIAN HEALTHCARE MARKET ......................... 31
  2.8.1 General Taxation .............................................................................................. 32
  2.8.2 Social Health Insurance Schemes ................................................................. 32
  2.8.3 User Fees ........................................................................................................ 34
2.9 PRIVATE PROVIDERS ............................................................................................. 34
  2.9.1 Private Healthcare Insurance and Prepayment schemes .............................. 36
  2.9.2 Prepayment Schemes ..................................................................................... 37
  2.9.3 Donor Funding ............................................................................................... 38
  2.9.4 Coordination .................................................................................................. 39
2.10 GOVERNMENT'S NEW HEALTH SECTOR REFORM POLICY .......................... 39
2.11 SUMMARY ............................................................................................................ 40
CHAPTER 3..........................................................................................................................41
LITERATURE REVIEW ..........................................................................................................41
  3.1 INTRODUCTION .............................................................................................................41
SECTION A: THEORETICAL FRAMEWORK ..............................................................................42
  3.2 WELFARE ECONOMIC THEORY AND REDISTRIBUTION1 ........................................42
    3.2.1 Introduction .............................................................................................................42
    3.2.2 Empirical Evidence ..............................................................................................45
  3.3 NATURE OF THE HEALTHCARE MARKET AND HEALTHCARE FINANCING ........46
    3.3.1 Characteristics of the Healthcare Market ..............................................................46
  3.4 THE CONCERN FOR EQUITY: EQUITY AND IDEOLOGY ...........................................49
    3.4.1 Theories of Justice ...............................................................................................50
  3.5 CONCERN FOR EQUITY IN HEALTHCARE FINANCING ...........................................54
    3.5.1 Vertical and Horizontal Equity ..............................................................................55
SECTION B: EMPIRICAL LITERATURE .....................................................................................59
  3.6 INTRODUCTION .............................................................................................................59
    Early Empirical Literature ..............................................................................................60
    Recent Developments ....................................................................................................61
SECTION C ..................................................................................................................................71
  3.7 BRIEF REVIEW OF PREVIOUS RESEARCH FINDINGS ON HEALTH INEQUALITIES ....71
  3.8 SUMMARY ......................................................................................................................78
  3.9 CONTRIBUTIONS OF THIS STUDY TO LITERATURE ....................................................78
CHAPTER 4................................................................................................................................81
METHODOLOGICAL FRAMEWORK .......................................................................................81
  4.1 INTRODUCTION .............................................................................................................81
  4.2 MEASUREMENT OF HEALTH STATUS .......................................................................81
  4.3 MEASURING INEQUALITIES IN HEALTH DISTRIBUTION ..........................................84
  4.4 MEASURING BIVARIATE DISTRIBUTIONS IN HEALTH .............................................85
  4.5 MEASURING THE REDISTRIBUTIVE EFFECT OF HEALTHCARE FINANCING ..........89
  4.6 KAKWANI’S INDEX OF PROGRESSIVITY ......................................................................91
  4.7 REYNOLDS-SMOLENSKY (RS) INDEX OF REDISTRIBUTION ......................................93
  4.8 DECOMPOSING THE REDISTRIBUTIVE EFFECTS .........................................................94
CHAPTER 5: THE SURVEY

5.1 INTRODUCTION ......................................................................................................................... 98
5.2 OBJECTIVES OF THE SURVEY ................................................................................................... 98
5.3 FIELDWORK METHODOLOGY .................................................................................................... 99
5.4 STUDY AREA ............................................................................................................................... 99
5.5 COVERAGE ................................................................................................................................. 100
5.6 SAMPLE DESIGN ......................................................................................................................... 102
5.7 SAMPLE SIZE ............................................................................................................................ 102
5.8 SELECTION OF THE EAs ........................................................................................................... 104
5.9 DESCRIPTION OF THE QUESTIONNAIRE .................................................................................. 106
5.10 TRAINING AND FIELDWORK .................................................................................................. 108
5.11 ENSURING DATA QUALITY ...................................................................................................... 109
5.12 WORK PLAN .............................................................................................................................. 110
5.13 DATA PROCESSING .................................................................................................................... 110
5.14 DATA AND VARIABLE DEFINITIONS ....................................................................................... 111
  5.14.1 Measurement Problems ........................................................................................................ 111
  5.14.2 Estimating Household Expenditure ...................................................................................... 111
  5.14.3 Accounting for Multistage Sampling Design: ................................................................. 113
    (a) Weighting ................................................................................................................................. 113
    (b) Stratification ............................................................................................................................. 114
    (c) Clustering ................................................................................................................................. 114
  5.14.4 Transformations on Data ..................................................................................................... 115
  5.14.5 Equivalence Scales and Scale Economies .......................................................................... 116
  5.14.6 Data Aggregation ................................................................................................................ 118
5.15 SUMMARY ............................................................................................................................... 119

CHAPTER 6: A DISTRIBUTIONAL ANALYSIS OF HEALTH IN NIGERIA

6.1 INTRODUCTION ......................................................................................................................... 120
6.2 SAH AND HSI AS HEALTH INDICATORS ................................................................................ 120
6.3 ANALYSIS OF HEALTH DISTRIBUTION IN NIGERIAN POPULATION ................................. 121
6.4 MEASURING INEQUALITIES IN HEALTH DISTRIBUTION ..................................................... 123
6.5 MEASURING SOCIOECONOMIC INEQUALITIES IN HEALTH ................................................. 128
6.6 INTRODUCING INEQUALITY AVERSION PARAMETER: EXTENDED CI ............................... 133
6.7 CONCENTRATION INDEX REGRESSION RESULTS ................................................................... 136
6.8 INTERPRETATION OF ESTIMATED PARAMETERS ................................................................. 138
6.9 SENSITIVITY OF CI TO CHANGES IN AVERSION PARAMETER ........................................... 140
6.10 THE DETERMINANTS OF HEALTH ....................................................................................... 144
6.11 ESTIMATED PARAMETERS OF THE DETERMINANTS OF HEALTH ..................................... 146
6.12 SUMMARY ............................................................................................................................... 149

CHAPTER 7: DESCRIPTIVE RESULTS

7.1 INTRODUCTION ......................................................................................................................... 150
7.2 PATTERNS OF HEALTHCARE EXPENDITURE ......................................................................... 151
7.3 Inequalities in Healthcare Expenditure .................................................. 156
7.4 The Incidence and Intensity of Catastrophic Healthcare Financing .......... 162
  7.4.1 Methods ..................................................................................... 163
  7.4.2 Results of the Estimations of Catastrophic Payments ....................... 167
7.5 The Impoverishing Effects of Healthcare Financing ............................... 171
7.6 Summary ...................................................................................... 175

CHAPTER 8 ....................................................................................... 176
THE ARONSON-JOHNSON-LAMBERT DECOMPOSITION FRAMEWORK ....... 176
  8.1 Introduction .................................................................................... 176
  8.2 The AJL Decomposition Method ..................................................... 178
  8.3 Estimation Procedure and Modeling Issues ....................................... 185
  8.4 The AJL Model Results .................................................................. 188
  8.5 Interpretation of $C_r$ .................................................................... 192
  8.6 Key Components of the Redistributive Effect ................................... 193
    Kakwani Index of Progressivity and Vertical Equity .............................. 193
    Horizontal Equity ............................................................................ 194
    Re-ranking Effect ........................................................................... 196
  8.7 Observations .................................................................................. 196

CHAPTER 9 ....................................................................................... 199
DUCLOS-JALBERT-ARAAR DECOMPOSITION FRAMEWORK ................... 199
  9.1 Introduction .................................................................................... 199
  9.2 The DJA Framework ....................................................................... 203
  9.3 The Modelling Issues in the DJA Framework ..................................... 210
  9.4 Social Determinants of Values of $\varepsilon$ and $\psi$ ................................. 214
  9.5 Decomposition Results from the DJA Framework ............................... 215
  9.6 Comparing Decomposition Results from AJL and DJA Frameworks ...... 220
  9.7 The Robustness of Decomposition Results ....................................... 223
  9.8 Stability of AJL Results ................................................................. 224
  9.9 Stability of the DJA Results ........................................................... 226
  9.10 Summary ...................................................................................... 227

CHAPTER 10 ..................................................................................... 230
SUMMARY OF FINDINGS AND CONCLUSION .................................... 230
  10.1 General Summary ........................................................................ 230
  10.2 Summary of Main Findings ............................................................ 231
  10.3 Policy Discussion and Conclusions ............................................... 240
REFERENCES ..................................................................................... 252
APPENDIX: QUESTIONNAIRE ............................................................. 252

  vii
List of Tables

Table 2.1 SAP and Obstetric Indices, Zaria, Nigeria 1983-1988 ................. 20
Table 2.2 Relative shares of various health care financing mechanisms in
Nigeria 1998-2002 .................................................................................. 38
Table 6.1 Distribution of Health as Measured by SAH and HIS Categories ........ 126
Table 6.2 Age and Gender Differences in Health by HIS ................................. 131
Table 6.3 Variable Definitions ....................................................................... 135
Table 6.4 Estimated Concentration Parameters ........................................... 136
Table 6.5 Sensitivity of CI to v of Yrsed ......................................................... 140
Table 6.6 Robust Estimation of the determinants of HIS ............................... 146
Table 6.7 Ordered Logit Estimation of Health in SAH ................................. 147
Table 7.1 Distribution of Gross and Health Expenditures by Quintiles ............ 151
Table 7.2 Incidence and Intensity of Catastrophic Payments ....................... 168
Table 7.3 Impoverishing Effects of Healthcare Financing ............................. 174
Table 8.1 Results from AJL Decomposition Based on Per Capita Income ...... 189
Table 8.2 Matrix of Gini coefficient for 5 percent ‘Income Equals’ ............... 194
Table 9.1 Decomposition Results Based on DJA Framework (ε = 0.4) ........ 216
Table 9.2 DJA Decomposition Results (ε = 0) ........................................ 217
Table 9.3 Results from AJL and DJA Decompositions Based on Per Capita ... 220
Table 9.4A Stability of AJL Results under Alternative Scales ..................... 224
Table 9.4B Changes in AJL Parameters under Alternative Scales ................. 225
Table 9.5A Stability of the DJA Results under Alternative Scales ............... 226
Table 9.5B Changes in DJA Parameters under Alternative Scales ............... 226
Table 10.1 Redistributive Effects of HCF in Fourteen Countries ............... 237
List of Figures

Fig 2.1 The Organizational Structure of the Nigerian Healthcare System..........................25
Fig. 2.2 Relative Shares of Various health care financing mechanisms in
Nigeria 1888-2002...............................................................29
Fig. 4.1 Lorenz Curve of Prepayment Income and Concentration of Postpayment
Income..............................................................................94
Fig 5.1 The Survey Area: Enugu State...............................................................101
Fig 6.1 Kernel and Normal Graphs of SAH..................................................124
Fig 6.2 Kernel and Normal Graphs of HIS..................................................124
Fig. 6.3 Kernel, Lognormal and Normal Densities.......................................125
Fig 6.4 Lorenz Curves of SAH and HIS.....................................................129
Fig 6.5 Different Values of Aversion Parameter.........................................141
Fig 6.6 Sensitivity of C to Changes in v.....................................................142
Figure 7.1 Lorenz Curve of Gross Exp and Concentration Curve of Payment.......157
Figure 7.2 Expected and Conditional Standard Deviation of Healthcare Expenditures........158
Fig. 7.3 Scatter Diagram of Gross Vs Expected Net Expenditure......................159
Fig. 7.4 Reranking of Household..................................................................160
Fig 7.5 Distribution of Healthcare Financing Burden Across income Groups........161
Fig 7.6 Pro-Poor Aversion to catastrophic Payments.....................................168
Fig 7.7 Concentration Curves of Hcat and Gcat............................................169
Fig 7.8 Concentration Curves of Hcat and Gcat (Cat. =10%) for Different values of v.169
Fig 8.1: Horizontal Equity and Reranking.....................................................178
Fig. 9.1 Cost of inequality using EDE (Duclos and Lambert 2000)....................206
Fig 9.2 V/H Plot for $\varepsilon = 0$ and $\varepsilon = 0.4$ ............................................218
## List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACPN</td>
<td>Association of Community Pharmacists of Nigeria</td>
</tr>
<tr>
<td>AJL</td>
<td>Aronson, Johnson, and Lambert (1994) framework for decomposing the Gini Index</td>
</tr>
<tr>
<td>ATP</td>
<td>Ability to pay</td>
</tr>
<tr>
<td>CBN</td>
<td>Central Bank of Nigeria</td>
</tr>
<tr>
<td>CI</td>
<td>Concentration Index</td>
</tr>
<tr>
<td>CIHI</td>
<td>Center for international health information</td>
</tr>
<tr>
<td>CMH</td>
<td>Commission on Macroeconomics and Health</td>
</tr>
<tr>
<td>COWAN</td>
<td>Countrywomen ass of Nig</td>
</tr>
<tr>
<td>CPI</td>
<td>Consumer Price Index</td>
</tr>
<tr>
<td>CSDH</td>
<td>Commission on social determinants of health</td>
</tr>
<tr>
<td>DAD</td>
<td>Distributive Analysis (software)</td>
</tr>
<tr>
<td>DALY</td>
<td>Disability adjusted life years</td>
</tr>
<tr>
<td>DJA</td>
<td>Duclos, Jalbert, and Araar (2003) framework for decomposing Gini Index</td>
</tr>
<tr>
<td>EA</td>
<td>Enumeration area</td>
</tr>
<tr>
<td>EDE</td>
<td>Equally Distributed Equivalent</td>
</tr>
<tr>
<td>FGN</td>
<td>Federal Government of Nigeria</td>
</tr>
<tr>
<td>FMoH</td>
<td>Federal Ministry of Health</td>
</tr>
<tr>
<td>FOS</td>
<td>Federal Office of Statistics</td>
</tr>
<tr>
<td>GNP</td>
<td>Gross National Product</td>
</tr>
<tr>
<td>HE</td>
<td>Horizontal Equity</td>
</tr>
<tr>
<td>HI</td>
<td>Horizontal Inequity</td>
</tr>
<tr>
<td>HIS</td>
<td>Health Status Index</td>
</tr>
<tr>
<td>HMB</td>
<td>Hospital Management Board</td>
</tr>
<tr>
<td>HRQoL</td>
<td>Health-related quality of life</td>
</tr>
<tr>
<td>HUI</td>
<td>Health Utility Index</td>
</tr>
<tr>
<td>IGR</td>
<td>Internally generated revenue</td>
</tr>
<tr>
<td>IMF</td>
<td>International Monetary Fund</td>
</tr>
<tr>
<td>LGA</td>
<td>Local government area</td>
</tr>
<tr>
<td>LGHDs</td>
<td>Local government Health Departments</td>
</tr>
<tr>
<td>MDG</td>
<td>Millennium development goals</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Square Error</td>
</tr>
<tr>
<td>NAFDAC</td>
<td>National Agency for Food Drug Administration and Control</td>
</tr>
<tr>
<td>NDHS</td>
<td>Nigerian Demographic and Health Survey</td>
</tr>
<tr>
<td>NEEDS</td>
<td>National (Nigeria) Economic Empowerment Program</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-governmental organization</td>
</tr>
<tr>
<td>NHIS</td>
<td>National health Insurance Scheme</td>
</tr>
<tr>
<td>NPI</td>
<td>National program on Immunization</td>
</tr>
<tr>
<td>OLS</td>
<td>Ordinary Least Squares</td>
</tr>
<tr>
<td>OPEC</td>
<td>Organization of Petroleum exporting countries</td>
</tr>
<tr>
<td>PHC</td>
<td>Primary Healthcare</td>
</tr>
<tr>
<td>PHI</td>
<td>Pseudo Horizontal Inequity</td>
</tr>
<tr>
<td>PSU</td>
<td>Primary sampling Unit</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>R</td>
<td>Reranking</td>
</tr>
<tr>
<td>RE</td>
<td>Redistributive Effect</td>
</tr>
<tr>
<td>RII</td>
<td>Relative index of inequality</td>
</tr>
<tr>
<td>SAH</td>
<td>Self-assessed health</td>
</tr>
<tr>
<td>SEG</td>
<td>Socioeconomic group</td>
</tr>
<tr>
<td>SII</td>
<td>Slope index of inequality</td>
</tr>
<tr>
<td>SMOH</td>
<td>State Ministry of Health</td>
</tr>
<tr>
<td>SWF</td>
<td>Social welfare Function</td>
</tr>
<tr>
<td>UNDP</td>
<td>United Nations' Development Program</td>
</tr>
<tr>
<td>VAS</td>
<td>Visual Analogue System</td>
</tr>
<tr>
<td>VAT</td>
<td>Value added tax</td>
</tr>
<tr>
<td>VE</td>
<td>Vertical Equity</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WLS</td>
<td>Weighted Least Squares</td>
</tr>
<tr>
<td>WTP</td>
<td>Willingness to pay</td>
</tr>
</tbody>
</table>
Summary

The economic reform policies of successive governments in Nigeria since the 1980s have moved the economy from state-dependent to market oriented processes. These reforms have involved the liberalization and deregulation of the economy and social services, withdrawal of subsidies in social services, many job losses and severe social and economic upheavals. It also aggravated the large income inequalities created by the elitist policies of the colonial days. The liberalization of the health sector, in particular, has resulted to an overwhelming dominance of the sector by private healthcare suppliers, direct financing of healthcare, and the introduction of user charges in public healthcare facilities. All these have moved the healthcare sector towards a competitive market model and shifted greater financial burdens on households. The consequence has been sharp drops in the use of medical facilities, re-emergence of old and emergence of new diseases, and worsening public health status.

The policy motivation for this research is, therefore, primarily, to investigate how in the absence of significant third-party financing mechanisms and government subsidies, direct purchase of healthcare affects the relative abilities of households to meet their other financial obligations after paying for cost of health services. In other words, this study aims to analyze the redistributive effect of direct healthcare financing in Nigeria. However, in order to have a more complete picture of the context of the redistributive analysis the study also examines the prior distribution of health in the Nigerian population as a prior condition for a complete distributional analysis of healthcare financing.

The sample data for the study was generated from a field survey in Enugu state in 2004 using a multistage survey design. Different methodological approaches are applied in order to achieve the objectives of the study. Two different frameworks for decomposing the Gini index into vertical and horizontal inequities, and reranking effects developed by Aronson, Johnson, and Lambert (1994), and Duclos, Jalbert, and Arrar (2003) are used to analyze the distributional consequences of direct healthcare financing in Nigeria. The
study examines also the convergence or otherwise of parameter estimates obtained using these two frameworks. The analysis of health distribution is done using the Gini index, the concentration and the extended concentration indices developed by Wagstaff et al. (1989, 1991) and Wagstaff (2002).

The study builds on existing literature on health distribution and redistributive effects of healthcare financing but also advances them in several important respects. For example, from the empirical perspective, almost all previous studies have been conducted in high income countries which have not only different epidemiological profiles but also different healthcare market structure, and healthcare financing methods that differ significantly from those prevailing in many developing countries like Nigeria. From a methodological perspective, the results obtained from two measures of health status: self-assessed health (SAH), and the health status index (HSI) using Visual Analogue Scale (VAS), are compared to see the extent of convergence between the two indices.

Borrowing from the tax literature, the study introduces aversion parameters that reflect the society’s aversion to inequality and aversion to uncertainty in post-payment income into healthcare financing analysis. Furthermore, previous studies used the Aronson, Johnson, and Lambert (1994) (AJL) method for decomposing the Gini index, this study in addition, introduces for the first time the Duclos, Jalbert, and Araar (2003) (DJA) method into the healthcare financing literature. It compares the results from the two frameworks showing the weaknesses of the previous estimations that used the AJL approach. The study also tests for stability of the estimated parameters of the two models under different assumption regarding intra-household allocation of resources and economies of scale.

From the analysis, the study finds that there are significant inequalities in health distribution in Nigeria, whether from the perspective of ‘pure’ health inequalities, or from the perspective of socioeconomic differences in health status. It also finds convergence between two indices of health status: SAH and HSI. Furthermore, the study finds large
incidence of catastrophic healthcare payments and impoverishing effects of direct healthcare financing in Nigeria.

With respect to the redistributive effects, the study shows that direct healthcare financing in Nigeria is generally proportional to income. For example, the study finds that households at the top 10 percentile of income pay relatively higher proportion of their income through direct health payments. But this could also imply that the poor tend to under utilize health services because they are unable to afford the cost of treatment. Yet, there are significant vertical and horizontal inequities as well as reranking effects in direct financing of healthcare. It should be noted, however, that the analysis of equity in healthcare financing does not give information about healthcare delivery or healthcare utilization. Such information can only be inferred indirectly.

Comparing the AJL and DJA frameworks for decomposing the redistributive effects, the study finds that the AJL method which has been the dominant method in literature has serious technical and conceptual weaknesses resulting in its underestimation of the components of the redistributive effects. But the components of the redistributive effect from the two models are shown to be unstable under different assumptions about intra-household resource allocation and household economy of scale elasticities.

The report is organized as follows. Chapter 1 sets out the policy motivations of the study. It states the problems the research intends to analyze and the objectives, significance as well as the limitations of the study. It is increasingly appreciated that medical care is only one, and not usually the main determinant of health outcomes. Chapter 2, therefore, provides elaborate profile of Nigeria, and the health sector in particular, emphasizing the structural defects in the society that generate the poor health and the healthcare financing outcomes. Chapter 3 looks at both the welfare economic theoretical grounds for distributional analysis of healthcare financing and makes a critical review of the existing empirical literature.
The methodological framework for empirical analysis is the concern of Chapter 4. It sets forth the basic frameworks for measuring health status and inequalities in health distribution. It also discusses the frameworks for estimating progressivity and redistributive effects of healthcare financing. Chapter 5 provides a full description of the survey methodology that was used to generate the data. Survey design issues, conceptual and measurement problems of variables are also discussed in the chapter.

Chapters 6 to 9 are devoted to problems of estimation and analysis of results. Chapter 6 deals with estimation and analysis of inequalities in health distribution. Chapter 7 presents information on healthcare financing distribution using various summary statistics. Chapter 8 reviews the AJL decomposition methodology and analyses the results obtained using this method while Chapter 9 reviews the DJA methodology, the underlying assumptions, and analyses the results, comparing them with those obtained from the AJL method. The final chapter reviews the entire study, summarizes the major findings and discusses the policy directions for healthcare financing in Nigeria as suggested by the results.
Chapter 1

Introduction

1.1 Background to the Study

Economic and social development policies in Nigeria since the mid 1980s have been marked by structural shifts from state-dominated to market-driven paradigms. While these shifts are considered necessary for improved economic performance, they have, nevertheless, accentuated the wider elitist motif that is operative in the general Nigerian political economy in which privilege and socioeconomic advantages have joined to create large inequalities between the affluent and powerful few, and the poor and powerless majority. In the provision of social services, this orientation, a carry-over from the colonial times, represents a complete dissonance between constitutional guarantees of rights to basic social services and the dominance of socioeconomic advantages as the most important parameters for access to these same services (Eze 1991; Alubo 1987). The end result is that majority of people is excluded from these services, or they must obtain them at great opportunity costs. This study focuses on the effects of these policy shifts in the health sector. In particular it analyzes how the policy of direct financing of healthcare based on ability to pay (ATP) that characterizes both the public and private supply of healthcare services affects the relative post-payment abilities of households to meet other basic needs. The study essentially aims to analyze the redistributive impact of direct healthcare financing by Nigerian households using survey sample data generated in Enugu state in 2004.

Socioeconomic reform policies in Nigeria have moved in tandem with global policy trends towards more market oriented economic and social reforms that have dominated the period of the 1980s and beyond. The essential theoretical underpinnings that sustain these reform programs are provided by neoclassical economics that gained momentum in
the 1970s. With its emphasis on the critical roles of the price system and competition in spurring economic growth, neoclassical economics provides the vision that underlies much of economic policy-making over the last two and half decades. The guiding economic principle is that the competitive market model provides the incentives and institutions that promote economic growth, discourage inefficiency and corruption that characterize large sections of the public sector. A key objective of economic policy-making in the neoclassical economic model is to ‘right size’ the state by freeing it from involvement in those economic and social processes that interfere with the workings of the competitive market ideals.

With the ascendance of conservative governments in the US, Britain and Germany in the early 1980s, the welfarist policies of previous decades gave way to the ideals of the competitive market. Multilateral agencies and institutions such as the International Monetary Fund (IMF) and the World Bank were in the forefront of the advocacy for this new policy approach to development. For many developing countries, including Nigeria, that were ridden with heavy debt burden and who approached the IMF for rescue, assistance was conditional upon the adoption of a reform package known as the IMF ‘Conditionalities’, or the Structural Adjustment Program (SAP). These conditionalities comprised a package of neoclassical economic policy instruments such as devaluation, privatization and liberalization of the economy, removal of import tariffs and all forms of trade restrictions, removal of all forms of economic subsidies and price distortions, rightsizing the state by retrenching workers in unnecessary state involvement in production processes, and creating more space for the competitive market. Many countries that adopted the policy package and were persevering in its implementation recorded significant economic improvement. But some others did not. But its initial impact was, almost everywhere, social and economic upheavals.

development agenda: reducing poverty, a program that has since remained in forefront of economic development research and policy ever since. The report proposed strategies for meeting the challenges of poverty: policies that promote all-inclusive growth and give value to the resources of the poor, and access of the poor to social services. After summarizing forty years of development experience, the 1991 report suggested that for poor developing countries to achieve economic growth and development, they had to adopt more market oriented policies. While the market should be allowed greater space in order to achieve economic growth the state still provides those public goods in which the private sector under-invest such as social infrastructure, education, health, legal framework for market regulation, improving the environment for the competitive market, liberalization of trade and so on.

The wave of health sector reforms that followed this new policy direction was basically motivated by the need for increased efficiency in the sector and to initiate new source of funding health services (World Bank, 1993, Akin et al. 1987). It also marked the beginning of the domineering influence of the World Bank in the health sector (Irwin and Scali, 2005). In the presence of severe resource constraints and rising cost of healthcare, particularly in developing countries, the need for greater efficiency in the utilization of available resources became imperative. In many industrialized countries alternative sources of healthcare funding including the policy of cost-sharing were introduced. In the poorer developing countries the policy of cost recovery including user fees were introduced.

In the 1990s, the effects of the introduction of these cost recovery measures were beginning to manifest. Policy-makers and researchers debated the likely effects of user fees on access to healthcare especially on the most vulnerable members of the population (Mwabu et al. 2003). The argument was whether user fees actually constrained access to healthcare utilization, particularly for the poor. For some on the one side of the debate, health is a basic need. Like other basic commodities it has a low elasticity of demand. Therefore, price increases in healthcare services could not be a barrier to access in consumption of healthcare services. Many on the other side of the debate, however,
argued to the contrary. A consensus seems to have emerged that user fees constrain access of the poor to healthcare service (Ichoku and Leibbrandt, 2003). Even the more recent World Bank publications now seem to acknowledge that prices could be impediments to access to healthcare services by the poor (see for example, Sachs et al., 2001). The other aspects of the health sector reform package that were inspired by the neo-liberal economic blue-print such as health sector decentralization have also come under negative critical appraisals (Homedes and Ugalde, 2005, Irwin and Scali, 2005, ). While a lot of research effort has been expended on the effects of price changes on healthcare utilization, what has hardly be considered in literature is the impact of these new financing policies on income redistribution. How do these new health sector reform policies square with the demands of equity in healthcare financing?

While research on the effects of prices change on healthcare demand was making progress, international literature was accumulating evidence on the nature of health distribution in developed economies with a long tradition of keeping records. The long held suspicion that a positive correlation exists between wealth and health was being supported with scientific evidence. There is uneven distribution of health itself and, in particular, that health distribution tends to follow the socioeconomic gradient. These studies suggest that there is a strong positive correlation between health and the socioeconomic gradient (The Black Report 1980; van Doorslaer et al., 1997; Macenbach et al. 1997; Lahema et al. 1997; Pritchett and Summers 1993; Vladimir et al. 1998; Macintyre, 1997; Arber, S. (1997; Valkonen, 1997; West, 1997; Marmot, M. et al. 1997; Lahelma, E. and Valkonen, 1989; Szalai, J. 1986; Hexel, and Wintersberger, 1986). Health outcome, whether measured as rate of mortality, self assessed ill-health (SAH) or life expectancy, and socioeconomic status, whether operationalized as educational level or occupational category or income level, has often shown a close association. It is purely a normative issue whether policymakers should be concerned with pure health inequalities (those resulting from biological differences between people) or merely with those aspects of it that are policy-responsive (those caused by differences in conditions of living). Literature on health inequalities has tended to follow one or the other of these two strands.
More research findings show that the world’s least advantaged population bears a disproportionate burden of ill-health and disabilities. For example, evidence show that in 1990 communicable diseases which characterize poor living conditions were responsible for 59% of deaths and 64% of disability adjusted lost years (DALY) among the poorest 20% of the world’s population while communicable diseases accounted for only 8% of deaths and 11% of DALY loss among the world’s richest 20% (Gwatkin et al. 1999; Lere et al., 1998). Disadvantaged populations not only suffer greater burden of disease, they also experience the onset of chronic illnesses more than their more affluent counterparts (Whitehead, 1992).

Furthermore, it has been generally recognized that the most important asset of the poor is their labour (Kanbur and Squire 2001, World Bank, 1990). Ill health can substantially devalue this asset (Okonjo-Iweala, 2001). Thus, there is also a nexus between poverty and health which mutually reinforce themselves (Thiede and Traub, 1997).

This evidence seems to suggest the conclusion that those at the lower rungs of the social ladder may not only bear greater disutility of illness, but also greater financial costs in terms of loss of productive resources (labour and leisure hours and non-monetary costs of access to healthcare), and financing the health sector under the new reform agenda. Thus, it would seem that an estimation of the redistributive effect of healthcare financing must also fundamentally explore the extent of inequality in health distribution in order to present a fuller picture of the distributional consequences of any financing arrangements. Not factoring in this prior disadvantage would seem to underestimate the implications of the redistributive effect of healthcare financing.

Inequalities in health distribution reflect the demand side of the healthcare market while payments for healthcare reflect the pattern of distribution of healthcare or the supply side of this market. It is important to consider these two distributions under a unified framework.
1.2 The Motivation of the Research

In the past, development economists thought that income inequality was necessary to facilitate economic growth (Lewis 1954) or at best a necessary outcome of growth (Kuznet, 1955). Huge economic inequalities, especially in developing countries, were therefore thought to be part of the growth process. However, many economists are today becoming increasingly less sanguine about this theoretical view point. Indeed, economic inequality is now seen more of a hindrance to growth than a complement to it because it tends to precipitate social conflicts and political instability. It has, therefore, become a central concern for many societies.

Huge inequalities in the distribution of and access to economic to resources are key to understanding the political economy of the Nigerian state. Income distribution has become a central policy issue in Nigeria, not only because it accounts to a large extent for the increasing scope and severity of poverty in the country [Federal Government of Nigeria (NEEDS), 2004] but has also accentuated the level of conflict in the country. There is increasing violence and threats to property rights which tend to deter investment and economic growth.

Fiscal policies which have traditionally been used as income redistributive instruments have been impotent in this regard on account of several factors including large informal sector, narrow tax base, and corruption (Nwosu, 2000, Anyanwu, 1997). Further more, much of the national income arises from state controlled oil exploration activities. The revenues from oil resources are channelled through the central government. This implies that much of the national income is liable, and indeed has been under capture by the privileged groups and individuals who have access to power. This has promoted what Joseph (1987) calls the “politics of prebendalism” defined as “the intensive and persistent struggle to control and exploit the offices of the state” It has created huge resource gaps between the rich who have access to state power (and those connected to them), and the poor who are powerless.
Furthermore, following the worldwide trend of economic reform, the Nigerian government has put in place several reform policies especially in the liberalization, and commercialization of public utilities, and the privatization of state enterprises. These reforms programs have tended to transfer greater economic resources to those who have the means. At the same time they have increased the responsibility for financing of social services such as education and health to households.

A major plank in the reform policy of the health sector since the late 1980s is the deregulation of the demand and supply of health services. The policy of cost recovery in public health institutions through user charges and the prevalence of for-fee private healthcare providers have tended to lessen government responsibility for healthcare and moved the healthcare sector towards a competitive market structure. But it has also increased the healthcare finance burden for households. It would seem, therefore, that the system of healthcare financing in Nigeria has potential to worsen the existing large income inequalities in the country by making relatively more financial and economic demands from the poor than the non-poor.

The major policy question this research intends to address is: what are the income redistributive implications of this pro-market healthcare financing policy? Put differently, does the distribution of income in the post-payment become more uneven than the prepayment income distribution? To what extent does healthcare financing feed into the national income inequality stream? In this respect, Wagstaff and van Doorslaer (2001) make a useful distinction between equity in utilization and equity in financing of healthcare. Here the concern is not so much with healthcare utilization (though this is also very important) as with the question of the impact of healthcare financing on net incomes or the relative abilities of households to finance other needs at the post-payment period.
1.3 The Policy Concern of the Study

One may suspect that the concern of policy is not limited to the distribution of healthcare utilization (who receives and who does not which is assumed to determine who gets cured and who does not in the event of illness) but also to the problem of redistribution of income in the post payment period. Indeed, a new official national health policy of the Nigerian Federal Ministry of Health (FMoH) explicitly emphasizes this concern. The policy states that eliminating avoidable inequalities in health through greater access of the poor to healthcare services (which are assumed to be restorative of good health) can significantly redistribute economic resources and reduce economic inequalities. It considers the ill-health variable as an important determinant of poverty. Therefore, addressing the health needs and the redistributive effects of healthcare financing ought to be a major component of poverty reduction strategy of the Nigerian government (Federal Ministry of Health, 2003)

Why is the post payment income distribution of concern to this study? The concern arises from the fact healthcare payments can eat so deep into the pocket of a household that it has little or nothing left to provide other basic life necessities such as food, and investments such as the education of the children. This obviously differs from the concern whether the household is able to utilize healthcare services

The basic presupposition is that if the distribution of health is skewed against those at the low rungs of the socioeconomic status, then the generally accepted principle of cross-subsidization and healthcare financing according to ability to pay (ATP) implies also pro-poor income redistribution if income is the measure of that ability to pay. In this case healthcare utilization decisions are independent of healthcare financing decisions. On the other hand, if healthcare is not financed according to ATP but according to need (in which case healthcare utilization decisions are not independent of healthcare financing), then the resulting income redistribution is likely to be pro-rich.
The concern is that social decision-makers may desire that healthcare payments exert an equalizing impact on income distribution. They want to ensure that the system of healthcare financing does not worsen the distribution of income. By implication the policy-maker wants to ensure that poor households are not pushed to the brink of existence because they have little or no income left to provide the other basic necessities of life. This concern is based on the fact that since ill-health is a stochastic and unpredictable event, it is important that the society should hedge particularly its most vulnerable members against the financial cost of such an adversity so as not to decrease the probability of their survival at the post-payment period (Wagstaff, 2002)

1.4 Objectives of the Study

The general objective of this research is to analyze the pattern of health distribution and income redistributive effects of healthcare financing policy in Nigeria. This broad objective is achieved through the following more specific objectives:

(a) To analyze the structure of health distribution among Nigerians and quantify the extent of inequalities in pure health, and those arising from socioeconomic differences;

(b) To determine the contributions of socioeconomic factors to over all inequality in health distribution and the sensitivity of socioeconomic inequalities in health to policy-maker’s aversion to inequality;

(c) To analyze the socioeconomic determinants of health;

(d) To provide a general description of the relationship between income and healthcare expenditure and assess the extent of catastrophic healthcare payments among Nigerians;
(e) To assess the extent of impoverishment that results from direct healthcare payments;

(f) To quantify the extent of income redistributive effect arising from direct financing of healthcare by Nigerian households;

(g) To decompose this total redistributive effect and establish the extent of progressivity, vertical equity, horizontal equity and reranking of households inherent in this system of healthcare financing in Nigeria using two decomposition frameworks;

(h) To examine the sensitivity of the components of total redistributive effect in (g) above to alternative definitions of 'prepayment equals';

(i) To establish the sensitivity of horizontal inequity in direct healthcare financing to policy-makers aversion to inequality;

(j) To establish whether the results obtained in (g) above are robust to household size and economies of scale;

(k) To compare the relative performance of different frameworks for decomposing redistributive effects.

To address these objectives, and in the absence of pre-existing data, a field survey was undertaken to generate suitable primary data for the analysis. Objective (a) was addressed by using non-parametric kernel graphs and the concentration index of inequality. Objective (b) was addressed using the extended concentration index which allows the possibility of incorporating an inequality aversion parameter that reflects the social decision-maker's level of pro-poor aversion to inequalities in health distribution. The Ordered Logit with Huber/White robust variance estimator was used to address objective (c). Objective (d) was addressed using summary statistics, tables and graphs. The extent
of catastrophic healthcare expending was investigated using extended concentration graphs, catastrophic headcount and weighted catastrophic payment gap indices while objective (e) was addressed using the conventional poverty headcount, poverty gap, and square poverty gap indices. Objectives (f) and (g) were addressed using the Aronson, Johnson, and Lambert (1994) (AJL) decomposition framework and the Duclos, Jalbert, and Araar (2003) (DJA) decomposition framework. Objective (h) was addressed using different definitions of income band in order to evaluate their effects on the components of the redistributive effect under the assumptions of the AJL decomposition framework. Objective (i) was addressed by evaluating the robustness of estimated inequity parameters to different values of inequality aversion parameter in the DJA framework. To address objective (j), different assumptions were made about intra-household resource distribution and household economies of scale and the estimated components of the redistributive effects under both the AJL and DJA frameworks were analyzed for differences in size and magnitude. Finally, objective (k) was addressed by analyzing the estimated components of the redistributive effects under the frameworks.

1.5 Justification of the Study

Empirical evidence shows that inequality in the distribution of resources is a major problem in developing countries and in Nigeria in particular. The research is particularly germane in the Nigerian context because highly uneven distribution of income is perceived to be inducing the socio-economic and political instability in the country (FGN – NEEDS, 2004; Egwu, 1998). This socioeconomic and political instability is evidenced in frequent riots, crime and violence, incessant strikes by labour unions, ethnic conflicts and youth restiveness (especially in the oil bearing regions of the country), ethnic accusations and counter accusations of economic and political marginalization. These have often burst the social fabrics at the seams, and they arise from unfair distribution of resources among different classes of people and geographical locations. Minimizing avoidable income inequalities is an essential factor in forming a stable polity and building a progressive economy and poses a major challenge to policy-makers.
More specifically, Nigeria lacks a universal public system of healthcare financing and is therefore likely to experience worsening income redistribution on account of the system of healthcare financing where 70% of the people live in poverty and yet over 70% of healthcare finance is through direct payment (Federal Ministry of Health 2003). A reformed healthcare financing system in Nigeria has, therefore, potential for pro-poor redistribution of income if healthcare utilization is according to need and payment is according to ability. By so doing the poor can use the money they spend on healthcare to provide themselves other households necessities.

Furthermore, it is generally accepted that almost every economic policy produces redistribution of income or wealth through direct or indirect channels. Redistributive effects of public policies often appear disguised and therefore require analysis to identify their consequences. Making the redistributive effects of policy options explicit also helps to distinguish between desirable and undesirable redistributions on the basis of equity (Acocela, 1998) and also to choose those policy options with least distortion on efficiency (Stiglitz, 1989).

1.6 Significance of the Study

The increasing global concern about the situation of the world’s poor and the need to redistribute resources in the face of increasing inequalities at national and global levels (Stern, 2001) have brought about a coalition of interests among governments, multilateral nongovernmental and donor agencies in an effort abolish extreme poverty among the world’s population (Sachs 2005). Many of these agents of development have found common ground in their commitment to programs and projects that are consistent with this common interest of poverty reduction and redistribution of resources.

However effective utilization of these resources and gains in synergies can only be fully utilized if there is sufficient information on the various aspects of deprivation and sources of inequalities that impoverish some and benefit others. This study would, therefore, be
of interest to public policy-makers and organizations that are interested in designing programs that attack the root of poverty and inequality in Nigeria.

One of the key concerns of the new health sector reform program of the Nigerian government is to ensure that health financing promotes even distribution of income. It is therefore, expected that the results of this research will be of interest to the ministry and all stakeholders in the healthcare provision and consumption.

1.7 Limitations and Assumptions of the Study

According the Federal Ministry of Health (FMoH) over 70% of healthcare utilization in Nigeria is financed out-of-pocket (Federal Ministry of Health, 2003). This is considered a conservative estimate since over the years healthcare spending has not been a priority of the government [Alubo 1987, 1994, Ogunbekun et al. 1999, World Health Organization (WHO) 2002]. It accounts for less than 3% of national budget and an insignificant portion of GNP (Anyanwu 1997; Nwosu, 2000). Under this circumstance, it seems right to assume that the estimation of the redistributive effect of healthcare financing is to be based on the amount paid for healthcare services. On the other hand, if healthcare services were provided free of charge or at significantly subsidized rates, the evaluation of the income redistributive effect of healthcare financing would require netting-off the gain (from utilization) from the loss (through payment). In that case the net income effect would show the change in income distribution that arises from health payments. To obtain information on how much different households benefit from subsidy in order to net it off the amount paid for healthcare would require a benefit incidence analysis. This study does not intend to carry out such a research. It rather assumes that in the absence of significant government subsidies on healthcare, and other third party interventions, that households must purchase healthcare at the going market prices. This implies that consumers have to determine the amount of healthcare to consume when they are deciding how much care they could pay for. In other words, it is assumed that in the
present context, the decision to utilize healthcare corresponds to the decision to finance healthcare.

It is, furthermore, important to acknowledge the partial equilibrium nature of the analysis of this study. Obviously, health financing decisions will modify other economic behaviour. A progressive healthcare financing for instance will impact on the incentive to supply labour, on household savings behaviour and even on relative market prices. Some of these influences will be direct. Others may be indirect. The study of the totality of such effect will require a general equilibrium analysis. Such a general equilibrium analysis may be pursued in a future work but the focus here requires the ceteris paribus assumption.

Furthermore, it need be emphasized that this study does not undertake a distributional analysis of all forms of healthcare financing in Nigeria. To do this would involve analyzing the distributional impact of healthcare financing through public expenditure, and perhaps also, the distributional impact of private and social health insurance. This study limits itself specifically to health financing in Nigeria through household direct out-of-pocket healthcare payments. In other words the study is a partial analysis of one (though the dominant) source of health financing in Nigeria.

It is also possible that there might be, in a few cases, some exemption policies and charities by private healthcare providers which may lead to greater cross subsidization. But such exemptions policies, if they exist at all, do not cover significant areas of health treatment costs. This study, therefore, assumes that households generally bear the cost of treatment when a member is ill.

In this first chapter, the attempt has been made to place this study in the context of the policies and research developments that motivated this study. The chapter tries to articulate the policy concerns that motivated the study and to set forth clearly defined objectives that the study intends to achieve. It also emphasizes the significance, the scope, the limitations and assumptions of the study. The rest of this report is organized as
follows. Chapter 2 focuses on the characterization of the Nigerian socioeconomic, political and policy environment under which the healthcare system with its financing policy operates. This is to enable the reader appreciate better the need for this study and assess its eventual outcome. A broad literature review is undertaken in Chapter 3. In the first section of this chapter an attempt is made to provide the economic theoretical framework for redistributive policies in general and in healthcare financing in particular. The other section of the chapter reviews existing international empirical literature on redistributive effects of healthcare financing. In Chapter 4 a general methodological framework for the analysis is provided but the details of the methodology are worked out in the relevant chapters and sections where the issues are analyzed. As earlier indicated, this study uses primary data. The process of generating the sample data, the survey issues, the problems and difficulties encountered as well as problems concerning variable definitions and measurements are taken up in Chapter 5.

Chapters 6 to 9 are devoted to empirical estimations and analysis. Chapter 6 is devoted to the analysis of inequalities in health distribution among the population. Chapter 7 provides different forms of summary statistics, and indices on healthcare expenditure in Nigeria. It also analyzes the problems of catastrophic payments and impoverishing effects of healthcare financing. In Chapter 8 the details of the AJL decomposition framework are presented along with the results and analyzed. The DIA framework is also described and results from this framework are assessed in Chapter 9. This chapter also compares the results from the two approaches and shows the underlying differences in assumptions that account for differences in estimated results. Chapter 10 summarizes the findings of the study and discusses the possible policy directions suggested by these findings.
Chapter 2

The Socioeconomic and Policy Contexts of the Study

"Do we not always find the diseases of the population traceable to defects in society?"

Rudolf Virchow (1821 – 1902)

2.1 Introduction

Since policies do not operate in a vacuum, it is seems appropriate to describe here the larger socioeconomic environment under which the health sector operates because that environment also determines the priority given to healthcare and the resources available to the health sector. In a background paper prepared for the Commission on Social Determinants of Health (CSDH) of the WHO, Irwin and Scali (2005) emphasized the close links between the socioeconomic environment and health in achieving the Millennium Development Goals (MDG). They note: "Without progress in fighting poverty, strengthening food security, improving access to education, supporting women’s empowerment and improving living conditions in slums, for example, the health-specific MDGs will not be attained in many low—and middle-income countries. At the same time, without progress in health, countries will fail to reach their MDG targets in other areas" (Irwin and Scali 2005:30). These inter-linkages between the socioeconomic environment and health underline the purpose of this chapter.

Human welfare in any society and at any point in time is largely a product of choices and decisions made in the preceding periods as well as of processes and institutions that determine individual and social responses to their environment which include psychological, economic, political, and other variables. It seems desirable, then, that the study of a national health financing and its impact on income redistribution should
military coups and several unsuccessful ones between 1960 and 1999. Thus, between 1960 and 2000 military leaders that came to power through coups had ruled the country for thirty-four years.

Politics even under democratic governments became fundamentally the struggle for the control of the enormous resources concentrated at the centre and often used for patronage by the government in power. Government increasingly became “a magnet for all facets of political and economic life, consuming the attention of traders, contractors, builders, farmers, traditional rulers, teachers, as much as that of politicians or politically motivated individuals” (Ayogu, 1999:170). Access to state resources or being 'politically connected', therefore, became a major determinant of income and wealth distribution in Nigeria. Those with access to state power corruptively enriched themselves and those connected to them while being poor defines lack of access to state power and resources. According to Alubo (1987:454) “This elite phenomenon is central to the analysis and the present operations of the Ministry of Health” Healthcare services are organized to suit the certain interests, those who have resources and those who are “politically connected”

Since much of health and healthcare utilization to a large extent result from the underlying structure of the economy (Sachs et al. 2001), it is important to place the Nigerian healthcare financing system in the larger context of the Nigerian economy. The Nigerian economy has undergone three major structural changes each of which has had profound effects on the health of Nigerians. The first of these changes was the switch from staple food production to cash crop production under British colonial rule. The effect of this was a switch of taste from local food to imported processed food items which were considered nutritionally ‘superior’ (Watts and Lubeck 1984:116). Such policies led to dependence on imported food and the consequent loss of ability of local communities to feed themselves. A different structural change took place at the end of the civil war (1967-70). The growth of the Nigerian economy was rapid. Between 1970 and 1976 the GNP grew at real annual rate of 7.4% and GNP per capita at 5.4% (Watts and Lubeck 1984). By 1973 Nigeria had become a major exporter of oil and was a beneficiary of the OPEC oil hike of 1973. Oil revenue grew from mere 3% in 1963 to 46% of GDP in
1974. The sharp rise in oil prices led the neglect of the neglect of the agricultural sector as people, especially the youth, deserted the rural areas and went to the cities in search of better life. By 1980 Nigeria had become a net importer of food and most traditional export cash crops had disappeared from the export list (Central Bank, 1982). As the cities burst at the seams, there developed the urban satellite ghettos to accommodate the poor city immigrants. But these suburbs were bereft of social amenities. Water, good roads, hospitals, recreational facilities etc. were in short supply in these areas. The situation created severe health problems.

Perhaps the most radical structural change in the Nigerian economy was the Structural Adjustment Program. The impact was deepest both in terms of suddenness and duration. Due to the crash in oil prices following the world economic recession and debt crises of the early 1980s, the country experienced severe fiscal crises. The fledgling democratic government (1979-1983) imposed economic austerity and cut down on social spending. The health sector budget was one of the worst hit. Social discontent and violent demonstrations led to the fourth military coup in 1983. The situation hardly improved with the new military government. Cuts in health spending led to one of the worst industrial strikes by the powerful Nigerian Medical Association (Ogunbekun, 1991). Another military coup took place in 1985. The new military government obtained the IMF loan with its severe ‘conditionalities’ Thus began the era of SAP in Nigeria.

2.3 Macroeconomic Policies and the Health Sector

The macroeconomic environment has serious implication for health policies and health outcomes (Sachs et al. 2001). In the first place the environment determines the demand for healthcare. In the second place it determines the extent of financial constraint on both public and private healthcare supply (Genberg, 1993). In Nigeria the dominant macroeconomic policy framework from 1986 to 1994 was the Structural Adjustment Program (SAP). As an economic policy SAP is based on the neo-classical paradigm of supply-side economics. It involves tight fiscal policy that includes sharp cut in government spending, cuts in social sector budget, removal of public sector subsidies on
The gains made with some preventable diseases during the period of economic growth were soon reversed, as some of the diseases like malaria, typhoid fever, tuberculosis, malnutrition etc soon made a come back (Pearce and Falola, 1994). A number of studies have been carried out to evaluate some of the impact of SAP on the welfare of Nigerians.

The impact of SAP on many developing countries that implemented is summarized by Irwin and Scali (2005:20) in the following words:

“The sudden layoffs propelled huge numbers of people into unemployment, with no safety nets and little chance of finding formal work in the private sector in many cases. The negative health effects for individuals, families and whole communities have been documented. In some countries, particularly in southern Africa, the resulting social destabilization and insecurity contributed to hunger, the propagation of armed conflict and the rapid spread of HIV/AIDS – with the poor, women and other socially disadvantaged groups bearing the brunt of the damage. As a result of SAPs and the global economic malaise, social sector spending in many countries plummeted during the 1980s, with negative effects on the health status of vulnerable communities. In the poorest 37 countries in the world, public spending on education dropped by 25% in the 1980s, while public spending on health fell 50%.”

A study that evaluated the socioeconomic impact of SAP in Nigeria noted:

“What comes out clearly from most of the studies is that SAP, rather than overcoming, has heightened the crises of the Nigerian economy and society. The consequence is that the country’s external debt profile, unemployment, inflation, industrial collapse and general economic hardship have risen to unprecedented levels during the years of adjustment. Some of these studies have shown that SAP has had ruinous effects on industry, agriculture and growth” (Egwu 1998:7).

With the coming of another military government in 1993, SAP was abandoned 1994 and the policy of “guided deregulation” was introduced. The few gains made under SAP especially in the agricultural and manufacturing sectors were soon reversed. For the rest of the 1990s the economy hardly recorded any appreciable growth. In fact for most of the 1990s the net growth of the economy was negative. Indeed by the time a new civilian
government was elected in 1999, about 70% of Nigerian households were living below the poverty line (defined as persons living on less than $1 per day) (Federal Ministry of Health, 2003)

Apart from these lasting effects of SAP and the political environment, another major source of macroeconomic instability is the national revenue and the inability of policy to mitigate the effects of unstable revenue flow. Oil revenue accounts for about 80% of export earning. However, since the price of oil is exogenously determined by world economic situation, it implies that government has little or no control over its fiscal and even monetary policies. Swings in the international oil prices determine to a very large extent the Nigerian macroeconomic environment. Annual budgets are based on the anticipated prices of oil which fluctuate frequently. Over the years, policies have not been right to smooth the effects of these fluctuations on national consumption. The consequence is that the economy is subject to frequent booms and bursts and this not only impact on health itself but also the amount of resources available for health needs.

2.4 Poverty and Income Inequalities in Nigeria

It is difficult to ascertain the exact level of income inequality in Nigeria (because of differing figures from different non-comparable data sources) but a number of studies show that inequality has worsened over time. A popular measure of extent of income or expenditure inequality is the Gini coefficient. World Bank 2001 estimates a Gini coefficient of 0.515. Using two national consumer surveys 1985/6 and 1992/3 Canagarajan et al (1997) showed that income distribution in Nigeria has been worsening over time. For example, they showed that the bottom five percent of the population had a large declining mean per capita expenditure of about 40.5 percent between the two periods while households in the second 5 percent had a decline of 20 percent. In contrast,

\footnote{The Lorenz curve plots the cumulative proportion of the population on the horizontal (in ascending order) and the cumulative share of income (expenditure) on the horizontal. For a perfectly egalitarian society, the resulting curve will lie on the 45\degree. Deviations from this diagonal measure the extent of inequality in income (or expenditure) distribution. The Gini coefficient is the summary measure which divides the area between the curve and the diagonal as a proportion of the area of the entire area below the diagonal. The Gini takes values that range from 0 (for a perfectly egalitarian society) to 1 for absolute inequality.}
households in the third decile had a mean per capita increase of 5 percent while those in the tenth decile had an increase of 47.5 percent between the two periods.

In a more recent and comprehensive study, Okojie et al. (2000), shows that inequality was exacerbated by the introduction of the Structural Adjustment Program (SAP) in 1986. The program drastically reduced government expenditure on social goods and services especially in health and education and thus accentuated the problem of inequality in the country (Central Bank of Nigeria/World Bank, 1999). The study shows that between 1980 and 1985 the national Gini coefficient for consumption expenditure declined from 0.502 in 1980 to 0.423 in 1985 and then worsened again between 1985 and 1992 to 0.507 before declining to 0.465 in 1996. These figures may be compared with those of other regions of the world. Latin American countries have average Gini coefficient of 0.5, developed countries have average of 0.32 and Asian countries have average of 0.41 (Suarez-Berenguela, 2000).

While the results of these different studies may not be directly comparable, they, nevertheless, give a picture of extent of inequality in the country. In general, empirical studies show that inequality in the distribution of resources is a major problem in Nigeria that warrants an in depth study of the channels that feed into the observed stream of inequality. This is consistent with an observation in the editorial of one the national newspapers:

"Nigeria is a rich country in which 75% are poor while the top 1% usually those with strong links to the corridors of power rival the affluent anywhere in the world. Income disparity between the top 1% and the bottom 75% in Nigeria is the worst in the world according to a U.N study"
(Vanguard Comment: August 8, 2003)
These social and economic developments have fundamentally shaped the national health profile. The next subsection describes the health outcomes under these prevailing socioeconomic conditions.

2.5 The Health Status of the Nigerian Population

The paucity of social support programs is reflected in the low priority given to social welfare programs in the public expenditure over the decades. Alubo (1987:453) puts it succinctly “… medical services for the generality of the people have remained a second rate priority of post-colonial governments, very much like the situation in colonial days. The care for state employees and other elites continues to take precedence” Health expenditure was mere 1.9% of total federal government expenditure in the 1980s (Orubuloye and Oni 1996, Ogunbekun 1991). The share of public health spending as a proportion of total budget improved to 2.55% in 1996, 1.96% in 1997, 2.99% in 1998, 1.95 in 1999 and 2.5% in 2000 (Central Bank of Nigeria 2001). Government healthcare funding is less than 0.2% of GDP (UNDP, 2000) as against the 1% benchmark set by the Commission on Macroeconomics and Health (Sachs et al 2001) and 15% of annual national budget stipulated in the Abuja 2001 Declaration on roll-back malaria program. This amounts to $2 healthcare subsidy per capita whereas healthcare expenditure per capita is $15. In other words, government is responsible for only about 13 percent of the entire healthcare expenditure in the economy. The WHO benchmark for public health spending in Low Income Countries is $34 per capita (WHO, 2002). This implies that social safety nets for the less advantaged members of the society are virtually non-existent. As Alubo (1987:1) summarizes it, “In practice, ..., status, power and privileges determine whether or not one gets Western medical services and of what type in contemporary Nigeria”

Consequently, life expectancy is very low. It is estimated to be about 52 years [Center for International Health Information (CIHI) 1996)]. Infant and child mortality rates are very high: 75 and 140 per 1000 respectively [Nigeria Demographic and Health Survey (NDHS), 2000]. NDHS (2003) paints even a more worrisome picture infant mortality rate
in Nigeria is compared with some other poorer African countries as shown in Figure 2.1. About 14 percent of Nigerian children do not survive their fifth birthday. The main causes of child mortality are malaria (30%) and diarrhoea (20%). The incidence of stunting, a measure of chronic malnutrition among children, is very high. In general, about 45% of Nigerian children are reported to be malnourished while 25% are severely malnourished (NDHS, 2000).

Maternal mortality rate (MMR) is estimated to be 800/100,000 live births. The prevalence rate of malaria in the general population is estimated to be 919/100000; dysentery, 386/100000, pneumonia, 146/100000, measles 89/100000. About 40 million are exposed to onchocerciasis while about 20 million are infected (WHO, 2002). The prevalence rate of HIV is not known but it is estimated that about 5.8% of the population is infected (WHO, 2002).

![Fig 2.1 Comparison of infant mortality rates across selected African countries](image)

Fig 2.1 Comparison of infant mortality rates across selected African countries

Source: Nigeria Demographic and Health Survey (NDHS, 2003)

Access to social services is generally very low, though worse in some places than others. The incidence of non-communicable diseases is also on the up-word trend. For example, a survey in 1989 showed that 3.5 million suffered mild hypertension and is now
estimated to affect about 10 million people. About 2.75% of the population is said to suffer diabetes mellitus (FMoH 1989)

This deplorable state of the health sector arising mainly from the gross under funding of the healthcare sector has led to exodus of trained medical personnel to other countries. Many of Nigerian doctors are now working in different countries in Europe, America and Saudi Arabia. Nigerian doctors constitute the greatest number of foreign doctors from a single country working in the US. Professor Ransome-Kuti, (Former Federal Minister of Health) gave the number of Nigerian medical doctors working in the US alone as 3000 as at 1997 (Ransome-Kuti 1997; see also Sachs et al. 2001)

The consequence of the exodus from public to private practice among Nigerian medical professionals is the crowding of the urban centres, especially in the southern states, with private hospitals and clinics. Ogunbenkun reports that between 1983 and 1987 alone there was an increase of 37% in the number of registered private medical establishments (Ogunbekun 1991). The rural areas are hardly getting attention. Most rural clinics and hospitals depend on fresh medical graduates on National Youth Service Corps (a compulsory one year national service for all Nigerian graduates) to staff their institutions. As the current serving members pass out of their service year, they are replaced by another batch of fresh medical graduates. In other words, these establishments must change their doctors every year.

The picture that emerges from this characterization of the Nigerian socioeconomic environment is the failure of existing policy framework to guarantee improvement in the social welfare of the majority of the people. In the midst of abundant resources, poverty is spreading and deepening. Health indicators are worsening due to inadequate funding. It is important at this point to examine the system of health financing in Nigeria which has brought about this deplorable state of the health sector. But since the fiscal arrangements that operate in the country have fundamental implications for healthcare financing it is important to describe briefly the nature of the fiscal responsibilities in the Nigerian system.
2.6 Fiscal Federalism in Nigeria

The political history of Nigeria shows the emergence of a nation from a combination of pre-existing sovereign jurisdictions that were amalgamated into one. This political history has necessitated the adoption of multiunit fiscal structure for the country. At present, Nigeria operates a federal fiscal structure with three tiers of government – the federal level, 36 states and 774 local governments. The major sources of government income such as petroleum tax and Value Added Tax (VAT) are paid into a common pool - the Federation Account - from where they are shared out to the various tiers of governments according to their levels of responsibilities.

The proportion of shares of the federation account going to the various tiers of government has varied over time. In the past two decades alone, the Federal government (FGN) share of national resources has varied between 41 – 56%. The states’ share has varied between 25 – 35% while the local governments have most times shared less than 20%.

These allocations are deemed to correspond to the level of constitutional responsibilities assigned to each tier. For example, FGN has responsibility to finance items on exclusive legislative list (defense, law and order, foreign affairs, internal affairs, inter-state or federal roads etc). It shares with states responsibility for items on the concurrent legislative list (education, healthcare, social infrastructure etc). The LGs have responsibility for primary education and Primary Healthcare (PHC) though in practice, the FGN and state governments have literally taken over these functions.

Each level of government complements its allocation from the federation account with its own constitutionally allowed supplementary sources of revenue or internally generated revenue (IGR). However, the efforts of the various tiers of governments in generating supplementary internal revenue vary. For example, it is clear that over the years most of the states and local governments have depended almost exclusive on their allocation from the federation allocation for their fiscal operations. For example, in 1994, IGR made up
41.38% of total revenue for Lagos state, 19.34 for Imo state, 0.36% for Yobe state and 3.6% for Oyo state (Calculated from CBN, 1995). The FGN fiscal operations, therefore, still dominates the operations of the other tiers of government put together. For example, in 2002 FGN expenditure as percentage of total national expenditure was 53.17%, States - 37.83% and Local Government Areas (LGAs) - 9%.

It is from these sources of revenue that each of the three tiers of government finances its own healthcare responsibilities. As an item on the concurrent legislative list, FGN has the main responsibility for formulating national policies on healthcare as well as for financing tertiary healthcare and institutions. State governments have responsibility for financing secondary and tertiary healthcare while the LGAs have responsibility for PHC. Yet given the fact that the entire economy is oil-dependent, and the instability of the international oil market, healthcare financing in Nigeria has not only been inadequate but also unstable.

2.7 The Structure of the Healthcare System

The organizational structure of the healthcare system has an important bearing on the level of efficiency in health service delivery and the distributional effects. Fig 2.2 shows the organizational structure of the Nigerian healthcare system. There are three levels of health administration corresponding to the three tiers of government in Nigeria, but with complex relationship among the tiers (WHO 2002, Ransome-Kuti 1999). The Federal Ministry of Health (FMOH) is accountable to the president, State Ministry of Health (SMOH) which is accountable to the state governor and Local Government Health Department (LGHD) which is accountable to the local government chairman. The FMOH is the ministry in charge of national health services. It is the apex health policy-making body for the country. In this capacity, it lays down guidelines and principles for health services in the country as a whole. FMOH provides technical assistance to SMOHs and LGHDS. It also oversees those tertiary health institutions – teaching hospitals of federal universities and colleges of health technology – that belong to the federal government. However, the teaching hospitals have separate management boards. The FMOH also
coordinates the activities of the 36 SMOHs and 774 LGHD but it has no direct control over them and they are functionally autonomous. For example, the FMOH cannot compel SMOH to implement a program if the latter is unwilling to do so.

Fig 2.2 The Organizational Structure of the Nigerian Healthcare System
There are also extra-ministerial agencies in the health system. These include the national Agency for Food and Drug Administration and Control (NAFDAC), National Action for Prevention of Prevention and Control of AIDS (NAPCA), National Program on Immunization (NPI) National Primary Healthcare Development Agency (NPHCDA). These are agencies of the FMOH, set up to deal with specific health problems.

The states have their own general hospitals and comprehensive health centres that provide curative and preventive care to their citizens. In addition, some states have teaching hospitals, schools of medical technology, nursing training schools etc. These institutions are regulated and supervised by SMOHs. Decentralization efforts in the 1970s led to creation of health management boards (HMBs) in each state to oversee the management of state hospitals. There have also been attempts to create district health boards but these have made little impact.

This is perhaps, as Ogunbekun (1991) notes, due to the fact that these DHBs control little resource. The SMOHs also provide technical assistance to the LGHDs. The LGHDs have the responsibility for the provision of primary healthcare in their areas. The organizational structure largely conforms to the constitutional principles of autonomy of federating units. It enables each level of government to identify its health priorities and pursue them with minimal intervention from the other levels but also makes planning and implementation difficult. It often leads to duplications and waste. It also makes healthcare delivery expensive. For example, in a recent address to the Pharmaceutical Society of Nigeria, the WHO country representative in Nigeria, Dr M. Belhocine notes "Verticalisation of programs, which is common in the health system often, results in sub-optimal drug management. Resources are utilised to evolve separate systems of drug management while structures for drug management in the public system are rendered non-functional and weak” (p.3)

It is important to note that the relationship between the FMOH and SMOHs is characterized by lack of cohesion which makes the establishment of a unified national
healthcare system difficult. FMOH and SMOH are accountable to different authorities. The federating states are constitutionally semi-autonomous from the federal government and consequently often have objectives that run parallel those of the federal. This arrangement makes consultation and cooperation difficult. In order to align health plans of the states and those of the federal, the National Health Council comprising representatives of the FMOH and the SMOHs meet once a year. But the council is only advisory and cannot compel SMOHs to a particular health plan. There are also no regular relationships between FMOH and other related federal ministries such as Education, Agriculture and Water Resources.

2.8 The Structure of the Nigerian Healthcare Market

The healthcare market is comprised essentially of two sectors: the public health sector and the private health sector. Both sectors supply healthcare to the consumers but it is the private sector that dominates the market (Ogunbekun et al. 1999; Alubo 2001). The public sector is more involved in regulatory activities, provision of primary healthcare for which the private sector has no incentive to provide, and the management of tertiary health institutions. At state level, the public sector also provides curative healthcare services such as the general hospitals and clinics. In this respect, the emphasis is often on the rural areas where the private sector does not also find sufficient profit reward to supply.

Four revenue bases or methods of financing healthcare have generally been identified: taxation (direct and indirect), social and private health insurance schemes, direct payments, and prepayment schemes. However, since the healthcare needs are generally high and no single one of these sources can adequately provide all the funds required, healthcare is typically financed from a mixture of sources. Though, one of these sources may be dominant in one particular system or over a particular period. The following are the main sources funds for the supply of public sector healthcare services:
2.8.1 General Taxation

Public healthcare financing in Nigeria over the years has relied mainly on direct and indirect general taxation. Direct taxation includes income tax, petroleum export tax, import tax etc. In actual fact, over the past three decades, revenue from petroleum extraction tax constitutes between 70% and 80% of total revenue of the government (CBN, 1995, 2003).

The sharing formula for the federation account revenue has been described above. But out of the total revenue allocation each tier of government receives and from the federation account and its own IGR it makes provision for its constitutional responsibility in the health sector. Since healthcare is financed from general taxation, it implies that the size of health sector expenditure in any particular year depends largely on the total amount of resources available to government that year. But in general, over the decades the Nigerian health sector has been grossly under funded as discussed earlier in section 2.5. A further break down of the application of budgeted provision for healthcare shows that on average 72% of the total health sector funds is devoted to recurrent component. Furthermore, grants and subventions to the specialist hospitals between 1989 and 1993 ranged from 47 to 73% of the total health sector annual budget. This indicates that the preponderance of attention to specialist tertiary healthcare at the expense of primary and secondary healthcare (Nwosu 2000). It is further observed that the main components of the yearly expenditures of the specialist hospitals are the overhead costs and salaries and wages which make up between 60 to 70% of total expenditure. Thus indicating that very little is left for the complementary healthcare inputs like drugs, equipment, drugs, and research.

2.8.2 Social Health Insurance Schemes

Many federal and state government establishments have clinics or some form of arrangement for healthcare of their staff. The employees of these establishments are offered free medical treatment, though the recent monetization policy of the government implies that these employees are now to get the equivalent of such medical provisions in
cash. Likewise some firms have retainer agreements with private hospitals for the healthcare of their workers and dependants.

The National Health Insurance Scheme (NHIS) was conceived by the Federal Ministry of Health and Social Services (FMOH) in 1992 as part of an effort to reform healthcare financing in the country. According to Decree 35 of 1999 setting up the scheme, the main objectives of the scheme among others include to:

(a) ensure that every Nigerian has access to good healthcare services
(b) protect families from the financial hardship of huge medical bills
(c) limit the rise in the cost of healthcare services
(d) ensure equitable distribution of healthcare costs among different income groups

In terms of structure, the decree also provides, apart from the central management, zonal branches determined by the central management for the purpose of administering the scheme. Under the scheme, those covered, with their dependants were to register with only one provider and thereafter would be entitled to comprehensive healthcare benefits which include preventive, inpatient, ambulatory care as well as family planning, diagnostic services, drugs and limited dental and optical services. The reimbursement system was to be by capitation. That is, the number of beneficiaries registered with a particular provider at the end of the given period. (Laws of the Federation of Nigeria Decree 35).

While the decree enunciates noble objectives, the mode of its implementation has been contentious as some medical professional associations have opposed and impeded its implementation in order to protect the interests of its members. As of July 25, 2005 the Association of Community Pharmacists of Nigeria (ACPN) was still threatening to stall the scheme’s take-off if the mode of payment which empowers primary providers to pay secondary providers was not reviewed (THISDAY, 2005). Thus while the launching of the scheme took place in March 2001, the scheme is yet to take-off due to various interest groups.
2.8.3 User Fees

Nigerians have never witnessed free medical services except for a brief spell of free medical care in the old western region of the country and a few special medical services. However, prior to the introduction of SAP in 1986, there were varying degrees of government subsidy of healthcare costs. The introduction SAP was followed by withdrawal of subsidies from social services including healthcare. The policy of cost recovery was introduced in all level of healthcare service. These subsidies were not restored even when government officially announced the abandonment of the SAP. Households pay for every social service they enjoy. Although there is no written policy of user charges in public health facilities, user charges cover such items as registration cards, drugs, laboratory investigations and in-patient care (Ekwempu et al. 1990). Such subsidy withdrawals from government funded health facilities implies higher healthcare financing burden on the households. Between 1985 and 1988, user charges in government owned health facilities increased by more than 200% at a time when people were experiencing greater economic difficulties. Some categories of patients that used to enjoy free treatment now had to pay user charges. This had a deterrent effect on the use of health facilities and consequent worsening health conditions as shown in Table 2.1. The funds generated from these sources are used to maintain and supply services through public healthcare institutions and other healthcare delivery channels.

2.9 Private Providers

The private healthcare market is crucial in the Nigerian healthcare arrangement. It is by far the dominant market (Ogunbekun et al. 1999). This section briefly characterizes the institutions and the operations of this sector of the market.

The supply side of the healthcare market is dominated by healthcare investments of the private sector of the economy. No data is available to show the level of private sector investment in healthcare but in view of the lean resources of government and rapid decline in government investment in health in the 1980s, the private sector is believed to
have grown rapidly (Ogunbekun et al. 1999). A number of factors helped to facilitate the rapid expansion of the private sector healthcare market in Nigeria. Apart from the oil burst of the early 1980s and the subsequent downsizing of the public sector, using the slogan ‘Be your own boss’ the government encouraged retrenched workers to become self-employed. For the health sector in particular, the mandatory five-year’s experience prerequisite for private practice for medical doctors was removed. Furthermore, the introduction of user fees added to the long queues in public health facilities contribute much to the rapid expansion of the private healthcare market (Alubo 2001). The private sector investments come from the private drug manufacturing firms, the large number of private hospitals and clinics, large number of pharmaceutical firms, and small drug retailers that dot every nook and cranny of the country (Brieger et al. 2004) as well as many traditional healthcare practitioners. It also includes investments by healthcare outfits supported by missions and NGOs.

A conservative estimate of the ratio of investments from these private operators to government investment must be about the ratio of 4:1. For example in 1996 Akwa Ibom State had only 20 public health facilities as against 149 of private facilities (Akwa Ibom State Department of Statistics, 1999). Apart from the missions and some NGOs investments in health, most of these private investors are driven by the profit motive. Most of them are solo-runs. Returns on investment are crucial for further investments. But it also leads to a lot of inefficiency in the sector. Facilities and human resources are often duplicated and are under-utilized particularly in urban areas where there is crowding of health facilities. The cost of these allocative inefficiencies is passed on to the consumer in the form high healthcare bills. At the same time the rural areas are underserved resulting in difficulties in access to care. For example, a survey showed that by 1987, 80% of private health facilities in Ogun state were located in urban areas. In Lagos and Oyo states, 75% and 83% respectively of private health facilities were located in urban areas (Ogunbekun et al. 1999). This expansion in private supply of healthcare, driven by the profit motive was interpreted as an evidence of willingness by consumers to pay.
Church organizations also contribute significantly to the supply of healthcare services particularly in the southern and middle belt regions of the country. Indeed, missionaries were among the pioneers of western medicine in Nigeria and continue to play important roles in the healthcare market. Unlike the other segments of the private health market, they tend to balance between urban and rural areas (Alubo 2001). Also, unlike the for-profit health facilities that are regarded by the government as competitors in the healthcare market, the church-based health facilities are regarded more as partners than competitors. It is, therefore, not uncommon to find government primary healthcare services, like immunization, being administered through church owned facilities.

Large and small drug dealers, patent medicine vendors or drug stores import and distribute drugs in the open market. Until recently when a few drugs were restricted by official control, it was possible to buy any drug or medical service over the counter as in any free market. This policy has given room for the importation and distribution of adulterated and expired drugs. In the last couple of years, however, the NAFDAC has been working hard to stem the excesses of these drug dealers. Every drug is now required to receive official approval before distribution. New regulations have not outlawed the existence of small drug dealers. They, along with the traditional healthcare practitioners still play very important role in the sector. In a recent study it was found that about 80% of consumers take as the first point of contact when illness occurs. It is only after the illness persists that they go for proper medical attention (Ichoku and Leibbrandt, 2003)

2.9.1 Private Healthcare Insurance and Prepayment schemes

Although arrangements exist for private insurance schemes, they cover a very limited number. The market for private health insurance is very small and just developing. There are four private firms marketing health insurance plans in Nigeria. Three of them are based in the cosmopolitan city of Lagos and one is based in Enugu. The number of people covered by each of these firms is relatively small compared with total population of the country. Ogunbekun (undated paper) estimates that none of these firms covers more than 18000 clients (including the beneficiaries). Altogether they cover less than
0.03 percent of the population. Those covered are mainly the rich who can afford the high insurance premiums.

Several factors account for the slow growth of the private insurance healthcare market in Nigeria. Some of these factors noted by Ogunbekun include lack of public confidence in the insurance industry, poor knowledge of the nation’s private healthcare market, inadequate technical capacity for underwriting health risks, absence of reinsurance backup and prevalence of infectious diseases, acute illnesses and injuries in the population. Given the poor population health, the probability of reporting ill and making claims on the insurance firm is high. This also implies that an actuarially fair premium must be relative high if an insurance firm is to breakeven and remain in business. But high premiums also mean that insurance would be beyond the reach of most people in the population given the low level of income. This, perhaps, explains why health insurance has not yet taken root in the country.

2.9.2 Prepayment Schemes

There are very few prepayment arrangements for financing healthcare in Nigeria, and even the activities of those existing are underreported (Ogunbekun, 1996). This makes the objective assessment of their potentials to contribute to the national healthcare system rather difficult. One such prepayment scheme is being managed by Country Women Association of Nigeria (COWAN). This organization is basically a rural cooperative society for women. Part of the activities of COWAN includes its Rural Integrated Health and Family Planning Program that implements a health credit fund – Health Development Fund (HDF) and a network of community based distributors of family planning products. Registered groups with membership of five to ten persons contribute fixed monthly amounts to the scheme. These monthly premiums entitle members to both agricultural credits and cost of catastrophic healthcare. As Ogunbekun, 1996 has rightly observed, the linkage of prepayment for healthcare with a credit scheme is a very interesting aspect of COWAN model of prepayment scheme particularly for those in the rural and informal sectors of the economy.
Apart from this, there are very few, less developed, unreported community healthcare prepayment schemes and other health risk sharing arrangements in the country. In other words a very insignificant proportion of the population is covered by such schemes.

2.9.3 Donor Funding

On account of the political instability in the country throughout the period of 1990s, Nigeria received very little aid from external donor agencies. This was particularly so during the days of General Abacha’s military dictatorship. Due to the anti-democratic posture of that regime and particularly the hanging of the human rights activist Ken Saro-Wiwa and his group in 1996, Nigeria became effectively a pariah state, deserted by most multilateral and bilateral donor agencies. It was only after the restoration of democracy in 1999 that some of these donor agencies started returning to the country. Even then, the level of donor funding especially in the healthcare sector still remains minimal when compared to other low and middle income countries. For example, Siddiq (2005) reports that Nigeria received less than two dollars ($2) per capita in aid between 2002 and 2003. This may be compared with about nine dollars per capita received by Bangladesh and Philippines, both of which are about the same population size as Nigeria.

Table 2.2: Relative shares of various health care financing mechanisms in Nigeria 1998-2002

<table>
<thead>
<tr>
<th></th>
<th>Total Health expenditure % GDP</th>
<th>Government expenditure as a % of total health expenditure</th>
<th>Private sector expenditure on health as % of total health expenditure</th>
<th>Households' OOP* as a % of private sector health expenditure</th>
<th>Prepaid and risk-pooling plans as % of Private sector expenditure on health</th>
<th>General government expenditure on health per capita at exchange rate</th>
<th>General government health expenditure as % of total government expenditure</th>
<th>Donors as % of Total expenditure on health</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>5.5</td>
<td>26.1</td>
<td>73.9</td>
<td>95.0</td>
<td>2.4</td>
<td>4</td>
<td>7.1</td>
<td>13.1</td>
</tr>
<tr>
<td>1999</td>
<td>5.4</td>
<td>29.1</td>
<td>70.9</td>
<td>94.8</td>
<td>3.4</td>
<td>5</td>
<td>5.4</td>
<td>13.8</td>
</tr>
<tr>
<td>2000</td>
<td>4.4</td>
<td>33.5</td>
<td>66.5</td>
<td>92.7</td>
<td>5.1</td>
<td>6</td>
<td>5.9</td>
<td>16.2</td>
</tr>
<tr>
<td>2001</td>
<td>4.5</td>
<td>31.4</td>
<td>68.6</td>
<td>91.4</td>
<td>6.5</td>
<td>6</td>
<td>3.4</td>
<td>5.6</td>
</tr>
<tr>
<td>2002</td>
<td>4.7</td>
<td>25.6</td>
<td>74.4</td>
<td>90.4</td>
<td>6.7</td>
<td>5</td>
<td>3.3</td>
<td>6.1</td>
</tr>
</tbody>
</table>

*Out-of-pocket payment.

Table 2.2 shows the relative contribution of each source of healthcare financing in Nigeria over the period 1998 - 2002. The table clearly indicates that the predominant source of healthcare financing is the private source, and the dominant payment system is the out-of-pocket payment. It also shows the relatively insignificant contribution to health financing in Nigeria by donor agencies.

2.9.4 Coordination

The Nigerian health system has been described as fragmented and uncoordinated. There is apparent disconnection between the Federal, State, and Private sector healthcare efforts. This is primarily because the FMoH and SMoH are responsible to different political authorities. There is also no effective framework for public-private partnership. Information does not always flow between the various ministries and departments. Cross referrals are not always smooth. This makes comprehensive planning difficult.

The picture that emerges from this characterization of the Nigerian socioeconomic environment is the failure of existing policy framework to guarantee improvement in the social welfare of the majority of the people. All these short-comings have coalesced to a system of healthcare that is unable to meet the basic health needs of the population especially the poor. In the midst of abundant resources poverty is spreading and deepening. Health indicators are worsening. Skewed resource distribution has been a major source of political and economic instability.

2.10 Government’s New Health Sector Reform Policy

In July 2003, Professor Eyitayo Lambo, formally a regional director of the WHO, was appointed as the federal minister of health. After a thorough review of the deplorable state of the Nigerian health sector he proposed a program of reform for the sector (Federal Ministry of Health, 2003) The new reform initiative is committed to undertaking a government-led comprehensive health sector reform aimed restructuring and strengthening the national public and private system in order to deliver qualitative health to the people effectively and efficiently. In respect of equity in access to healthcare, the
reform document further states: "The health interventions should be provided in such a way that will not only enhance the financial and cultural access of the people, particularly the poor, but also their physical access which in some parts of many African countries is currently a big problem" (Federal Ministry of Health 2003: 5).

These objectives and principles indicate clearly that efficiency and equity are central policy concern to the new Nigerian health sector reform being initiated by government. There is need not only for effective and efficient application of scarce resources but also for the financing system to be pro-poor and equitable if the healthcare system is to serve as a major component of the poverty reduction effort of the government.

2.11 Summary

This chapter has attempted to place the study in social and policy contexts. The purpose is to provide the reader with basic background information that underpins policy concern and the concern of this study with the redistributive consequences of health financing in Nigeria. The socioeconomic environment has been declining over the past two and half decades. Macroeconomic indicators have been falling too. The healthcare system has similarly witnessed a systematic decay leading to very unfavourable health outcomes. Health indicators show that the health of the Nigerian population is poor. The healthcare market leans heavily towards the competitive ideals. In the next chapter attention will be focused on the theoretical and conceptual underpinnings of the study. A critical review of empirical studies from international literature on health inequalities and redistributive effects of healthcare financing systems will also be undertaken.
Chapter 3

Literature Review

3.1 Introduction

The preceding chapter has tried to situate this study in the social and economic policy context of Nigeria. It tried to highlight the prevailing social and economic factors that influence health and the healthcare financing mechanisms in the country. In this chapter, we establish the economic theoretical and empirical grounds for redistributinal policies in healthcare financing. The chapter comprises three main sections. Section A attempts to present a theoretical justification for redistributive policies based the fundamental assumptions of welfare economics. Section B critically examines the empirical literature on redistributive effects of healthcare financing and shows the contributions this study intends to make to existing stock of knowledge in this area of research. Section C summarizes the existing literature in health inequality to provide a framework for subsequent analysis of socioeconomic inequalities in health in chapter 6.
SECTION A

Theoretical Framework

3.2 Welfare Economic Theory and Redistribution

3.2.1 Introduction

The twin concerns of healthcare financing are efficiency and equity. In general, efficiency is a concern for optimization - the maximization of output for a given level of input or the minimization of costs for a given level of output. Equity on the other hand is a concern for distribution, a fair share of the pie for everyone. Unfortunately, from the viewpoint of neoclassical economics, there appears to be considerable scope for conflict between the two objectives, a conflict recognized as the trade-off between growth and distribution. What this implies in practice is that a government policy (such as in healthcare financing) that is directed to achieving equity or distribution is unlikely to achieve the goal of an efficient healthcare system which would require that individuals consume healthcare only according to their ability to pay for the services they receive.

Under the neoclassical framework, inequality in distribution is seen as a necessary consequence of economic growth as suggested by Kuznet's U curve (Kuznet, 1955) which implies that distribution gets worse-off before it gets better-off in the process of economic development. In fact Lewis' theory of economic development (Lewis, 1954) explained inequalities in income distribution as facilitating economic growth. This position is also sometimes supported from a Keynesian perspective. It is argued that since individual's savings rate increases with increases in income, therefore, redistribution of income from the rich to the poor will lower aggregate savings and so lowers investment.

---

1 Although distribution and redistribution will often be used interchangeably in this report, it is important to keep in mind that distribution refers to allocations generated by the market process, redistribution, on the other hand, refers to policies that aim at redistributing the distribution generated by the competitive market.
and economic growth. Therefore, a rise in inequality will increase investment and economic growth (Barrow, 2001).

Furthermore, there has been considerable debate whether distributional concerns in general and in health in particular lies within the legitimate province of the interests of the economist since the ‘economic man’ is essentially a utility maximizer. Whatever variable that does not enter into his utility maximizing function, such as the welfare of another person, is considered as an externality. These debates seem to cast doubt on the relevance of redistributive policies as an economic concern (see for example Harberger 2001).

Several counter arguments to justify distributional concerns in health have been based largely on sympathy (see for example, McGuire et al. 1988, Chapter 4). Equity in healthcare financing or redistribution through healthcare financing is seen as a case of specific egalitarianism. As McGuire et al. (1988: 55) state:

“In essence concern with equity involves the wish to abandon a strict willingness to pay criterion since such willingness must inevitably be based on ability to pay, which in turn reflects the existing distribution of income and wealth. If ability to pay is for some individuals or groups judged to be in some way or other ‘impaired’ then for various reasons this may mean that society generally, or those unimpaired in this respect, may wish to ensure that the distribution of certain relevant commodities (in this case health) is not based solely on willingness to pay....., this type of specific egalitarianism may well include healthcare”

While these counter-arguments are generally valid, it has, however, been shown that beyond the appeal to sympathy and commitment, distributional concerns or equity in healthcare financing also enters as an independent argument in the social maximization function. This view is suggested by new attempts to clarify the conditions for the operation of Adam Smith’s ‘invisible hand’ as expounded by Arrow-Debreu equilibrium. The core of this new economic theory rests on the existence of market imperfections, information asymmetries and social costs of inequality that suggest that the relationship
between efficiency and equity is not, after all one of substitute but complement (see for example, Hoff 1994; Hoff and Stiglitz, 2001).

To substantiate this argument, it is noted that under the assumptions of the Arrow-Debreu conditions\(^2\) the competitive market is generally considered to lead to Pareto efficiency. However, real life is characterized by the existence of monopolies, non-convexities, oligopoly, imperfect information, and other forms of market imperfections such that equilibrium conditions are not guaranteed. The assumptions of the fundamental theorems seem, therefore, too restrictive that by the very nature of things, they are bound to be violated, at least in some sectors of the economy. The implication is that there is need for Pareto improving market interventions. This new focus arises from the deeper analysis of the implications of the fundamental theorems of welfare economics and the conditions under which the Arrow-Debreu equilibrium may be attained\(^3\).

Once the Arrow-Debreu equilibrium conditions are relaxed, the separation between efficiency and equity ceases to exist. The issue of equity or distribution becomes inseparably tied to the issue of efficiency. For example, Vohra (1998) shows that under increasing returns the issues of income distribution and equity cannot be divorced from efficiency concerns. Hoff (1994) shows that with imperfect information and under certain conditions, lump-sum redistributions of endowments can, indeed, improve efficiency. According to Hoff, the ‘Second Theorem of the Second best’ implies “a breakdown of

---

\(^{2}\) Arrow and Debreu (1954) highlight the very restrictive assumptions underlying the working of Smith’s ‘invisible hand’ as formally defined under the first fundamental theorem of welfare economics: no increasing returns, no monopolies, complete set of markets for spot and futures markets, convex technologies complete insurance, symmetric information, the existence of lump-sum transfer instruments among others. It is characterized by tangency between marginal rates of substitution in consumption and marginal rates of technical transformation in production with price as parametric. It is basically the Walrasian equilibrium for state contingent commodities (Mas-Colell et al. 1995).

\(^{3}\) Much of the inspiration in current literature in the understanding of the relationship between efficiency and distribution is due to the clarifications and extensions of the Arrow-Debreu model highlighting the strident conditions for general equilibrium to be attained. However, driving the Arrow-Debreu equilibrium conditions to their logical limits, Joseph Stiglitz and George Akerlof and others showed that relaxing any of these assumptions such as the assumption of full and symmetric information the competitive equilibrium was not even constrained Pareto efficient. That is to say, in the presence of information asymmetric, an allocation cannot be Pareto improved (Hoff 1994)

---
the neat dichotomy between distribution and efficiency in the neoclassical model" (p.235). In the words of Banerjee (2001:464) "The second welfare theorem notwithstanding, few now believe that the questions of efficiency and growth can be resolved independent of issues about distribution"

This is the unfortunate outcome (for advocates of the inviolability of the market) of the theory of the second best that specifies that "if there are a number of constraints preventing the satisfaction of optimal conditions it will not generally be desirable to have these conditions hold in the rest of the economy" (Killick, 1981:18). But this dichotomy may also breakdown for other reasons such as incomplete markets and externalities. Thus, it would seem that the efficiency-oriented and equity-improving healthcare financing need not after all be contradictory but actually mutually reinforcing

3.2.2 Empirical Evidence

To show that a healthcare financing policy that is directed at improving efficiency and one that is directed towards redistribution are not contradictory, a number of recent empirical studies both at the micro and macroeconomic levels have shown that in practice, redistribution can increase efficiency in a variety of ways (Banerjee, and Newman, 1993; Galor and Zeira, 1993; Curie, 1995). For example, Currie finds that reducing malnutrition that inhibits the development of human capital may increase the social product. If credit constraints prevent the poor from having access to productive investment, inequalities in wealth distribution may have negative impacts on economic growth (Bruno, Ravallion and Squire, 1996). In this sense, distribution and efficiency may be considered as complements rather than substitutes. Once the restrictive assumptions underlying the Arrow-Debreu model that give rise to the separability of efficiency and distribution breakdown, then the outcome will be quite model specific (Kanbur and Lustig, 2000).
3.3 Nature of the Healthcare Market and Healthcare Financing

While the foregoing section suggests that the goals of equity and efficiency are not mutually exclusive as they seem initially when things are taken at their face value, in the subsection it is also argued that because of the very nature of the commodity itself – health, there is also market failure such that achieving the objective of efficiency may not also be guaranteed. The objective here is to highlight those features of the healthcare market that impose special difficulties in realizing the ideals of market efficiency which would have obtained if the healthcare market were to operate like any other commodity market.

3.3.1 Characteristics of the Healthcare Market

The healthcare market is characterized by risks and uncertainties, economies of scale or externalities, and information asymmetry. These features of the healthcare market make it susceptible, more than many others, to failure. While many other markets may violate one or two of these assumptions of the welfare theorem, the healthcare market violates almost every aspect of these assumptions and thus makes the extensive intervention of other non-market institution necessary.

In the healthcare industry, there are ethical restrictions on the activities of the physician. These activities are characterized by compassion for the well being of the patient. The physician is assumed to play his part in the treatment process based on the considerations for the actual needs of the patient and not on any financial considerations. The production of health depends on the activities of both the physician and the patient and the relationship between them. Fuchs (1983) emphasizes this vital relationship and contrasts it with the type of impersonal relationship in the competitive market:

Health is the outcome of a process that involves patients and health professionals working together; mutual trust and confidence contribute greatly to the effectiveness of that process. (p.viii)
This collective responsibility for health outcome is an important characteristic of the healthcare market that sets it apart from other types of commodity markets. Demand and supply of healthcare are not independent as is assumed under the competitive market model. It also eliminates advertising and overt price competition that characterizes the ordinary commodity market (Arrow, 1963). The interdependence in the medical market is not limited to the physician and the patient but also manifested in the general concern for other people’s health than is generally shown to their welfare. In other words there are interdependences in utility functions implying that one person’s health condition enters into other peoples’ utility functions, which contradicts the assumption of independent utility functions.

One of the most obvious distinguishing characteristics of the healthcare market is that its demand is rather stochastic. Illness is unpredictable and the consumption of healthcare cannot be planned. Not only is the probability of becoming ill unpredictable, there is also the associated uncertainty about how long it is going to last. The probability of full recovery or permanent deformity or even the ultimate demise from an illness cannot be estimated with any certainty. This risk is apart from the pecuniary and non-pecuniary costs of being ill. This uncertainty affects the production of healthcare since it is a service that cannot be stored. Its production will also vary according to diagnostic results that are heterogeneous and customized (McGuire et al. 1988). This customization implies some uniqueness of every case in the healthcare industry and very low cross elasticities because of lack of substitutability between treatment for one case and another. In other words the assumption of many buyers and sellers does not strictly hold, as every transaction is unique.

A presupposition of the market system is the presence of many buyers and sellers. Such an environment creates the required competitive pressures that in turn generate the market-clearing price between demand and supply (Atkinson and Stiglitz, 1980). On the other hand, where there are limited number of sellers and buyers, as is the case with the medical market, this competitive pressure may be absent, giving rise to considerable transaction costs in terms of resources spent in arriving at equilibrium price. Such extra
resource spending is a measure of market inefficiency. If in addition to the context of limited market there is also information asymmetry then there is potential for opportunistic behaviour (McGuire et al. 1988:188)

Finally, the vulnerability of the healthcare consumer may often place her in a weak position to negotiate a market-clearing price even if she has the required information. Specifically, the atmosphere of the healthcare consumption is an integral part of the treatment that leads to the desired outcome. Such considerations may lead to the separation of the transaction process from the purely therapeutic process because the transaction process may affect the utility of the consumer.

Asymmetric information represents a significant evidence of the violation of the basic assumptions of the welfare theorems in the healthcare market. In this market the consumer relies almost exclusively on the information gained from the diagnosis performed by the healthcare provider concerning the nature of the illness, the appropriate medical intervention and the expected outcome of such intervention.

Beyond these, it is observed that the medical market is characterized by externalities and undefined property rights. There are substantial spill overs in the consumption of healthcare. For example, many communicable diseases like tuberculosis constitute dis-utilities that can spill over into other peoples’ utility functions even though such dis-utilities fall outside the latter’s preferred set of goods and services. The individual’s demand for a cure, though with a view to maximizing personal utility, also maximizes others’ utilities. There are many public health interventions that often assume the status of public good because the marginal cost of such interventions is zero. Such, for example, is the case with wide area disease control and radio and television healthcare information. Obviously, these are evidences of market failure in the healthcare market requiring non-market interventions.
In view of these shortcomings of the healthcare market it would seem that all the factors that lead to market failure are present to some degree in healthcare. Given the violations of the conditions for optimal efficiency of the market in this sector it would seem that market failure is gross enough to warrant the intervention of non-market agents.

The presence of market failure in the healthcare sector implies that there may be substantial reductions in welfare if the society leaves the market to allocate resources in this sector. The presence of such allocative inefficiencies then requires the intervention of other social institutions and policies that attempt to provide some remedy for market failure in the healthcare sector.

These considerations on the peculiar nature of the healthcare commodity and the healthcare market lead to very important conclusions. They suggest that Pareto allocative efficiency should not be the sole criterion for financing and distribution of healthcare. Society does not generally require that healthcare be financed according affordability or ability to pay (ATP). Rather, the fortuitous nature of illness, non-market characteristics of healthcare, and the general disposition of the society to hedge people, especially the poor, against unplanned healthcare expenditure require that access to healthcare should depend on need and this entails equity considerations in the distribution of health and healthcare financing.

One of the major goals of a health system is, therefore, to respond to the challenges posed by the failure of the market system to function effectively in the healthcare market, and the peculiar social expectations of fairness in healthcare. A healthcare financing system should not only be sensitive to the needs of its client but also ensure that everyone is protected against the costs of ill-health (Preker, 2005).

3.4 The Concern for Equity: Equity and Ideology

Since the competitive market does not secure the social objective of redistribution, for the social decision maker to secure that objective it is important to establish the criterion
upon which redistribution could be carried out. As was noted earlier distribution and redistribution are issues that involve inter-personal comparisons. Unfortunately economics does not seem particularly equipped to analyze interpersonal comparisons of welfare because it often involves incomplete ordering\(^4\), and so cannot provide the criterion for redistribution. Such an analysis must borrow the tools of other disciplines to provide the theoretical basis for redistribution. The essential theoretical framework here is provided by political philosophical theories of economic justice.

3.4.1 Theories of Justice

Traditional theories of justice could be characterized as forming three major overlapping ideological tendencies. The left is dominated by socialist tendency that promote equality of individuals with economic reward system based on need. The right is dominated by capitalist structures that promote the freedom of the individual and market capitalism with economic reward system based on contribution to the production process. The former are the socialists and egalitarians and the latter are the libertarians. In between the extremes are the liberals who believe in a mixture of equality and freedom and promote welfare state capitalism (Kymlicka, 1990). This section examines how a society or a policy-makers inclinations towards these ideological groups would affect the organization of the health sector and the method of financing healthcare as well as the extent the society would want to pursue redistributive policies.

Utilitarianism

Utilitarianism basically holds that a policy change is warranted if it brings about an increase in the sum total of utilities (welfare) in the society. In other words, the right action is that which maximizes total welfare. Equal amounts of utility matter equally. More precisely, the utilitarian SWF may be expressed as: \( W(u) = \sum u_i \), where utilities are

\(^4\) A complete ordering is one where a logical sequence is maintained. For example, if A is preferred to B and B is preferred to C, then, A must be preferred to C. If A is preferred to B and B is preferred to C but C is preferred to A, then, it is an incomplete ordering. Interpersonal comparisons are often characterized by the latter form of relationship.
symmetric. On the other hand, if utilities are nonsymmetric, then \( W(u) = \sum \alpha_i u_i \).
(Where \( \alpha \) = non-negative weight). In this case an increase (decrease) in the utility of the individual will translate into an increase (decrease) in social welfare. The objective function of the utilitarian is to maximize the social welfare function. It follows then from the above utility functions that utilitarianism is indifferent to inequality in distribution. In that case a redistribution of wealth is justified on the basis of declining marginal utility of wealth to the rich. Unfortunately, as Kymlicka (1990:47) points out modern utilitarianism 'does not immediately identify any set of policies as distinctively superior'. But in the healthcare sector, the utilitarian relying on the principle of marginal utility in allocation of resources or maximizing total utility (an efficiency principle) would probably opt for allocation of medical care in accordance with the probability of getting cured.

**Libertarianism**

The libertarian theory of justice is best exemplified in Nozick's classical work: *State, Anarchy and Utopia*, (Nozick, 1974) where he espouses his entitlement theory. Basically, libertarians defend all sorts of negative freedom – individual freedom, market freedom, and unrestricted capitalism. They emphasize the limits of the state as an apparatus for social policy. They strongly hold that what an individual is, results from his choices and they therefore leave no room for the rectification of unequal circumstances. Indeed for libertarians the outcome of a free market is just. Redistributive policies are, therefore, unjust because they constitute a violation of individual rights. Nozick draws his inspiration on the individual right to property from Kant's practical imperative: no individual should be treated as a means to an end, and Locke's theory of labour as the primary title to property. In 'Second treatise of Civil Government and on Letter on Toleration' Locke insists that labour is the key to entitlement: "Though the water running in the fountain be everyone's, yet who can doubt but that in the pitcher is his only who drew it out? His labour hath taken it out of the hands of nature, where it was common and belonged to equally to all her children, and hath thereby appropriated it to himself" (qtd. Copleston 1964:139-140). For Kant human beings are ends in themselves and no one should be treated as a means to an end, and his own concept of self-ownership.
Accordingly Nozick (1974:ix) argues: 'Individuals have rights, and there are things no person or group may do to them (without violating their rights). So strong and far-reaching are these rights that they raise the question of what, if anything, the state and its officials may do (Kymlicka, 1990). People have right to what they own and the state has no right to interfere in deciding how they dispose of what they own. Any such interference is equivalent to slave labour. By inference a libertarian healthcare system is one more motivated by efficiency considerations than redistribution. It would be a healthcare system that operates on the principles of competitive market system where healthcare utilization is closely linked to ability to pay.

The Liberals

The liberal theory of justice attempts to balance between equality and freedom. Unlike the libertarians for whom justice is only ambition-sensitive and who thus pay little regard to inequalities arising from the very nature of things, liberals believe that justice must be both ambition-sensitive and need-sensitive. The liberal tradition is perhaps best exemplified by Rawls for whom justice is fairness. According to Rawls (1972:102)

"...no one deserves his greater natural capacity nor merits a more favourable starting place in society ...it does not follow that one should eliminate these distinctions. There is another way to deal with them. The basic structure can be arranged so that these contingencies work for the good of the least fortunate. Thus we are led to the difference principle if we wish to set up the social system so that no one gains or loses from his arbitrary place in the distribution of natural assets or his initial position in society without giving or receiving compensation in return."

The implication then is that those who are naturally talented deserve more shares of social goods only in so far as they benefit those who are disadvantaged by nature. In the Rawlsian sense, people are treated as equals not by eliminating all sources of inequalities but by removing those sources of inequalities that place some people at disadvantage position. Commenting on Rawls' position Kymlicka, (1990:53) notes that, if some inequalities promote the common good, then no one need bother about removing such inequalities. If giving more money to someone who is richer than I am enhances my own
interest, then such inequality need not be removed. In other words, inequalities in society are only allowed if it improves the initial equal share of everyone.

**Radical Egalitarianism**

Like the liberals the radical egalitarians insist on fundamental equality of everyone. But they go beyond the liberals in pursuit of this goal in advocating a status-less society. A society where there are no extensive differences between persons based on income, power and authority and prestige. Nielson (1985) represents the views of the radical egalitarians. According to him:

"The goal we are seeking is an equality of basic conditions for everyone. Let me say a bit what this is: everyone as far as possible should have equal life prospects, short of genetic engineering and the like and the rooting out any form of the family and undermining of our basic freedoms. There should where this is possible be an equality of access to equal resources over each person's life as a whole, though this should be qualified by each person's varying needs.... This equal access to resources should be such that it stands as a barrier to there being the sort of differences between people that allow some to be in a position to control and exploit others; such equal access to resources should also stand as a barrier to one adult person having power over other adult persons that does not rest on the revocable consent on the part of the persons over whom he comes to have power" (p. 56)

Like the utilitarians the radical egalitarians also insist on the maximization of the over all societal welfare. But unlike the utilitarians who do not care whether this maximization is achieved through the rich getting richer and the poor getting poorer, the radical egalitarians believe that this maximization must be for everyone under the condition of equality.

In summary then, historical injustices demand under Nozick’s entitlement principles a redistribution of wealth from the rich and affluent individuals to the poor. The egalitarians of the liberal bent and the radical bent all justify redistribution on the basis of some equality. Since inequality arises from natural contingencies that no one has control over, the difference principle requires a redistribution that benefits those who have less.
In this particular case the poor and those who bear the greater burden of disease and ill-health need be favoured in a system of redistribution.

Furthermore, the common denominator of the Rawlsian, egalitarian and utilitarian social welfare functions would suggest that redistribution is warranted even by the differences in ability. Redistributive policies should be used to redistribute incomes so that those who are disadvantaged (e.g., in health) could meet their basic needs and would be able to achieve basic functionings. Therefore, socioeconomic inequality in health status demands a redistribution of income in favour of those disadvantaged by their health status. But the conflicting theoretical perspectives on social justice highlight the scope for conflict between efficiency and equity.

3.5 Concern for Equity in Healthcare Financing

The objective of healthcare financing is not only a concern for efficiency (ensuring that benefits of healthcare financing are maximized and costs are minimized) but also a concern for equity. Indeed, our earlier analysis of Pareto efficiency shows that the single-minded pursuit of efficiency may lead to suboptimal economic outcomes that tend to lower social welfare. However, the various theories of theories of justice remind us that there is little unanimity about what constitutes justice and, hence, equity in healthcare. Yet the policymakers are often as much concerned about equity as they are about efficiency. They may want to ensure that people contribute to healthcare financing in an equitable way.

Of the theories of justice described above, the more influential ones in the healthcare sector are the libertarian and egalitarian theories. Libertarian healthcare systems emphasize the value of consumer sovereignty while the egalitarian systems emphasize the equal rights of all citizens to healthcare services (Donaldson and Gerard, 1993). More specifically, libertarian systems are not bordered with the existence of unequal access to healthcare services which they believe should exist by the very fact that resources are not equally distributed, but would not, based on sympathy, oppose a policy that provides minimum healthcare for those who otherwise would not be able to pay for the services.
For libertarians a fair healthcare system is one in which people are free to purchase the type of care they want depending on their ability to pay. Inequality in health resulting from unequal access to healthcare services which in turn results from unequal distribution of ability to pay is part of the society's reward for merit. Healthcare financing system should not be used to redistribute income. In the context of the libertarian theory, the pursuit of equity is a very limited objective: only to ensure the availability of minimal healthcare for the poor.

Egalitarian healthcare systems, on the other hand, are concerned to ensure that healthcare payments are completely de-linked from healthcare utilization and linked to ATP. Healthcare utilization should be based on need while payment should be according ATP. This is in sharp contrast to the libertarian view that basically considers access to healthcare as an extension of the society's reward system.

3.5.1 Vertical and Horizontal Equity

Conventional analysis of equity often distinguishes two aspects of equity: vertical and horizontal equity. Put simply, vertical equity is the "unequal treatment of unequals" while horizontal equity is the "equal treatment of equals" (Mooney 1983:179). The later is concerned with the fair treatment of people who are alike in every respect. In the case of healthcare financing, horizontal equity requires that households or individuals who have equal ability to pay should pay equal amounts for equal healthcare needs. Vertical equity, on the other hand, requires that households with unequal ability to pay should pay unequal amounts for the equal needs. As Mooney (1983) notes, it is conceptually straightforward to understand the principles of vertical and horizontal equity. In practice, however, problems may arise when one attempts to give empirical contents to the definitions. In what, for instance, would equal treatment consist in? What exactly is meant by "equals"?

A number of suggestions have been made in order to give empirical contents to the terms "equal" and 'unequals' as used in reference to horizontal and vertical equities. Mooney
(1983) and Donaldson and Gerard (1993) for instance proposed the following empirical contents: equality of expenditure per capita, equality of utilization for equal need, equality of opportunity of access for equal need. Mooney adds other criteria that include: equality of inputs for equality of inputs of resources, equality of input for equal need, equality of marginal met need and equality of health. These empirical contents point to the desire for people with equal ability to be treated equally in healthcare financing irrespective of their attributive characteristics. There is a high tendency for inequity in healthcare financing if for instance chronically ill people are required to pay more than their non-sickly counterparts who have the same ATP.

Vertical equity has been similarly operationalized as: unequal treatment for unequal need (utilization) and progressive financing based on ATP. The further question is: how is ATP to be measured in practice? And how it is to be related with healthcare payment? The clue from the tax literature is that ATP could be measured through the concept of comprehensive income. By this is implied current consumption plus net capital gains (Acocela, 1998).

In the healthcare economic literature, ATP has played a crucial role in the controversy about the introduction of user fees. A number of demand studies (Heller 1982; Akin, 1981; Akin et al. 1986) found evidence that financial constraints were not sufficiently binding to deter the poor from the use of healthcare services. Thus, they provided justification for the introduction of the policy of user fees. Some later studies (Gertler and van der Gaag, 1990; Yoder, 989; Ellis and Mwabu, 1991) found contrary evidence to show that what seemed like widespread willingness to pay (WTP) among the poor was actually at great opportunity cost of other household necessities and, therefore, was not the same as ATP.

In the main, ATP has been defined from the short and long-run perspectives. The short run perspective defines ATP in terms of health expenditure-income ratio or in terms of the ratio health expenditure to non-food expenditure (Wagstaff and van Doorslaer 2001). Information from health expenditure surveys in many developing countries shows that
most households spend between 2 -5% of their income on healthcare (Russell 1996). Therefore, a 5% healthcare expenditure to income ratio became the benchmark for measuring ATP. Beyond the 5% threshold a healthcare expenditure was regarded as catastrophic. While the flaws in such an indicator as a measure of ATP has been pointed out (Russell 1996) its use has become widespread in the health economic literature.

The long run perspective of ATP is considered in terms of opportunity costs of healthcare payments. As Russell summarized it,

In the long-run, the costs of accessing healthcare can be considered affordable when consumption of and investment in essential items such as health and education does not fall to levels which threaten health and future earning capacity, and when health expenditures do not threaten productive assets, livelihood sustainability, capacity to pay for minimum needs in the long run and ultimately health (p.226).

Obviously, the difficulties of measuring the long run opportunity costs imply that empirical research has still to fall back on the short-run perspective. In this research, the ATP is related to household total expenditure because it provides a window into the opportunity cost to a household of financing healthcare and also because of the arbitrary nature of judging what may considered basic necessities and what may not.

The operational definition of ATP enables one to relate it to healthcare payments. The question here is what should be the actual functional relation between ATP and healthcare payment? That is, granted that healthcare financing should be according to ATP and granted also that ATP is proxied by income (or expenditure), should those with higher ATP pay more in absolute terms or should payments be made in proportion to income? Should healthcare payment be progressive or regressive? If progressive, what is the extent of progressivity that is demanded by equity?

The practical answers to these questions depend on the society’s ideas about equity but it is also important in practice to consider the form of healthcare financing under consideration. For example, vertical equity may more easily be achieved if healthcare is
financed through a general taxation that depends on income tax and less likely to be attained through a healthcare finance system that depends on private health insurance scheme.

This section has shown that theoretically, the most efficient method of financing healthcare is through the competitive market because in that case the market outcome is suppose to be Pareto efficient. Healthcare choices both on the part of suppliers and consumers will always tend to minimize cost and maximize output However, the section also shows that there are major factors that make efficient allocation and financing through the market system inapplicable in the healthcare market: the extensive market failure in the healthcare market and the ethical consideration for equity in healthcare. These also establish the ground for a policy of redistribution away from the distribution that results from the competitive market. Section B critically reviews the results from some empirical studies of redistributive effect of healthcare financing.
SECTION B

Empirical Literature

3.6 Introduction

Empirical studies in health inequalities could generally be classified into three main parts: equity in healthcare financing, equity in delivery/utilization of health resources, and equity in health status. Literature on equity in healthcare financing usually analyzes the equity implications of the different forms of healthcare financing including general taxation, private and social health insurance, user fees, and pre-payment schemes all of which were discussed in chapter 2. Literature on equity in delivery/utilization of health services considers equity in access to healthcare services. It considers constraints to equity in delivery and utilization which are raised by geographical distribution of the population, locations (urban and rural), distance to health facilities, provide behaviour that impede access as well as information problems that create impediments to healthcare utilization. Literature on equity in health status is usually concerned with inequalities in health itself whether as generated by biological factors or mediated by the interactions of socioeconomic and behavioural variables.

Although in most studies, these various aspects are analyzed separately, there may be links between them. For example, there are links between equity in financing and equity in access to health services. The method of healthcare financing is likely to have far reaching equity implication for access to healthcare. For instance, healthcare financing that depends mainly on general taxation is likely to be more equitable than one that depends on private health insurance or direct payments. The links between equity in financing and equity in health status are indirect with several possible confounding factors which may be facilitated through social institutions (Mackintosh, 2001). The health status of the poor may depend largely upon the extent the healthcare financing system imposes constraints on their claim to health services. Methods of healthcare financing and risk-sharing are likely to determine the amount of resources available (both
at aggregate and individual levels) to finance healthcare, and therefore, determine also health outcomes [see for example, Carrin et al. (2004)].

The empirical literature on the impact of healthcare financing on income redistribution is very recent. The first genre of empirical literature started in the mid 80s with simple methods of computing differences in healthcare payments made by different income groups in order to ascertain the relative healthcare financing burdens of the different income groups. The literature has now gone beyond these simple methods. It has borrowed extensively from the tax literature to estimate redistributive effects and to decompose differences in the prepayment and net payment Gini indices into their different components. The review follows the pattern of development in the literature. It starts with the early empirical literature and proceeds to the more recent stage.

**Early Empirical Literature**

While there is a considerable amount of theoretical literature on the redistributive effects of taxation, very few empirical researchers have attempted to investigate the income redistributive effects of healthcare financing. Many of those who did have used simple tabulation methods to analyze the redistributive effects of healthcare financing [see for example, Hurst (1985), Cantor (1988), Wolfe and Gattschale (1987), Gattschale et al. (1989), Baker and van der Gaag, 1993, Paninarunothai and Mills (1997)]

In comparing the progressivity of healthcare payment systems in the UK, Canada and US, Hurst (1985) used tables to indicate how average payments for healthcare varied according to income groups in each country. Surprisingly, these payments were analyzed in terms of there absolute values rather than in terms of proportion of the household income spent on healthcare services. This approach excludes the possibility of comparing the progressivity of healthcare payments across the countries included in the study. Cantor (1988) represents substantial improvements on the preceding study by Hurst. Using the US data and tabulation method, this study showed that proportions of income contributed to healthcare financing by various income categories declined as one moved
from the lowest to the highest income categories. These studies show that there are substantial income redistributions in the US health financing system. This redistribution is, unfortunately, pro-rich implying that the poor who bear the greater burden of ill-health are also likely to make greater contribution to the healthcare system.

Wolfe and Gottschalke (1987) and Gattschalke et al. (1989) like Hurst (1985) did cross country comparative analysis of income redistribution arising from the methods of financing healthcare services. Wolfe and Gottschalke studied the patterns of healthcare contributions among the elderly in the USA using the simple healthcare contribution to income ratio as basis for progressivity. Their result showed extensive regressivity in healthcare contributions. For instance they found that elderly people in the lowest income decile contributed about 64 percent of their pretax income to the healthcare system. Gottschalke et al (1989) analyzed the progressivity of healthcare payment in UK, the Netherlands and US using the same tabulation method as Wolfe and Gottschalke. Their findings confirmed the results from Cantor (1988). They showed that US had a regressive healthcare financing system while the other two countries generally had progressive healthcare systems. However the simplicity of these methods does not permit a more in-depth analysis of the redistributive effects of health financing. For example none of the methods used in these studies could inform policy about the extent of horizontal equity or reranking inherent in these healthcare systems. They do not provide information as to the degree of progressivity or regressivity of different healthcare financing systems.

**Recent Developments**

Recent methodological improvements in the analysis of the redistributive effects of health financing have been largely due to a group of European researchers under the aegis of COMAC project on Equity in the Finance and Delivery of Healthcare. These recent studies have gone beyond the descriptive analyses of the previous studies to develop a single measure of progressivity. They have relied extensively on Kakwani (1977) who developed the mathematical model for analyzing tax progressivity, and Aronson, Johson and Lambert (1994) who developed the method for decomposing redistributive effects of
taxation. Wagstaff and Van Doorslaer have extended and applied these models in the field of healthcare financing. Wagstaff et al (1989) used the Kakwani index of progressivity to confirm the descriptive results of the previous studies based in the USA, UK and the Netherlands using 1981 data. They showed that total contributions to healthcare financing in USA and Netherlands were regressive, but more regressive in USA (-0.15) than in the Netherlands (-0.06). It was progressive in the UK (0.03). Out-of-pocket payment was even more regressive in the USA (-0.39) but was surprisingly found to be progressive in the Netherlands (0.12).

Van Doorslaer and Wagstaff (1993) could be said to represent a threshold in the analysis of the redistributive effects of healthcare financing. The study basically attempted to compare progressivity of different conventional forms of healthcare financing among ten OECD countries, namely, Denmark, France, Ireland, Italy, Netherlands, Portugal, Spain, Switzerland, UK and US. Instead of tabulating proportions of income spent on healthcare payment by different categories of income earners as in the previous studies, the authors applied the Kakwani and Suit’s indices of progressivity which made it possible to compare the progressivity of healthcare financing across countries.

The Suit Index, unlike the Kakwani index of progressivity plots the cumulative percentage of pretax income beginning with the lowest to the highest, against the cumulative percentage of pretax income which, therefore, coincides with the 45° diagonal and thus serves as the benchmark against which the progressivity of tax is to be measured. It then plots the cumulative proportion of pretax income against the cumulative proportion of tax payments made by the different categories of households (van Doorslaer and Wagstaff, 1993). If the tax payments lie below the diagonal, then we concluded that the financing system is progressive. If the opposite is the case, then the tax system is regressive and if the tax concentration curve coincides with the diagonal, then the tax payment is said to be proportional.

Applying these methods to the different forms of healthcare financing in the countries, the study found that direct tax revenues as means of financing healthcare were
progressive in all the countries included in the study but the degree of progressivity differed from country to country. Not surprisingly, indirect tax revenues were in most cases characterized by regressiveness but the degree of regressivity again varied across countries. In three of the countries where social health insurance predominated, it was found that this method of financing healthcare was largely regressive because contribution tended to be proportional to income up to an income ceiling (p.40) In some others, the social insurance system was found to be progressive due to the structure of the contributions. The progressivity of private insurance method of healthcare financing was found to depend on a number of other variables. If private insurance is taken out against public sector co-payment, its progressivity depends on the level of co-payment. Where the level of co-payment is low, it tended to be progressive but regressive otherwise. Where private insurance was taken out as supplement to social insurance, it tended to be regressive in character (since in this case, it would be considered as a luxury good). However, among individuals who took private insurance as the sole source of cover, healthcare financing tended to be regressive.

In almost all the countries considered, the study found that direct payments tended to be regressive. Yet the level of regressivity varied across countries. The degree of regressivity was influenced by the prevalence of other forms of payment. For example, where there was general public healthcare cover, such that direct healthcare payment could be considered a luxury good which only the rich could afford, direct payment would tend to be strongly regressive but weakly so if conditions were otherwise.

Klavus and Häkkinen (1996), and Gerdtham and Sundberg (1996) analyzed the income redistribution consequences of different forms of healthcare financing in Finland and Sweden respectively. In Klavus and Häkkinen (1996), the effects were analyzed in terms municipally provided services and national sickness insurance based services. In contrast to some other studies which analyzed redistributive effects simply in terms the differences between pre- and post-payment income like Van Doorslaer and Wagstaff (1993), this study included the income effect of healthcare utilization. The argument for the inclusion of income effect from healthcare utilization is that its exclusion could be
theoretically valid where the decision to utilize healthcare is identical with the decision to finance healthcare. This would be true, for example, in a market oriented healthcare system. But if the two decisions are distinct, then it is important to include the income effect of utilization since some may lose more or less from healthcare financing and gain more or less through service utilization. This would be particularly the case where general taxation and social health insurance separate healthcare financing and healthcare utilization. In terms of methodology, this study like other studies reviewed earlier, used the Kakwani index to measure progressivity of redistribution of net income at the post-payment period.

Gerdtham and Sundberg (1996) analyzed redistributive effect of healthcare financing from four different sources in 1980 and 1990. These sources include the county council taxes, payroll taxes, direct payments and state grants. The decomposition method was the Aronson, Johnson, Lambert (1994) framework. This enabled the authors to disentangle the different components of redistribution effect: vertical and horizontal equity as well as the reranking effect. This study showed that while these financing sources tended to be largely weakly progressive, the major source of horizontal inequity and reranking is the county council taxes varies geographically.

Unlike the situation in Sweden, Klavus and Häkkinen (1996) showed that healthcare financing in Finland was substantially redistributive in character. There were significant transfers of income from the rich (through healthcare financing) to the poor (through healthcare utilization). In other words the rich contribute more to healthcare financing than the poor, but the poor utilized healthcare services more frequently than the rich. This is consistent with the principle of cross subsidization in healthcare financing and healthcare financing according to ability to pay and utilization according to need.

However, the level of redistributive effect varied between the extremes of exclusion and inclusion of healthcare utilization. The redistributive effect arising from excluding healthcare utility was minimal while the bulk of the effect arose from its inclusion. It should be noted, however, that like in many of the other empirical studies reviewed above
which were based on developed countries’ experience of health financing, out-of-pocket or direct payment for healthcare services makes almost an insignificant contribution to the overall healthcare financing system in most of the developed countries where forms of healthcare payment are divorced from healthcare consumption (Van Doorslaer and Wagstaff 1993) The greater proportion of healthcare financing is accomplished through revenues from general taxation, social health insurance and private health insurance. This is in stark contrast to the situation in most developing countries where social and private healthcare insurance are almost non-existent.

Two companion papers, van Doorslaer et al. (1999) and Wagstaff et al. (1999) extended the scope of international comparison of the redistributive effects of healthcare financing to twelve and thirteen OECD countries. Van Doorslaer et al analyzed the effects of average proportion of income spent on healthcare, the progressivity of health financing, the extent of horizontal inequity and reranking on income redistribution using the AJL decomposition method. It showed that progressivity interacts with these three other variables to determine the impact of financing system on redistribution. Wagstaff et al (1999) on the other hand focused on just the progressivity factor in healthcare financing. The paper developed further the methodology used in van Doorslaer et al (1993). Instead of using country-specific equivalent scales and equivalized income, the paper adopts a common equivalence scale. In addition to using equivalent income scales, it also equivalized healthcare payment. In this way, the paper was able to achieve a higher degree of comparison among the countries studied.

The paper also presented evidence of shifts in mixes of financing systems’ progressivity in these countries. While it is generally acknowledged that there were shifts from public to private sources of financing healthcare, the paper analyzed the trends in mixes within the public and private sources of healthcare financing and progressivity consequences of such shifts in mixes.

The results from this particular study showed that using the Kakwani index of progressivity in all the thirteen countries, the direct taxes used to finance healthcare were
progressive although the degrees of progressivity differed from country to country reflecting, perhaps, the degree of income tax progressivity in each country. Indirect taxes were, as expected and as shown in other similar studies, regressive. Overall, however, the Kakwani indices for general taxation (which the weighted average of the indices is for direct and indirect taxes) were positive, implying the dominant effect of direct over indirect taxation. Social insurance was found to be progressive in all but four of the thirteen countries included in the study. This again reflects the level of participation of workers and exemptions or minimal contributions to the scheme by pensioners and the unemployed. Private insurance was progressive in nine of the thirteen countries and regressive in others. The main influence generating this result was the extent of public co-payment in private healthcare insurance. This basically tallied with the earlier findings from Van Doorslaer and Wagstaff (1993): where there was reasonable co-public payment, the tendency was for a wide spread coverage especially for the poor, and hence, higher level of progressivity.

In general direct payments or out-of-pocket payments were found to be regressive. However, the degree of regressivity was also found to vary across countries. For example, while direct payment was found to be regressive in Sweden, Switzerland and France, in some others like Germany, Netherlands and Italy, the regressivity of direct payment was highly attenuated and in fact was progressive in Spain.

The factors accounting for differential levels of progressivity of direct payment include the levels of public cover for the better off and the choice of insurance with low premium and deductibles by those privately insured. If persons of low income were liable to full out of pocket payment, the probability would be that the payment system would be highly regressive.

The concern for progressivity by Wagstaff et al (1999) seems to have overlooked the other potential redistributive effects of financing healthcare. A financing system might well be very progressive on average but still contain a high degree of horizontal inequity due to dispersions at the different income levels. These dispersions may arise from
differences in income taxes which may vary according to geographical location (e.g. urban – rural), tax deductibles, differences in contribution rates for health insurance schemes, variations in risk rating of private insurance premiums, variations in the extent of liability for direct payments, age groups etc (Van Doorslaer et al. 1999). The later paper therefore, considered the full range of effects listed above. Like its companion paper, it used the AJL decomposition method to analyze the redistributive effects of healthcare financing in twelve OECD countries.

The results of these decompositions largely confirmed the results of previous analyses that used data from the same OECD countries. The progressivity of direct taxes and regressivity of indirect taxes were confirmed by the study. The progressivity of social and private health insurance was also found to be conditional upon other attenuating circumstances as discussed earlier. With respect to horizontal inequality, the study found that in all the twelve countries in the sample, differential treatment of equals in direct taxation reduced actual redistributive effect below its potential maximum by a very small amount. In fact, in the absence of horizontal inequity, direct taxes would be only about 5% more redistributive. Thus, almost all the redistributive effect associated with direct taxes arose from vertical redistribution.

Surprisingly, the results showed that indirect taxes have similar redistributive effect as direct taxes in respect of horizontal equity. Indirect taxes would be only 5% more pro-poor if there were no differential treatments at each given level of income. Thus, most of the redistributive effects of indirect taxes seemed to be generated from vertical redistribution. Not surprisingly, social insurance and private healthcare insurance produced more horizontal inequity than direct taxes. This implies that the structural designs of these forms of healthcare financing do not lend them to as much equity as direct taxes.

Direct payments for healthcare had large redistributive effects in many of the countries in the sample. These large effects were attributed, in part, to the random nature of illness but also, in part, to variations in liability for co-payments at each level of income (p.305) In
other words, the redistributive effects of direct payment was heavily modified by differences in the extent of private and social health insurance coverage. The full redistributive impact of direct payment can only be ascertained in a financing system where everyone or at least majority of the population is subject to this system of payment, that is, where the attenuating impact of other forms of healthcare financing on direct payment are minimal. Direct taxes and social insurance seemed in general, to have positive vertical redistribution effect. Direct payment private health insurance schemes in most cases were associated with pro-rich income redistribution.

The main conclusions to be drawn from these results are that the degree of progressivity of healthcare financing system determines, to a large extent, its redistributive effect. But redistributive effect depends on vertical equity and the proportion of household's income taken up by healthcare financing. It also shows that vertical equity is more important than horizontal equity when we are considering redistributive effects of healthcare financing.

All the empirical work reviewed above was based on the experience of OECD countries which could be said to constitute a homogenous group of countries at the same level of economic and social development. Developing countries are more heterogeneous. Social and economic conditions and institutions prevailing in these countries differ in fundamental ways from those existing in the developed countries and even differ widely among themselves. In this sense, Wagstaff and van Doorslaer (2001), Baker and van der Gaag (1993) and Pannarunothai and Mills (1997) who analyzed the redistributive effects of healthcare payment in Vietnam, five developing countries, and Thailand respectively could be considered as remarkable break from the other previous studies. Baker and van der Gaag (1993) and Pannarunothai and Mills (1997) in contrast to the study of Wagstaff and van Doorsaler used simple tabulation method to analyze equity in healthcare financing. While Baker and van der Gaag found evidence of progressivity in the five countries considered, Pannarunathai and Mills found contrary evidence in Phitsanulok municipal area of Thailand. Much of this regressivity owed to the regressivity of the Thai tax system. Less than one fifth of the general tax revenue came from income taxation while the rest derived from general and specific sales tax. There was also evidence of
horizontal inequity in healthcare financing as those who were covered by the health insurance system did not make contributions to the scheme but were rather subsidized from general tax revenue.

The study by Wagstaff and van Doorslaer used data from Vietnam Living Standard Surveys (VLSS) for 1992-93 and 1997-98. This study, unlike Baker and van der Gaag (1993) and Pannarunthai and Mills (1997) mentioned above used the AJL decomposition method to obtain the various components of the total redistributive effect of direct payment. This decomposition allowed the researchers to quantify the extent of vertical inequality in healthcare payment, the extent people with different ATP end up paying similar proportions of ATP toward healthcare financing, the extent of horizontal inequality and the extent of re-ranking associated with the move from prepayment to post-payment income distribution.

The results from this study showed that total redistributive effect of direct payment were generally negative for both years. In other words, the redistribution of income at the post-payment period favoured the rich. However, the redistributive effect fell between the two time periods. Healthcare financing contributed immensely to achieving a more egalitarian Vietnam between 1993 and 1998. The reduction in the negative redistribution effect was accounted for largely by a fall in the contribution of direct payment to healthcare financing in that country. Out-of-pocket payment absorbed increasingly smaller share of prepayment income in Vietnam. The study noted that much of the reduction in redistributive effect with respect to prepayment income was due to reduced vertical income redistribution. For example, the redistributive effect fell from 47% in 1993 to mere 6% in 1998 indicating increased concern for vertical equity.

While the above studies have used largely the AJL method of decomposition to achieve a high degree of analytical clarity, Klavus (2001) introduced a new perspective to the analysis of income redistributive effects of healthcare financing by using the concept of progressivity dominance. Dominance relations provide a useful extension to summary statistics. Dominance relations have found extensive applications in the theory of income
distribution and poverty analysis especially when there is need to compare the extent of inequality in two or more distributions. Its usefulness lies in its ability to reveal some features of inequality that are not often explicit in analysis based on summary measures of distribution. As the study confirms, the distributional outcome associated with the entire curve do not always conform. For example, at certain income levels the dominance analysis may reveal redistributions quite opposite that shown by summary statistics.

This study used two sets of data from the Finnish healthcare financing system to demonstrate the redistributive effects of health financing and the effectiveness of dominance analysis in this area. The data sets were the healthcare mixes of 1987 and 1998. The period represented an era of severe economic recession in that country. The recession led to structural changes in Finish economy which in turn led to changes in financing structure, incidence of taxes and payments as well as changes in income distribution. All of these affected the progressivity of health financing during the period.

The results from using this method were not remarkably different from those obtained using summary statistics. Healthcare financing through general taxation was shown to be generally progressive but more so at certain nodal points than others. State taxes were shown to be regressive at low income levels but progressive at high income nodal points. In general healthcare financing between the periods in Finland was regressive but the levels of regressivity varied across different components of the financing system. The dominance analysis revealed that income distribution worsened between 1987 and 1996 as the Lorenz curve of inequality for 1996 clearly dominated that of 1987. Indirect taxes and direct payments were generally regressive. But again this regressivity varied according income levels and also between the two periods. Households’ direct payments for healthcare services were however, less regressive at lower income levels and more regressive at higher income brackets. Local taxes and sickness insurance payments were, on the whole progressive in both periods.
SECTION C

3.7 Brief Review of Previous Research Findings on Health Inequalities

Though there were research results prior to the Black Report in 1980 linking social structure and health (Irwin and Scali 2005), the publication of the report could be said to have formally inaugurated the debate in literature about the nature and extent of health inequalities caused by differences in socioeconomic status. Essentially, the Black Report investigated differences in mortality rates among socioeconomic classes in Britain since the establishment of the National Health Services (NHS) in 1948. Using as methodology the simple range, the report concluded that in spite of NHS, inequalities in mortality rate among occupational classes in Britain had been increasing since the 1930s (Black et al. 1980). Since the publication of the report there has been a growing number of international literature in the area of inequalities in health and mortality.

Most of these studies have, however, concentrated on international comparison in social economic morbidity and mortality differentials in the industrialized countries of Europe and America. These studies include Pamuk (1985, 1988); Pans and Lillard (1995); Van Doorslaer et al. (1997); Mackenbach et al (1997); Lahelma et al. (1999); Marnot et al. (1997); Hanstrom (1999); Solie and Johnson (1999); Larer et al (1998); Kennedy, Kawachi and Prothrow-Stith (1996); Wilkinson (1997); Judge (1995); Stechel (1988); Waldmann (1992); Cibu Study Team (1992); Barrera (1990); Thomas et al. (1996); Guilkey and Riphahn (1998); Franklin (1999); Pal (1999); Martorell et al (1981); Mackenbach and Kunst (1997), Kawachi and Kenedy (1999); MacIntyre (1997), Lynch et al (1998); Judge, Melligan and Benzeral (1998); Shikolnikov et al (1998), Humphries and Van Doorslaer (2000); Wagstaff (2002); Van Doorslaer and Wagstaff (1993); Wagstaff et al (1991); Wagstaff (2001, 2002); Wagstaff, Paci and Joshi (2001). These studies reflect this general concern with inequalities in health and cross country comparisons in Europe and America.
Many of the studies have however differed remarkably in their conceptualization of the indicator of social position along which health inequalities should be modeled. In other words, the studies have differed in the way they categorize or rank the members of the population in analyzing inequalities among them. The ranking variable has often been income or wealth (van Doorslaer et al 1997; Humphries and van Doorslaer 2000; Pritchett and Summers 1993). For example, van Doorslaer et al (1997) using the concentration index (Wagstaff et al, 1991) as a measure of inequality compared variation in self rated health along the income gradient among nine industrialized countries of Europe and America. Their results suggest that income related health inequalities exist in all the nine countries included in the study. There were substantial variations in the survey which made comparison difficult and rather unreliable. Pritchett and Summers (1993) studied the relationship between per capita GDP and two measures of health - infant mortality and life expectancy - using instrumental variable specification model. They combined both cross sectional and time series data of both developing and developed countries in effort to eliminate potential biases. They established strong links between growth in per capita GDP and infant and child mortalities. Income elasticity of infant and child health is higher than overall life expectancy perhaps indicating that infant mortality and child health respond to changes in economic indicators than life expectancy. Humphries and van Doorslaer et al (2000) using same methodology but with Canadian data, surprisingly found that Canada with almost similar income distribution as Europe shows a degree of income related health inequality closer to that of US than Europe.

A number of studies have used the education variable as an indicator of the social economic gradient (Shkolinkov et al 1998 and Valkonen et al 1997; Lahelma et al (1997). Lahelma et al (1997) also analyzed variations in health (indexed by limiting long standing illness and self-assessed health) along several social gradients including education, regional structures, employment status, gender, and age in Finland. A multivariate logistic regression model was used. Unfortunately, in attempt to render the results amenable to easy interpretation they had to adopt the Simple Odd Ratio method, implying that variables with multiple categories were simply dichotomised. This
The dichotomization process led to a loss of information. In any case, the study found that women were more likely than men to report health as below good. Differences in age were also found to be levelling off over time for both men and women. Regional patterns of health distribution were also flattening off, but differences in health along the educational gradient were quite distinct.

A similar result is reported by Shkolnikov et al (1998) who using cross sectional data for Russia in 1979 and 1989, found that in spite of the socialist policies in pre reform Russia, differences in adult mortality along the educational gradient were as high as those found in capitalist West. Valkonen et al (1997) surprisingly found that in Finland life expectancy with disability is shortest among the educated, while longest disability-free life expectancy was higher for the educated than for the less educated. The same is also true for men and women. Women had a higher prevalence of disability, but longer life expectancy because life expectancy with disability was higher for women than for men. Mackenbach et al (1997), Lahelma and Valkonen (1990) are related studies in differential health inequalities along social classes in European countries. All the countries surveyed showed the same close association between mortality and education.

The Black Report was based on health differences arising from the occupational class distribution in England and Wales. It used the range method, which consists in simply comparing the health distribution of social class I with social class II. It found the social patterning of health along the class gradient while the result of this study has been confirmed by several others, since, the method – the range method has been largely discarded in preference for more sophisticated method that put into consideration the experience of the whole population rather the restrict comparison to two classes at a time.

Pamuk (1985) is one of the studies that advanced the method used by the Black Report. In his study of social class inequality in mortality in England and Wales from 1921 to 1972, Pamuk used the slope index of inequality (SII). This measure will be reviewed more extensively under methodology, but his finding indicated that social class inequality in health in England and Wales has been in an upward trend, though much of the increase
in health differentials may be due to changes in behavioural patterns across social classes. Blairie (1986) while confirming the results that mortality and disease distribution are patterned along social classes attempted to explain the channel of transmission; there are social class differences in exposure to factors implicated in the aetiology of diseases. Marmot et al (1997) emphasized that both educational and occupational gradients are important in explaining the differences in health, though they are not accounted for by parental social class or by intelligence, but rather are mediated through behaviour such as smoking, and the features of psycho-social environment.

A number of studies have also focused exclusively on gender differences in health in developed counties (Arber, 1977; Fernandez et al 1999; Diaz 2002; and Haavio-Mannila, 1986). There is near unanimity in these studies, irrespective of method of investigation, showing that women have a higher morbidity rate than men, but also that men have higher mortality rate than women. Diaz (2002), for example found reverse in health scores for men and women at younger and older ages. Women are worse-off at a young age, but better of than men at older ages. Hummer et al (1998) observed a similar inversion in what they referred to as an “epidemiologic paradox and racial mortality cross over” (p.557). The former refers to lower adult mortality rates exhibited by Hispanics in US despite socioeconomic profiles that favour whites.

A large part of literature on health inequalities based on the social gradient has been devoted to inequalities and variation in child health. Studies in this category include Pal (1999); Barrera (1990); Guilkey and Riphahn (1998); Cebu Study Team (1992); Skoufias (1998); Van Den Brock et al (1993); Franklin (1999); Pans and Lillard (1995); Thomas et al (1996); Habicht et al (1974); Steckel (1995) Mantorell et at (1981); Bielicki et al (1981), Wagstaff (2000); Oni (1988) Currie (1995); and Thomas et al (1990). The dominant child indicators used by many of these studies are the anthropometric outcomes and child nutritional status. Pal (1999) for example used the probit and ordered probit to model child nutritional status in six villages in India. This study found significant negative association between the weight-for-age and the gender of the child. Boys were more likely to be well nourished than girls. The likelihood of a girl being well nourished
decreased as birth order increases and as age increases. Surprisingly, increased girl literacy seems also to worsen the level of malnutrition. For boys the likelihood of being well nourished, as expected, increased with per capita income of households.

Ityavyar (1988) discussed different forms of inequalities in distribution of health services in Nigeria, particularly regional inequalities and blames the imbalances in the development of pre and post-colonial capitalism. Oni (1988) carried out a similar study, but restricted to a Nigerian city – Ilorin urban. He applied the indirect demographic estimation technique – the Brass method to estimate child mortality rate and then used regression analysis to estimate child mortality differentials along the socio economic gradient. His findings were similar to those of Caldwell (1979) who found that parental education was a critical factor in mortality decline in Nigeria (Caldwell 1979). However, unlike Caldwell (1979) who found maternal education was the crucial factor in child mortality decline, this study found a significant independent effect for father’s education.

Hillier and Serviens (1986) studied differences in health along ethnic groups in the UK. They emphasized the susceptibility of ethnic groups to different diseases and the impact of differing experiences of ethnic groups in their perception of illness, and the use of health services. This inequality in health is found to vary according to ethnic groups. While the results from Oni (1988) and Caldwell (1979) conform to intuitive expectations, Skoufias (1998) seems to present a different picture. Skoufias, in studying the determinants of child health during the economic transition in Romania, used growth attainment of children measured by weight-for-height and height-for-age as indicators of short run and long run nutritional status. He used the reduced form regression model and found significant differences among children across gender, rural, and urban locations. Surprisingly and in contrast to some of the earlier studies reviewed above, he found that in urban areas income and the father’s level of education do not correlate with growth attainment for either boys or girls. The educational level of the mother also does not correlate with long run nutritional status of children in rural areas although it is associated with poor growth attainment of boys and higher growth attainment for girls in urban areas. Barrera (1990) suggests that the mechanism of transmission from the mother’s
education to child health is through the productivity of child health production input and lowering the information cost in this production process. Kovsted et al. 2003) who studied the relationship between child health and mortality on one hand and the mother’s education on the other in Guinea-Bissau argue that this transmission is accomplished specifically through the mother’s health knowledge.

Guilkey and Riphahn (1998) also confirm the importance of maternal levels of education (and breast feeding) in reducing child mortality even after controlling for the direct determinants of child mortality. The findings by Guilkey and Riphahn (1998) confirm similar results from the Cebu Study Team (1992) who used exactly the same data set from Philippines. The Cebu Study Team (1992) in analyzing the child health production functions found strong positive effects on child survival of biological determinants (breast feeders), and the negative effects of mechanisms such as diarrhoea and respiratory infections.

Steckel (1988) took a historical perspective to the search for socioeconomic variables that influenced child and women mortality. He used the American national census data for 1850 to 1860 and employed the logistic regression model to obtain results that showed mortality among women and children during the period varied by region, number of siblings. Infant mortality declined with occupational status, and the socioeconomic characteristics of households, though less than found in modern survey data. He concluded that “when and how socioeconomic conditions began to exert a strong influence on health remain open questions” (p. 345).

In two related contributions, Heller and Drake (1979), and Thomas et al (1990) confirm the interdependence of a child’s poor nutritional status and poor health. Heller and Drake emphasized that the level and quality of nutrient intake clearly vary with parental income and education. But they did not specify which is more critical to the nutritional status of the child - mother’s or father’s level of education. Inequalities among a household’s children are influenced more by the limited ability of households to even out available
resources over time among the children. But it is also possible that inequalities in the welfare of children especially in developing countries may be a result of traditional patterns of investment on children in a household, rather arising from inability to even resources over time.

Wagstaff (2000) compared the levels of inequalities in child mortality across income quartiles in nine developing countries. He used stochastic dominance of concentration curves to carry out pair-wise inter-country comparisons. The results showed that the child mortality was higher for the poor than for the well-off. It was also found that inequalities were higher in some countries such as Brazil, Nicaragua, and the Philippines and low in Ghana, Pakistan, and Vietnam. However, the study did not explain why the countries varied in their rate of child mortality, or why inequality is in favour of higher income groups.

A new dimension to the health inequality literature has sought to explain how inequality in income distribution (rather than average income) can give rise to inequalities in health. Leading this research group in the UK is Richard Wilkinson, who in a series of articles in the 1980 and 1990s, proposed the view that income inequality is the critical determinant of variation in health among developed countries (see, for example Wilkinson 1997, 1995). In the US, the views of this group were advanced by Kaplan et al (1999), Kawachi and Kennedy et al (1996), Lynch et al (1998), and Lynch and Kaplan (1997), among others. The consistent evidence from these studies is per capita or average income distribution is a more critical determinant of inequalities in health distribution.

While some other researchers have tried to question the quality of the data, the method of aggregation and the evidence from these researches (Judge, 1995; Gravelle et al 2002; Wagstaff and van Doorslaer 2000) others have tried to improve the methodology of the research in this area (Wildman, 2003) the argument has progressed beyond accumulating evidence to the specific pathways or channels through which inequality in income distribution negatively impact on health distribution.
Some have pointed to the fact that income inequality may influence health through the mediation of under investment in social goods like education and healthcare and description of social cohesion and erosion of social capital. Kawachi and Kennedy, (1999); Kawachi et al (1997); Kawachi et al (1999), Lynch and Kaplan (1997) emphasize the possible role the social structure under which individuals play out their daily lives have on the health of individuals. Thus while not neglecting individual characteristics; social structures may be critical determinants of health. This is consistent with the considerations of the Commission on Social Determinants of Health (2005) that emphasizes the need to attack the social structures that give rise to inequalities in health. The next subsection analyzes the evidence from Nigeria.

3.8 Summary

In summary, it seems evident from a review of available empirical studies that different forms of healthcare financing have different implications for income redistributive effects. Some methods of healthcare financing have demonstrably strong pro-poor effects on redistribution while some others have pro-rich redistributive effects. But in all, it seems correct to say that the redistributive effects of a financing system depends to a large extent on existing social, economic and institutional set up in the economy. Even a strongly pro-poor financing system such as direct income taxation depends on the structure of taxation adopted by the policy-makers. Current literature on socioeconomic inequalities in health shows clearly that in general, health distribution is, in part, determined by distribution in some socioeconomic variables.

3.9 Contributions of this Study to Literature

Our research builds on and extends the frontiers of the existing studies in a number of directions. In the first place it is important to observe that results from these empirical studies depend to a large extent on the underlying assumptions of the study. Econometric estimation results are liable to a wide variety of data generation processes (Leibbrandt and Woolard, 1999). For example, assumptions about household size, equivalent scales,
the presence or absence of household economies of scale etc. are likely to influence the household or individual's income classification and consequently the presence or absence of equity in resource distribution. This research will establish the robustness or other wise of the effects of alternative assumption on household economies of scale and composition on the estimated redistributive effect.

Furthermore, the results from the above studies derive basically from healthcare systems where healthcare consumers have wide range of choices of methods of paying for health their healthcare cost. Cross elasticities of demand are likely to be significant. For example, in almost all the country studies, consumers have alternatives to direct payments, significant subsidies through general taxation or one form of third party intermediation or the other (Social health insurance, private health insurance, pre-payment schemes etc). These available alternative forms of financing through third-party arrangements are likely to influence the over all impact of healthcare payment on a household or individual and its consumption of other basic necessities. This is different from a situation where direct payment has no significant alternative payments system and where there is little or no significant public subsidization of healthcare consumption like the situation in Nigeria. Redistributive outcomes are more likely to be different in a healthcare financing system or market structure that is more competitive market oriented than a system that has substantial public and third party interventions. The study provides concrete evidence of redistributive effect of healthcare financing from a distinctively developing country perspective.

It is also to be noted that most of the studies have concentrated on the industrialized countries with epidemiological features that are radically different from those of developing countries. In the former, diseases associated with old age and chronic health problems predominate while in the latter, acute illnesses and infectious and parasitic diseases are prevalent. Therefore, financing outcomes and policies are not likely to be the same. None of these studies related the prior distribution of health to income redistributive effects. This study provides this necessary background to more meaningfully relate health distribution and income redistribution.
Furthermore, apart from the studies by Gerdtham and Sundberg (1996), van Doorslaer et al. (1999), and Wagstaff et al (1999) none has decomposed inequalities in healthcare financing using the Gini index. Even these later studies had the problem of defining 'income equals', a weakness that is clearly manifest in every application of the Aronson, Johnson, and Lambert (1994) (AJL) method of decomposition whether in the tax literature on the healthcare finance literature. Moreover, they have been unable to estimate the standard errors that are necessary for post estimation analysis. The present study takes care of these weaknesses by not only estimating the standard errors. This research will not only a statistically determined band width (based on the kernel function) to replace the arbitrarily determined bandwidth of the previous studies. This will impact significantly on the estimated components of the redistributive effect. The study will also advance the research in this area by introducing from the tax literature an ethically flexible parameter of the policy-maker's aversion to inequality.

In terms of methodology, it will compare two frameworks for decomposition – the AJL and Duclos, Jalbert, and Araar (2003) (DJA) frameworks and compare the relative performance. It also contributes to the health inequality literature by comparing the relative performance of SAH (Self-assessed Health) which is popular in general surveys and the HSI (Health Status Index) that is new and remains largely untested.
Chapter 4

Methodological Framework

4.1 Introduction

The foregoing chapter situates the research in the context of its economic theoreic foundations and the previous empirical literature in this area of research. It is shown that this study emanates from and builds on the previous efforts. This chapter sets out the main outlines in the development of the methodological framework for analysis for health distribution and income redistribution which are the key issues in the subsequent chapters. Since there are several different objectives requiring the employment of different analytical tools at different sections of this research, only the general frameworks will presented here. The details of these will be presented at appropriate points in the report. This is to enable the reader have immediate touch with the relevant tool of analysis at each point in time and to compare the results generated with the method used in generating it. Two methodological outlines will be presented. The first is the outline on issues concerning the measurement of health and health distribution. The second focuses on the development of the methods for income redistribution.

4.2 Measurement of Health Status

Choosing the appropriate indicator of health status and how to measure it are critical to the study of health distribution. Several indicators have been suggested and used in empirical literatures as measures of health status (see McDowell and Newell 1996) These include child mortality, adult mortality rate, self assessed health, anthropometric outcomes, health utility index (HUI), hospital visits, health status index (HSI) and
presence of chronic illness among others. The data available for this study allows an analysis of the robustness of transforming the categories of the frequently used indicator of health status: self assessed health status (SAH) into a continuous variable. This is made possible by the inclusion in the questionnaire of a more objective and cardinal measure of health, the health status index (HSI) in our data set. The HSI is derived from EuroQol group's recently developed instrument for measuring health-related quality of life (HRQoL).

This instrument was first developed in 1987 by the European Quality of Life Group (McDowell and Newell 1996) but has now been translated into several languages (Coons, 2003) for population health surveys. Structure-wise, EQ-5D has two parts: the descriptive part that describes 243 possible health states, and a visual analogue scale (VAS) with (0, 100) endpoints. The 0 describes the “worst imaginable health state” while 100 is the “best imaginable health state” It has 5 dimensions of self rated health in respect of Mobility, Self care, Usual activity, Pain/Discomfort and Anxiety/Depression. Each of these has three levels. The scores on the 5 different dimensions produce a health index scaled on the (0,1) interval and the VAS gives a self-rated assessment of health status (McDowell and Newell, 1996). The descriptive part and the visual analogues scale are alternative to each other. The questionnaire included both parts. However, knowing the inclination of the survey population to visual presentation, the main interest was on the response to the visual scale.

SAH has been standardized and adopted by WHO (World Health Organization 1996). Respondents in health survey are usually asked the question “How would you generally rate your state of health today?” Respondents are required to evaluate their health on a likert scale: excellent, very good, good fair, poor. Although SAH as a measure of health status appears simple and subjective, it has been shown to be a very powerful predictor of mortality and onset of disability (Van Doorslaer and Koolman 2002)

Furthermore, this indicator has been shown to be homoscedastic as it does not vary systematically with socioeconomic status (Burstrom and Fredlund 2001) It is assumed
that individual's health status is determined by a health production function with a vector of exogenous variables, $X_s$. Some of the $X_s$ would naturally be deterministic (e.g. biological constitution of an individual) and so not likely to be influenced by policy. Others may be socioeconomic characteristics such as income which can be modified by policy.

Since SAH is basically a categorical variable, it provides ordinal ranking of individual's perception of their health status. This is a major limitation for the use of SAH as health indicator.

In order to use SAH for empirical investigations, researchers have devised several methods of transforming the categories into a cardinal measure in order to overcome this data limitation. One way is to apply OLS to the ordinal data but this has severe econometric limitations associated with linear probability models (see for example, Aldrich and Nelson 1984). Another approach to this transformation is to dichotomize the responses into two categories-good health and less than good health and then apply logistic regression. But this obviously leads to loss of information and the cut-off point might be considered rather arbitrary (van Doorslaer et. al. 1997). It has been shown also to lead to rank reversal when comparing health inequality among populations or over time (van Doorslaer and Koolman, 2002).

A more recent approach to overcoming the categorical data problem is to impose some sort of scaling assumptions on the data. One way this has been achieved by researchers is to assume that underlying the categorical responses is a latent $SAH$ variable which follows a standard lognormal distribution (Wagstaff and van Doorslaer 1994, Van Doorslaer et al 1997). This enables the generation of a cardinal scale on $SAH$ and thus allows for the construction of concentration indices for estimation of health inequality. Unlike the method of dichotomizing, this method allows for comparison of inequality in different populations but its main disadvantage is the strong assumption of identical latent health distribution in different populations. Furthermore, the latent variable does not lie
on the [0,1] scale that should characterize a health utility score (van Doorslaer and Jones 2003).

Another method of imposing cardinal scale on the ordinal responses is to estimate Ordered Probit or Logit regression models. In this case, it is assumed that underlying the SAH responses is a latent variable with an Ordered Probit or Logit distribution. This is typically used in modelling categories with ranks such as SAH (Stata Reference Manual, 1997). This is the model that will be used in estimating the determinants of health in SAH.

4.3 Measuring Inequalities in Health Distribution

The development of methodological models for the estimation of health inequalities has paralleled the developments in the income distribution literature from which many of the health inequality indices have been adapted. Summaries of measures of income distribution inequalities and their theoretical fundamentals are provided in Sen (1973) and Sen and Foster (1997) while Wagstaff et al (1991), and Mackenbach and Kunst (1997) provide good summaries of the indices of health inequalities. Some of the health inequality indices that have found empirical applications in the health economics literature include: the Lorenz curve and associated Gini index, the generalized Lorenz curve and associated generalized Gini index, Pseudo Lorenz curve and associated Pseudo Gini Index, the range, percentiles, population attributable risks (PAR), GMF indices, index of similarity and a class of regression-based indices (Wagstaff et al. 1991; Mackenbach 1997).

A major debate surrounding these health indices is the normative question of what should policy concern itself? A strand of the health inequality methodology and indices has concerned itself with the measurement of pure inequality in health (health per se) as a univariate distribution (e.g. Le Grand 1987, Gakidou et al. 2000) Others have focused on the joint distribution of health and any of the indices of socioeconomic status. The later form of health inequalities is assumed to be amenable to policy changes. In this study the health inequality both those amenable to policy changes and those that are not (e.g. arising from age and gender) will be measured according to their variable sources. If
inequalities in health per se were the object of measurement, then the appropriate index is
the Lorenz curve and the associated Gini index just as in the measurement of income
inequality. If the joint distribution of health with some other indicator of socioeconomic
status is the object of measurement (a bivariate distribution) then the parallel measure
may be the concentration curve and the associated concentration Index.

4.4 Measuring Bivariate Distributions in Health

Three measures of inequality in socioeconomic distribution of health have been shown to
possess the following desirable properties required of a good measure of health
inequalities (Wagstaff et al. 1989; Wagstaff et al., 1991; van Dooslaer et al., 1997;
Kakwani et al. 1997)

(i) It must reflect the experiences of the entire population
(ii) It must be sensitive to the distribution of the population across socioeconomic
groups (SEG)
(iii) It must reflect the socioeconomic dimension to health inequalities.

The three measures that possess these properties are the slope index of inequality (SII)
and its associated relative index of inequality (RII) (Pamuk, 1985; 1988) and the ill health
Concentration Index (C) (Wagstaff et al., 1991). SII is defined as the “slope of the
regression line showing the relationship between a social class health status and its
relative rank (Rj) in the socioeconomic distribution” (Mackenbach and Kunst, 1997). It is
the average decline in the standardized ill-health as one moves from the lowest rank to
the highest (Pamuk, 1985). It measures the 'absolute' advantage of moving a step higher
in SEG.

The relative index of inequality is measured by dividing the slope by the mean health. It
indicates the proportionate decline in health over the population when ranked by social
class (Pamuk, 1985). It is important to note that the difference between SII and RII is that
while the former is a measure of absolute difference (and therefore mean-sensitive) the
latter is a measure of relative difference (and therefore mean insensitive). Therefore, here again the choice is a normative issue: whether one considers absolute or relative differences as the matter of social concern.

To estimate the SII, if the data were grouped data (e.g. if the population is structured according to income group, level of education attained, or occupational group), the distribution of the random error term would tend to be heteroskedastic, implying that the use of OLS estimators is, in this case, inefficient though they are still unbiased. The remedy is to recourse to generalized least square (GLS), in particular, the Weighted Least Squares (WLS). In this case the functional model may be expressed as:

\[ \frac{\mu}{\mu} \sqrt{n_i} = a \sqrt{n_i} + \beta R_i \sqrt{n_i} + u_i \]  

(4.1)

Where: \( \mu = \left( \frac{1}{n} \right) \sum_{i=1}^{n} x_i \) or the overall mean level of health (or ill-health score)

\( \mu_i \) = Health score of ith class (or ill-health Score).

\( n_i \) = the size or weight of the SEG

\( R_i \) = relative rank of the SEG

\( u_i \) = homoscedastic error term

The WLS Estimator of \( \beta \) is the RII.

The Concentration curve (Wagstaff et al. 1991; van Doorslaer et al. 1997; Kakwani, Wagstaff and van Doorslaer et al 1997) plots the cumulative proportions of the population in ascending order of advantage (on the x-axis) against the cumulative proportions of health (on the y-axis). Health categories are assumed to be transformable into continuous variable. If the health of the least advantaged person in the population is as good as that of the most advantaged persons, that is, if health status is completely egalitarian in distribution, then the concentration curve (\( L_{oo} \)) will coincide with the
diagonal. However, if the curve lies above the diagonal, then inequality in ill health exists and favours the more advantaged members of the population.

In comparing health inequalities across two populations A and B, if the cumulative distribution of ill-health in group A lies everywhere closer to the diagonal than the cumulative distribution of B, then population A’s concentration curve is said to stochastically dominate that of population B in which case it would be concluded that the level of inequality in population A is unambiguously less than that in population B (Fields and Fei, 1978). However, if the cumulative distribution of ill-health in populations A and B cross, then it is difficult to say without ambiguity that A dominates B or vice versa. In this case, a numerical measure can be used to quantify the relative inequalities in the two populations. Unfortunately, such dominance ordering does not give a complete ordering of the distribution of interest: while some pairs of distribution can be unambiguously ranked some provide incomplete or partial ordering or ranking.

A numerical measure that has been used in such cases is the ill-health concentration index C. It measures the area between the concentration curve and the 45-degree line as a fraction of the total area under the 45-degree line. That is

\[ C = \frac{M}{M + L} = 2L = 2\left(\frac{1}{2} \int L\right) \]  

(4.2)

Where \( M \) is the area between the egalitarian line and the concentration curve and \( L \) represents the whole area under the egalitarian line.

\[ C = 1 - 2\int_0^L p - L(p) \, dp, \quad \forall p \in [0,1] \]  

(4.3)

Where \( L(p) \) is the fraction of total health received by the lowest \( p \)th percent of the population and satisfies the conditions (Kakwani, 1980)
(a) For \( p = 0 \) \( L(p) = 0 \)

(b) For \( p = 1 \) \( L(p) = 1 \)

Thus, \( C \) is one minus twice the area between \( L(p) \) and the diagonal. The minimum and maximum values of \( C \) for continuous distributions are \(-1\) and \(+1\) respectively. When \( C = -1 \), this implies that the population’s ill health is concentrated on the most disadvantaged member of the population and when \( C = +1 \), ill-health is concentrated on the most advantaged member of the population. When \( C = 0 \), that is when \( C \) coincides with diagonal, ill-health is equally distributed among all the members of the population (Wagstaff, Paci and Van Doorslaer, 1989)

For individual level data, Kakwani et al., (1997) show that \( C \) can be alternatively calculated as:

\[
C = \frac{2}{n \mu} \sum_{i=1}^{n} y_i R_i - 1 \tag{4.4}
\]

Where \( y_i \) = health (or ill-health) score of the \( i \)th individual

\( n \) = the sample size

\( \mu = \frac{\sum_{i=1}^{n} y_i}{n} \) is the mean level of ill health in the population

\( R_i \) = the relative rank of the \( i \)th person

However, Wagstaff (2005) shows that when the variable of interest is binary with mean \( \mu \), \( C \) has a minimum and maximum values given by \( C = \mu - 1 + (1/n) \) and \( C = 1 - \mu + (1/n) \) respectively. In large samples, that as the sample size \( n \to \infty \) these bounds reduce to \( C = \mu - 1 \) and \( 1 - \mu \) respectively since in these cases the last term in both equations vanishes.

For grouped data, equation (4.4) becomes
\[ C = \frac{2 \sum_{t=1}^{T} f_t \mu_t R_t}{\mu} \]  \hspace{1cm} (4.5)

Where \( \mu_t = t(t = 1, 2, \ldots, T) \) is the morbidity rate of the \( t \)th SEG.

\( f_t = \) \( t \)th SEG share or weight in the entire population

\[ R_t = \sum_{\gamma=1}^{t-1} f_\gamma + 1/2 f_t \]  \hspace{1cm} (4.6)

\( R_t \) indicates the fractional rank or cumulative percentage of the population ending at the midpoint of each group interval.

4.5 Measuring the Redistributive Effect of Healthcare Financing

If the poor have heavier disease burden, then, they are likely to suffer greater impoverishment due to either catastrophic health spending or more frequenting of healthcare facilities if the healthcare financing system is not designed to be sufficiently progressive. Healthcare cost may be sufficiently high to threaten the ability of the household to provide other household needs such as food and education. In that case, they are likely to be worse off in relative terms than the well-off in the post payment situation. In other words, the intention here is to measure the income equalizing or dis-equalizing effects of healthcare payment system. Put another way, what is the extent of progressivity of the health payment system? Does the payment system redistribute income in a more or less egalitarian way? The following subsections show the derivation of the critical components of the redistributive analysis.

Measuring Progressivity and Redistributive Effects of Healthcare Financing

As the development of the methods for measuring progressivity and redistributive effects have borrowed largely from the theoretical insights from the tax literature, it is on the latter that this section will depend. Analogous to tax progressivity, the progressivity of a healthcare payment system is the "extent to which payments for healthcare rise or fall as
a proportion of a person's income as his or her income rises or falls" (van Doorslaer and Wagstaff, 1993). A system of healthcare payment is progressive if healthcare payment rises by a higher proportion in response to a proportional increase in income. If the percentage or proportionate increase in payment is equal to the percentage increase in income, then healthcare payment is a proportional system. On the other hand, if the percentage increase in healthcare payment is less than the percentage increase in income then the system is regressive. This implies that for a progressive healthcare payment system the lower income group bears a proportionately lesser cost of financing the healthcare system. Conversely, the topmost class bears a disproportionately higher cost of healthcare financing.

One simple way to measure the progressivity of healthcare payment is to compare the share of each income decile with its share of healthcare payments. However, while the method is sufficient to indicate whether or not progressivity is present, it is hardly useful for estimating the degree of progressivity especially when we need to compare the progressivity of healthcare payment in two different economies. A measure of progressivity that has become both standard and conventional is the one developed by Musgrave and Thin (1948). They defined a measure of "effective progression" as an index that measures the "extent to which a given tax structure results in a shift in the distribution of income towards equality". This implies

\[ RE = G_x - G_{x-T} \tag{4.7} \]

Where \( RE \) = Redistributive effect; \( G_x \) = Gini index of prepayment income; \( G_{x-T} \) = Gini index of post-payment income. That is the redistributive effect is twice the area between the pre- and post-payment Lorenz curves. If the value is positive then the health payment system is progressive. It reduces income inequality in the after-payment period. If the value is zero, then the payment system is proportional while a negative value will suggest a regressive healthcare payment mechanism.
4.6 Kakwani’s Index of Progressivity

Kakwani (1977) defined progressivity in terms of the elasticity of the tax function $T(x)$ with respect to income $x$. It is derived from the principle of Lorenz curve also. Let $Lx(p)$ be the Lorenz for pre-payment income. Let $Lc(p)$ be the payment concentration curve obtained by plotting the cumulative percentage of the population ranked according to pre-payment income on the $x$-axis, and the cumulative percentage of healthcare payments on the vertical axis. For a proportional healthcare payment system, the $Lx(p)$ curve and $Lc(p)$ curve must coincide. Progressivity is then measured by departures of $Lc(p)$ from $Lx(p)$. That is, the index of progressivity measures the extent the tax function departs from proportionality. Thus the Kakwani index of progressivity of healthcare payment on prepayment income may be defined as:

$$K = Lc(p) - Lx(p) \text{ or } K = 2\int_0^1 Lc(p) - Lx(p)dp \quad (4.8)$$

In other words, $K$ may simply be defined as $K = C_{tax} - G_x$. The Kakwani index shows that progressivity of healthcare payment is measured by twice the area between the payment concentration curve and the prepayment income Lorenz curve. $K$ is positive if the payment elasticity is greater than unity for all levels of income. It is regressive if payment elasticity is less than unity for all $x$. It is proportional if payment elasticity is constant at every level of $x$.

Since, unlike the Gini index that takes a value in the [0,1] range, the concentration curve takes values in the [-1,1] range. This implies that the minimum value of $K$ is $-2$. This occurs when one person receives all the income so that $G_x = 1$ while the poorest person pays all the taxes so that $C_{tax} = -1$. Then, $K = 0$ if payment elasticity is unity at all incomes. This implies that progressivity does not only depend on $F(G_x - G_{x-\tau})$ as is the case in Musgrave and Thin (1948) specification, but also on payment elasticity. This is clearly seen when it is recalled that
\[ G_{x-T} = G_x - \frac{gK}{(1 - g)} \]  \hspace{1cm} (4.9)

Where: \( g \) is the average tax rate and the other symbols remain as previously defined. Equation (4.9) shows that the value of the post-payment Gini index is a function of pre-payment Gini, average tax rate and payment progressivity, K. But average payment is independent of progressivity, and therefore can change without changing the level of progressivity. The elasticity of post-payment income inequality with respect to average tax rate is given by (Kakwani, 1977):

\[ \eta_z = \frac{gK}{(1 - g)^2 G_{x-T}} \]  \hspace{1cm} (4.10)

Equation (4.10) suggests that for a given payment elasticity and payment progressivity, the higher the average proportion of income devoted to healthcare payment, (i.e., average payment rate, \( g \)) the lower the lower the level of inequality in the post-payment income. The implication is that a policy of healthcare payment that reduces everyone’s income may still be consistent with a policy of redistribution.

The elasticity of post-payment inequality with respect to K is given by

\[ \eta_K = -\frac{gK}{(1 - g)G_{x-T}} \]  \hspace{1cm} (4.11)

Equation (4.11) similarly suggests that for a given level of payment rate, the higher the progressivity of healthcare payment on prepayment income, the lower the level of inequality in the post-payment income, and vice versa. This shows that increases in K leads to a fall in the post-payment income inequality. However, the ratio of the two elasticities \( \eta_t, \eta_K \) is greater than 1 in absolute terms indicating that post-payment income inequality is more sensitive to average payment rate than to payment progressivity.

From the foregoing, it follows that
\[(G_x - G_{x-T}) = K \text{ iff } g = 0.5\]  

(4.12)

4.7 Reynolds-Smolensky (RS) Index of Redistribution

Redistributive effect may be conceptualized as the equalizing or disequalizing effects associated with a transition between the pre- to post-payment periods. In this sense, a progressive payment system is essentially redistributive. As Lambert (1993: 180-181) notes, “a progressive income tax can be redistributive in its own right, regardless of what is done with the revenue” (See also Podder, 1993: 53). Jakobsson (1976) also proved this theoretical result.

A measure of redistributive effect that has gained a substantial attention in literature is the Reynolds-Smolensky index (Reynolds and Smolensky 1977). They defined redistributive effect as

\[\pi^{rs} = 2 \int_0^1 \left[ L_{x-\pi(p)} - L_{x(p)} \right] d\phi = G_x - C_{x-T}\]  

(4.13)

Where \(\pi^{rs}\) = Reynolds-Smolensky redistribution index defined as twice the area between the Lorenz curve for pre-payment incomes \(L_x(p)\) and the concentration curve for post payment income \(L_{x-T(p)}\). \(G_x\) is the prepayment Gini index while \(C_{x-T}\) is the post payment concentration index. This is shown in fig 4.1.

The index is positive if the pre-payment Lorenz distribution curve lies below the concentration curve for post payment concentration curve. In this case payment reduces income inequality on prepayment income distribution. If the pre-payment Lorenz lies above the post payment concentration curve, then, \(\pi^{rs}\) is negative and in this case healthcare financing system worsens inequality on prepayment distribution.
Fig 4.1 Lorenz Curve of Prepayment Income & Concentration of Payment

The Kakwani progressivity index and Reynolds-Smolensky redistributive index are linked by the following mathematical relationship:

\[ \pi^{rs} = \frac{g}{1-g} K \]  

(4.14)

Where g = is again the average tax rate or proportion of prepayment income taken up healthcare payment.

Equation (4.14) suggests that the Reynolds-Smolensky redistribution index is an increasing function of Kakwani index of progressivity. The higher the proportion of income individuals and households spend on healthcare the more income equalizing will be the payment system and the higher the value of the K the more redistributive the policy is.

4.8 Decomposing the Redistributive Effects

Reynolds-Smolensky measure of redistribution indicates that total redistribution is made of two components – average payment rates and the index of progressivity. However, there is a fundamental assumption underlying the Reynolds-Smolensky index. It assumes either that horizontal equity (i.e. equal treatment of equals) does not exist or, based on
normative stance, that horizontal equity is not a separate concept of equity. Everyone at a particular income level is assumed to contribute the same amount to the financing of healthcare system. But this can hardly be said to be a plausible assumption since in practice households at the same level of income may vary widely in their healthcare payments due to the stochastic nature of illness.

Horizontal inequity is more likely to be the norm than the exception. But this is not to say that there is any unanimity about the place of horizontal equity in the distributional spectrum. While some authors would readily agree that horizontal equity is a concept of equity with content, some insist that horizontal equity is not an independent evaluative principle. One of the most consistent opponents of this principle is Kaplow. In furtherance of this opposition he writes: “I have suggested that there is in fact no good argument for viewing horizontal equity as an independent evaluative principle and that pursuing horizontal equity is in conflict with the very core of welfare economics, the Pareto principle” (Kaplow, 2001:91). Obviously from the neoclassical economic point of view this argument is valid. Horizontal equity considerations may reduce the aggregate social welfare if this welfare is considered from the point of view of commodities. It was shown a short while ago that reduction in average income is consistent with a decrease in overall inequality. But it is equally true that society might want to forego higher average income to attain more equality, to reduce crime for instance. This should also be deemed as increasing social welfare.

Another ground of objection against the principle of horizontal equity is that it considers groups as insular. One the other hand, if a single relevant group were to be defined, then it would be inclusive of everyone in the population, so that horizontal equity becomes a concept without content. It is well known, however, that people generally evaluate changes in their welfare with that of those of relevant equals or those they consider themselves to be in equal status with prior to the change. This is a key issue in the theory of relative deprivation. It is also is clearly demonstrated in the famous observation by Runciman (1966) about the postwar British welfare state: “the reference groups of the recipients of welfare were virtually bound to remain within the broadly delimited area of
potential fellow-beneficiaries. It was anomalies within this area which were the focus of successive grievances, not the relative prosperity of people not obviously comparable” (p. 71). Horizontal inequity in certain respects, look more damaging than vertical inequity and thus deserves no less attention than vertical inequity. It, therefore seems reasonable that the pursuit of horizontal equity in healthcare financing is important because people who have suffered financial and welfare set back on account of ill-health often become psychologically alienated (Durant and Christian 1990).

Furthermore, there is the problem of re-ranking that is not considered by the Reynolds – Smolensky index. In many developing countries it is a common experience to see that catastrophic healthcare payments may push an average income family below the poverty line, behind families that it ranked higher than before illness occurred (See for example, Sauerborn et al 1996; World Bank, 2001).

This re-ranking effect leads to people having different ranks in the pre- and post-payment periods (Felstein 1976, Plotnick 1981). Yet these authors consider re-ranking as part of horizontal inequity. But subsequent authors have argues against this by treating re-ranking as referring to treatment of unequals rather that of equals (Kaplow 1989, Aronson et al., 1994, Duclos et al 2003) Thus, it is included as an independent component in the decompositions to follow. Unfortunately, the Reynolds-Smolensky index assumes there is no horizontal inequity and no re-ranking occurs in the transition from pre- to post-payment income periods. This makes the model unsuitable for the purpose of this research. The two models that will be employed in the decomposition analysis are the Aronson, Johnson, and Lambert (1994) decomposition model and the Duclos, Jalbert and Araar (2003) model. Both models contain the essential components presented above. Their results are therefore comparable.

This section on methodology has described the main features of the tools of analysis to be used in subsequent chapters. The details of these will be presented at the relevant sections. The next chapter examines the details of the survey method that generated the data for this analysis. The survey design, the survey instrument, the details of the
Chapter 5

The Survey

5.1 Introduction

The key consideration has been to provide the basic analytical infrastructure for the study. From now, the focus turns to the empirical results. This chapter provides an entry into the empirical sections by describing the data for the study. The data used for this study is primary data generated from a cross-sectional survey. It was planned and executed between May 2003 and August 2004.

The drafting of the instrument started in June 2003 and went through three stages of redrafting and refinement before it was approved by the Research Ethics Committee of the University of Cape Town. After draft zero was prepared, experts at the University of Nigeria, Nsukka were consulted. The corrections and suggestions from them were incorporated into the next draft. This draft was approved by my supervisors at the University of Cape Town with further corrections and suggestions. The third draft was sent to the research consultants of the African Economic Research Council (Nairobi). The corrections and suggestions from the research consultant were unified with those from my supervisors before the final copy was prepared. During this period also, the permission of the Enugu State Government was requested for to enable the conduct of survey in the State. This permission was eventually granted in a letter of State government approval signed by the Permanent Secretary, Enugu State Ministry of Health and dated February 3, 2004. The final draft was sent to the Research Ethics Committee of the University of Cape Town. The approval was granted on February 11, 2004.

With these approvals and with financial assistance from AERC, the logistics for the survey was prepared. Since the Federal Office of Statistics and the National Population Commission were already involved in different forms of survey, it was considered convenient to use their facilities. Their assistance in this respect was generously given. The area mappings of the Enumeration Areas, the numbering of the households and stratification between rural and urban areas were made available. This greatly facilitated the task that would have otherwise been involved in fresh mappings and listing of households. The most current sampling frame being used by FOS was therefore used for the survey. Their expertise and familiarity with the entire state topography and survey experience were invaluable to the survey team in spite of our relatively good knowledge of the State. This will be further highlighted when we discuss the field work and training. The actual field survey lasted from April to August 2004.
5.2 Objectives of the Survey

A number of factors motivated and guided the process that generated the data. A close examination of existing data in Nigeria showed they did not have some of the key variables that are necessary to undertake this study, in particularly, those relating to health. It was considered absolutely necessary to generate a set of fresh data of high quality for the study. Since the cost of resources required to generate a nationally representative data would be rather prohibitive, the survey had to focus on a state – Enugu state - with sizeable population (3 million) and which mirrors all the developmental features of the Nigerian environment. It has a dual economy with both metropolitan industrial and rural agrarian economy. This was to make the results obtained from the study as much as possible generalizable.

The following thematic steps will guide this presentation: Fieldwork methodology, the study area, sampling frame and data collection process, survey instrument, sample size, weighting, training of interviewers and pre-test of instrument, data processing and management, conceptual issue about income, expenditure, measuring household expenditure, definition of variables and transformations on variables.

5.3 Fieldwork Methodology

The overriding need for high quality data that was representative of the population was the main motivating factor in planning and executing the fieldwork survey. This is because for an empirical econometric study to achieve its purpose, it is clearly important that the underlying data-generating process has to be captured as accurately as possible in the sample data. The accurate description of such a process (that is the goal of econometric model-building) depends largely on the extent the sample data set employed mirrors the behaviour of the underlying economic and social process.
5.4 Study Area

A nationally representative sample would be the ideal for a study as this. But because of the problems already mentioned above, the survey was carried out in Enugu State of Nigeria. It was decided that for effectiveness of coverage and sharper focus, the data collection would be limited to one state. This, it was believed, would improve the accuracy and quality of the data making it more representative of the underlying population. While such a decision raised the question of generalizability of the results to the rest of the country, it was felt that the gain in precision more than compensates for the loss in generalizability of results. More so because as indicated in Chapter 2, more of healthcare policies are made at the state level than at the federal level and most states in Nigeria have the same epidemiological experience, healthcare financing policies and socioeconomic problems.

The choice of Enugu state as the study site was not random. It was rather based on the researcher's relative prior information about the state population. Such prior information often enhance the accuracy of the data and the econometric estimates since it can assist the researcher in designing an effective method that strengthens the efficiency of the statistical inference about the population under study (Deaton 1997). The choice of the study site is predicated on the following considerations:

(a) Cost: it is considered that concentrating on a state will provide a sharper focus, effective supervision of the process and limit the cost of expanding the study to several states

(b) With an estimated population of 3 million people and diverse sub-ethnic groups, Enugu state to a large extent mirrors the problems of development in Nigeria as a whole. This implies that the results could easily be generalized.

(c) The area is relatively well known to the researcher. The cultures and sub-cultures of the state as well as the topography of the area are relatively well known.
(d) Recent studies show that the research of this kind that focuses on small area analysis and samples more intensely tends to yield better results than those that attempt a large coverage (Diaz, 2002)

Fig 5.1 Survey Area: Enugu State

Enugu state is one of the 36 states of Nigeria. It is located in the South Eastern part of the country and lies approximately between latitude $5^\circ\,55'$ and $7^\circ\,08'$ north of the equator and $6^\circ\,55'$ and $7^\circ\,08'$ east of the Greenwich Meridian. By the end of the 2003 the state had a projected population of about 3 million people (extrapolated from the National Census of 1991). The population density is about 273 persons per $km^2$. Enugu state
comprises 17 local government areas (LGAs). It has three sub-ethnic groups: Udi, Nkanu and Nsukka, corresponding to the three senatorial districts of the state. Its metropolitan city is Enugu (the coal city) which also serves as the administrative capital of the state and used to be the administrative capital of the whole of the South Eastern Region of Nigeria, comprising nine states in the present 36 states structure. It has also a number of semi-urban centres as well as numerous rural village communities. Majority of the urban dwellers are civil servants, coal mine workers, traders, transporters and artisans. Most of the rural dwellers are subsistence farmers and petty traders. Industrialization is growing.

5.5 Coverage

This survey covers both the urban and rural areas and all the 17 local government areas of Enugu state. A total of 100 enumeration areas (EAs) constituting the primary sampling units (PSUs) were selected and covered in the survey. Information was collected on the following: Education, occupation, health and healthcare expenditure, gender, household assets, and household expenditure etc. The survey spanned the period April to August, 2004.

5.6 Sample Design

The survey used the standard two-stage stratified self-weighted design to generate population representative sample. This implies that each household in the state has equal probability of being selected into the sample. The current enumeration areas (EAs) used by the Federal Office of Statistics (FOS) for its regular household surveys were used to construct the sample frame. There are currently about 212,079 non-overlapping enumeration areas nationwide and about 5000 of such EAs demarcated in Enugu State. These EAs were drawn along village clusters in rural areas and quarters in urban areas. The EAs were stratified into rural and urban strata. In this context an urban EA was defined as one within an area comprising more than 20000 of population. A rural EA on the other hand was an EA in a locality of less than 20000 of population.
In addition, the EAs were stratified into three senatorial districts that make up the state. These EAs formed the primary sampling units (PSUs). Sampling was carried out separately in each stratum.

5.7 Sample Size

In a household survey framework, sample sizes usually vary depending on the purpose and size of the population. While it is often possible to work out the optimal sample size from the confidence interval of parameter estimates that a researcher is willing to accept, in practice more practical considerations tend to dominate the criteria for choice of a sample size (Hoddinot, 1994). Deaton (1998) suggests that in national population surveys, sample to population ratio of 1:500 or even 1:2500 are admissible.

However, there are standard techniques for determining the optimal sample size. The more conventional method (with different variants) is given by the formula

\[ n = \frac{z^2 \sigma^2}{\mu^2} \]  

(5.1)

Where \( n \) = required sample size  
\( z \) = level of significance  
\( \sigma \) = standard deviation  
\( \mu \) = population mean

The alternative to this is the Taro Yamane formula (Cited in Otto, 2004) that may be specified as:

\[ n = \frac{N}{1 + N(u)^2} \]  

(5.2)

For \( N \) = population size and \( u \) = margin of error

The Taro Yamane formula has an advantage of being simpler to apply. It was therefore used in this case to calculate our optimal sample to population ratio. In standard statistical test of significance it is usual to allow for five or one percent margin of error. This survey
allowed for two and a half percent margin of error in its calculating the optimal sample size (i.e., \( u = 2.5\% \)). Noting that the population size (N) in this case is about 3 million the estimate sample size is calculated as:

\[
n = \frac{3000000}{1 + 3000000(0.025)^2} = 1599
\]

This was approximated to 1500. Thus the number of households interviewed during the survey was 1500.

### 5.8 Selection of the EAs

Sufficient care was taken to avoid over or under sampling of any stratum or EA that would warrant weighting the population to make up for shortfalls or over-sampling of any group. This objective was to have a self-weighted sample in respect of both the strata and the clusters. This was meticulously adhered to knowing full well that the implications for the point estimates of parameters and statistical inference of incorrectly weighted sample. This objective was achieved in the final result. There is particularly a balance in the number of rural and urban EAs that were surveyed. Thus the only weighting required in the computations is the household weights. Altogether 19 EAs were sampled from the urban and 81 from the rural areas. This largely reflects the ratio of rural to urban population of the state.

The number of EAs to be selected \((a)\) was therefore given by

\[
a = \frac{k}{r}
\]

Where \( k = \) the fixed number of households to be sampled (in this case 1500)
\( r_i = \) the number of households to be selected from each EA.

From the literature, the optimal number of households to be selected from each EA [in trading off the effects of stratification and clustering (Deaton 1997)] is variable, but
usually lies between 15 and 25. This study in preference for wider coverage of the EAs but a smaller number of households, selected 15 households from each EA. The reason for this is because in the study area, each EA was on the average made up of approximately 20 households, such that sampling more than 15 households in each EA may imply complete enumeration of the households. However, two additional households (to the 15) per EA were provided for as replacements in the event that some of the original 15 households could not be interviewed. The value of \( r \) in this case is, therefore, 15. The value of \( a \) was then computed as:

\[
a = \frac{1500}{15} = 100\text{EAs}
\]

(5.5)

Thus, in the first stage of sampling, 100 EAs were selected randomly using systematic sampling probabilities proportional to the estimated number of households [Probability Proportional to Size (PPS)].

In each stratum, the probability of selecting an EA is given by

\[
P_{ii} = \frac{b_i H_i}{\sum H}
\]

Where \( P_{ii} \) = the probability of sampling ith EA in the first stage.

\[
b = \text{allocated number of EAs to be selected from the stratum}
\]

\[
H_i = \text{the total number of households in the ith EA}
\]

\[
H = \text{total number of households in the stratum.}
\]

At the second stage, a complete household listing was carried out. For this purpose, the Federal Office of Statistics staff provided the required assistance in respect of identification of the EAs selected in the first stage and to indicate the buildings within the EA. This list of households provided the sampling frame for the second stage of the selection.
At the second stage, 15 households were selected from each EA in all the strata. The probability of selecting a household at this second stage was given as follows:

\[ p_{2i} = \frac{m}{M_i} \text{ for each EA} \]

Where \( p_{2i} \) = probability of sampling a household in the second stage

\( m \) = number of households selected from each EA

\( M_i \) = total number of households in the ith EA

Thus, the product of the probability of the first and second stages gave the overall probability of selecting any given household. This may be specified as:

\[ F_i = p_u p_{2i} = b \frac{\sum H_i}{H} \cdot \frac{m}{M} \]

(5.6)

It is, therefore assumed that every type of household in the population is proportionately represented in the sample data. The implications of survey design issues about weighting, clustering and stratification and their implication for point estimates and statistical inference will be discussed under the subsections on weighting, stratification and clustering.

5.9 Description of the Questionnaire

The main instrument for data generation is the household questionnaire (shown in the appendix). The questionnaire basically attempts to generate information on the relevant variables of the study, specifically on health and household expenditure and consumption, and other socioeconomic and demographic variables such as education, employment, occupation redistribution effects of healthcare financing.

Sections A and B of the questionnaire seek general information about the household. Section C concentrates on educational attributes of the household. Section D seeks information on the health variables of the household. In particular, it seeks information
on indicators of health status such as, self-assessed health (SAH), health status index, healthcare utilization, and the presence of chronic illness. However, the central concern were health status index and self assessed health which has been shown to be a good predictor of mortality and the on-set of disability.

To obtain SAH, question D.15 asks: How would you generally rate your state of health now? The response options are in five categories: Excellent, Very good, Good, Fair, and Poor. This is considered a subjective measure of health. It is simple but it has been shown to be a powerful indication of true health status and has been used in this formulation in several health surveys. Although the responses are categorical, it can be converted to a continuous scale. The usual practice is to dichotomize the responses by choosing an arbitrary cut-off point between those whose health is good and those whose health is poor.

Questions E14 to E16 are based on the health status index. It has a pictorial index in the form of a calibrated thermometer scaled from zero (0) to hundred (100). It was adapted from the EuroQol group’s EQ-5D Virtual Analogue. Hundred, (100) represents the best imaginable state of health while zero represents near-death state or the worst imaginable health. Only adults of 16 years and above were eligible to respond to this question and to rate their health status on this 100-point scale. The reason for its inclusion is to ascertain whether there is convergence between the two SAH indices that are often used in literature. The descriptive part of the EQ-5D was also included though the main interest for the present analysis is the Virtual analogue scale which we believe speaks more to our population than the descriptive part. Other salient features of the health questionnaire are information on healthcare utilization, healthcare expenditure and healthcare seeking behaviour. These features are not included in most household surveys.

Section D of the questionnaire seeks information on the occupational variable of households'. Occupation, like education and income, has often been used as an index of the household's socio-economic status. Most of the other sections of the questionnaire deal with issues related to households' expenditure and consumption. In particular
information on food, water, energy consumption, household durables and other non-food expenses were also obtained. Question on the income variable was excluded altogether because of the heavy biases usually associated with collecting information on. Usually most household are unable to tell how much they earn in a given period because of the wide variability in their source of income. There is plenty of empirical literature suggesting that household expenditure is a better proxy for household permanent income than household transitory income captured in cross section surveys. This study will therefore be based on household expenditure rather than income.

5.10 Training and Fieldwork

Ten teams, each comprising two interviewers were used for the actual survey. There were also two field supervisors and two editors that went round to crosscheck the work of the teams and carried out on the spot editing of their work although one of the interviewers still doubled also as field editor. The duties of the supervisors included observing the interviews and checking that the quality of work done by the interviewers conform to instructions. The researcher coordinated the field activities and provided logistics.

The field staffs were trained for a period of one week on the technicalities of the interview process and how to obey the skip instructions. The training was facilitated by the fact that only field workers who had had several years of experience in field work were engaged for this survey. They knew each EA by the name of the number one household in the EA. All the field workers participated in the Core Welfare Indicators Questionnaire Survey (CWIQ) - a collaborative effort of the World Bank, UNDP, UNICEF and ILO – that was conducted in the state between July and August 2002. They had also participated in several other surveys conducted by other agencies including the Federal Office of Statistics surveys. For the CWIQ survey, they were all given a 12-day training workshop. Furthermore, the interviewers were familiar with the shading method adopted in designing the questionnaire for the survey. Making use of experienced fieldworkers ensured that the survey was properly done and with minimal hitch.
At the end of the training program, a pilot study was conducted in order to assess and improve the validity of the instrument. Again the experience of the fieldworkers was brought bear in this regard. The results from the pilot study, which was conducted with 30 households, suggested the reframing and reordering of some of the questions that had potentials to generate confusion in the minds of the respondents. It also facilitated the complete mastery of the process of questionnaire administration and ensured that the prospective respondents responded appropriately to the questions. After the pilot exercise, the questionnaire was finalized to accommodate the pre-test experience. The main fieldwork then began. Call back interviews were made as often as necessary.

5.11 Ensuring Data Quality

To enhance the quality of data, data collected by interviewers were randomly crosschecked by supervisors. Errors identified were corrected during the call back visit to households. Data entries were planned to proceed simultaneously with the data collection. This was carried on in the initial phase of the survey. However, it was found later that the entry could not keep pace with returns from the field. The data entry process proved to be slower than was anticipated. However, since the returns were considered to be very good and editing was being carried on at the same time, it was considered needless keeping entry at pace with the returns from the field. Data entry was finished a month after the conclusion of the field work.

To facilitate identification and cross-checking of data, households were linked to their EAs and stratum through household identification number. Each household was given a unique identification number constructed from the combination of household number, EA number and stratum number. This was the first thing the interviewer was required to fill in. This system of numbering also facilitated the merging of files during the data entry and analysis phases.
5.12 Work Plan

As stated earlier, ten teams of two field interviewers each, two field editors and two editors were used to conduct the survey. Each team was required to complete a maximum of ten questionnaires a day. This was to ensure that they did good quality work. The teams were assigned different EAs and each was expected to cover 10 EAs altogether, which is, 150 households selected randomly from the list of households in the EA. The teams did not work everyday. In fact, there were some weeks of heavy rain when they did not work at all. But the last team submitted its last return in the first week of August 2004.

5.13 Data Processing

The questionnaire was clearly coded to make entries easy. Three competent data processing assistants were used for data entry, although one had to drop out due to other pressing needs. They worked long hours for weeks to complete the screening. The basic software for this exercise was the Excel Spreadsheet. By the second week of September 2004 the data entry was successfully concluded. Data cleaning then started. The rate of return was hundred percent. Exactly 1500 households were interviewed from the 100 EAs, 15 households from each EA and the strata balance was maintained. Information was collected altogether 5814 individuals living in these households. In very few cases (about 12 out of 5814) information was not available on one or two variables. To make up for these variables in these observations, the ‘impute’ in Stata was used to make up the appropriate values of the variable for these individuals. The Stata *impute* command uses the scores of the candidates based on other variables to compute what the individual would have scored on this particular variable in which ‘she’ is missing.
5.14 Data and Variable Definitions

5.14.1 Measurement Problems

Recall Period: In terms of the period of coverage, most of the questions were focused on household health and economic decisions within the four weeks 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 empirical studies have shown that household expenditures and consumption tend to diminish the longer the recall period (Scott and Amenuvegbe 1990; Deaton 1998:25). As much as possible, care was taken in the field to check the responses. This method was adopted after considering the pros and cons in the normative issue about recall period and the implied trade off between length of recall period and accuracy of information. While the counter argument based on seasonal variations in household expenses is valid the accuracy of recall was considered a more overriding objective. Household healthcare expenditures may vary from month to month, from season to season. Illness being a random phenomenon, a household that spends on healthcare this month need not necessarily spend in the next. Nevertheless, it was considered that the advantages of recalling the exact amounts spent or consume (particularly for a society where records are not often kept) outweighs the counter argument for stretching the recall period. As Wagstaff et (1999) point out, while long recall periods provide a window to pick up variations in household spending, it also frequently leads to recall error which could be more severe. And it is hazardous attempting to convert one to the other.

5.14.2 Estimating Household 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 welfare income than short run income (Stevenson et al
There are conceptual and empirical grounds for preferring household expenditure as a measure of household welfare than income or other available measures of welfare. The basic argument is that household consumption does not generally fluctuate with fluctuations in transitory income. Consumption expenditure indicates not only the extent of household command over resources based on its current income, but also its level of access to credit, or savings from past incomes (Hentschel and Lanjouw, 1996). When transitory income is zero, the household can dis-save or borrow to consume. Furthermore, while it is known that household expenditures are often limited to a narrow range of commodities households' total income (particularly in a developing country like Nigeria with predominant informal sector), come from several sources. Consumption expenditure is, therefore, often regarded by researchers as better reflection long-run welfare than income. 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 measure on the agricultural season. Rural households tend to consume more during harvest periods than during the planting season.

In order to minimize these difficulties, several strategies were adopted. Firstly, 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. Secondly, 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, etc.

5.14.3 Accounting for Multistage Sampling Design:

(a) 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 behaviour 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. Since observations are selected through a random process, they may have different probabilities of being selected into the sample. To make up for this, weights are attached to observation in the inverse of the probability of their being selected into the sample. Thus, a weight of \( w_i \) suggests that \( i \) is representing \( w_i \) of the population from which the samples was selected.

An estimation problem arises if the method of weighting the sample does not mirror the population distribution in respect of the phenomenon being studied. In the first place it is obvious that such improper sample weighting will generate miss-leading or biased point estimate of the parameter which will never converge to the through population parameter. Thus wrong weighting may lead to biased parameter estimate. In the second place, the weights will affect the estimated standard errors and therefore impair statistical inference from sample to population.

In generating the sample used for this study, particular attention was paid to the issue of representativeness of the sample. This objective was largely achieved in designing the survey (and in actual sampling) as was indicated above. There was therefore, little or no need for attaching weights to reflect the population weights, as the sample was already
self-weighted. However, since households differed in size, it was still necessary to weight sampled households according to size. The weighting variable in our estimation is the household size.

This survey interviewed exactly 1500 households with 5814 individuals. However, in the process of cleaning the data three households were dropped. One household was dropped because of inconsistent information. The other two were dropped because their expenditures were far off the next nearest incomes. They were therefore deemed to be outliers whose inclusion could exact disproportionate impact on the final outcome of the exercise. Altogether, therefore, the final sample size was 1497 instead of 1500 that was collected.

(b) Stratification

While weighting affects both the point estimate of the parameter, the neglect of the stratification and clustering effects of the survey design affect the estimated standard errors and therefore, affect inference, although the point estimates are not affected. Stratification, (say into urban and rural strata as in this survey) will affect the estimated standard errors of the point estimates if the presumed independence between urban and rural strata does not hold. This survey was designed to achieve this independence. In order to do this, it was noted earlier that sampling was done in each stratum independently with the boundaries between the urban and rural strata fixed in advance (see Stata User's Guide) However, in the estimations the strata variable was specified.

(c) Clustering

The EAs represent the clusters in this survey. Clustering has been known produce an effect on the estimated standard errors if the clustering effect is not taken into account in the estimation process. This is because observations in the same cluster are not really independent. Households in the same cluster are likely to share many things in common and behave alike. Thu, if one were to use estimators that assume this independence when
in fact the observations are not independent, then this will affect the standard errors. The standard errors will be too small (Stata User's Guide), implying that one may likely accept an estimated parameter to be statistically significant when in actual fact it is not. An estimator that takes into account this dependence in cluster observations produces 'more honest' standard errors (Stata, User's Guide). To take care of this dependence in the estimations, estimators that do not assume the independence of observations in clusters were employed.

5.14.4 Transformations on Data

There is also the problem of price variation between the rural and the urban area which had to be accounted for. It is generally acknowledged that the cost of living is usually higher in the urban cities than in the rural area. This suggests that equal expenditure in the city and rural area would fetch different amounts of welfare. To take care of these price differences prices were collected on a basket of the commonest 20 consumption items used for the constructing the national consumer price index (CPI) in each of rural and urban areas. The Laspeyres' index was then used to deflate the nominal values of expenditures of the rural EAs to bring them on par with the urban cost of living. This implied multiplying the rural nominal expenditures by a factor of 1.255

To obtain household total expenditure, the household expenditure on various goods and services (including all healthcare related expenditures) were aggregated. From the total expenditure, the values of the per capita expenditure were obtained by simply dividing these values by the household size. This gave the prepayment income variable. But other transformations were carried out to obtain economies of scale of households. The basic concern in welfare studies of this nature is not necessarily with income or expenditure as such, since the household gross expenditure had to be converted into utilities or welfare. The household healthcare expenditure and the net household expenditure were similarly converted into utilities under different assumptions about intra-household resource allocation and size economies as shall be discussed shortly.
To compute household expenditure on healthcare services the sum total of all direct expenditures relating to the production of treatment for an illness episode, including the cost of drugs, cards, consultation fees, transportation, and so on, in the four weeks preceding the interview were used to generate the healthcare expenditure variable. Zero was recorded for a household that did not spend any amount on healthcare services.

The post-payment income variable was obtained simply by subtracting the healthcare cost from the gross expenditure. A fortiori, for households that did not make any expenses on healthcare within the period their net income or post-payment income was the same as their gross income.

5.14.5 Equivalence Scales and Scale Economies

Since the data for our study are household survey data, it is important to make transparent the assumptions underlying the data-generating process and the current developments in literature in this respect. The complex structure of households, their demographic composition, decision-making processes, and division of labour and income earning activities all affect the dynamics of intra household distribution of resources (Mattila-Wiro, 1999) and consequently the welfare of the members. The choice of indicator and method of estimating household welfare need, therefore, take these considerations on board especially today when the inadequacies of the neoclassical unitary (the consumer as an individual) model of the household has become apparent. Chiappori (1992) notes that the neoclassical model of the household misleads policy by its concentration on inter-household distribution while neglecting intra-household distribution of welfare when the latter is central to individual welfare.

In dealing with the effects of intra-household distribution of resources on the individual welfare there are two key issues that come into question: the intra household distributional impact of the demographic composition and the existence or otherwise of
scale economies in the household. Neoclassical economics does not consider these issues. It simply treats the consumer as an individual and proceeds to estimate welfare on per capita basis. Recent findings (Buhmann et al. 1988, Coulter et al. 1992, Lanjouw and Ravallion, 1995), however, suggest that welfare measures and inequality are sensitive to the treatment of household structure and size and, therefore, should be taken into account in welfare studies. This is so for data from developing countries but particularly so for African households where large and extended households predominate and intra household relationship is built upon deference and cooperation in order to get more out of available resources. This suggests the likely presence of scale economies in consumption.

The presence of large households also indicates the likely predominance of children in the household composition. This further implies, assuming that children consume less of everything than adults, that both household composition and size can effectively serve as expenditure deflators to increase the welfare of the household members. Our data makes it possible for us to include both consideration and to check the stability of our models’ decomposition results to different assumptions about intra-household resource distribution and different size elasticities.

A general formula for deflating for household variable structure or differences in household needs and household economy of size may be specified as:

\[ S = (n_a + \rho n_c)^\epsilon \]  \hspace{1cm} (5.7)

Where \( n_a \) and \( n_c \) represent the number of adults and children respectively. The parameters \( \rho \) and \( \epsilon \) represent household need elasticity or weights and household size economy respectively and may be assumed to vary between 0 and 1. As \( \rho \neq 1 \), it implies that children consume as much as adults. Similarly as \( \epsilon \neq 1 \), it implies there is no economy of scale. The adjusted household welfare may, therefore, be specified as:

\[ W_i = Y_i/(n_a + \rho n_a)^\epsilon \]  \hspace{1cm} (5.8)
Where $Y_i$ is the gross expenditure of the $i$th household. While the value of the scale economy parameter has often been fixed arbitrarily, this value can also be estimated by the coefficient of log of gross expenditure on log of household size. Regressing log of per capita expenditure on log of household size using our data gave a coefficient of $-0.32$ implying that that economy of scale is about 0.68. This means that 10 can live as cheaply as 7 (Deaton 1997:243). The values of $e=0.7$, 0.3 and 1 are therefore used to check the stability of the decomposition of redistribution effects of healthcare financing under two models: the Aronson-Johnson-Lambert (AJL) model and the Duclos-Jalbert-Araar (DJA) model. For convenience we have used the following notation to indicate household composition and economy of scale:

Eqvs. 1-1 $e=1$ implies equal or per capita distribution of resources between adults and children without any economy of scale. Eqvs. 1-0.5 $e=0.7$ implies that an adult receives a full adult allocation of resources, a child receives 0.5 and $e=0.7$ implying that 10 people living together can obtain equal welfare as seven people living separately from the same amount of resources.

5.14.6 Data Aggregation

It should be noted that in much of the analysis to follow the unit of analysis is the individual representative of the household. In other words there were 1500 observations in the study, although after cleaning three of these were dropped so that the number of observations in the study was 1497. This was generally the case particularly for those parts of the analysis requiring information on household expenditure, in particular the aspects dealing with income redistributive effects of the healthcare financing. For example, to obtain the level of welfare of the representative individual in a household, the total household expenditure was shared either on per capita basis or based on some intra-household sharing formula (equivalent scales) as the case may be and then imputed to the individual. The health expenditure and net expenditure were computed on similar basis. The observations were then weighted by the household size. However, in analysis
involving health distribution, only observations on individuals of 16 years and above were used. There were about 3673 such observations.

5.15 Summary

Chapter 5 has tried to provide a comprehensive account of how the data for this analysis was generated. It details the objectives, design, planning and execution of the field survey and how far the objectives were achieved. It went further to provide information on precautions that were taken to avoid biases and inaccuracies that are often encountered in data generation. The chapter also provides information on some conceptual problems about some variables and how these are treated in this study and then describes some of the transformations that were made on the data. The next and subsequent sections will present and analyze the results from the data. It will begin with a descriptive analysis of health distribution in Nigeria. This is intended to provide an entry point into a higher level analysis of inequalities in health distribution which will follow.
Chapter 6

A Distributional Analysis of Health in Nigeria

6.1 Introduction

The Constitution of the World Health Organization provides what has been generally considered a complete definition of health. It defines health as “a state of complete physical, mental and social well-being” (WHO, 1948). The distribution of health is an important argument in the social welfare function as it is a precondition for the enjoyment of other goods and services. It seems, therefore, appropriate that the analysis of the distribution of health in a population must precede the analysis of the income distributional impact of the system of healthcare financing in that population since the purpose of the latter is to improve health and health distribution. This chapter looks at the empirical distribution of health in the study population. It aims to provide the background for analyzing the implications of the policies that determine who bears the cost of ill-health. The subset of observations used in analysis in this section is adults, implying in this case, persons above fifteen years of age. Children of 15 years and below are excluded because we felt that most of them could not reasonably provide responses to this part of the questionnaire that require a high level of personal judgment.

The chapter begins by introducing previous findings on the social determinants of inequalities in health distribution. It then proceeds to relate the outcomes of health measurement using two key measures of health that have been developed in the health measurement literature, namely, a discrete measure of self assessed health (SAH) and a continuous measure of health (Health Status Index – HSI), based on the EQ-5D described
above. While SAH has been generally used in survey data to collect information on the health of individuals, HSI is a rather recent development in field surveys. The chapter assesses health distribution by using both the Gini and the concentration indices, introducing in the process an ethically flexible measure of the policy-makers aversion to uneven distribution of health. A decomposition of the Gini is performed in order to disaggregate and unpack the total inequality in health into pure inequality in health and socioeconomic contributions to uneven health distribution. It then uses the Ordered Logit to assess the socioeconomic determinants of health in the Nigerian population.

6.2 SAH and HSI as Health Indicators

In most general surveys, the use of self assessed health questions to indicate the health level of the individual is common. Respondents are usually required to choose from four or five health categories (poor, fair, good, very good and excellent) the option that best describes their health state generally. In econometric analysis the use of ordinary least squares method to run such regression models has been found to be unsatisfactory. Some of the short-comings include the fact that the probabilities are not constrained within the 0-1 probability range. The problem of converting such categorical variables into a continuous measure of health arises, as is often the case, when computing health inequality indices. How does the researcher impose cardinality on ordinal measures of health? The response has sometimes been to merely dichotomize the variable into two groups, healthy and non-healthy. But this approach leads to loss of information about the differences within each category. This is apart from the fact that such dichotomization might be deemed arbitrary since their may be no clear criterion for deciding the cut-off point between the two groups.

To obviate this problem, Wagstaff and van Doorslaer (1994) developed a method of imposing cardinality on the categories. Their method assumes that underlying the categories of SAH is a continuous latent variable with standard lognormal distribution. But it is not certain that health in every population follows the lognormal distribution. However, this method was applied by van Doorslaer et al (1997) in their study of income-
related inequalities in industrialized countries. They found health to follow the income
gradient in all the countries. Subsequent studies in the industrialized countries including
Gerdtham et al (1999), and Humphries and van Doorslaer (2000) confirmed the earlier
findings based on Wagstaff van Doorslaer method of cardinalizing SAH categories.
Crossley and Kennedy (2002) using Australian data raised doubts about the stability of
responses to the SAH categories.

However, van Doorslaer and Jones (2003) developed another method of imposing
cardinalization on the SAH categories that relies on the interval regression. The use of
interval regression (or group regression) presupposes knowledge of what the cut-point of
the various categories are. This would require some external information about the actual
cut-off points of the SAH categories. To obtain this information they combined the use of
SAH categories with the McMaster Health Utility Index Mark III (HUI). By comparing
the empirical distribution of HUI with that of SAH from the same population they ‘read-
off’ the end values of the SAH. Again, this study used data from an industrialized
country, Canada, which was included in all the three previous studies.

The main claim by van Doorslaer and Jones (2003) is that interval regression which maps
the empirical distribution function of HUI into SAH leads to greater precision in
estimation of parameters. However, as they noted, such claim only confers internal
validity to their method, but not external validity. As described earlier in Chapter 6, the
survey that generated the data for this study includes a simple pictorial thermometer for
assessing one’s health status developed by the Euroqol group. The thermometer was
calibrated between 0 (for dead) and 100 (for best possible state of health). This is
indicated here as health status index (HSI). This affords the opportunity to compare the
outcomes of both methods of measuring health.

In a sense, SAH could be said to be a subjective measure of health while HSI could be
regarded as an objective measure of health state. HSI may be regarded as a social
valuation of health states since it depicts the views of individuals in a given society about
different health states. Being a subjective measure, SAH could suffer from cut-points bias
since different sub-populations may have different conceptions about the categories. A highly educated person might say her general state of health is *fair* where a less educated person might say she her health is *good*. But both may still be describing the same general state of health. In other words, the possibility of language bias between sub-populations within the same population may introduce bias in the measurement of health as measured by SAH.

6.3 Analysis of Health Distribution in Nigerian Population

In order to analyze the nature of health distribution in the study population (Enugu state) only the adults in the sample, (i.e. persons above age 15) were used. Out of a total number of 5814 individuals in the entire sample, only 3673 were included in this part of the study. This is because persons below age 16 were considered not mature enough to respond reasonably to questions involving self assess health status. To compare the distributional similarities and dissimilarities between the indices of health, empirical distributions of SAH (Fig 6.1) and HSI (Fig 6.2) are graphed using a nonparametric function, the kernel function. The normal distribution is imposed to visualize the extent these distributions deviate from the standard normal distribution. The kernel density estimator, like the histogram, and being also non-parametric, allows an approximation of the density of the empirical distribution of a variable \( x \). In kernel estimators, unlike in histograms, the range is divided into intervals or windows that overlap. The density of the distribution of the \( x \) variable within each window is estimated at the centre of that window. The details of the method are discussed in Chapter 9 of this report.

The distance to the centre from the boundary of the interval is therefore, the half width. Also unlike the histogram, instead of counting the number of observation points within each interval, a kernel function is used to weight the data appoints by assigning weights to the observations according to their distance from the centre of the interval. These weighting functions include the Epanechnikov, the Biweight, the Gaussian etc. But the weights in most kernel functions are usually in the 0-1 range with data points furthest from the centre of the window in most of the function getting least weight. These weights
are then summed to determine the density of the distribution within that window. (Silverman 1986, Yatchew 1998, *Stata Reference Manual*) It should be noted that the smaller the window width, the more details about the contours of the distribution we obtain but the wider the window width the smoother the curves. Also, the unit of the half width is the same as $x$.

Fig 6.1 Kernel and Normal Graphs of SAH

Fig 6.2 Kernel and Normal Graphs of HSI
The kernel function used in plotting the SAH and HSI graphs is the Epanechnikov function which is the default option in Stata. The half window width used in plotting the SAH is 0.5 while the half width used in plotting the HSI is 10. These window widths seem to have given the optimal result from the process of minimization of the mean integrated square errors. In Fig 6.1 the categories of SAH are represented along the x-axis just as HSI in Fig 6.2 is plotted along the x-axis while densities of the distributions are on the y-axis. In Fig 6.1 also, SAH is increasing in good health as we move away from the origin along the horizontal axis and excellent health is depicted as 5. Similarly in Fig 6.2, moving away from the origin along the horizontal axis, HSI is increasing in good health where the best possible health is 100.

It is interesting to note the similarities between the two distributions depicted in both figures. Each of them seems to tell the same story: health distribution is skewed to the left and nearly normally distributed. The reverse side of the story is that ill-health is skewed to the right and follows the lognormal distribution as shown by Wagstaff and van
Doorslaer (1994). Figure 6.3 shows the lognormal distribution (skewed to the right) along with the kernel function of SAH and the normal distribution.

Table 6.1 Distribution of Health as Measured by SAH and HSI Categories

<table>
<thead>
<tr>
<th>HSI UPPER CUT-OFF POINTS</th>
<th>CORRESPONDING TO SAH</th>
<th>30</th>
<th>50</th>
<th>70</th>
<th>90</th>
<th>100</th>
<th>Sub-totals</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td></td>
<td>13</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>0.93</td>
</tr>
<tr>
<td>Fair</td>
<td></td>
<td>6</td>
<td>119</td>
<td>57</td>
<td>2</td>
<td>2</td>
<td>186</td>
<td>5.09</td>
</tr>
<tr>
<td>Good</td>
<td></td>
<td>4</td>
<td>16</td>
<td>892</td>
<td>388</td>
<td>5</td>
<td>1305</td>
<td>35.44</td>
</tr>
<tr>
<td>Very Good</td>
<td></td>
<td>0</td>
<td>1</td>
<td>69</td>
<td>1426</td>
<td>197</td>
<td>1693</td>
<td>46.11</td>
</tr>
<tr>
<td>Excellent</td>
<td></td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>25</td>
<td>419</td>
<td>455</td>
<td>12.43</td>
</tr>
<tr>
<td>Sub-totals</td>
<td></td>
<td>26</td>
<td>157</td>
<td>1026</td>
<td>1841</td>
<td>623</td>
<td>3673</td>
<td>100.00</td>
</tr>
<tr>
<td>%</td>
<td></td>
<td>0.93</td>
<td>5.09</td>
<td>35.44</td>
<td>46.11</td>
<td>12.43</td>
<td>3673</td>
<td>100.00</td>
</tr>
</tbody>
</table>

In Figure 6.1 as in 6.2, it is observed that most people reported 'good' to 'excellent' health. The proportion of the population in each of the five categories of SAH is shown in Table 6.1 along with the corresponding distribution of HSI. To obtain the upper cut-off points for SAH intervals in the HSI scale, the intervals were first mapped into an empirical distribution function of HSI. Following van Doorslaer and Jones (2003), the cumulative frequency of observations for each SAH category was computed. The number of observations in HSI distribution corresponding to the cumulative frequency in SAH distribution was then computed. This implies that in a regression analysis, in which the categories of SAH are mapped into a continuous variable like HSI, the unit of SAH must be in HSI scale but the predictions also assume the linearity and continuity of the HSI. A further advantage in doing the mapping is that individual heterogeneity which can make significant difference in the estimated results is regained.
Inspecting the distribution of the health in the population, it is easily noticed that the mapping from SAH to HSI is not perfect. The responses tend to overlap in moving from one index to another. For example, 119 people reported their health to be fair according to the SAH index which would correspond to the 50 percent upper benchmark in the HSI index. Yet, there are two people in that category of SAH who estimated their health to be 100 percent. However, in general there is a good correspondence between the responses individual respondents gave in respect of the two indicators. Except for those who rated their health as ‘Poor’ in the SAH index, where majority of the respondents actually fail in the 50% cut-off point, in all the other categories, the greater majority of the SAH responses fail in the appropriate boundaries of HSI. It is even observed that greater majority of the responses do not go a step higher or below their appropriate rate in the HSI.

This leads us to the conclusion that there is some consistency in the population’s evaluations of its health states. A simple regression was carried out to ascertain the level of correlation between responses to both indicators. The simple $r^2$ was found to be 0.94 suggesting a very high degree of association between the two indices. The Spearman’s rank correlation test rejects the null hypothesis of independence and implies a high level of association between the responses obtained from the two measures of health. This clearly indicates that there is a valid mapping from SAH to HSI. This suggests also that in health surveys, one would expect the result from using one or the other of these measures of health indicator to proxy the other.

Furthermore, Table 6.1 indicates sharp increases as one moves from the lower to higher levels of health. Notice for instance the increase in the distribution from 0.93% for ‘Poor’ to 5% for ‘Fair’. Then, suddenly there is a steep rise from 5 percent for ‘Fair’ to 35% and 45% for ‘Good’ and ‘Very Good’ respectively. Finally, there is a decline from 45% for ‘Very Good’ to 12% for ‘Excellent’. These structural characteristics emphasize the skew to the left exhibited by the graphs of the kernel functions. It seems to suggest that in social evaluation of health states, there is a conceptual gap between a state of ‘poor’ and ‘fair’ health. However, this gap is even wider between saying “My health is ‘fair’” and
"My health is ‘Good’". Not much the same gap exists between "Good and ‘Very Good’. From the responses, ‘Excellent’ seems to be a state of health attainable by only a few. Generally, however, over 90% of the population reported their health to be ‘Good’, ‘Very Good’ or ‘Excellent’.

This result compares with the responses obtained from the Canadian National Population Health Survey 1994-95 as reported in van Doorslaer and Jones (2003). The results of the survey are based on the five categories of SAH. There are obvious differences in the distribution of SAH in the Canadian population and that of the Nigerian population. For example the proportion of the Canadian population that responded to be in the state of ‘Excellent’ health almost doubles that of the Nigerian population. However, it seems the basic structures of health distribution in the two populations as measured by the SAH index are the same. The considerable increase from ‘Poor’ to ‘Fair’, the sharp rise from ‘Fair’ to ‘Good’ and the decline from ‘Very Good’ to ‘Excellent’ that was observed in Table 6.1 also characterize the structure of the responses in their results.

6.4 Measuring Inequalities in Health Distribution

So far, our focus in this chapter has been to establish the correspondence between two measures of health used in this study, SAH and HSI by analyzing the distributional structure of the health of the population from the responses obtained from instruments using both indices. The clear indications from the foregoing are that the two indices of health are good proxies for self and social evaluation of different states of health, and that health is unequally distributed in the population. It is important to analyze in further details the essential structure of inequalities in health distribution in our population in order to provide information for policy.

The primary question that is to be addressed here is: what is the extent of inequalities in health in Nigerian population? In order to address this question, it is important to recall the earlier indication that there are two basic tendencies in literature on inequalities in health as noted by Wagstaff and van Doorslaer (2002).
While one tendency seems to consider inequalities in health occasioned by differences in socioeconomic circumstances which are amenable to policy influence as the key issue to be considered and measured (Pamuk, 1988, Kunst and Mackenbach 1994), the other tendency considers any form of differences in health, no matter how it arises as a problem for scientific study, and policy attention (Gakidou et al 2000, Le Grand 1987). Both forms of inequalities in health will be analyzed in this section using the Gini index, the concentration index, and econometric regression.

The Lorenz curve and the associated Gini index is a useful tool for analyzing the extent of inequality in a distribution. The curve plots the cumulative proportion of the variable being analyzed on the y-axis and the cumulative proportion of the population ranked from the least advantaged to the most advantaged on the x-axis.

Fig 6.4 is the Lorenz curve of total inequality in health as measured by the HSI. It shows the extent of deviation of health distribution from the line of perfect equality, that is, the
diagonal. The curve simply tells us that health is not equally distributed but this inequality seems to be most pronounced in the lowest category of the health distribution because at this level the curve is almost parallel with the horizontal axis.

The Lorenz curve, though visually appealing, does not give an index the magnitude of inequality in the distribution of health in this population. To obtain this information the Gini index was computed using the covariance method. The covariance method for the computation of the Gini index of health inequality is specified as (Jenkins 1988):

\[
Gini = \frac{2 \text{cov}(y, R_i)}{\bar{y}}
\]

Where \( y_i \) is the \( i \)th individual's score in the relevant health index, \( \bar{y} \) is the mean health of the sample and \( R_i \) is the fractional rank of the individual in the distribution of health.

Using this method to estimate inequality in health as measured by HSI gave the Gini index of 0.1214 on the index of scale of 0-1. Using the SAH index the value of the Gini coefficient was 0.1143. It should however, be noted that the lesser value obtained using SAH is due to the loss of individual heterogeneity in the categories. Thus, by using the categories rather than a continuous distribution like HSI, there is a loss of the individual differences which the Gini index fails to capture. The Gini in this case measures the inequality between groups but does not include the inequalities that exist within each group. With either measure, however, there is still significant level of inequality in health distribution.

Although it is difficult to compare two measures of different scales, the these figures when juxtaposed with estimates from some other previous studies which used either the SAH or related index tend to give an idea of the level of inequality in health distribution in Nigeria. For example, Van Doorslaer and Jones (2003) using the Health Utility Index and SAH estimated 0.0675 and 0.0461 respectively from the Canadian data. Retana and Jones (2003) estimated Gin index of 0.123 for England using the HSI.
Beyond estimating the Gini index for the entire sample, it is also useful to look at differences in distribution of health between subgroups. The Gini index (using HSI) for women was estimated to be 0.1223 and 0.1139 for men showing that inequality in health within the women group is about 7.4% higher than for men. This is consistent with the results of Lahelma et al (1997) for variation in health in Finland and Sehili et al (2005) that investigated inequalities in self-reported physical health in the US and concluded among other things that inequality in health among women in the US is greater than that of men.

Tables 6.2 in a simple form present inter differences in health distribution by tabulating the responses of different age groups and sexes to HSI.

Table 6.2 Age and Gender Differences in Health by HIS

<table>
<thead>
<tr>
<th></th>
<th>Females</th>
<th></th>
<th></th>
<th>Males</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age Groups</td>
<td></td>
<td></td>
<td>Age Groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16-40</td>
<td>41-60</td>
<td>60+</td>
<td>16-40</td>
<td>41-60</td>
<td>60+</td>
</tr>
<tr>
<td>No of obs.</td>
<td>1216</td>
<td>665</td>
<td>194</td>
<td>832</td>
<td>585</td>
<td>181</td>
</tr>
<tr>
<td>% of Pop</td>
<td>33.11</td>
<td>18.11</td>
<td>5.28</td>
<td>22.65</td>
<td>15.93</td>
<td>5.0</td>
</tr>
<tr>
<td>Mean HSI</td>
<td>75.32</td>
<td>71.29</td>
<td>61.30</td>
<td>77.29</td>
<td>73.31</td>
<td>65.05</td>
</tr>
<tr>
<td>STD</td>
<td>14.29</td>
<td>15.63</td>
<td>20.35</td>
<td>14.42</td>
<td>14.86</td>
<td>18.46</td>
</tr>
</tbody>
</table>

6.5 Measuring Socioeconomic Inequalities in Health

The preceding sub-section attempts to quantify the level of total inequalities in the distribution of health among Nigerians. This sub-section will estimate the level of inequalities in health arising from economic and social differences of individuals. To do this we make use of concentration curves described in Chapter 5. Essentially, while the Gini index is used for measuring inequality in univariate distributions concentration curves are useful for analyzing bivariate distributions (Wolfson and Rowe 2001, Wagstaff and van Doorslaer, 2002)
A normative issue here, which is by no means trivial, is whether or not in analyzing such bivariate distributions more weight should be attached to the disadvantaged individuals or groups. For example, in analyzing income related inequalities in health should more weight be attached to the health of individuals at lower ranks in the income parade income or should the measurement should be neutral to inequalities in the distribution of income. To frame the problem another way, is equality in distribution of income, for instance, an important argument in the social welfare function? This is a matter of the social decision maker’s ethical or ideological orientation to considerations about inequality and the social welfare function the social decision maker chooses. For example, in Atkinson’s type of utilitarian social welfare function what is important is not the rank of the individual in the distribution but the aggregate utility.

The concentration curve and its associated concentration index, like the Lorenz curve and its associated index, the Gini, has also implicit ethical judgment about inequality aversion. The concentration index, like the Gini index measures the average deviation between cumulated proportion shares of the population ranked by the socio-economic variable and the cumulated shares of health or ill-health. It is simply defined as

\[ C = 1 - 2 \int_0^1 L(p) \, dp \] \hspace{1cm} (6.1)

That is, 1 (the area of a unit square) minus two times the area below the concentration curve. Or alternatively as (Kakwani, Wagstaff and van Doorslaer 1997)

\[ C = \frac{2}{n \mu} \sum_{i=1}^{n} y_i R_i - 1 \] \hspace{1cm} (6.2)

Where \( n \) is the number of observations, \( \mu \) is the mean level of health, \( y_i \) is the health score of the \( i \)th person and \( R_i \) is the \( i \)th person’s fractional rank in the socioeconomic variable of interest.

The concentration index follows the Gini index in implicitly weighting all share deficits from the line of perfect equality equally and goes ahead to compute the average deviation
between the cumulated population shares in one variable and the cumulated shares in the ranking variable. However, there are other possible weighting schemes. That is, it is possible to allow the social welfare to be more averse to inequality by attaching more weight to proportions of the population in the lower ranks of the distribution. There is already a variety of such weighting schemes in the income distribution literature in connection with the Gini index (see Yitzhaki 1983, Donaldson and Weymark 1983, Atkinson1970). Following this example in the income distribution literature, Wagstaff (2002) showed the possibility of using similar flexible parameters of inequality aversion in the construction of concentration index. Here, in consonance with our general orientation of allowing for the policy-maker's level of aversion to inequality in distribution, an ethically flexible parameter of inequality aversion is adopted. The general framework for such weighting schemes is already explained in Chapter 5. It is merely applied here in the context of analysis of socioeconomic inequalities in health distribution. The details of this application are considered in the following subsection.

6.6 Introducing Inequality Aversion Parameter: Extended CI

As indicated in Chapter 5, linear inequality measures often apply percentile dependent weights to deviations from the line of perfect equality. In this respect, different weighting functions may be used. One such one-parameter weighting function is given by

\[ w(p, v) = v(v - 1)(1 - p)^{(v-2)} \]  

(6.3)

Where as before, \( p \) is the cumulative proportion of the people ranked according to the socioeconomic indicator. \( v \) is a parameter that measures the policymaker's aversion to inequality. Aversion to inequality increases monotonically as \( v \) increases from 0 to infinity. The higher is the value of \( v \), the higher the implied level of pro-poor inequality aversion.

If \( 0 \leq v \leq 1 \), it implies the policy-maker prefers inequality to equality. For \( w(p, v = 2) \equiv 2 \). In other words, when \( v = 2 \), then weighting scheme reduces to that implicit in the standard concentration index. Deviations from equality are weighted

133
equally at every rank of the socioeconomic variable. The policy-maker has no aversion as to whether inequality occurs at the low or high ranks of the distribution. For $1 < v < 2$ relatively more weight is given to deviations from equality occurring at, say, higher levels of income. However, when $v > 2$, it implies that he policy-maker has pro-poor aversion to inequality. As $v \to \infty$, it becomes the case of Rawls social welfare function of maxi-min.

It is from this general framework that Wagstaff (2002) following Lerman and Yitzaki (1984) develops what he calls the extended concentration index which may be specified as:

$$C(v) = 1 - v(v - 1) \int_0^1 (1 - p)^{v-2} L(p) dp, \quad v > 1$$

(6.4)

Where: $C(v)$ is the extended concentration index that varies with the level of aversion to inequality.

When specified in the form of (6.2), Wagstaff (2002) proves that the extended concentration index is given by:

$$C(v) = 1 - \frac{v}{n \mu} \sum_{i=1}^{n} y_i (1 - R_i)^{(v-1)}$$

(6.5)

Or, equivalently

$$C(v) = 1 - \sum_{i=1}^{n} (y_i / n \mu) w_i (R_i, v)$$

(6.6)

Following Lerman and Yitzaki (1984) formulation of the extended Gini index, the extended concentration index then becomes:

$$C(v) = \frac{-v \text{cov}[y_i (1 - R)^{(v-1)}]}{\bar{y}}$$

(6.7)

Where $\bar{y}$ is now the mean of the health indicator variable.
Since it may it may be of interest, as is often the case in post estimation analysis, to carry out econometric tests on the estimated parameters, Kakwani, Wagstaff and Van Dooslaer (1997) show that the concentration index could be specified in the convenient regression form:

\[
2\sigma_h^2 \frac{y_i}{\bar{y}} = \alpha_1 + \beta_1 R_i + u_i
\]  

(6.8)

Model (6.7) may then be specified in the convenient regression functional form

\[-\nu \text{var} \left( (1 - R)^{v-1} \right) \left( \frac{y_i}{\bar{y}} \right) = \alpha_1 + \beta_1 (1 - R)^{v-1} + u_i \]

(6.9)

Kakwani et al (1997) show that the concentration index (C) is the estimator of \( \beta_1 \), while \( \alpha_1 \) is the constant intercept of the regression model. It should be noticed that that the left hand-side variable in the model is the proportion of the ith individual’s health score to the mean health score weighted by the product of the variance of the complement of rank and the aversion parameter. The right hand side variable is the exponentiated complement of the ranks, the exponent being the aversion parameter minus one. The \( u_i \) is the error term that is assumed to be random.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>gpc</td>
<td>This variable is the household per capita expenditure, gross of healthcare expenses</td>
</tr>
<tr>
<td>Years of education</td>
<td>yrsed</td>
<td>This is defined as the number of years one received formal education.</td>
</tr>
<tr>
<td>Age</td>
<td>Age</td>
<td>Defined as the number of chronological years attained at the last birthday preceding the interview.</td>
</tr>
</tbody>
</table>
In the functional form specified in (6.9) it is possible to compute the standard errors and the confidence intervals and thus carry out statistical tests on the estimated coefficients. This is the model that will be used here in computing the concentration indices. The effect of varying the aversion parameter on the extended concentration index will be demonstrated using the income variable.

6.7 Concentration Index Regression Results

Table 6.4 Estimated Concentration Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\hat{a}_i$</th>
<th>Se</th>
<th>$\hat{\beta} = C$ Index</th>
<th>Se</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gpc</td>
<td>0.162)</td>
<td>(0.0046)</td>
<td>0.0091</td>
<td>(0.0075)</td>
</tr>
<tr>
<td>Yrsed</td>
<td>0.144</td>
<td>(0.0033)</td>
<td>0.0372**</td>
<td>(0.0043)</td>
</tr>
<tr>
<td>Age</td>
<td>0.181</td>
<td>(0.0030)</td>
<td>-0.0301**</td>
<td>(0.0035)</td>
</tr>
<tr>
<td>Loc</td>
<td>0.073</td>
<td>(0.0029)</td>
<td>-0.0009</td>
<td>(0.0053)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.119</td>
<td>(0.0021)</td>
<td>0.0061**</td>
<td>(0.0017)</td>
</tr>
<tr>
<td>emplostat</td>
<td>0.085</td>
<td>(0.0019)</td>
<td>0.0099**</td>
<td>(0.0022)</td>
</tr>
<tr>
<td>Disabl</td>
<td>0.008</td>
<td>(0.0005)</td>
<td>0.0049**</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>Ilistat</td>
<td>0.071</td>
<td>(0.0054)</td>
<td>0.0290**</td>
<td>(0.0022)</td>
</tr>
</tbody>
</table>

** Estimated $\hat{\beta}$ statistically significant at 5% critical level
To determine the extent of inequality in health distribution along the socioeconomic gradient, eight separate regressions of model (6.9) were first carried out using the ordinary least squares in SPSS. The results obtained with this method were unsatisfactory as it turned out (on testing for heteroskedasticity) that the model was afflicted with the problem of heteroskedasticity and did not take into account the clustering design that was used in the survey. To correct for these short-comings we resorted to the Huber/White robust estimator of variance in Stata’s ‘areg’ (Stata Reference Manual, 1997). This produced consistent standard errors which also took into account the clustering design of the survey.

The results obtained from the robust regression method are shown in Table 6.4. As the table shows, except for two of the variables, the coefficients on the other six are significant at 5% critical level. Yrsed, the variable representing the number of years of formal schooling of the respondent, turned out to have the largest concentration index (CI) value. In this context, this implies that the concentration curve generated from inequality arising from differences in levels of education deviates most from the line of equality. Since the estimated parameter is positive, it also means that this inequality favours the more educated. The importance of educational attainment as an index of social standing in Nigeria is not surprising. Right from the colonial days, educational certificate has always determined who gets recruited into the elite ranks in Nigeria. As Alubo (1987:454) notes “As in the colonial times, education continues to be an important determinant of recruitment, mobility and life changes. It determines whether one gets a job and what job and rewards”

The coefficient of age is also considerable in size. But unlike education, it varies inversely with age. This is not surprising as younger people are generally expected to feel healthier than older members of the population. Whether a person reported ill or not during the interview is naturally expected to show up in her health status index. People who did not report any illness during the interview are more likely to rate their health higher just as the result indicates.
Gender also shows up as a significant variable such that men seem to rate their health higher than women. Inequalities in health distribution resulting from age and gender will be given further consideration below. The coefficient of location is also not statistically different from zero implying that there is no difference in self rated health between urban and rural dwellers. Surprisingly, the coefficient of income turns out not to be statistically significant in this case. In other words, income does not appear to contribute very much to unequal distribution of health as evaluated by this sample of the population. This is possibly because poorer appear more optimistic about their health than the rich (Macintyre et al 2005). In that case health inequalities may not fully follow the income gradient as the health optimism of poor people may partially offset the differences in health arising from differences in income.

6.8 Interpretation of Estimated Parameters

It is important to notice in Table 6.4 that income has greater magnitude than gender, yet the coefficient of gender is significant but income is not. This is simply because while our constructed dependent variable, and some of the exogenous variables, like income and years of education, are measured as continuous variables, some of the regressors, like gender, location, employment status, disability status and illness status are included in the model as categorical variables taking the value of 0 or 1. This calls for a different type of interpretation than that used to interpret the coefficients of the continuous variables. While coefficients of continuous variables are interpreted as the direct impacts of changes in that variable upon the dependent variable, the same is not true of the coefficients of categorical variables. The coefficients of dummy variables measure the difference in the response of the dependent variable to the values taken by the dummy. The coefficients, therefore, refer to the differences in response of the dependent variable to changes in the categories of the independent variable."

Following Palmquist and Halvosen 1980) the coefficient on the dummy can be interpreted by taking exponential of the coefficient and subtracting one from it and multiplying by 100 to obtain the percentage difference in the response of the regresand to
the values taken by the regressor. This implies that in the gender variable, for example, the difference between male (the category coefficient shown as $\hat{\beta}$ that appeared in this regression) and females (the category with the suppressed coefficient) is given by $\exp(0.0061) - 1 \times 100\%$. This value gives us 1.45%. This is the size of the difference between the concentration curve for males and females and this is to the advantage of males. The other coefficients of dummy variables in the model could similarly be interpreted.

The concentration index is derived from the Gini, though unlike the latter, it measures inequality in bivariate distribution. That is to say, it estimates the inequality in one variable when ranked by another variable. Like the Gini, however, it also presents some difficulties with respect to straightforward interpretation in natural units since redistribution in the variable of interest, say health, need not affect the other by which it is ranked.

Koolman and van Doorslaer (2003) have developed a method of interpreting the concentration index in the manner of the Schutz index. To develop this interpretational approach to concentration index, they showed that the convenient regression of (6.8) can be written as:

$$C = \frac{2}{ny} \sum y_i R_i - 1 - \frac{1}{n}$$  \hspace{1cm} (6.10)

From (6.10) they derived a redistribution of $y$ in (6.8) which reduces $\beta_1$ to zero. This redistribution method is referred to as linear redistribution scheme and is specified as:

$$R_i = \frac{100}{4} \beta_1 = 75C$$  \hspace{1cm} (6.11)

Equation (6.11) implies that the percentage of the variable $y$ to be redistributed is given by the estimated concentration index multiplied by 75. Thus, the estimated concentration index of the variable ‘Yrsed’ is 0.0372. This distribution scheme suggests that in actual fact, the lump sum amount of health to be redistributed from the healthy to non-healthy based on the variable ‘Yrsed’ is 75% of 0.0372 (≈0.0279). Furthermore, Koolman and
van Doorslaer (2003) redistribution scheme makes explicit the underlying ratio scale property of the concentration index. If the \( y \) variable has ratio scale, and the ranking of \( y \) is linear, then if the concentration index doubles then inequality in \( y \) also doubles.

6.9 Sensitivity of CI to Changes in Aversion Parameter

Table 6.5 Sensitivity of CI to \( \nu \) of Yrse

<table>
<thead>
<tr>
<th>Est. CI</th>
<th>( \nu = 1.5 )</th>
<th>( \nu = 2 )</th>
<th>( \nu = 3 )</th>
<th>( \nu = 5 )</th>
<th>( \nu = 9 )</th>
<th>( \nu = 12 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\alpha}_i )</td>
<td>-0.4005</td>
<td>0.1441</td>
<td>-0.2492</td>
<td>-0.2784</td>
<td>-0.1957</td>
<td>0.1239</td>
</tr>
<tr>
<td>Se</td>
<td>(0.0145)</td>
<td>(0.0033)</td>
<td>(0.0039)</td>
<td>(0.0043)</td>
<td>(0.0030)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>( \hat{\beta} )</td>
<td>0.0912</td>
<td>0.0372</td>
<td>0.0552</td>
<td>0.0719</td>
<td>0.0759</td>
<td>-0.0688</td>
</tr>
<tr>
<td>Se</td>
<td>(0.0215)</td>
<td>(0.0043)</td>
<td>(0.0058)</td>
<td>(0.0081)</td>
<td>(0.0093)</td>
<td>(-0.0086)</td>
</tr>
</tbody>
</table>

To illustrate the impact of increasing aversion to inequality on the extended concentration index Table 6.5 shows how changes in the value of \( \nu \) affects the coefficients, particularly \( \hat{\beta} \) while Fig 6.5 illustrates the implied structure of these weights.

It is to be recalled that the value of \( \nu \) between 0 and 1 implies that the policy-maker has in fact preference for inequality. In general, it could be noticed that the values of the estimated parameters move in cyclical form. Since HSI variable is increasing in good health, it is expected that the concentration index, the value of \( \hat{\beta} \), should lie in the 0-1 range, though in general, in the limits, it can take values in the range of 1 to negative 2. As explained earlier in a previous chapter, the closer the value is to 1 the more inequality in this distribution of health to the advantage of those higher in the social ladder.

However, if the variable of concern is a 'bad' such as illness, when it takes a negative value, implying that the curve lies above the diagonal, then it plies that its distribution is negative to the disadvantage of the those at the lower levels of social status.
The estimated $\hat{a}$, though sometimes without any meaningful economic interpretation, is often taken to absorb the effects of omitted variables. However, in this regression, it moves in a rather explosive but antithetical direction to $\hat{\beta}$, from negative to positive value, then back to the negative quadrant and as the values of $v$ increases to 12 it moves back into the positive quadrant. Figure 6.6 helps to show the movements in the values of these parameters in response to changes in value of $v$. The value of $\hat{\beta}$ starting from $v = 1.5$ remains positive, though oscillating, until it enters the negative quadrant as $v = 12$. The implication is that using education to proxy socio-economic status, the more the policy-makers are sensitive or averse to inequality in health that disadvantages those at the low rungs of the social ladder, the more inequality they perceive in the distribution of health and, thus, the concentration index becomes the more left-skewed. However, the negative value of $\hat{\beta}$ in Table 6.5 as $v$ takes the value of twelve seems to suggest that there is a limit to this kind of Rawlsian policy since it can also reverse the situation against those who were initially advantaged.
It is important to analyze further the reason for the reverses in the value of the concentration index since the structure of the function $\hat{\beta} = f(v)$ in the graph seems to run against expectation. To do this, it seems necessary to re-examine the nature of the convenient regression model that yielded the results. The basic model as specified in (6.8) indicates that $\hat{\beta}$ is the magnitude of the average response of the proportion of the $i$th individual’s health to mean health of the population following a change in individual $i$’s relative ranking in the population. However, following Milanovic (1997) Koolman and Van Doorslaer (2003) show that, analogous to the Gini, the concentration index can be written as

$$C = \frac{12\sigma^2_i^2}{\sqrt{3}} \frac{\sigma_y}{y} \rho(y, R) \quad (6.12)$$

Since from the convenient regression specification $\hat{\beta}$ is the estimator of $C$, it implies that $\hat{\beta}$ can replace $C$ in model (6.12). Thus equation (6.12) suggests that the value of $\hat{\beta}$ is a function of three variables: the variance of ranks in the ranking variable — income, the coefficient of variation of $y$, that is $\frac{\sigma_y}{y}$, which measures the absolute variation in $y$ as a
percentage of its mean, and the correlation coefficient $\rho(y_i, R_i)$ which measures the strength of linear association between health and the ranking variable, income.

This relationship seems to provide the key to understanding the movements in the estimated value of $\beta$. While the correlation coefficient or strength of association of the health score $(y_i / \bar{y})$ and the relative rank $R$ in income may be very higher suggesting high inequality of health in respect to income, there are other components that are also exerting influence on $\beta$. For example, if the variance of $y$ is very high, suggesting that there is a lot of inequality in the distribution of $y$ itself, and further, implying a high value of the ratio $\frac{\sigma_y^2}{\bar{y}}$, then it implies that the estimated value of $\beta$ may be significantly altered by the impact of coefficient of correlation. Now, in the case in point, the rank used in computing the concentration index is the complement of the fractional rank of the individual in the distribution of the socioeconomic indicator of interest. Looking at the extended concentration model as specified in (6.9) it becomes obvious that as the value of $v$ keeps increasing exponentially, the value of the variance of $(1-R)$ keeps shrinking in proportion such that even when it is multiplied by (1-$v$) to obtain the left hand side of (6.9) the strength of the association between $(y_i / \bar{y})$ and $(1-R)$ on the right hand side is considerably reduced or at least weakened. Thus, in this particular model, an increasing pro-poor inequality aversion does not seem synonymous with increasing pro-rich inequality distribution in health.

This result also seems also to be consistent with Podder (1993) which showed that for a variable to increase the level of inequality, it is not sufficient that it is a significant proportion of total income, it must also be distributed unevenly. Putting this in the context, strong association between health score and relative rank in the socioeconomic variable is important in determining the size of inequality in health but if the variance in the distribution of health is very small, the impact of that strong association on inequality will not be large.
6.10 The Determinants of Health

In the foregoing analysis, the effort has been to assess some of the socioeconomic variables that contribute to inequalities in health distribution. These are the factors whose unequal distribution tends to enhance uneven distribution of health across the population. It was noted that for a variable to influence inequality in health the variable itself must be unevenly distributed. A factor may be strongly related to health by having a high coefficient of correlation with health, yet it need not necessarily contribute to health inequality if that variable it is evenly distributed in the population. Thus, our basic concern here is to briefly analyze those socioeconomic factors that contribute significantly to health. For this purpose, two regression models are used. In the first place we use the robust regression model, with HSI as the dependent variable, to explore the socioeconomic determinants of health. Secondly, the Ordered Logit with the categories of SAH as the dependent variables is used to check if these socioeconomic variables are consistent in their impact on health indicators. The models are, of course, not directly comparable but the signs and statistical significance might provide some useful information about their behaviour under both indices.

In both robust and Ordered Logit models the intention is to estimate a model specification of the form:

$$y' = \beta' x + u$$

(6.13)

Where $y'$ is the dependent variable and $\beta' x$ is a vector of regressors and associated parameters. The difference is that unlike in the ordinary regression where $y$ is assumed to be continuous, in the probability or discrete choice models to which Ordered logit belongs, $y$ takes discrete values. The same set of determinants may also be specified for both models. Also in the discrete choice models, the estimated parameters in themselves are usually of limited value since the interest is on the probability of a given outcome given the vector of variables. The prediction of any given outcome is interpreted in the form of probabilities rather than direct impacts of the individual exogenous variables on the dependent variable as is usually the case with in OLS approach.
The robust estimation technique is essentially the Huber/White function for the estimation of variances in the presence of heteroskedasticity. Using the robust regression enhances the efficiency of the estimation and the accuracy of the estimated standard errors. As noted earlier on, it also takes into account sampling weighting, clustering and stratification effects introduced by the design of the survey in survey data. These survey variables are included in present data and so warrant the use of robust estimation model.

The Ordered logit is a regression model that estimates the relationship between ordinal dependent variable and a vector of independent variables. A variable may be said to be ordinal if it is a categorical variable that is ordered. For example, the responses to SAH is typically ordered: Poor, Fair, Good, very Good, Excellent. Therefore the Ordered logit or Ordered probit model fits perfect well the model being analyzed here. The error term $u$ in the Ordered logit and Ordered probit models follow different distributions. While the error term in probit model follow the normal probability distribution, the error term in the Logit model is assumed to follow the logistic distribution and so both models use different computational approaches but lead to essentially the same results. Of course, because they follow different distributions their coefficients are not directly comparable.

In Ordered logit estimations the assumption is that an underlying score $y^*$ is estimated as a function of the set of linear parameters and a set of cut points (see Stata Reference). The probability of observing any of the outcome categories (this is estimated by the Maximum Likelihood function) is given by

$$\Pr(y^* = i) = \Pr(k_{i-1} < \beta' x_j + u_j < k_i)$$

(6.14)

In this case, $u_j$ is assumed to follow the logistic distribution while \( k_1, k_2, \ldots, k_{1-1} \) are the cut-points and $i$ represents the number of cut-points. $k_0$ is assumed to range from $-\infty$ to the cut point of the first category ('Poor' in this case). The highest cut point is taken as $+\infty$ ('Excellent' in this case) Both robust and Ordered logit estimations are carried out using Stata.

145
6.11 Estimated Parameters of the Determinants of Health

The results of the robust and Ordered logit regression models are presented in Tables 6.6 and 6.7 respectively. In the first of these tables, it is observed that the variables included in the model explain only about 21% of the variations in health across the population. This is of course not unexpected as there are many more factors that contribute to differences in health of people other than socioeconomic factors. A great percentage of these variations may be attributable to biological factors. The omission of such factors also sows up in the high level of the estimated constant parameter.

Table 6.6 Robust Estimation of the determinants of HSI

<table>
<thead>
<tr>
<th>Dep var HSI</th>
<th>Coef.</th>
<th>Se</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inc</td>
<td>0.0005</td>
<td>0.0002</td>
<td>2.06</td>
</tr>
<tr>
<td>Agegrp2</td>
<td>-1.1873</td>
<td>-0.8747</td>
<td>-1.36</td>
</tr>
<tr>
<td>Agegrp3</td>
<td>-6.0826</td>
<td>1.8446</td>
<td>-3.30</td>
</tr>
<tr>
<td>Sex1</td>
<td>-0.9427</td>
<td>0.5223</td>
<td>-1.80</td>
</tr>
<tr>
<td>Emplstat2</td>
<td>-1.4005</td>
<td>1.0273</td>
<td>-1.36</td>
</tr>
<tr>
<td>Disabl2</td>
<td>-12.3826</td>
<td>6.0542</td>
<td>-2.05</td>
</tr>
<tr>
<td>yrsed</td>
<td>0.5457</td>
<td>0.1124</td>
<td>4.86</td>
</tr>
<tr>
<td>_cons</td>
<td>72.2093</td>
<td>1.8900</td>
<td>38.21</td>
</tr>
<tr>
<td>Adj Rsq</td>
<td>0.2100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is noticed that income now surfaces as a significant determinant of health, albeit, marginally significant. In this model, as in Table 6.2, the population was divided into three age groups: 16-40, 41-60 and 60.

In consonance with regression methods that include dummy variables, the reference category, 16-40, age bracket has been dropped to avoid perfect collinearity and identification problem in estimation. However, as explained previously, the coefficients of the Agegrp2 (41-60) and Agegrp3 (60+) are shown. Looking at either the standard errors or the t-statistics it is easy to see that the difference in health, as indicated by HSI, between the age groups 1 (the omitted category) and 2 is insignificant but the difference between age groups 1 and 3 is statistically significant.
Table 6.7 Ordered Logit Estimation of the Determinants of SAH

<table>
<thead>
<tr>
<th>Dep var SAH</th>
<th>Coef</th>
<th>Se</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linc</td>
<td>0.0001</td>
<td>.00004</td>
<td>3.02</td>
</tr>
<tr>
<td>yrsed</td>
<td>0.0649</td>
<td>0.01494</td>
<td>4.34</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0233</td>
<td>0.0062</td>
<td>-3.77</td>
</tr>
<tr>
<td>gender</td>
<td>0.2545</td>
<td>0.08463</td>
<td>3.01</td>
</tr>
<tr>
<td>Loc</td>
<td>-0.4595</td>
<td>0.3473</td>
<td>-1.32</td>
</tr>
<tr>
<td>ilstat</td>
<td>-2.95325</td>
<td>.2554015</td>
<td>-11.56</td>
</tr>
<tr>
<td>disable</td>
<td>-1.294367</td>
<td>.8323683</td>
<td>-1.56</td>
</tr>
<tr>
<td>emplst</td>
<td>-4.360729</td>
<td>.2056545</td>
<td>-2.12</td>
</tr>
<tr>
<td>Pseudo R sq</td>
<td></td>
<td>0.130</td>
<td></td>
</tr>
<tr>
<td>Log pseudo-likelihood</td>
<td>-273.098</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ancillary Parameters**

<table>
<thead>
<tr>
<th>_cut1</th>
<th>_cut2</th>
<th>_cut3</th>
<th>_cut4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6.936784</td>
<td>.4583477</td>
<td>-9000756</td>
<td>.3746202</td>
</tr>
<tr>
<td>-4.358528</td>
<td>.3329522</td>
<td>1.7396</td>
<td>0.4277</td>
</tr>
</tbody>
</table>

1 Pr(xb+u<_cut1) | 0.0061 |
2 Pr(_cut1<xb+u<_cut2) | 0.0483 |
3 Pr(_cut2<xb+u<_cut3) | 0.3613 |
4 Pr(_cut3<xb+u<_cut4) | 0.4630 |
5 Pr(_cut4<xb+u<_cut5) | 0.1212 |

This difference can again be calculated in percentage as \[\exp(-6.0826) - 1\]100% which gives a 99% probability that the a person of age group 1 is likely to report better HSI than a person in age group 3. Similarly, it can be said that a disabled person is likely to report worse health than a non-disabled person. Number of years of education here, again, shows up as an important parameter in measuring health. Health is increasing in years of education, though to a limit. The difference in health between sexes may be assumed to be marginally significant. We find, however, that by and large, the same factors that drive inequality in health distribution are also influencing health status.

Table 6.7 reports both the substantive and ancillary parameters. The substantive parameters tell about the relationship between the categories of the dependent variable and the set of independent variables. It is to be noted however, that the individual parameter estimates are only of secondary value. That is to say, the primary interest is not
in the value the latent variable \( y^* \) takes given the \( x \) values because \( y^* \) itself is an abstract concept (see Wooldridge 2002). However, the signs of the estimated parameters tell us about the direction of effect, that is \( \Pr( y = 1 \mid x ) \). The \( \Pr( y = 0 \mid x ) \) moves in the opposite direction with the sign of the coefficient while the sign on \( \Pr( y = 1 \mid x ) \) goes in the same direction with the estimated coefficient. That is to say, as \( \beta \) increases in magnitude, it implies the variable in question is increasing in the direction of ‘Excellence’.

The auxiliary parameters which correspond to the \( k \)'s assist the interpretation of this relationship. ‘_cut1’ for instance, gives us the upper boundary for ‘Poor’, ‘_cut2’ gives us the upper limit for ‘Fair’, ‘_cut3’ gives the upper boundary for ‘Good’ and ‘_cut4’ gives upper threshold for ‘Very Good’, while the upper threshold for ‘Excellent’ is positive infinity. These cut-point have also standard errors. The second group of auxiliary parameters gives us the observed probabilities which tallies more or less with the percentage score of each category in Table 6.1. It is s important to recall that the probability of observing a given out-come, say, ‘Good health’ in SAH, is given by the probability that the estimated function including the error term lies in the range of the cut points estimated for ‘Good’.

In summary then, Table 6.7 shows that the variables that have positive influence on health using the HSI indicator also influence health as measured by SAH, to a more or less degree. There it is observed again that ‘yrsed’ and income influence health positively. Increase in the number of years of education increases the probability of scoring ‘Excellence’ in SAH. In contrast, getting older, being ill or unemployed increase the probability of observing ‘Poor’ in SAH. The levels of significance of some of the variables do not remain the same in both the robust and the Ordered logit models, but in general, there is consistency in the behaviour of the variables in the two indicators and under the two estimation models.

This chapter has highlighted a number of issues. It has shown that there is a considerable level of inequality both in pure health distribution (measured by the Gini index of health)
as well as in health as determined by socioeconomic differences of life (measured by Concentration Index). This uneven distribution is robust to the two indices of health measurement: HSI and SAH. Conversely, by estimating the degree of inequality in health distribution using two indices helped to establish the near-equivalence of the results obtained from both indices.

6.12 Summary

The chapter also highlighted some of the socioeconomic variables that drive inequalities in health and those that determine health. In both cases, the results show that these factors are consistent in their influence on health and health distribution. A new approach to interpreting the estimated coefficient of concentration developed by Koolman and van Doorslaer (2003) index is applied. This new approach gives some sense of the magnitude of transfer to be made from those above the mean to those below the mean in order to achieve equality. Finally, a new method of estimating inequality in health that considers the policy-maker's aversion to inequality was implemented. Establishing the presence of large inequalities in health distribution, particularly those arising from socioeconomic differences in life, strongly suggests that policy has a major role to play in removing these avoidable inequalities. These results provide a solid background for the subsequent chapters where the problem of the redistributive effect of healthcare financing will be analyzed.

A well accepted principle in healthcare financing is the principle of cross-subsidization: that the rich and the healthy should subsidize the poor and the sick. Furthermore, it is generally accepted that healthcare should be financed according to ability to pay, this ability being measured by the level of income. Does the system of healthcare financing in Nigeria promote or deviate from these ideals? How far does it? These questions are the focus of analysis in the next and subsequent chapters.
Chapter 7

Descriptive Results

7.1 Introduction

The preceding chapter clearly indicates significant levels of socially determined inequalities in health distribution in the Nigerian population. Health correlates positively with measures of socioeconomic status suggesting that there are different levels of healthcare needs among the population. In particular it points to the fact that those at the lower levels of social and economic distributions have higher probability of needing healthcare. If healthcare were costless, social welfare would be improved by giving everyone unlimited access to healthcare. This would tend to equalize health for all assuming healthcare is the source of health restoration. If, on the other hand, healthcare has costs (as is always the case), but it is publicly provided, it seems social welfare would be better improved by giving greater access (in terms of removing barriers that impede consumption of health services) to those who are less advantaged in the health distribution. In this case both Pareto efficiency and vertical equity would be satisfied because in that case, improvements in the health of the ‘poor’ would not deteriorate the health of the ‘rich’ [assuming there is no envy (Varian 1975)]. But the government has to raise the resources for the health sector by taxing the rich and the healthy to finance the poor and the sick through the provision of healthcare. From the utilitarian, liberal, and egalitarian perspectives, this redistributive policy would be both efficient and vertically equitable.

However, in the absence of significant government intervention and subsidies in healthcare financing and the dominance of for-profit healthcare as is the case in Nigeria, households must have to bear the cost of their health needs whether they are rich or poor.
In this case, those who are poor and are also more likely to be sick would have to pay more because they are more likely to need healthcare. If they could not pay, they would be excluded from healthcare consumption. This situation would give rise to pro-rich income redistribution. If everyone, both rich and poor alike, buys healthcare in proportion to their level of income, then there would be no income redistribution arising from the financing of healthcare. In this and the next two chapters we analyze the redistributive effects of healthcare financing. Unlike in Chapter 6, where analysis of health distribution was restricted to a sub-sample, the adult respondents, the analysis in the rest of the chapters include all the observations in the sample. In this section we provide some descriptive results before presenting in the next subsection the analytical results of the decomposition. The purpose is to provide some background information to the analytical results.

7.2 Patterns of Healthcare Expenditure

Table 7.1 Distribution of Gross and Health Expenditures by Quintiles

<table>
<thead>
<tr>
<th>Quintile</th>
<th>No. obs</th>
<th>Mean Inc</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gross exp</td>
<td>300</td>
<td>741.75</td>
<td>233.86</td>
<td>133.87</td>
</tr>
<tr>
<td></td>
<td>H.Exp</td>
<td>300</td>
<td>52.89</td>
<td>91.51</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Gross exp</td>
<td>299</td>
<td>1547.88</td>
<td>240.73</td>
<td>1138.91</td>
</tr>
<tr>
<td></td>
<td>H.exp</td>
<td>299</td>
<td>104.70</td>
<td>215.10</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Gross exp</td>
<td>300</td>
<td>2496.01</td>
<td>323.40</td>
<td>1988.28</td>
</tr>
<tr>
<td></td>
<td>H.Exp</td>
<td>300</td>
<td>152.01</td>
<td>320.09</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Gross exp</td>
<td>299</td>
<td>3861.11</td>
<td>483.76</td>
<td>3081.03</td>
</tr>
<tr>
<td></td>
<td>H. exp</td>
<td>299</td>
<td>210.24</td>
<td>505.35</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Gross exp</td>
<td>299</td>
<td>9299.82</td>
<td>6432.82</td>
<td>4852.67</td>
</tr>
<tr>
<td></td>
<td>H. Exp</td>
<td>299</td>
<td>1065.43</td>
<td>2882.82</td>
<td>0</td>
</tr>
</tbody>
</table>

The sample data reveals that of 1497 households retained in the study, 522 representing 35% percent of the total sample reported financing healthcare in the two weeks preceding the survey. This indicates a high incidence of disease and poor population health as described in section 2.5. In general, the sample shows that 38% of households with income below the poverty line of 2900 naira (N2900) (see section 7.5 below) reported financing healthcare while the 30% of household above the poverty line reported financing healthcare during the period. However, the mean amount spent on healthcare
financing by the poor and non poor differed widely. The poorest 75% of the households spends on average N342.38. The top richest 5% of households spends on average N2389.40 or 700% of the amount by poorest 75% of households. Similarly, poor households that financed health during the period spent on average N255.00, the non poor spent on average N2021.71. The top richest 1% of household that financed healthcare spent on average N8429.76 or 3,300% of the amount spent by the poorest 60% of households.

Table 7.1 shows that in terms of per capita expenditure, the distributions of the gross expenditure and healthcare expenditure are skewed. This distribution of healthcare payment and income tally well with those reported in Table 5 in Gotsadze et al (2005). The mean income among the top income quintile is N9300 which is about 10 times the mean income of N742 of the lowest quintile. Even among the lowest quintile, there is a wide variability among the members as indicated by the standard deviation (SD) which increases for the rest of the quintiles. Similarly healthcare expenditure follows the slope of the gross expenditure as we observe the healthcare expenditure increasing with gross expenditure in all the quintiles. This increase in healthcare payment is particularly steep going from the 4th to the 5th quintile showing clearly that those on the 5th quintile spend relatively more on health than other quintiles. But again, the variation in this group is wide, obviously indicating the difference between those who paid and those who did not pay for healthcare in the period covered by the survey. In every quintile there are households that paid nothing for healthcare, as indicated by the zero minimum payment. Table 7.1 indicates clearly that households below the 5th quintile seem to under-utilize healthcare services, probably because they could afford it. This inference is consistent with the highly skewed income distribution and high rate of poverty in the country. [It is estimated that over 60% of the population is poor (Ogwumike, 2001)].

The over all average payment for healthcare is N261 per visit. The median payment is N225.9. The mean payment is about 35% of the mean income (total expenditure) of the lowest quintile. This again raises the problem of affordability and exclusion of the poor from the health system. A poor household with the mean income of the lowest quintile
cannot afford a visit to a health facility because it would have little or nothing left after the visit to take care of other basic needs such as food and shelter. For the second and third quintiles, the average cost per visit to a medical facility takes about 17% and 11% respectively of their mean income. Such level of expenditure would be considered catastrophic by most standards. Again, this seems to emphasize the point of exclusion of the poor from healthcare utilization on account of the high opportunity cost of buying healthcare.

There is no significant disparity in gender between males and females in reporting illness. According to the data, about 8.9% of women had one form of visit or the other to a health facility while 9.8% of men visited a health facility during the period. However, mean payments between men and women differ significantly. The mean payment by men is N1128.81 while the mean payment for women is N785.13. This may suggest that there are inequalities in allocating health resources between genders in the household, or that men are more likely to report severe illnesses than women. The data defines severity in terms of number of days in the last one month that a person was unable to attend to normal duties. On this score men who reported sick had more severe illnesses of about 6 days while females had average severity score of 5 days. Males also had greater average number of visit to facilities than women. Among males that reported illness, the average number of visits to healthcare facility is 3.7 while for females it is 3.4. It is not certain that the higher severity scores and higher number of visits by men could have accounted for large difference in the mean payments made by the two genders. This raises the suspicion that there might be discrimination between men and women in healthcare financing.

According to the data, the mean payments for illnesses differ according to facility visited. Private hospitals and clinics have the highest average charge of N1869.38. Public hospitals and clinics have the next highest average charge of N1438.24. The difference in prices in the public and private health facilities in the data is consistent with what was reported in Ogunbekun et al (1999) for Lagos and Ogun states. However, this difference does not translate to much if the extra non-monetary costs of seeking health in public
facilities is factored in. These include long queues that translate into longer waiting time, abuses from staff in public health facilities, and ‘os’ (out of stock) syndrome that characterize the public health facilities (Alubo 2001). Church owned hospitals and clinics have mean charge of N1375.31. Traditional medicine has average cost of N1169.86 while drug stores have an average charge of N231.03.

In terms of preference of facilities for treatment, the data shows that about 295 or 46% of reported cases were treated in the pharmacy/drug-vendors category of healthcare providers. About 23% and 21% were treated in private hospitals/clinics and public hospitals/clinics respectively while about 6% of health seekers consulted the traditional medicine practitioners. About 4% of those who reported sick said they were treated at home. Altogether, only about 44% of those who reported sick were treated in either a public hospital/clinic or private hospital/clinic. That only about 21% of the population seek healthcare in public health sector tells the story of the relative strength of this segment of the healthcare market. This is despite the fact that prior to oil burst of the early 1980s the government had monopoly power and was virtually the sole employer of labour in the health sector.

The conclusion that can be made from all these statistics is that the cost of seeking healthcare is beyond the means of the average household in Nigeria. Calculated in monetary terms, there is significant difference in price of healthcare in public and private hospitals and clinics. But this difference can easily be overlooked when ones factors in other non-monetary considerations such as non-availability of prescribed drugs which the patient has to purchase eventually from private drug stores. Further more, the facts emphasize that most Nigerians are excluded from seeking appropriate health for their illnesses because they cannot afford it. The evidence is clear from the fact that over 55% percent of the population seek healthcare from the largely informal suppliers where the prices are cheaper even if even this means postponing the evil day when illness becomes very acute (Alubo 2001). In general, the rapid growth of the private health market sector since the early 1980s following the oil burst and subsequent deregulation of the sector seems to have worked in the direction of the World Bank policy of creating more space
for the market in the healthcare sector. The growth of private medicine seems to have fulfilled and important gap in the supply of healthcare in Nigeria.

The policy implications are clear from all this. The importance of the private sector healthcare service delivery (a large portion of which is the informal healthcare market) cannot be disregarded by any policy-maker. Drug vendors, traditional medicine suppliers all play significant roles in the healthcare market and are widely patronized. There is, therefore, the need for a policy that integrates them formally into the national healthcare system so that they can be reformed and regulated instead of allowing them to operate at the informal fringes of the healthcare market where they can engage in all manner of unethical practices. At present the only official engagement with this large army of ‘informal’ healthcare suppliers is giving them the permit to operate and the official concern for their involvement in all manner of unethical behaviour. Lack of adequate control over their activities is a major limitation that underlines the whole private healthcare sector in the country. Driven purely by the profit motive, many of the private healthcare practitioners and suppliers engage in unethical conduct. These unethical conducts include the importation and distribution and administration of fake, substandard and expired drugs. Many engage in illegal administration of injections with its serious implications for the spread of HIV/AIDS, hawking and selling of drugs without regard to appropriate dosage. Patients are often required to deposit money before they could be treated and treatment is discontinued when the deposited amount is exhausted (Alubo, 2001). The combination of high cost healthcare products and services, inadequate control, and the profit motive has often led to tragic healthcare incidences such as drug poisoning. Attributing the death of 109 children through fake paracetamol in 1990 to this pursuit for profit Alubo (1994:97) notes “... this tragedy, and others before it, are more fully understood within the context of a seemingly free-for-all drug scene on the one hand, and a health care system which excludes the majority on the other”.

There are also cases of private practitioners pilfering drugs and medical equipments from the public healthcare facilities, where they work as full time staff, to support their own private facilities (Ransome-Kuti, 1998). There are also reports of healthcare cartels in
some parts of the country whereby some hospitals gang up to fix high prices thereby denying patients even the market determined prices of care. Moreover, these high prices do not often seem to provide sufficient incentive for higher quality care, rather where quality is high the price regime is even higher. This severely limits access, especially for the poor (Alubo, 2001). All these abnormalities in the private healthcare practice constitute severe limits both in respect of the access of the population to healthcare and the extent to which unfettered private healthcare market can be trusted as an alternative source of healthcare for the population. Furthermore, the well-known information asymmetry and supplier moral hazard that exists in the healthcare market is amplified in the private sector segment of the market. This suggests the need for greater government and third-party involvement in the provision and financing of healthcare services particularly in developing countries like Nigeria.

7.3 Inequalities in Healthcare Expenditure

Figure 7.1 plots the Lorenz curve of prepayment and the concentration curve of payment to provide a visual representation of the extent of inequality in the gross and healthcare payments based on per capita. Both the Lorenz and the concentration curves clearly indicate significant levels of inequality in the surveyed population. Comparing the Lorenz and the concentration indices in Fig 7.1, it is also found that the concentration index deviates more from the Lorenz at the higher levels of income than at the lower levels. This implies that there is greater variability in healthcare expenditure among the top income brackets than among the lower income brackets. This is, of course, not expected since illness is a stochastic shock to household income and so does not vary systematically with income. Poorer households have very limited scope to increase their expenditure on healthcare on the event of illness so that at lower income levels, the difference between households that spent on healthcare and does that do not may not be very wide. This is unlike the situation among the richer income classes where, in the event of an illness, such households can devote considerable amount to healthcare implying greater differences among rich households that paid for healthcare and those that did not.
This divergence begins to widen significantly at the about the 40th percentile of the population arranged in increasing order of income suggesting that the variability between health expenditure at the lowest percentiles of income is very small. This divergence between income and healthcare payment becomes widest at the uppermost percentiles suggesting that in fact the richer people may be spending relatively higher proportion of their income on health than do poor people.

Fig 7.2 plots the expected and conditional standard deviation of healthcare expenditures. The expected value is obtained simply by plotting the non-parametric function of expenditure. The expected values were generated by running a non-parametric kernel regression of gross expenditure on payment using the distributive analysis software DAD. As can easily be observed, both the conditional and expected standard deviation tell the same story of greater variability among the higher income earners than among the lower income earners.
What effect does healthcare payment have on the distribution of income at the post-payment period? To answer this question a scatter-diagram of the gross income is plotted against the post-payment net income to obtain Fig 7.3. This figure shows that the distribution of post-payment net income on prepayment income is restricted to the lower, right hand triangle. Notice that if there were no payments made, the plot of the net income against prepayment income would be a diagonal line. If every household contributes to the health system in exact proportion to its prepayment income, the diagonal line would shift downward by the exact percentage of the contribution to the gross income, just the same effect as would a lump-sum tax.

However, none of these is the case here. What is seen in Fig 7.3 is that while many households that did not contribute to healthcare financing maintained their original positions on the diagonal, many others who made contributions did not. Thus, the scattered residuals represent households who fell below the diagonal implying that their
net income is below their prepayment income. Such payments could indeed be catastrophic by taking a large proportion of the household income and diverting it to healthcare and thus leaving such households with very little to spend on other basic needs. This is also a major source of policy concern (Wagstaff and van Doorslaer, 2001). This phenomenon will be considered in greater details in subsequent chapters as the incidence of horizontal equity and reranking effect.

Figure 7.3 Scatter Diagrams of Gross vs. Expected Net Expenditure

A similar graph is obtained in Fig 7.4 by plotting the difference between the prepayment income rank distributions of households against the post-payment rank distributions. Households are first ranked according to their income in the prepayment income parade. Then the amount they paid for healthcare is deducted from their gross incomes. Households are then ranked again according their post-payment income. For many households that experienced ill health and therefore had to finance healthcare, their ranks in the post-payment income dropped from their rank in the prepayment income. Households that did not make payments either retained their prepayment income rank or moved ahead of their counterparts and income superiors that made payments.

\[4\] For many households that did not experience ill health at all and so did not pay for health services the amount deducted from the gross is zero.
Thus, assuming for example, households A, and B had exactly the same income of N100, while household C had N90 at the prepayment income. Assuming further those households A and C did not experience any ill-health and so did not make any contributions to the healthcare system. Household B did and had to pay an amount N20 to finance cost of treatment. At the post payment income households A and C would still have N100 and N90, respectively, while B now has N80. We notice immediately that the payment system has introduced both horizontal inequity and reranking. Horizontal inequity occurs because healthcare finance system has treated A and B unequally, and reranking because B has now fallen behind C in income rank at the post-payment income distribution.

Fig 7.4 therefore suggests that out-of-pocket financing of healthcare among the population generates both horizontal inequities and often leads to loss of ranks by households that make healthcare payments. Households that experience severe illness and consequent large healthcare expenditure may not only be impoverished if there is no
other third-party financing of health but could also see themselves falling behind their prepayment income equals or even their prepayment income inferiors. As noted earlier this could lead to psychological feelings of alienation and resentment. This reranking in the transition from pre- to post-payment income distributions is depicted in Fig 7.4 by the dots below the diagonal.

Fig 7.5 shows healthcare expenditures as a percentage of total expenditure for 10 deciles of income. It indicates that the graph of the financing burden is not a straight line with consistent slope. It rather shows that a declining slope up to the median percentile and stagers off on an upward trend. Contrary to expectation, the 10th decile seems to spend a relatively larger proportion of income on health. It spends about 14% which is about the proportion of healthcare to total income of the average household as reported in Gerdtham et al (1996) for Sweden.

![Relative Healthcare Financing Burden](image)

**Fig 7.5 Distribution of Healthcare Financing Burden Across income Groups**

Other deciles of income spend on average about 7.5% of their gross income on healthcare. This would seem to suggest, given the general evidence that health follows the a positive gradient, that under self financed healthcare system, majority of low income earners would tend to avoid using health facilities on account of cost, or under-
spend in healthcare probably because they have to trade-off healthcare expenditures with other basic needs like food, shelter and schooling.

It may also suggest that available public healthcare expenditure is so well targeted that the poor benefit the most. This would be the case for instance, if the public health facilities are so poorly run that the poor crowd out the rich who would prefer to spend more on private healthcare facilities to obtain higher quality care. In that case, public healthcare facilities would become a sort of catastrophic health insurance because the poor would avail themselves of the poor quality care in public care services where they could pay minimum amount at point of consumption.

It is, however, more likely that at lower deciles, income is so low that any amount of healthcare expenditure is bound to constitute a significant part of the gross income. Many poor households would avoid the use of healthcare services. At the highest income deciles households are fully able to finance their healthcare. However, in the absence of significant subsidies healthcare must be paid for at the going market prices. It then implies that these richer households must devote an increasing proportion of their incomes on healthcare needs.

The level of burden imposed by healthcare payment at all levels of income as shown in this graph emphasizes the case for policy intervention in the healthcare financing arrangement. *Health being a merit good (Mwabu et al., 2003) deserves greater public sector financial contribution than is the case in the present financing arrangement in the country.*

### 7.4 The Incidence and Intensity of Catastrophic Healthcare Financing

It was emphasized in section 1.3 above that healthcare financing could sometimes absorb a large proportion of a household’s income so that it becomes also a source of policy concern. This is clearly shown in section 7.3. For households that are living on the
solely on direct payment, and one that depends on general taxation will require different level of contribution from one that depends on private health insurance. In these cases it is reasonable to set different catastrophic thresholds. In this study we set a multiple of thresholds to see the extent of catastrophic effect varies with thresholds.

To articulate the estimation model, we assume that a household has total expenditure (gross of healthcare expenditure) $Y$. We define the catastrophic threshold to be $Z_{cat} \cdot$ Where $Z_{cat}$ is $x\%$ of $Y$. Following the poverty measurement literature, Wagstaff and van Doorslaer (2001) proposed the following indices of catastrophic healthcare expenditure:

(a) **Catastrophic Payment Headcount Index** $H_{cat}$. This is defined as the percentage of individuals in the sample population whose healthcare expenditures as a proportion of their income exceed the threshold $Z_{cat} \cdot$

Let $T_i$ be the healthcare payment of the $i^{th}$ household, and $Y$ the total household expenditure, as defined before.

Let

$$\left( \frac{T_i}{y_i} \right) - Z_{cat} = \begin{cases} 0 & \text{for } \left( \frac{T_i}{y_i} \right) > Z_{cat} \\ s_i & \text{otherwise} \end{cases}$$

(7.1)

$s_i$ is the excess of healthcare payment-income ratio over the catastrophic threshold when $T_i/y_i > Z_{cat}$.

Also let

$$E_i = 1, \quad \forall S_i > 0.$$  

(7.2)

Then, borrowing from the poverty measurement literature, catastrophic headcount ($H_{cat}$) may be defined as (Wagstaff and van Doorslaer, 2001):

$$H_{cat} = \frac{1}{N} \sum E_i = \mu_s$$

(7.3)
\( N \) is the sample size and \( \mu_n \) is the proportion of the population whose healthcare payment exceeds the catastrophic threshold or the catastrophic payment headcount index.

The society or the policy-maker may not just be concerned with the proportion of the population whose healthcare payments exceed the catastrophic threshold. Policy may also be concerned about the height by which those who exceed the threshold do in fact exceed it. Analogous to the severity of poverty index or poverty-gap index, the intensity of catastrophic healthcare payment may be defined as

\[
G_{\text{cat}} = \frac{1}{N} \sum S_i = \mu_s
\]  \hspace{1cm} (7.4)

Where \( \mu_s \) = mean 'overshoot' (Wagstaff and van Doorslaer, 2001) or the average amount by which those who exceed the catastrophic payment threshold actually exceed it.

It is natural that the policy-makers may not only be interested in the catastrophic headcount and the intensity of the catastrophic payment, they are likely to have an aversion for catastrophic healthcare payment, that is, whether it is more prevalent among the poor or among the rich. There is, therefore, the need to weight \( E_i \) according to the income rank of the household. Catastrophic healthcare payment for a poor household is likely to be weighted more than for a rich household. The weighting factor appears to be a normative issue depending, in the main, on the policy-maker’s aversion to catastrophic payment among poorer households.

There are many possible weighting schemes or income rank-sensitive parameters of aversion to catastrophic payment that may be developed to take this into account. Wagstaff and van Doorslaer (2001) suggest the weighting scheme

\[
w_i = \frac{2(N + 1 - r_i)}{N}
\]  \hspace{1cm} (7.5)
Equation (7.5) limits the maximum relative weight for the poorest individual to 2, and this declines by factor of \(2/N\) for each additional increase in rank and declines to \(2/N\) for the most advantaged person. But (7.5) may be considered ethically rigid in the sense that it does not allow for decision-makers' different degrees of income-rank aversion to catastrophic healthcare payments and it is not normalized to 1 as would be expected of a weighting variable. Moreover, (7.5) declines at a constant rate as rank increases which may not be a desirable property, at least in some cases.

To obviate these short-comings we borrow from the inequality literature (Duclos et al 2003, Duclos and Lambert 2000) to apply a weighting scheme that has the following desirable properties: flexibility, non-negative, rank-dependent and normalized to unity.

Let \(r\) be the proportional rank of the \(i\)th person or household in the income parade of households with \(E_i = 1\). Following Duclos et al (2003) we define a non-negative weighting variable \(w(r)\) such that

\[
\int_0^1 w(r)dr = 1
\]  

(7.6)

A popular single parameterization of (7.6) is the S-Gini (Donelson and Weymark 1983, Kakwani, 1980, Duclos et al. 2003). The functional specification of (7.6) also ensures the Dalton transfer principle such that:

\[
w(r_i) \leq w(r_j) \text{ for } r_i \geq r_j
\]  

(7.7)

Introducing a rank-dependent, non-negative parameter that ensures ethical weights are given to individuals according to their income level for the subgroup with \(E_i = 1\), \(w(r_i)\) may be specified as (Duclos et al 2003):

\[
w(r; \psi) = \psi (1-r)^{(v-1)}, \quad v > 1
\]  

(7.8)
Where: the parameter \( v \) is an index of pro-poor aversion to catastrophic healthcare payment. The larger the value of \( v \) the faster the fall in \( w \) suggesting the social decision maker’s greater pro-poor concern in healthcare financing.

Given (7.8), we can then specify an index of weighted headcount catastrophic payment as:

\[
wh_{cat} = \frac{1}{N} \sum E_i w(r, v)
\]  

(7.9)

In this case, if those who are poor tend to pay more ‘catastrophically’ for health than the rich, we expect index value of (7.9) to be higher than (7.3).

The weighted catastrophic gap index may similarly be expressed as

\[
wG_{cat} = \frac{1}{N} \sum S_i w(r, v)
\]  

(7.10)

If the poor have the greater tendency to ‘overshoot’ the catastrophic threshold than the rich, then (7.10) will register a higher value than (7.4).

### 7.4.2 Results of the Estimations of Catastrophic Payments

The adoption of a flexible parameter that captures the policy-maker’s aversion to inequality implies different weightings that reflect the extent of this concern for catastrophic payment among the poor. Figure 7.7 shows these different weightings reflected in the slope of the different values of the aversion parameter \( v \). It is easily noticed that the slope of \( w \) increases as we move from 2 to 3 and 5 implying that increases in \( v \) reflect increasing concern for catastrophic payment among the poor.
When these weighting factors are applied to the catastrophic headcount and catastrophic gap payments (when the indicator of catastrophe is defined at 5% and 10% levels of income) we obtain the concentration curves shown in figures 7.7 and 7.8 respectively. The concentration curves for most parts lie above the diagonal indicating that the inclusion of these concerns imply that more of the poor have greater tendency for catastrophic healthcare financing than the rich, and this tendency increases with the level of the social decision makers concern for catastrophic payment among the poor.

Table 7.2 Incidence and Intensity of Catastrophic Payments

<table>
<thead>
<tr>
<th>Index</th>
<th>Est. Value (H)</th>
<th>Index</th>
<th>Est. Value (H)</th>
<th>Index</th>
<th>Est. Value (H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{ca},5%$</td>
<td>29.15</td>
<td>$H_{ca},5,\nu=2$</td>
<td>31.50</td>
<td>$H_{ca},10,\nu=2$</td>
<td>22.16</td>
</tr>
<tr>
<td>$G_{ca},5%$</td>
<td>5.67</td>
<td>$H_{ca},5,\nu=3$</td>
<td>33.15</td>
<td>$H_{ca},10,\nu=3$</td>
<td>22.45</td>
</tr>
<tr>
<td>$H_{ca},10%$</td>
<td>21.75</td>
<td>$H_{ca},5,\nu=5$</td>
<td>34.67</td>
<td>$H_{ca},10,\nu=5$</td>
<td>22.12</td>
</tr>
<tr>
<td>$G_{ca},10%$</td>
<td>4.40</td>
<td>$G_{ca},5,\nu=2$</td>
<td>4.68</td>
<td>$G_{ca},10,\nu=2$</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$G_{ca},5,\nu=3$</td>
<td>4.43</td>
<td>$G_{ca},10,\nu=3$</td>
<td>3.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$G_{ca},5,\nu=5$</td>
<td>4.07</td>
<td>$G_{ca},10,\nu=5$</td>
<td>2.68</td>
</tr>
</tbody>
</table>
Fig 7.7 Curves of Hcat. and Gcat. (Cat. = 5\%) for Different Values of \(v\)

Figure 7.8 Curves of Heat and Gcat (Cat. = 10\%) for Different values of \(v\)
In general, we observe a high incidence of catastrophic healthcare financing in the population. For example, it is noticed that prior to introducing the policy-maker's aversion to catastrophic spending among the poor, the incidence of catastrophic headcount is about 29% at the 5% threshold. It is natural to expect that as the catastrophic threshold increases, increasingly less proportion of the people will be found to finance healthcare catastrophically. This is evident from Table 7.2. It is noticed from column 2 that more people are caught up in catastrophic financing when the threshold is 5% than when it is defined as 10% of gross expenditure. This holds for both catastrophic headcount and catastrophic gap indexes. If the social decision maker's pro-poor aversion parameter (ν) is raised to 2, this increases the incidence of catastrophic headcount and catastrophic gap marginally to 31.50% and 4.68% respectively. As the value of ν increases to 5 this is accompanied by significant increases in the headcount reduction in the gap indexes. The behaviour of the indexes as the threshold is raised to 10% is shown in the last column of the Table 7.2. There is a general decline in the values of the indexes though it is important to observe that at this threshold, the impact of the ν is less pronounced. In fact, there seems to be a gradual reversal of this impact as we move from the ν = 2 to ν=5 for the headcount index while the decline in the gap index continues over the range.

The important conclusion from the above analysis is that the unweighted concern for catastrophic healthcare financing does not provide the complete picture of the impact of healthcare financing on households' finances. The level of the social concern for the impact of healthcare financing on the poorer segments of the society is also important. In other words, a cent from the poor to the health system is not the same as a cent from the rich and, therefore, an equity-oriented healthcare financing system must be sufficiently progressive on prepayment income. This concern arises from social empathy and concern when ill-health which is a random event takes relatively more from the poor than from the rich. This important result forewarns that an analysis of redistributional effects of welfare financing, analogous to equity concerns in taxation, must anticipate different degrees of ethical concerns for social, but particularly, economic inequality. In other
words, a consideration of the redistribution impacts of social programs needs an
environment of ethical flexibility, by which is meant a scope for incorporating social
ethical desires into inequality measures.

7.5 The Impoverishing Effects of healthcare Financing

In this last section of this chapter we briefly consider the impoverishing effect healthcare
financing has on households in the population. The basic economics behind this,
following Grossman-Wagstaff model, is that the cost of healthcare is the amount of other
household other basic consumptions that must be foregone in order to purchase
healthcare. In other words, the shadow prize of healthcare financing in this case where
households are assumed to bear the full cost of their treatment are the other basic needs of
the household. It is therefore, hypothesized that financing healthcare could push
households just above the poverty line into poverty and those already in it deeper into
poverty (Wagstaff and van Doorslaer 2001).

To investigate the extent of this phenomenon in the population of study, we use the
simple head count, the poverty gap, and the square poverty gap indices. The head count
poverty gap index is defined as the fraction of the population below the poverty line
(Deaton 1997). The poverty-gap index is the measure of the population per capita deficit
of all incomes below the poverty line while the poverty gap square index like the poverty
gap index measures the population per capita deficit of all incomes below the poverty line
but is more sensitive to the distribution of income among those below the poverty line.

To derive the poverty line, we use the poverty lines constructed by Aigbokhan (2000) for
Nigeria, which used the Cost of Basic Needs approach to set the poverty line for the
entire country and for specific regions in the country using 1997 data from the Federal
Office of Statistics (FOS). The poverty line specific to the South-East region (to which
Enugu state belongs) is in terms of the national currency, the naira 1379.39 (N1379.39).
In order to take into account the rate of inflation, from 1998 to mid year 2004, the
poverty line was adjusted by the rate of inflation derived from the national consumer
price index (CPI) for the years (FOS, 2004): 10%; 6.6%; 6.9%; 18.9%; 12.9%; and 14% for the respective years. This gives a poverty line of approximately N2900 for the region.

In order to estimate the headcount poverty index \( P_0 \) the poverty line may be denoted by \( z \) while our measures of welfare are the pre- and post-payment incomes denoted as \( x^b \) and \( x^a \) respectively. The headcount index for pre-payment poverty is then specified as:

\[
P_0 = \frac{1}{N} \sum_{i=1}^{N} 1(x^b_i \leq z)
\]  

(7.11)

Where \( 1(\cdot) \) is an indicator function taking the value of 1 if \( x^b_i \leq z \) and 0 otherwise. Equation 7.11 gives the proportion of people falling below the poverty line. This is what is captured by \( P_0 \). A similar index could be constructed for the post payment headcount index by simply replacing \( x^b_i \) with \( x^a_i \). This is done for equations (7.12) and (7.13)

The main problem with \( P_0 \) as a poverty measure is that it is violates the transfer principle. It does not tell how deep those that are below poverty line are really below it so that if policy is transferring resources from the poor and giving to the rich \( P_0 \) does not change

The poverty gap index attempts to remedy this weakness in poverty head count by taking note of the level of contribution of each household below the poverty line to the aggregate poverty. The poverty gap index is specified as:

\[
P_1 = \frac{1}{N} \sum_{i=1}^{N} \left( 1 - \frac{x^b_i}{z} \right) 1(x^b_i \leq z)
\]

(7.12)

Thus, \( P_1 \) could easily be interpreted as the sum of all income short falls below the poverty line divided by the total population. The deeper the \( i \)th household is in poverty, the larger its contribution to the overall level of poverty. Therefore a policy that transfers income from a poor household to a non-poor household increases the population level of poverty.
Finally, for this purpose, the poverty gap square index which derives from Foster, Greer, and Thorbecke (1984) (FGT index) generalization of poverty gap in (7.12) is specified as

\[ P_\alpha = \frac{1}{N} \sum_{i=1}^{N} \left( 1 - \frac{x_i^\alpha}{z} \right)^\beta 1(x_i^\alpha \leq z) \]  

(7.13)

Where, for the poverty gap square index, \( \alpha = 2 \). The index specified in (7.13) is sensitive to poverty deficits for \( \alpha \gg 1 \). The higher the value of \( \alpha \) the more sensitive is \( P_\alpha \).

To reveal the level of poverty induced by direct healthcare financing in Nigeria (as exemplified in Enugu state) using each of these measures, we simply estimate each of the poverty indices in terms of before and after the healthcare payment. We assume that the difference represents the level of poverty induced by healthcare payment. To see also the sensitivity of the impoverishing effect to household composition and economies of scale, each of the indices is applied also to three assumptions about income utility as described in subsection 5.15.2 above.

**Results: Impoverishing Effects of Healthcare Payments**

Table 7.3 summarizes the findings from this estimation exercise. The table shows that the proportion of the people that lived in poverty before healthcare payment is about 0.57 (or 57%). Healthcare financing induced further increases in headcount index from 0.57 (i.e., under per capita assumption of household income distribution) to about 0.61 (or 61%). This gives an increase of 7%. For all the three indices, headcount, poverty gap, and poverty gap squared indices, healthcare financing worsens the rate of poverty, irrespective of the welfare measure used. This is shown in the last column of the table where we see that with head count index healthcare payment worsens the level of poverty by 7% under per capita distribution of household resources. The index is worsened still if assume there are intra-household differences is allocation of resources with children getting half the size of adult equivalences and if we also assume there are economies of scale in household consumption.
Table 7.3 Impoverishing Effects of Healthcare Financing

<table>
<thead>
<tr>
<th></th>
<th>Pre</th>
<th>Post</th>
<th>Diff</th>
<th>% Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Poverty Head-Count Index $P_0$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per Capita</td>
<td>0.5725</td>
<td>0.6139</td>
<td>-0.0414</td>
<td>-7.23</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.7</td>
<td>34.87</td>
<td>39.55</td>
<td>-4.68</td>
<td>-13.42</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.3</td>
<td>20.11</td>
<td>23.71</td>
<td>-3.6</td>
<td>-17.90</td>
</tr>
<tr>
<td><strong>Poverty Gap Index $P_1$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per Capita</td>
<td>0.27107</td>
<td>0.299935</td>
<td>-0.02887</td>
<td>-10.65</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.7</td>
<td>0.135468</td>
<td>0.159544</td>
<td>-0.02408</td>
<td>-17.77</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.3</td>
<td>0.0637</td>
<td>0.080208</td>
<td>-0.01651</td>
<td>-25.92</td>
</tr>
<tr>
<td><strong>Square Poverty Index $P_2$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per Capita</td>
<td>0.163423</td>
<td>0.185497</td>
<td>-0.02207</td>
<td>-13.51</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.7</td>
<td>0.069813</td>
<td>0.08522</td>
<td>-0.01541</td>
<td>-22.07</td>
</tr>
<tr>
<td>Eqvs 1-0.5 e=0.3</td>
<td>0.02887</td>
<td>0.036695</td>
<td>-0.00783</td>
<td>-27.11</td>
</tr>
</tbody>
</table>

Proceeding to $P_1$ and $P_2$ indices it is noticed again that healthcare financing contributes significantly to poverty in Nigeria under the three scale assumptions. It could be seen that even under the moderate scale elasticity of 0.7, healthcare financing worsens the poverty gap index by as much as 18% while square poverty gap index is worsened by as much as 14% even under assumption of per capita income distribution in households.

The obvious conclusion from these figures is that healthcare financing is a major source of impoverishment amount the Nigerian population. Together with information on the incidence of catastrophic financing discussed in section 7.4, it is clear that the healthcare market and healthcare financing arrangement needs urgent policy reform as part of the poverty reduction strategy of the country.
7.6 Summary

In general, this chapter has tried articulating some of the salient concerns that will be the focus of the next two chapters. The chapter aims at providing some 'soft' information about the issues to be discussed subsequently. It provides basic statistics from the data about healthcare payments, its distribution and how this correlates with income (proxied by gross expenditure) of households. It also provides some information about the 'market powers' of the various forms of healthcare providers, the common practices that act as limitations in the healthcare market, the problems of catastrophic financing and impoverishing effects of healthcare payments in the population. In the next two chapters, the core issues about equity and redistributive effects of the healthcare system in Nigeria will be analyzed using more complex models.
Chapter 8

The Aronson-Johnson-Lambert Decomposition Framework

8.1 Introduction

The analysis in this section assumes that healthcare payments impose expenditure burdens on households in the same manner as indirect commodity taxes do. As This appears to be a plausible assumption since, like the imposition of taxes a household has little or no room to escape from the obligations of financing its healthcare needs when illness strikes. This is even more so in the absence of third party interventions that may provide households some degree of freedom to exercise preferences with respect to the method of financing healthcare obligations. In this case, the household is constrained to a trade-off between healthcare needs and other goods (Grossman 1972, Wagstaff 1986). As Arrow (1963:959) notes, medical expenditure acts as random deductions from the household income and it is the net income at the post deduction that is of interest. It is, therefore, reasonable to use the same tools of economic analysis used in the tax literature in the analysis of the effects of healthcare financing on income distribution.

Like the imposition of tax, healthcare expenditure may bring about redistributive effects in the sense that inequality in the post tax or post-payment income may deviate significantly from the level of inequality in the pretax income. Households' income ranking may alter on account of the imposition of tax just as in financing healthcare. In the absence of risk pooling mechanisms such as social or private health insurance or substantial healthcare subsidies guaranteed by the government, there are several reasons to expect changes in the ranking of households resulting from unequal contributions of households to the healthcare system. For one thing, illness is a random shock to the household that is not evenly distributed across households, and so requiring differences in
healthcare expenditure among households. For another, even if illness is equally
distributed, individual household preferences with respect to means of restoration to good
health would still differ. Thus, it is possible that two households of similar characteristics
and on the same level of income at prepayment income may end up on different income
levels at the post-payment income. A household may move up and another moves down
the income ladder at the post-payment period. It was shown in section 4.5 above that
Musgrave and Thin (1948) developed a simple measure of this redistributive effect
represented by

\[ RE = G_x - G_{x-T} \]

In section 4.7, however, it was shown that the size of the average tax rate also determines
the extent of redistributive effects. This resulted from the Reynolds-Smolensky
decomposition of redistributive effect as:

\[ RE = \left[ \frac{g}{1 - g} \right] K \]

This suggests that redistributive effect is a function of average tax and Kakwani measure
of progressivity (discussed in section 4.6), but it does not take account of the fact that
due to differences in characteristics (for example, marital status, presence of children, and
other civil considerations), the tax system may impose different tax liabilities on two
individuals on exactly the same income level so that, in the final analysis, the tax function
may be specified as: \( T_x = T_x(x) + \epsilon(x) \). The last term on the right hand side is a
disturbance term that should have a zero mean at each income level. However, due to
heterogeneity within classes of prepayment equals, it turns out to be a measure of the
deviation of the tax liability of an individual from the average tax liability. If \( \epsilon(x) \neq 0 \)
then it implies the presence of horizontal inequity which violates the maxim that requires
that equals be treated equally. Above all, it may also happen that pretax subgroups may
overlap at the post payment period due essentially to tax liability causing some people to
slip below their prepayment income inferiors, and some to rise above their prepayment
income superiors in the ranking after tax. This is the key to the Aronson-Johnson-Lambert (AJL) framework for decomposing the Gini coefficient.

8.2 The AJL Decomposition Method

\[ x - T(x) - z(x) \]

![Diagram showing the AJL Decomposition Method]

Fig 8.1: Horizontal Equity and Reranking: Source: Aronson et al (1994)

The underlying concern in the AJL framework is articulated in Lambert and Aronson (1993) where they state: “When the Gini coefficient of income inequality is decomposed into between and within-groups contributions, a troublesome and little-understood residual term arises if the subgroup income ranges overlap” (p.1221).

Figure 8.1 shows the appearance of horizontal inequity and reranking effects that follow the unequal treatment of prepayment equals. The spread or fans show the appearance of horizontal inequity among prepayment income equals. The over-lapping show “those troublesome and little-understood residual terms” (reranking effects). The figure illustrates the case of three income groups or income bands. The post tax incomes of
households in a given prepayment income band are grouped around the mean given as \( x_i - T(x_i) \) which is indicated by a horizontal line from the function \( x - T(x) \) to the vertical axis. According the Aronson et al (1994), the appearance of the fans in the post-payment distribution shows unequal treatment of prepayment equals. The fans are induced by the payment system. That is to say, the model assumes that prepayment equals were really equals. This point is very pertinent because in practice it is difficult to find in a microdata distribution two individuals that are exact equals, though this is not unlikely in a large population.

The vertical redistributive effect arises from the fact that tax liability \( T(x) \) is an increasing function of \( x \) such that: \( 0 \leq T(x) \leq 1 \). A marginal tax rate that is greater than unity would obviously lead to reranking since it would imply that at the top of the income distribution marginal tax exceeds marginal income.

In an important theorem Aronson et al (1994) established that

\[
RE = V - H - R
\]  

(8.1)

That is, redistributive effect \( (RE) \) is determined by vertical equity and classical horizontal equity, and reranking effect. But vertical equity is in turn composed of two separate effects – average tax level \( (g) \), and Kakwani index of progressivity \( K \). It follows that in the absence of horizontal inequity and re-ranking, redistributive effect reduces to

\[
RE = \left( \frac{g}{1 - g} \right) K
\]  

(8.2)

However, in the presence of classical horizontal inequity and reranking effect, the redistributive effect may be more fully expressed as (Aronson et al 1994)

\[
RE = \left( \frac{g}{1 - g} \right) K - \sum a_x G_{F(x)} - (G_{X-T} - C_{X-T})
\]  

(8.3)
The first term on the right of equation (8.3) estimates the level of inequality reduction that would have obtained had everyone within a given income bracket made equal contribution to the healthcare financing system, (that is, the counter factual reduction in inequality under equal payment). This counterfactual reduction can easily be computed by noting that (following van Doorslaer et al., 1999)

\[
\frac{RE}{RE} = \left(\frac{g}{1-g}\right) K - \left[\sum a_i G_{i(x)} - \left(G_{X-T} - C_{X-T}\right)\right]\]

\[
\Rightarrow \frac{RE}{RE} = \frac{V}{RE} - \frac{(H + R)}{RE}
\]

\[
\Rightarrow V^{100} = 1 + \frac{(H + R)}{RE}
\]

The expression \(V^{100} = 1 + \frac{(H + R)}{RE}\) makes it clear that \(\frac{(H + R)}{RE}\) is the percentage by which the payment system would have been more redistributive if everyone at a given prepayment income bracket had made equal contribution to the healthcare system. That is, the payment system would have been more redistributive by \(\frac{(H + R)}{RE}\)% if everyone within an income band paid the same amount. This counterfactual reduction in income inequality is composed of progressivity of payment \(K\) (Kakwani 1977) and the average level of taxation \(g\). The Kakwani index is obtained as the concentration index (CI) of healthcare payment using the average payment made by individuals in a pre-payment income class rather than each individual’s actual payment and then plotting it against the fractional ranks of individuals in the prepayment income. The CI can also be obtained using the OLS ‘convenient’ regression (Jenkins 1988, Kakwani et al. 1997) given by:
\[
\frac{2\sigma^2_y}{\bar{y}} \bar{y}_i = \alpha_1 + \beta_1 r_i + \varepsilon_1 i;
\] (8.5)

Where \( \sigma^2_y \) is the variance of rank \( r \), \( \bar{y} = \) mean \( y \), and \( r \) is the fractional rank of the ranking variable (e.g. income), \( \hat{\beta}_1 \) is the OLS estimate of CI.

The second term in (8.3) is an index of classical horizontal inequity:

\[
H = \sum a_x G_{f(x)}
\] (8.6)

Where \( f(x) \) denotes class of equals at point \( x \) in the distribution of the prepayment income. If it is assumed that income equals at point \( x \) at the prepayment income were really equal then unless they make equal payments to the tax or health system, they would experience horizontal inequity within the ‘band’ at the post-payment distribution. This within-class inequity at the post-payment income distribution is local HI. Aggregating these local HI for all the groups will give rise to global HI. The problem however, is: what should be the nature of this aggregating function? It would seem that the ‘appropriate’ aggregating function would require some form of normative judgment in deciding the weighting scheme for aggregation of local HI across the levels of \( x \). This is a decision the policy-maker has to take depending on the social aversion to inequality at the different levels of the income distribution. If the policy-makers feel, for example, that HI among the low income groups is more damaging to the social fabric because it would create greater resentment among this group, they may place more weight on local HI arising at the lower income individuals. They may, on the other hand, feel that HI among the middle class is more damaging to the economy because it would dampen work–effort among this important segment. Which ever is the case, the weighting function has to reflect these social preferences.

In the AJL framework, the weighting function assumes the implicit weight of the Gini index as shall be discussed very shortly. The term \( G_{f(x)} \) is the Gini coefficient that measures inequality that arises in the post payment distribution solely from the fact that
individuals at the same pre-payment income level are now less equal in the post payment than they were assumed to be at prepayment period. This is because households or individuals at the same pre-payment income level have contributed unequally to finance the healthcare system. The weight $a_x$ is the product of the population share of $i$th prepayment income band and the post-payment income share of the band. Thus, in this model, global HI is the summation of $G_{l(x)}$ weighted by $a_x$, at every level of $x$.

Obviously $a_x$ is not the only feasible weighting scheme. Nor is the Gini index the only feasible measure of inequality that could be used. It is for example possible to use the population weighted Gini or Atkinson Index of inequality. But this is the one chosen by the AJL framework to aggregate the local HI indices. Implicit in the choice of Gini as the index of inequality under the framework is an underlying ethical preference. It is well known that the Gini index measures the deviation of the distribution of $x$ from perfect equality. In fact, the Gini is often defined as the average difference between all possible pairs of incomes in the population expressed as a percentage of the total income (Cowell 1995:23). However, in doing so it does not matter where the deviation occurs, whether at the lower or higher income levels. All deviations are equally weighted irrespective of where they occur in the distribution. In other words, the model constrains HI within any given income bracket to induce the same level of loss in social welfare. This implies that the Gini is an ethically neutral measure of inequality. Policy-makers that decide to use the AJL framework for policy guidance must, therefore, also assume the implied ethical content. If not, they must have to replace the social evaluation function implied by the Gini with one they consider more appropriate. However, as Auerbach and Hasse (2001) suggest, if HI must have a different content from VE, then it seems that attitude to it must be different. That is, it does matter at what level of income inequality occurs.

The choice of the Gini as the index of the inequality under the AJL framework also implies that that total redistributive effect cannot be straightforwardly decomposed into the between-group and within-group contributions to overall inequality without the ‘troublesome and little-understood’ residual referred to above. This is the additional term $R$ in equation (8.1) which measures the extent of reranking consequent upon payment.
$$R = G_{X-T} - C_{X-T}$$  \hspace{1cm} (8.7)

Where: $G_{X-T}$ is the Gini coefficient of the net income distribution. $C_{X-T}$ is the concentration index of the net income where the ranking variable is the prepayment income. Reranking occurs because illness is a random phenomenon. A severe illness may warrant such spending that forces a household to move down in rank below its prepayment income inferior as discussed in section 7.4 above. The extent of re-ranking is shown by the difference between the Gini index for post payment income and the concentration index for post-payment income. Where $R$ takes the value of zero, then we know that no re-ranking actually took place in the transition from pre-payment to post payment periods.

Van Doorslaer et al (1999) observe that positive values of $H$ do not necessarily imply positive values of $R$. This observation is correct only in so far as it is assumed that 'prepayment equals' are actually equals and that they are only made unequal by the payment system at the post payment distribution. In other words, it is being assumed that there are no heterogeneity among 'equals' or that the data is grouped data. If, however, they are not actually equals or there is presence of heterogeneity in the prepayment distribution within 'equals' and now change ranks at the post-payment distribution, then it would seem to suggest that every instance of positive value for $H$ would also imply a positive value for $R$.

The appearance of $R$ in the post-payment distribution and whether it should be regarded as a separate index of HI or whether it is only an extension of the later has generated considerable debate in literature and is deeply rooted in the theories of justice discussed in Chapter 3. Feldstein (1976) lists non-re-ranking as one of the important properties of equitable tax. He argues that the tax system must preserve the pre-tax utility rankings of individuals. The underlying ethical assumption here is the libertarian argument that the pre-payment income distribution generated by the market is fair and the policy-maker should not use the tax or healthcare payment system to redistribute income. However,
Lerman and Yitzhaki (1995) argue that post-payment distribution is a better reflection of the distribution desired by the policy-maker and should therefore be seen as the benchmark for measuring the extent of reranking implied by the payment system. According to Plotnick (1981) rank reversals indicate the extent of HI in the transition from pre- to post-payment distribution.

The two terms $H$ and $R$ are non-negative. For example a tax rate in excess of 100% may lead to re-ranking even without horizontal inequity (Wagstaff and Van Doorslaer 2001). However, since such taxes do not operate in reality, horizontal inequity will lead to reranking. But again, it may be noted that in grouped data, income levels may be demarcated as bands of income. This implies that positive value of $H$ may not necessarily imply a positive value of $R$ in grouped data since healthcare payment differences between two individuals within the same income band may not be sufficient to re-rank them. For example, if household A has N530 and B has N540 at prepayment income distribution and both are within the income band defined as N500 to N550. Assuming A incurs a health finance cost of N10 and so has N520 at the post-payment and B incurs a health cost of N30 and so has N510. Within the AJL framework horizontal inequity has occurred but no reranking has taken place. From equation (8.4) it is clear that model provides two degrees of freedom. Having computed the value of $V$, we can either compute $H$ or $R$ first, or compute the other as residual.

To compute $K$, and therefore, of $V$ under the AJL framework, it is important to determine the question: who are ‘income equals’? This is identification problem and it is obviously a critical issue since the estimated local and global HI will be, among others, a function of the definition of ‘income equals’. At the same time it is easy to see that a discrete distribution with little gaps between income groups could easily lead to a situation where a little change in the income of an individual leads to a jump in the membership of income group. In other words, the values of $V$ which depend directly on $K$, and $H$ and $R$ which depend indirectly on $K$ are not likely to be robust to alternative definitions of income group. The robustness of the value of $V$, $H$ and $R$ will be tested using different definitions of income bands or ‘income equals’.

184
It is clear from the AJL decomposition method that the total contribution of healthcare payment system to income inequality can be decomposed into four main components: (i) vertical equity component represented by the degree of progressivity of the healthcare financing system. If payments are progressive on prepayment income then we conclude that the healthcare payment system exerts equalizing effect on post-payment income given the average payment. That is there is greater parity in households' ability to purchase other basic life needs at the post-payment period. (ii) The percentage of total income that is used to finance healthcare. This is represented by g. Following Podder (1993) a rise in proportion of income devoted to healthcare payment raises the degree of vertical equity and fall in proportion of income devoted to healthcare similarly decreases inequality. (iii) When households are grouped according to their income categories, the horizontal inequity (i.e. within category inequity) arising from the health financing system is estimated by the level of inequality in the post-payment income.

This within category inequity in post-payment income is estimated by the weighted within-group Gini coefficients. The population-weighted sum of the within-group Gini coefficients gives the level of horizontal inequality $H$ in the post-payment distribution. Thus the $H$ has a positive sign (since within-group Gini coefficient at the post-payment period cannot be less than zero, some groups may have positive Gini in the post-payment distribution) so that $H$ necessarily reduces the redistributive effect just as $R$.

8.3 Estimation Procedure and Modelling Issues

The redistributive effect, (RE) is computed as the difference between prepayment and post-payment expenditure Gini indices ($G_x - G_{x,T}$). For the purpose of calculating the Gini indices the 'convenient covariance' method as suggested by Jenkins (1988) is used. This is given by

$$G = 2 \text{cov}(x_i, R_i) / \bar{x}$$
Where: $x_i$ is the income (expenditure) of the $i$th unit and $R_i$ is the fractional rank of that unit on the income parade, and $\bar{x}$ is the mean of $x_i$ variable. The same computation could be carried out using the integration method or, more straightforwardly, by using the DAD software. But using the later would not allow us to obtain some components of the AJL decomposition. Moreover, DAD has no program for computing the components of redistributive effect (K, V, H, and R) based on the AJL method. However, the DAD package was used to obtain the Gini indices. This makes it possible to obtain standard errors (Se) and confidence intervals (CI) of the estimates of $G_X$ and $G_{X-T}$.

In order to obtain the value of vertical inequality $V$ in the model as specified in equation (8.1) above the values of $g$ and $K$ are first computed and then $V$ computed as: $V = [g/(1-g)]K$. Where $g$ is simply the ratio of healthcare expenditure to total expenditure and $K$ is the Kakwani (1977) index of progressivity discussed earlier. The value of $K$ is obtained from the model $K = C_T - G_X$. Where $C_T$ is the concentration curve for payment computed on the assumption that each household within a defined ‘income equal’ pays the same amount for healthcare. The concentration curve is obtained in analogous manner as the Gini using the ‘convenient covariance’ method. Thus $C_T = 2\text{cov}(f_{i(x)}, R)/\bar{T}$

Where, $f_{i(x)}$ is the average healthcare payment made by ith pre-payment equals and so varies according to prepayment income or expenditure band. $\bar{T}$ is the overall mean payment by all prepayment expenditure bands. The fractional rank $R_i$ in this case refers to the ranking of the individual unit by their prepayment income. The difference between $C_T$ and $G_T$ gives the Kakwani measure of progressivity. The product $[g/(1-g)]K$ gives the extent of vertical inequity in the distribution of payment.

---

1 For the implementation, ranking is done using the RANK command in SPSS. To calculate the covariance, the Correlation command with the Covariance option in SPSS is used to generate the covariances. Similar procedure is used to compute the concentration indices.
For the estimation of the horizontal inequity (H), it is recalled that 
\[ H = \sum a_c G_{f(c)} \]

Where:

\[ a_c = \text{[proportion of population in ith prepayment income class]} \times \text{[proportion of post-payment income attributable to the ith prepayment income class].} \]

Obtaining the post-payment income attributable to the ith pre-payment income class is sorted out using the *Stata* software. Depending on the particular width of the income class, prepayment incomes are first grouped according to this definition of class. The classes are sorted and used to tabulate the post-payment income in order to obtain the post-payment income attributable to the ith pre-payment income class. \( G_{f(c)} \) is the Gini coefficient of post-payment income distribution in the ith prepayment income class. The summation of the products of \( a_c \) and \( G_{f(c)} \) gives the extent of the horizontal inequity engendered by healthcare payments in this framework. The re-ranking variable \( (R) \) was then calculated as a residual but could also be calculated as shown in equation (8.7). In this case, any of H or V could be computed as a residual.

It is worth emphasizing that since point estimates of parameters are affected by the weighting of observation in the sample, precaution was taken during the design and implementation of the survey to ensure that the sample was self weighted. This was achieved and so in the estimation the only weights applied are the households’ weights. The stratification and clustering designs of the survey as noted in section 5.14.3 do not affect the point estimates but they are liable to affect the test statistics like the standard errors, t-values, and confidence intervals. In the estimations carried out under the AIL framework standard errors could only be computed for \( G_X \) and \( G_{X-1} \) because they are, in this case, computed using the *DAD* program. The rest of the parameters in the model are computed ‘manually’ as described above and only the point estimates are provided. The
confidence intervals of $G_x$ help to establish the statistical significance of the quantity 
$G_x - G_{x-r}$ and, therefore, also of the total redistributive effect $RE$.

8.4 The AJL Model Results

The first set of results from AJL model is presented in Table 8.1. It is computed based on 
the assumption that household resources are distributed on per capita basis and that there 
are no economies of scale in household consumption. These assumptions would be 
relaxed subsequently when intra-household resource allocation and economies of scale 
are allowed to vary. The parameter estimates are obtained under different definitions of 
prepayment 'income equals' or income bands. Since in the empirical distribution of 
prepayment there may actually be no exact prepayment income equals, the AJL 'income 
equals' may be more appropriately be regarded as only 'close equals'. Under the model 
there are no clear cut parameters for defining these income classes although ideally the 
band should be as narrow as possible and, in the limit, approach zero.

In the results set out in Table 8.1, the decomposition parameters are estimated under 
different definitions of 'income bands'. Column 2 presents results estimated using a 
definition of income as 10 percent or deciles of the population incomes. In this case, 
given the total number of observations as 1497, there are, on average, about 150 
observations under a given income band. Column 3 shows results estimated using 5 
percent income band. In this case there are, on average, about 75 observations in each 
income band. The 3.3 and 1 percent income bands imply there are about 50 and 15 
observations, respectively in each income band.

The results shown in Table 8.1 clearly indicate a large amount of income inequality in the 
state. The prepayment Gini Index ($G_x$) is about 0.45 under income per capita and 
assumption of no economies of scale in household consumption. It can be seen that this 
value is not affected by the definition of income equal under the frame work. This is not 
the case for the estimated core parameters of the distributive effect: $V$, $H$, $R$ as will be
discussed shortly. The value of $G_x$ is, therefore, constant across the different income bands. The post payment income Gini index ($G_{x-T}$) is also constant across income bands. Likewise the parameters $RE$ and $g$. These parameters are computed prior to the estimation of $K$ where the definition of income band begins to exert influence.

The estimated prepayment Gini Index is 0.45 on the $[0, 1]$ interval. It is statistically different from zero as shown by the estimated standard error. It is however lower than the unweighted estimate of 0.47.

Table 8.1 Results from AJL Decomposition Based on Per Capita Income

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Income-bands: 10 percentiles</th>
<th>Income-bands: 5 percentiles</th>
<th>Income-bands: 3.3 percentiles</th>
<th>Income-bands 1 percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_x$</td>
<td>0.4474**</td>
<td>0.4474**</td>
<td>0.4474**</td>
<td>0.4474**</td>
</tr>
<tr>
<td>Sc</td>
<td>0.0111</td>
<td>0.0111</td>
<td>0.0111</td>
<td>0.0111</td>
</tr>
<tr>
<td>Cl</td>
<td>0.4691</td>
<td>0.4691</td>
<td>0.4691</td>
<td>0.4691</td>
</tr>
<tr>
<td></td>
<td>0.4258</td>
<td>0.4258</td>
<td>0.4258</td>
<td>0.4258</td>
</tr>
<tr>
<td>$G_{x-T}$</td>
<td>0.4482**</td>
<td>0.4482**</td>
<td>0.4482**</td>
<td>0.4482**</td>
</tr>
<tr>
<td>$RE$</td>
<td>-0.0008</td>
<td>-0.0008</td>
<td>-0.0008</td>
<td>-0.0008</td>
</tr>
<tr>
<td>$G$</td>
<td>0.0883</td>
<td>0.0883</td>
<td>0.0883</td>
<td>0.0883</td>
</tr>
<tr>
<td>$C_T$</td>
<td>0.5754</td>
<td>0.5875</td>
<td>0.5875</td>
<td>0.5893</td>
</tr>
<tr>
<td>$K$</td>
<td>0.1280</td>
<td>0.1401</td>
<td>0.1416</td>
<td>0.1419</td>
</tr>
<tr>
<td>$V$</td>
<td>0.0124</td>
<td>0.0136</td>
<td>0.0137</td>
<td>0.0138</td>
</tr>
<tr>
<td>$H$</td>
<td>0.0086</td>
<td>0.0070</td>
<td>0.0042</td>
<td>0.0009</td>
</tr>
<tr>
<td>$R$</td>
<td>0.0045</td>
<td>0.0074</td>
<td>0.0103</td>
<td>0.0137</td>
</tr>
</tbody>
</table>

NB ** Estimated parameter statistically significant at 5% level

It is comparable to the estimated national Gini Index of 0.50 reported in other studies (see subsection 2.4). The result is not surprising considering that the state is culturally and economically less heterogeneous than the entire country. The result indicates that there is a high level of income inequality among the population. The estimated Gini coefficient is also high relative to that for developed and most developing countries even in Africa. The
Gini index for the former group of countries lies within the range of 0.20 and 0.30, and the latter countries within the range of 0.30 and 0.40.

The table also shows that the value of $G_{x-T}$ lies within the lower and upper bounds of the confidence interval of $G_x$. This suggests that the difference $(G_x - G_{x-T})$ is not statistically different from zero. Hence the redistributive effect, though indicating a regressive financing system (pro-rich redistributive effect) is nevertheless statistically insignificant so that the sign, though negative is actually indeterminate.

The conclusion from this is that under the assumptions of the AJL framework, (particularly under the implied assumption of inequality neutrality of the Gini index of this framework), the direct method of paying for healthcare in Nigeria does not lead to any significant income redistribution. Rather households tend to finance their healthcare needs according to their level of income.

A comparable result was also observed by Wagstaff, van Doorslaer et al (1999) among some EU countries where they found low degree of regressivity in out-of-pocket payment in Germany, the Netherlands and Ireland. But this result was there explained as reflecting the absence of or incomplete social health insurance cover for insurance for the better-off. It would now seem, in the light of this further evidence from a predominantly market-driven health market that out-of-pocket financing tends to force households to buy health only in proportion to their income without much of redistributive implications. In any case, the important conclusion from this result is that healthcare financing in Nigeria is largely determined by market forces and the principle of effective demand. Households seem to purchase healthcare services in direct proportion to their income or their ability to pay for those services if this ability is proxied by household income.

The parameter $g$ is analogous to the average tax and indicates the proportion of households’ income that finances healthcare services. Here $g = 0.9$, indicating that an average household in the population spends about 9% of its income on healthcare
services. This, again, is a very high proportion when compared with similar estimates (Wagstaff and Van Doorslaer, 2001) of 5.7% and 5.9% taken up by out-of-pocket payment in Vietnam in 1993 and 1998 respectively. Gerdtham and Sundberg (1996) using the Swedish data for 1980 and 1990 found the value of $g$ to be 12% and 14% respectively. However, the parts of these sums borne directly by the individual household were only 0.26% and 0.19% respectively. Thus, the issue is not that the value of $g$ per se is very high. Indeed it is small in relative terms but the issue is that this proportion is coming through direct financing. The obvious implication is that, as was suggested in Chapter 2, in the absence of significant and effective government intervention in healthcare financing the average Nigerian household bears almost entirely the full cost of its health needs. This finding also strengthens the evidence of large incidence of catastrophic healthcare financing as discussed in Chapter 7.

The relatively high value of $g$ not only confirms the virtual dependence of households on direct financing of their health needs but also signals that health financing is a potential source of impoverishment among the population. Healthcare financing is a big burden to most households in Nigeria (see also Gotsadze et al (2005). And the situation is made worse in a cash economy where, in the absence of third party intervention and significant subsidies, household must deposit money for any form of medical care. But as Wagstaff and van Doorslaer (2001) suggest, the policy-maker’s concern is not only in respect of access of the poor to healthcare utilization but also the impact healthcare expenditure has on the household’s ability to provide itself other basic needs. The large value of $g$ coming from out-of-pocket is a measure of the battle households have to wage in the trade-off between health and basic needs in the face of meagre income.

While the parameters $G_x, G_{x,t}, RE$ and $g$ are invariant to the definition of prepayment equals, payment concentration index is not. $C_r$ is a measure of inequality that indicates the extent of inequality in one variable in relation to ranking in another. It is computed on the assumption of effective payment schedule; that is, the inequality in average payment charged on each group of prepayment equals (see Aronson et al. 1994)
in relation to ranking of individuals according to prepayment income. In other words, this method of computing \( C_T \) eliminates both the influence of horizontal differences within each pre-payment income group (since each household within the group is charged the average payment of the group) and all measurement errors. The estimated \( C_T \) will therefore vary according to the definition of prepayment income equals adopted. The estimated \( C_T \) for decile definitions of pre-payment equals is about 0.57 and increases rather slowly as the size of the income band shrinks.

### 8.5 Interpretation of \( C_T \)

Koolman and van Doorslaer (2003) provide an interpretation that helps give an intuitive meaning to the estimated \( C_T \) index in natural units (see section 6.8 above). According to this interpretation, our index value of 0.57 implies that a lump sum of redistribution of 57% of the total amount of healthcare payment is required from the richest half of the payment parade to the poorest half of the parade in order to achieve equality of payment. Noting that the total healthcare expenditure for our sample of 1497 households is about N1588158.00, under the assumption of per capita distribution and absence of economies of scale, this would imply a lump sum of N905250.06 available for redistribution among the poorest half of the payment parade.

This sum appears to provide a good insight into the magnitude of redistribution required though it does not provide information on details of how this sum is to be distributed within the lower half of the income parade, nor does it assume that such a transfer from the richer half to the poorer half will automatically reduce \( C_T \) to zero since their will still be inequalities even within the recipient poorer half. However, the authors suggest that a linear scheme for that ensures that \( C_T = 0 \) could be obtained by \( \hat{y}_i - \bar{y} \) for every individual; \( \bar{y} \) being the rank-predicted value from an OLS regression but, of course, in this case we are dealing with predicted rather than actual values of the distribution.
8.6 Key Components of the Redistributive Effect

Kakwani Index of Progressivity and Vertical Equity

The Kakwani index ($K$), the measure of progressivity of healthcare payment, is computed straightforwardly as the difference between $C_r$ and $G_x$ (Kakwani, 1977). Noting that the latter is computed based on the individual level data and the $C_r$ based on group observations, Aronson et al (1994) suggest that $K$, in this case gives the counterfactual Kakwani index indicating the reduction in inequality that would have obtained in the absence of horizontal inequality in healthcare financing. In the absence of horizontal equity and re-ranking considerations, $RE$ of healthcare financing is simply the vertical equity ($V$) and tells us how much income inequality is reduced on account of the progressivity (or regressivity) of the payment system since everyone faces the same payment schedule.

The level of vertical inequity is a function of the effective payment schedule $g$ (i.e., average payment rate) and $K$. In other words, vertical equity is not just a function of the progressivity of the payment schedule but also of the proportion of average households' income used to finance healthcare. A high value of $g$ may imply a high value of $V$ even if $K$ is small. In Table 8.1 as in other subsequent tables displaying the AJL results, we find that the value of $V$ is small though it may be considered high in relation to its source: household direct payment. That is, the value of $g$ is relatively high in relation to other components of household expenditure but is small in relation to what ought to be the average healthcare expenditure per capita.

The estimated value of the parameter $K$ ranges between 0.128 to 0.1419 for the 10 decile and 1 percentile income bands respectively. This is may be considered relatively low considering that the value of $K$ lies within the [-2, 1] range. The result is that the value of $V$ which varies with income bands between 0.0124 and 0.0139 is only moderate under
the AJL assumptions. This suggests that healthcare in Nigeria, under the present financing arrangement, does not contribute effectively to income redistribution given the proportion of household income it absorbs. This reinforces our earlier observation that households seem to purchase healthcare in proportion to their income not necessarily in proportion to their need and this calls for a major policy response.

**Horizontal Equity**

Turning now to the horizontal equity component $H$, we first note that this component measures the classical horizontal inequity as different from the Atkinson/Plotnick index of re-ranking. It captures the unequal treatment of individuals in equal need. It measures the post-payment unequal treatment of pre-payment equals. It is, as noted earlier, the weighted sum of the Gini coefficients (see Table 8.2) in the post-payment income of households grouped according to prepayment equals ($G_{f(x)}$).

The estimated values of $H$ vary from 0.0009 (for the 1 percentile band) to 0.0149 (for the 10 percentile band). That the values of $H$ are greater than zero imply that the financing system would have been more redistributive without horizontal inequity.

**Table 8.2 Matrix of Gini Coefficients for 5 Percent ‘Income Equals’**

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{f(x)}$</td>
<td>0.105</td>
<td>0.073</td>
<td>0.089</td>
<td>0.064</td>
<td>0.080</td>
<td>0.054</td>
<td>0.058</td>
<td>0.076</td>
<td>0.064</td>
<td>0.077</td>
</tr>
<tr>
<td>$a_{x}G_{f(x)}$</td>
<td>4.1E-5</td>
<td>4.3E-5</td>
<td>6.09E-5</td>
<td>5.5E-5</td>
<td>7.9E-5</td>
<td>6.5E-5</td>
<td>7.7E-5</td>
<td>1.2E-4</td>
<td>1.0E-4</td>
<td>1.4E-4</td>
</tr>
<tr>
<td>Percentiles</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>$G_{f(x)}$</td>
<td>0.051</td>
<td>0.080</td>
<td>0.084</td>
<td>0.046</td>
<td>0.078</td>
<td>0.072</td>
<td>0.069</td>
<td>0.084</td>
<td>1.258</td>
<td>0.349</td>
</tr>
<tr>
<td>$a_{x}G_{f(x)}$</td>
<td>1.1E-4</td>
<td>1.8E-4</td>
<td>2.E-4</td>
<td>1.3E-4</td>
<td>2.4E-4</td>
<td>2.5E-4</td>
<td>2.7E-4</td>
<td>3.8E-4</td>
<td>6.8E-3</td>
<td>3.1E-3</td>
</tr>
</tbody>
</table>
It is however difficult to say, in this case by much the system would have been more redistributive. The formula specified in (8.4) does not seem strictly applicable in the case where RE is negative.

If the variations within the net income of each prepayment income band are high, we expect \( H \) to be high. This means that as the size definition of the income band increases the value of \( H \) also increases. This is reflected in the estimated results where the wider income bands have high estimated values of \( H \). But as the size of the income band approaches zero the value of \( H \) also approaches zero (see for example to value of \( H \) at income band of 1 percentile).

It is also to be noticed from Table 8.2 that within-group or within-band inequality is highest among the poorest and highest groups of the population. This in itself provides useful information for policy aimed at targeting the most vulnerable segments of the society through provision of social programs or to tackle intra class inequalities in the top income classes where the amounts involved are highest. The within group Gini coefficients contributed more to the size of \( H \) than the weighting factor \( \alpha \), although when the income bands are wide the latter contributes considerably to the size of \( H \) as noticeable in Table 8.1.

But there are several other factors that could contribute to a high value of \( H \) through \( G_{F(s)} \). One is the differences in amounts that households are required to pay in the event of an illness. Others may include age and differences in ease of access to healthcare facilities. But, perhaps the most important factor would be the random nature of illness itself. It would, therefore, seem that some method of protecting households against unpredictable illness is desirable. The policy of social health insurance which is still in infancy in the country or some form of prepayment scheme seems to require urgent development.
Re-ranking Effect

The final component of the AJL decomposition is reranking $R$ which measures the extent of overlapping of flaps between income bands. Conceptually, horizontal inequity and re-ranking, though distinct, are very closely related and in practice are inversely related as shown by the estimated results in Table 8.1. There, it is easily seen that as the size of the pre-payment income band broadens, $H$ increases while $R$ decreases and vice versa indicating a trade off between $H$ and $R$. As the band width approaches zero, the value of $H$ approaches zero as all horizontal inequity now turn into reranking effect. All these point to the fact that the estimates of $V$, $H$, and $R$ are sensitive to the definition of income band.

It is possible that a policy-maker might be more concerned about re-ranking than with horizontal inequity resulting from healthcare financing. This is because losing one's income status among prepayment peers has a potential to generate a psychological feeling of alienation, an observation that has since been formalized into theory of relative deprivation. Thus, ethically, a social decision maker might consider re-ranking a more serious problem than horizontal inequity and this requires cushioning the effects of healthcare financing through medical subsidies particularly for those groups of the population that are more prone to catastrophic healthcare spending.

8.7 Observations

The key components of the AJL decomposition vary across definitions of prepayment income equals. Evidently, as the class size shrinks, $V$ and $R$ components increase while $H$ decreases. The movement in opposite directions between $H$ and $R$ seems intuitive. An increase in number of income bands is likely to increase the chances of overlapping between income classes at the post-payment distribution. On the contrary, the smaller the number of income bands the higher the variability within classes and, hence, increase in the size of $H$. This implies that in the context, not only $H$ and $R$ are sensitive to the definition of income equals but also $V$. We recall that in the AJL model $H$ is computed
from the Gini indices of post-payment groups who were deemed equal before payment. By reducing the number of income bands, we are also reducing the size of vertical inequity because we thereby also reduce differences between observations that may be close equals though they belong to different income bands. This observation seems to be confirmed by Lambert and Ramos (1997) but they nevertheless proceeded in very much the same way as Aronson et al. (1994) by developing the concept of 'pseudo horizontal inequity'. The significant difference between the AJL approach and Lambert and Ramos approach being the latter's introduction of a new weighting scheme, a pure weight rather than a weighting scheme contaminated by vertical considerations as in Aronson et al. In Lambert and Ramos the weighting scheme is only the population share of each prepayment income group rather than population share multiplied by post-payment income share as done in AJL approach.

This highlights, on one hand the importance of appropriate definition of pre-payment 'income equals' for policy purposes and, on the other hand, seems to emphasize the limitations of this approach to estimating redistributive effects as the AJL model leaves this important question of deciding the optimal pre-payment 'income equals' unresolved.

Furthermore, it is noticed that the level $g$ has important implications for the level of $V$, though not for $K$. If policy or changes in healthcare financing is able to alter the payment structure, for instance, this would affect the value of $V$. But does the level of $g$ also affect the levels of $H$ and $R$? This problem was first noticed by van Doorslaer et al. (1999) but still remains unresolved. It seems that $g$ enters the computation of $H$ and $R$ indirectly since any of these components can be calculated as a residual of the decomposition equation $RE = V - H - R$. Moreover, $H$ is computed from net income (after subtracting healthcare costs) which implies that the level of $g$ could as well influence the level of $H$ indirectly. It seems, therefore, that altering the average level of healthcare payment made by households could also alter the level of $H$ and $R$. It seems then, that within the AJL framework, there is no way of isolating the effect of $g$ on $H$ and $R$. If this were possible, it would assist policy in separating pure horizontal equity and re-ranking
effects from the effect arising from a general rise in the value of $g$. Both aspects of horizontal and re-ranking effect would seem to require different policy instruments to reduce.

However, the AJL model provides an important framework for thinking clearly about the different components of redistributive effect and showing that each of these components could become a separate policy objective requiring a different set of policy instruments to tackle it. In the next chapter we consider an alternative framework for analyzing redistributive effects of taxes and other social welfare financing systems. This is the Duclos et al (2003) framework. The details of this decomposition methodology is first laid out and then applied to the same data set used to generate the AJL results analyzed in this chapter. The results are then compared.
Chapter 9

Duclos-Jalbert-Araar Decomposition Framework

9.1 Introduction

The AJL approach to decomposition of total redistributive effects of taxation and transfer programs into vertical, horizontal and re-ranking effects has dominated the literature since it was first developed\(^1\). But it has also certain obvious short-comings that raise analytical concerns which were highlighted in Chapter 8. We highlight and discuss these concerns briefly to enable us situate the recent attempt by Duclos, Jalbert, and Araar (2003) (DJA) to re-establish this decomposition on a different analytical grounds. This will enable us make useful comparisons between the results obtained from both methods.

In the first place, it is noticed that the AJL model implicitly assumes a rigid ethical social welfare function that is insensitive to the policy-maker’s concern or level of aversion to inequality. To show this, we note that the Gini which is the main framework for their decomposition, as well as for DJA framework, may be the written as:

\[ G = 2 \int_0^1 p - L(p) \, dp \] \hfill (9.1)

Equation (9.1) suggests that the Gini is 2 times the area covered by the deviations of the shares of \(100p\) proportion of the population from what would have been their actual share were there no inequality in the distribution. In other words, the Gini measures the average deviation or distance of the shares of the population from the expected mean.

Note that this deficit from $p$ is not weighted, implying that the Gini is not sensitive as to whether the deviation from $p$ is among those below or above the mean income. Every deviation is equally weighted. This is analogous to the method of the mean absolute deviation (MAD), whereby each residual in the regression function enters the population regression function with its absolute distance or deviation from the expected mean. In this sense, the Gini index could be considered an ethically neutral measure of inequality. The Gini index could be said to be set in a politically rigid environment. Consider, for example, that a policy-maker who faces social uprising from the masses on account of inadequate social services that are otherwise available to the economic elites, might be more averse to inequality in distribution of social amenities as it affects the poorer segments of the society. In this case, the Gini index, and its implied neutrality about whether inequality occurs at the top or below, might be of little avail.

Second, the outcome of the redistributive effects using the AJL approach depends largely on one's definition of `income equals` or income bands. This is the identification problem which has dogged attempts to measure horizontal inequity in literature. A grouping of incomes into quintiles is as much a definition of `income equals` as a grouping based on one percentile. In other words, the definition of `income equals` seems arbitrary. The process can generate `income equal` that are clearly not equals at all, prompting Plotnick (1985, cited in Duclos and Lambert 2000) to dismiss the classical methods of measuring horizontal inequity and insisting that grouping of data in sparse samples in order to create close equals is an unsatisfactory solution to the identification problem. Yet the definition will determine to a large extent the level of HI and re-ranking effects with their consequent implications for policy. Aronson et al. themselves acknowledged this difficulty but seem to have left the solution to be determined by the researcher.

Before considering the DIA decomposition framework, it is apposite here to briefly describe other attempts that have been made to provide solutions to these two observed shortcomings of the AJL framework. Lambert and Ramos (1997a), instead of estimating HI as the sum of weighted local inequities in the Gini (as the AJL), estimated HI as the sum of weighted local mean logarithmic deviations. But the weights are simple
population or 'pure' weights. Like AJL horizontal inequity index, the Lambert-Ramos HI index is also obtained by first estimating the counterfactual vertical equity (VE).

To deal with the identification problem, Lambert and Ramos (1997a) proposed the concept of pseudo-horizontal inequity (PHI) as the spreading-out of incomes of close prepayment equals in the transition from pre- to post-tax income distribution (Lambert and Ramos, 1997b; Ramos and Lambert 2003). Denoting within local inequality in the ith income band before tax as $I_i^b$, and inequality in the same ith income band after tax as $I_i^a$, and using the mean logarithmic deviation (MLD) then the increase in local PHI induced by the payment system is defined as

$$h_i = I_i^a - I_i^b \tag{9.2}$$

The global measure of PHI is the given as

$$PH_m = \sum_i \pi_i h_i \tag{9.3}$$

Where $\pi_i$ indicates the proportion of the population in the ith income band. Like in the AJL approach, the closer the income-band width is to zero, the closer the estimate estimated PHI approaches the true population parameter. In order to determine the 'optimal' income band a viewing window of fixed width is passed across the pre-tax income distribution range in a non-overlapping manner. But the choice of the window width in the Lambert and Ramos (1997a) and Ramos and Lambert (2003) framework, like the choice of size of income band in AJL approach, is arbitrary. In the Lambert and Ramos approach, the PHI is given a normative interpretation: 'near equals' should be treated 'near equally' which seems to suggest that vertical redistribution within groups is precluded in the concept of PHI.

Auerbach and Hasset (2002, 2001) provide another perspective to the weighting and identification problems. These authors start there consideration of the problem of weighting by adopting the Atkinson (1970) utility function that reflects von Neumann-
Morgenstein preferences under uncertainty. Define $M$ to be the number of income bands (classes or groups), and $N$ to be the number of individuals in the $i$th income band, their social welfare function that imposes aversion to inequality on the after-tax income is specified as:

$$w = \left[ \sum_{i=1}^{M} \sum_{j=1}^{N} (x_i - T_{ij})^{1-v} \right]^{\frac{1}{1-v}}, \quad v \geq 0$$

(9.4)

Where $x_i$ represents the pre-tax income of individuals in the $i$th group. $T_{ij}$ is the tax liability of $j$th individual in the $i$th income band and $v$ is a measure of aversion to inequality. The greater the value of $v$, the more averse the social decision maker is to inequality. At the extremes, for $v = 0$ the social decision maker is unaverse to inequality so that the social welfare is simply the summation of individual incomes. As the value of $v \to \infty$, the social decision maker adopts the Rawlsian attitude so that the sole criterion for social welfare is the welfare of the person with lowest income in the society. As specified, (9.4) incorporates both the desire for vertical redistribution and the concern for HE. However, it constrains local HI in the $i$th income class to induce the same loss of social welfare as local HI in the $n$th income class.

In order to give independent content to HE as opposed to VE [individuals tend to compare their welfare with those of their peers (Runciman (1966))], equation (9.4) is modified as:

$$w = \left[ \sum_{i=1}^{M} N_i \left( \frac{1}{N_i} \sum_{j=1}^{N_i} (x_i - T_{ij}) \right)^{1-v} \right]^{\frac{1}{1-v}}$$

(9.5)

In (9.5), $v$ represents inequality aversion within bands, while $\rho$ represents inequality aversion between income bands. This suggests that the policy-maker has different tastes for inequality within groups and inequality across income groups. For $v = \rho$, that is, if
the policy-maker is neutral between within and between groups inequalities, then (9.5) reduces to (9.4).

To address the identification problem, Auerbach and Hassett (1999, 2001) recognized that any group boundary would impose arbitrary discontinuity on the data distribution with the result that a small change in the value of an individual’s income may imply a change in the membership of the income group as was noted in Chapter 8. In order to obviate this problem, they adopted a less restrictive definition of income band. Income bands are defined in terms of 'overlapping’ windows with normal density distribution rather than non-overlapping windows as in Lambert and Ramos (1997a). In this case, an individual could belong to several contiguous overlapping income bands. A density function that applies over the ith income band defines the distance over the band. Their index of HI is then defined by:

\[ H_i = \left[ \sum_k b_{ki} \cdot \frac{1}{N_k} \sum_j \left( \frac{1 - t_{kj}}{1 - \bar{t}_i} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}} \]  

(9.6)

Where \( t_{kj} = \frac{T_{kj}}{x_k} \) refers to the average tax rate of the ith individual in group i. The term \( b_{ki} \) is the weighting scheme that normalizes to 1 over k income group and \( \bar{t}_i \) is the weighted average of tax rates in group i. Equation (9.6) shows that HI is a function of local inequality aversion \( v \) but does not depend on the global inequality aversion parameter \( \rho \). However, it could be noted that the weights assigned to individuals in a given reference group, \( b_{ki} \) depend on the comparison groups (Auerback and Hassett 2001 p.53)

9.2 The DJA Framework

The DJA approach while based on the insights of the AJL improves and generalizes the framework of the decomposition to accommodate ethically flexible parameters that measure the level of social aversion to inequality as shown in Auerback and Hassett (1999, 2001) framework. However, unlike the Auerback and Hasset framework that uses overlapping windows, the DJA framework obviates the apparent arbitrariness in the determination of
"income-equals" by using non-parametric regressions to first generate expected values of net income conditional on a given level of pre-tax/pre-payment income. We now consider this approach in order to provide a methodological framework for the empirical estimation process and to pinpoint the essential differences between the AJL and DJA approaches.

To overcome the problem of ethical rigidity associated with the AJL model, the DJA model proposes the introduction of a flexible ethical parameter that is percentile-dependent into the index of inequality. This is similar to Yitzhaki (1983) generalization of the Gini index. The DJA single parameter, ethically sensitive weighting scheme uses differences in income rank of individuals.

Let \( p \in (0,1) \) be the fractional rank of an individual in the income distribution. A non-negative weight \( v \) may then be defined such that

\[
w(p,v) = v(1 - p)^{(v-1)}, \quad v \geq 1
\]

(9.7)

Where \( v \) is a parameter reflecting social preferences for equality or inequality aversion using differences in the ranks of individuals in the income parade. Equation (9.7) has the following restrictions:

(1) \[
\int_0^1 w(p)dp = 1
\]

(9.8)

That is the weights are normalized to 1. Furthermore,

(2) \[
w(p_i) \leq w(p_j) \text{ for } p_i \geq p_j
\]

(9.9)

In other words, the weights are concave in relation to the individuals' rank in the income distribution and obey the Pigou-Dalton transfer principle which requires that inequality must decrease if a transfer is made from an individual on higher income rank to an individual on lower income rank. As noted in Chapter 6, for \( v = 2 \), (9.7) reduces to the standard Gini Index.

204
The purpose of the analysis dictates the DJA choice of welfare function: a concave utility function that is additive and linear in levels of income and that can generate relative inequality indices. This is specified as:

\[ W_X(\varepsilon, u) = \int_0^1 U_\varepsilon(X(p)) u(p, v) dp \]  

(9.10)

Equation (9.10) is the SWF of the gross income (X). Note that \( U_\varepsilon \) is Atkinson (1970) concave utility function defined as:

\[ U_\varepsilon(y) = \begin{cases} \frac{y^{1-\varepsilon}}{1-\varepsilon}, & \text{for } \varepsilon \neq 1 \\ \ln y, & \text{for } \varepsilon = 1 \end{cases} \]  

(9.11)

The parameter \( \varepsilon \) is an index of aversion to uncertainty in post-payment income of those in the \( i \)th income level. In other words, \( \varepsilon \) may be interpreted as a measure of aversion to horizontal inequity. Analogous to (9.10) a SWF is also defined for post-payment income as:

\[ W_N(\varepsilon, u) = \int_0^1 U_\varepsilon(N(p)) u(p, v) dp \]  

(9.12)

Let the expected post-payment income of those individuals having rank \( p \) in the distribution of the pre-payment income be denoted by:

\[ \bar{N}(p) = \int_0^1 N(q | p) dq \]  

(9.13)

Where \( q \in [0,1] \) denotes percentile or post-payment income rank. Equation (9.13) suggests that if the payment system were horizontally equitable, then those in the fractional rank \( p \) in the prepayment distribution would obtain mean income of \( \bar{N}(p) \) in the post-payment period. \( N(q | p) \) is the post-payment income of an individual with post-payment income rank \( q \) among all the individuals with prepayment income rank \( p \). The resulting social welfare from (9.13) is
\[ W_N^E(\epsilon, v) = \int_0^1 U_\epsilon(\tilde{N}(p)) w(p, v) dp \] (9.14)

However, if instead of obtaining \( U_\epsilon(\tilde{N}(p)) \) as in (9.14) individuals at pre-payment income rank \( p \) are given the mean post-payment income utility \( \bar{U}_\epsilon(p) \) the resulting social welfare would be:

\[ W_N^P(\epsilon, v) = \int_0^1 \bar{U}_\epsilon(p) w(p, v) dp \] (9.15)

Since \( U_\epsilon \) is a concave utility function (Fig 9.1), it implies that \( U_\epsilon(\tilde{N}(p)) \geq \bar{U}_\epsilon(p) \), (see Lambert 1993:90) and by implication \( W_N^E(\epsilon, v) \geq W_N^P(\epsilon, v) \) which indicates the local cost of post-payment income uncertainty among those at rank \( p \) in the pre-payment income distribution.

![Graph](image)

**Fig. 9.1 Cost of inequality using EDE (Duclos and Lambert 2000)**

Following Atkinson (1970) DJA introduced the concept of equally distributed equivalent (EDE) in order to analyze the cost of inequality. Let \( \xi_x(\epsilon, v) \) be the EDE that would give
the same social welfare as the mean \( \mu \) of unequally distributed pre-payment income \( X \) such that

\[
W_X(\varepsilon, v) = \frac{1}{U_c(\xi_X(\varepsilon, v))} w(p, v) dp = U_c(\xi_X(\varepsilon, v))
\]

(9.16a)

This is illustrated in Fig 9.1 where on account of concavity of the utility function reflecting aversion to inequality, expected utility of EDE equals expected utility of average post-payment income \( E[U_c(\xi_X)] = E[U_c(\mu_X)] \). Since EDE is determined by the level of \( \varepsilon \), the inequality index can be expressed in the form (Atkinson 1970, Blakorby and Donalson, 1978):

\[
I_X = 1 - \frac{\xi_X}{\mu_X}
\]

(9.16b)

Equation (9.16b) expresses the cost of inequality in terms of proportion of total income.

"It is the percentage of total income that could be spent in removing inequality with no resulting loss in social welfare" Duclos et al (2003:10) This implies that if risk or uncertainty in net income increases (reduces) the ratio \( \frac{\xi_X}{\mu_X} \) falls (rises) and thus

\( I_X \) (inequality) rises (falls) requiring less (more) level of equally distributed income (relative to the mean of the actual distribution) to achieve the same level of social welfare as before. This concept of EDE is central to the DJA model estimation of \( H \) and is captured by the \( \varepsilon \) parameter.

Inequality can similarly be defined for post-payment income to obtain \( (I_N) \), for expected post-payment income \( (I_N^E) \), and for expected net income utility \( (I_N^P) \). We note the following inverse relationships between the welfare measures expressed in (9.12), (9.14), and (9.15) and their associated inequalities: \( W_N \leq W_N^P \leq W_N^E \) while \( I_N \geq I_N^P \geq I_N^E \). Note that \( I_N \) is the level of inequality in the post-payment income. \( I_N^P \) can be interpreted as the coefficient of concentration of post-payment income when the ranking variable is
the prepayment income. Similarly, \( I^p_n \) may be interpreted as the index of concentration of the \( expected \) post-payment income \( E(N \mid X) \) when the ranking variable is the prepayment income.

These relationships provide the key to the DJA decomposition framework:

\[
\Delta I = I^X - I^N = I^X - I^E_N - \left( I^P_N - I^E_N \right) - \left( I^N - I^P_N \right)
\]

Like in the case of the AJL framework, the decomposition has three main parts: the vertical equity, horizontal equity and reranking effect. \( V \) is the index of progressivity of payment \( X(p) - \bar{N}(p) \) for individuals at rank \( p \). The HI index (H) aggregates the over all increase in inequality that arises due to unequal payment by those at the same level of prepayment income. This may be illustrated using Fig. 9.1. It is assumed that at any fixed point in the distribution of prepayment income, say point \( x \), there is a group of individuals having exactly income \( x \) at prepayment distribution who can be denoted as \( \Omega_x \). This is the group of prepayment equals at point \( x \). HI at point \( x \) is the measured level of inequality induced among the group \( \Omega_x \) located at point \( x \) by the payment system.

Referring to Fig 9.1, the horizontal axis is assumed to measure the post-payment incomes of pre-payment equals. They were equals at the prepayment income distribution because they had income \( x \). However, the payment system has introduced inequality among them at the after payment income. An individual who belonged to this class of prepayment income may now find her income at point \( x_1 \) or point \( x_2 \) due to the fact that the payment system has treated equals unequally. This is a source of uncertainty and risk factor. The average post-payment income among the set of prepayment equals with \( x \) is \( \mu_x^n \). However, a risk averse individual may prefer to have \( \xi^n \) with certainty rather than take a gamble that may give her \( x_1 \) at the worst and \( x_2 \) at the best with respective probabilities. Thus, the H index measures the risk premium:
\[ H_x = \mu_x^s - \xi_x^s. \] (9.18)

For the policy-maker, \( H_x \) represents the amount that has to be given up among those with post-payment income \( \Omega_x \) so as to remove the uncertainties in the post-payment income of this group without loss of welfare.

To obtain the global measure of \( H \) involves aggregation of the form: \( \sum H_x, \forall x \in R \).

However, the local \( H_x \) is weighted in order to account for differences in population at the different \( x \)s. In the DJA framework, like in Lambert and Ramos framework, the weight chosen is pure weight: the proportion of the population at the given \( x \). In other words the weighting factor is independent of the income at point \( x \). This is important because it ensures that HI at point \( x \) is not contaminated by vertical considerations as is the case with the AJL framework where the weighting factor is the product of the population and income shares at point \( x \).

R index, as in the case of AJL model, measures the reranking effect. However, while in the AJL framework, R arises because individuals may move out of their prepayment income class to other classes, (rank-switching) due to effects of payment, in the DJA model, R arises as a result of any changes in rank induced by the fiscal policy or financing system. Thus, in the AJL model, a change in rank between two people in the same prepayment class and who nevertheless remain in the same post-payment class, is not reckoned as reranking, and R will not capture such a rank switch. However, this would be fully recorded as re-ranking under the DJA model. In other words, denoting \( R_{DJA} \) as the R of DJA framework, and \( R_{AJL} \) as the reranking effect of the AJL framework, then \( R_{AJL} = R_{DJA} \) iff the income bandwidth of \( AJL \to 0 \).

Furthermore the \( H \) in AJL and DJA frameworks seems to measure different concepts. While for the AJL model \( H \) measures the extent of post-payment income heterogeneity introduced by the payment system at the post-payment period among people of the same
prepayment income band, which is then weighted by the product of the population and post-payment income shares at a particular income band and then summed across the bands, the DJA approach estimates $H$ as loss in social welfare arising from the fact that the payment system generates uncertainty and is a source of dis-utility. Inequality itself is a loss of utility to the social welfare because it generates resentment and a feeling of deprivation among peers. It is the cost of this resentments that is captured by $H$.

9.3 The Modelling Issues in the DJA Framework

The estimation of $W_X, W_N$ and $W_N^P$, and their corresponding inequality indices $I_X, I_N$ and $I_N^P$ in the DJA model proceeds by replacing the population income distribution with the empirical distribution. This implies replacing the integration signs in (9.10), (9.12) and (9.15) with summation signs to sum across sample observations. In the case of $W_N^P$, the post-payment income of individuals at the pre-payment income level $x_i$, for example, is replaced by the predicted level of their post-payment income, say, $\hat{n}_i$ at that level. The predicted value is estimated using non-parametric technique.

The major problem, however, is the difficulty involved in the estimation of $H_x$ with micro-data, that is, the identification problem that was encountered earlier in the previous frameworks. In general, the joint population income distribution of pre- and post-payment income ($x^p$ and $x^n$ respectively) will tend to be continuous but the sample joint distribution drawn from the populations may not be. To solve this problem, the Duclos et al (2003) adopts a non-parametric approach to estimating $H_x$. Since in reality it is difficult to observe the joint distribution of the pre- and post-payment incomes ($x^p$ and $x^n$) of those at income level $x$, it is nevertheless possible to sample distributions of pre and post- income distributions to estimate the continuous population distributions of $x^p$ and $x^n$ using the non-parametric approach. The areas between the estimated $x^p$ and $x^n$ are then integrated to yield estimates of the $H$ index and the progressivity index.
Using the non-parametric technique which was briefly introduced in section 6.4, a consistent estimator \( \hat{f}(x^n | x) \) of the conditional density function at point \( x \) is then used to generate estimators \( \hat{\xi}_s^a \) and \( \hat{\mu}_s^a \) of \( \xi_s^a \) and \( \mu_s^a \) respectively and, hence, a consistent estimator \( \hat{H}_s \) of \( H_s \) using the concept of EDE discussed above.

For the prepayment equals group \( \Omega_s \) at point \( x \), the estimator \( \hat{\xi}_s^a \) is given as

\[
U_s(\hat{\xi}_s^a) = \int_{\Omega_s} U_s(x^n)\hat{f}(x^n | x)dx^n
\]

(9.19)

The estimator of \( \mu_s^a \) is given by

\[
\hat{\mu}_s^a = \int_{\Omega_s} x^n\hat{f}(x^n | x)dx^n
\]

(9.20)

Following from (9.18), the estimator of \( H_s \) is given by

\[
\hat{H}_s = \hat{\mu}_s^a - \hat{\xi}_s^a
\]

(9.21)

The estimator \( \hat{H} \) of the global \( \bar{H} \) is then obtained by integrating \( \hat{H}_s \) over all \( x \)s

\[
\hat{H} = \int_{0}^{\infty} \hat{H}_s\hat{f}(x)dx
\]

(9.22)

The basic idea underlying nonparametric estimations is that information about an observation at any given point in a smooth function can be obtained by using information from the observations in the neighbourhood of that observation. Yatchew (1998) suggests alternative ways of estimating the sample or empirical formulation of (9.22) non-parametrically. One is to use either the local averaging estimators whereby the scatterplot is divided into vertical bands and the local means or medians are used to approximate the regression function or to slide windows or bands (as in Lambert and Ramos, 1997) along the x-axis and then compute the moving average. The wider the
window width is, the smoother the estimate at the given point. He notes however, that with this method it is difficult to impose additional structure (e.g. additivity) on the estimating function. The alternative is to use optimization estimators that are more amenable to incorporating other structures (p.680)

Although it is possible to use any of the non-parametric estimation techniques, Duclos et al. (2003) following Duclos and Lambert (2000) used the kernel estimation procedure with the window-width chosen to minimize the mean integrated square error. Kernel functions basically attempt to approximate the density \( f(x) \) from observation on \( x \). The difference between the density estimator \( \hat{f} \) and the true population density \( f \) at any particular point in the distribution is measured by the mean square error (MSE) which is defined as (Silverman 1986:35):

\[
MSE_2(\hat{f}) = E\left\{ \left( \hat{f}(x) - f(x) \right)^2 \right\}
\]

Or, (by properties of mean and variance)

\[
MSE_2(\hat{f}) = E\left\{ \left( \hat{f}(x) - f(x) \right)^2 \right\} + \text{var} \hat{f}(x) \tag{9.23}
\]

Kernel density estimate of a series \( X_i \) at a given point \( x \) is obtained by summing (or as in 9.22 above, integrating) the weighted values computed as

\[
\hat{f}_K = \frac{1}{Nh} \sum_{i=1}^{N} K \left[ \frac{x - X_i}{h} \right]
\tag{9.24}
\]

Where \( N \) = number of observations, \( K \) is the smoothing function that integrates to 1 while \( h \) is the window width or the smoothing parameter that controls the smoothness of the estimated density at point \( x \). The larger the window width is, the smoother the estimated function. While the \( K \) controls the shape of the weight (which normalizes to 1), \( h \) determines the magnitude (Yatchew, 1998). In the estimation of the density of the distribution of the prepayment income at a given point \( x \) located at the centre of the
window, weights are attached to observations according to their nearness to $x$. A large value of $h$ implies that a relatively greater weight is being put on observations that are far from $x$. In other words, increasing window width implies that information about the shape of the density function $f(x^a | x)$ is being sought from observations that are increasingly dissimilar to $x$ in the computation of expected post-payment income since it is the information drawn from these neighbouring observations that are used to make inference about the shape of expected post payment income density at $f(x^a | x)$. This introduces bias in the estimation of the true population density at point $x$. However, if $h$ is narrowed so as to eliminate bias, it increases integrated mean error squared shown in (9.23) which may introduce systematic error in the estimation. Larger bandwidths estimate the densities and the cost of horizontal inequalities along the percentiles of prepayment distribution more smoothly but this means that the estimated $\hat{H}_v$ becomes less representative of the local inequality at point $x$. This suggests the need for the selection of the optimal window width that balances the trade-off between eliminating bias and introducing systematic error into the estimation. For the choice of optimal bandwidth, Duclos et al (2003) use the robust bandwidth defined by

$$h = 0.9 A N^{-0.2}$$  \hspace{1cm} (9.25)$$

Where $A = \min\{\text{standard deviation, inter-quartile range}/1.34\}$

However, alternative definitions of the bandwidth could be specified in most statistical packages (including $DAD$ which will be used for this estimation).

There are several kernel functions ($K$) that could be used for the estimation. These include Epanechnikov, Biweight, Triangular, and Gaussian kernel functions (Silverman 1986 p.43; Yatchew 1998 p 685). The DJA framework is based on the Gaussian kernel because of the well known properties of the Gaussian distribution. However, it seems that, in general, the choice of the kernel function is not as important as the choice of the window width (Yatchew, 1998, Silverman 1986, Stata Reference Manual 2002)
By adopting this statistical approach, therefore, the DJA framework transfers the normative decision of determining income equals from the decision maker to statistical exercise: the choice of the window width, though arbitrary, is now determined by the optimal trade-off between bias and minimization of the squared mean error. The only assumptions required are statistical assumptions such as the smoothness and continuity of the joint distribution of gross and net incomes.

About the consistency of kernel estimators, Silverman (1986:71) shows that the conditions for the consistency of the kernel estimators are mild though the rate of convergence to true density may be slow. However, if the conditions of smoothness and continuity are satisfied then the kernel estimators are generally asymptotically consistent. This is a further advantage of the DJA model. The parameter estimators in the model could be assumed to be asymptotically consistent estimators of the true parameter values and this consistency is obtained under mild conditions.

9.4 Social Determinants of Values of $\epsilon$ and $\nu$

What are the social bases for determining the values of $\epsilon$ and $\nu$? Duclos et al (2003) suggest that the empirical values of these parameters can be deduced from dead weight loss involved in social transfer of resources from the rich to the poor. These include administrative and efficiency costs. It would also include the level of transparency and accountability in public institutions and the rate of compliance with the payment system (e.g., compliance with tax laws, compliance with regulations about payments for social services, etc). Such costs are suggested to lie between 0.25 and 1.0 and therefore, implying that $\epsilon$ may lie between this range (i.e. $0.25 \leq \epsilon \leq 1$). More efficient fiscal system would have the value of $\epsilon$ near to the minimum while inefficient system would have values of $\epsilon$ close to 1. On the same basis, the social value of $\nu$ is suggested to range between 1 and 4. In the estimation that follows we follow the Duclos et al (2003), example where they used the value of $\epsilon = 0.4$. But we also estimate for $\epsilon = 0$, which amounts to the assumption of the null hypothesis: $H_{DJA} = 0$; where $H_{DJA}$ refers to the value of horizontal inequity obtained under the DJA framework. The results obtained
from the latter assumptions, (i.e. \( \varepsilon = 0 \Rightarrow H_0 : H_{DJA} = 0 \)) is comparable to the results estimated from the AJL framework when the income bandwidth in the AJL framework approaches zero.

As was noted earlier on, the DJA generalizes the value of \( v \) which in AJL is implicitly fixed at \( v = 2 \). This means that the results of the DJA are only comparable to those of the AJL when \( v \) in DJA assumes the value of 2. For the following estimations we use a broad range of values of \( v \) from 1.5 to 5.0. It is to be noted that \( v = 1 \) implies that the weight in the social welfare equations \( w(p, v) \) is equally distributed across individuals so that the social welfare function is simply the summation of the utilities of all individuals implying that ranking is no longer necessary (i.e., for \( v = 1, R = 0 \)) Higher values of \( v \) imply that more weight is given to the "reranking resentment of the poorest " (Duclos et al 2003, p.19).

As in the estimation of the AJL model, we take into account the multi-stage design of the survey that generated the sample, namely the stratification, clustering and weighting (see section 5.1.4.3 above). Since the Distributive Analysis software (DAD) contains program for estimation of the DIA model and provides, on request, information about test statistics for all estimated parameters, the bootstrapped asymptotic standard errors are also included in the table.

Having outlined the basic conceptual issues underlying the DIA decomposition framework, the next subsection presents the estimated results which are subsequently compared with the estimated results from the AJL model. Finally we consider the issue of the robustness of the estimated results to alternative assumptions about intra-household resources distribution and household economies of scale.

**9.5 Decomposition Results from the DJA Framework**

Results of the DJA Model decomposition based on an assumed value of aversion to uncertainty in net income after payment \( \varepsilon = 0.4 \) and \( v = 1.50, 2.0, 3.0, 5.0 \) are
presented in Table 9.1. The results are also based on the assumption that there are no intra-household inequalities in resource distribution (resources are distributed on per capita basis), and there are no economies of scale. In a subsequent subsection we relax these assumptions to examine the robustness of the decomposition results to changes in these variables.

Table 9.1 Decomposition Results Based on DJA Framework (ε = 0.4)

<table>
<thead>
<tr>
<th>Scale Assumption: Per Capita</th>
<th>ε = 0.4, v = 1.5</th>
<th>ε = 0.4, v = 2.0</th>
<th>ε = 0.4, v = 3.0</th>
<th>ε = 0.4, v = 5.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_X$</td>
<td>0.3807**</td>
<td>0.4966**</td>
<td>0.6145**</td>
<td>0.7150**</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.1173)</td>
<td>(0.1498)</td>
<td>(0.1849)</td>
<td>(0.2169)</td>
</tr>
<tr>
<td>$G_{X-T}$</td>
<td>0.3829**</td>
<td>0.4999**</td>
<td>0.6201**</td>
<td>0.7224**</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.1182)</td>
<td>(0.1503)</td>
<td>(0.1871)</td>
<td>(0.2199)</td>
</tr>
<tr>
<td>RE</td>
<td>-0.0021</td>
<td>-0.0033</td>
<td>-0.0056</td>
<td>-0.0075</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.0055)</td>
<td>(0.0040)</td>
<td>(0.0054)</td>
<td>(0.0071)</td>
</tr>
<tr>
<td>$V$</td>
<td>0.0284**</td>
<td>0.0334**</td>
<td>0.0398**</td>
<td>0.0509**</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.0102)</td>
<td>(0.0120)</td>
<td>(0.0133)</td>
<td>(0.0158)</td>
</tr>
<tr>
<td>H</td>
<td>0.0202**</td>
<td>0.0253**</td>
<td>0.0342**</td>
<td>0.0473**</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.0059)</td>
<td>(0.0076)</td>
<td>(0.0104)</td>
<td>(0.0144)</td>
</tr>
<tr>
<td>R</td>
<td>0.0103**</td>
<td>0.0114**</td>
<td>0.0111**</td>
<td>0.0111**</td>
</tr>
<tr>
<td>B. $\beta_b$</td>
<td>(0.0031)</td>
<td>(0.0036)</td>
<td>(0.0047)</td>
<td>(0.0046)</td>
</tr>
</tbody>
</table>

NB. ** Statistically significant at 5% level

Column 1 of Table 9.1 lists the parameters and their bootstrapped standard errors (shown in brackets). Since the definition of income band is statistically determined in the framework it obviates the need for the arbitrary definition of income bands. The table does not include some of the intermediate parameters: $g$, $CI$, and $K$ as was done in the AJL results because the DAD program used to achieve this decomposition does not provide such parameters explicitly. We note, however, that while the value of $g$ will be the same under both frameworks, we cannot obtain the values of $CI$ and $K$ in the DJA framework (unless they are estimated separately) since in this case it is statistically determined by the value of $h$ – the smoothing parameter (see Silverman 1986:15 and eqn 5 – in DJA methodology).
Table 9.2 DJA Decomposition Results ($\varepsilon = 0$)

<table>
<thead>
<tr>
<th>Scale Assumption : Per Capita</th>
<th>$\varepsilon = 0, v = 1.5$</th>
<th>$\varepsilon = 0, v = 2.0$</th>
<th>$\varepsilon = 0, v = 3.0$</th>
<th>$\varepsilon = 0, v = 5.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_X$</td>
<td>0.3104**</td>
<td>0.4474**</td>
<td>0.5830**</td>
<td>0.6970**</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0909)</td>
<td>(0.1349)</td>
<td>(0.1750)</td>
<td>(0.2088)</td>
</tr>
<tr>
<td>$G_{X-T}$</td>
<td>0.3105**</td>
<td>0.4482**</td>
<td>0.5863**</td>
<td>0.7025**</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0915)</td>
<td>(0.1349)</td>
<td>(0.1764)</td>
<td>(0.2108)</td>
</tr>
<tr>
<td>RE</td>
<td>-0.0001</td>
<td>-0.0008</td>
<td>-0.0033</td>
<td>-0.0056</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0054)</td>
<td>(0.0054)</td>
<td>(0.0061)</td>
<td>(0.0056)</td>
</tr>
<tr>
<td>V</td>
<td>0.0186**</td>
<td>0.0250**</td>
<td>0.0314**</td>
<td>0.0430**</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0092)</td>
<td>(0.0111)</td>
<td>(0.0119)</td>
<td>(0.0150)</td>
</tr>
<tr>
<td>H</td>
<td>0.0056**</td>
<td>0.0119**</td>
<td>0.0221**</td>
<td>0.0372**</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0019)</td>
<td>(0.0036)</td>
<td>(0.0068)</td>
<td>(0.0112)</td>
</tr>
<tr>
<td>R</td>
<td>0.0131**</td>
<td>0.0139**</td>
<td>0.0126**</td>
<td>0.0113**</td>
</tr>
<tr>
<td>B. Se</td>
<td>(0.0045)</td>
<td>(0.0054)</td>
<td>(0.0044)</td>
<td>(0.0044)</td>
</tr>
</tbody>
</table>

Table 9.1 also shows that under a constant value of $\varepsilon$, the estimated values of $G_X, G_{X-T}$ and the absolute values of RE increase consistently as the value of $v$ increases from 1.50 to 5.0. In none of the values of $v$ within the tabulated range is the redistributive effect significant. However, the increasing negative value of RE shows that as more emphasis is put on the poorer segments of the society, the redistributive effect becomes more pro-rich. In other words, the more sensitive or ‘equality-minded’ the policy-maker is the more pro-rich the redistributive effect appears. Similarly vertical and horizontal inequities each increase with increases in the value of $v$ keeping the value of $\varepsilon$ constant. In all cases, vertical inequity is higher than horizontal inequity and reranking effect. This indicates that vertical equity is a more serious problem in the population. But the rate of increase in the ratio V/H decreases with increase in the value of $v$ (keeping $\varepsilon = 0.4$), though this decrease is not very steep. This is shown in Fig 9.2.

It is to be noted that contrary to what the theory suggests under the DJA assumptions, namely that when $\varepsilon = 0, H \rightarrow 0$, we still find that the estimated values of $H_{DJA}$ under
the different values of \( v \) are still statistically significant. Thus, it seems that there is a
difference between what is predicted by theory and what is obtained in actual
implementation of the model (at least as estimated by the DAD software). This requires
further investigation.

![Graph](image)

**Fig 9.2 V/H Plot for \( \varepsilon = 0 \) and \( \varepsilon = 0.4 \)**

Comparing Tables 9.1 and 9.2, we find that for all values of \( v \), the value of \( H \) is greater in
Table 9.1 with \( \varepsilon = 0.4 \) than in Table 9.2 with \( \varepsilon = 0 \). This is obviously so because the
cost of uncertainty in post-payment income distribution given prepayment rank \( N(q \mid p) \)
increases with increases in the value of aversion to uncertainty in the post-payment
income measured by \( \varepsilon \). This result can easily be seen from equation (9.16) and by noting
the inverse relationship between \( \varepsilon \) and \( \xi \). That is, the higher the level of uncertainty in
net income the less the EDE. From equation (9.21) it follows that higher values of \( \varepsilon \)
would be associated with higher values of \( H \). Furthermore, setting \( \varepsilon = 0 \) is test of the
null hypothesis that there is no \( H \). This constraint is rejected as can be seen from Table
9.2. It is to be noted that the constraint becomes increasingly non-binding as the value of
\( v \) increases. This perhaps explains why the value of the ratio \( V/H \) in Table 9.2 declines
more dramatically than in Table 9.2 (see also Fig. 9.2). The relations between \( V \) and \( H \) is
exhibited when the value of $\epsilon$ is constrained to zero as in Table 9.2. The divergence between the values of $V$ and $H$ is even more significant at lower values of $v$ for $\epsilon = 0$. However, it seems likely that in the DIA increases in the value of $\epsilon$ must be accompanied by not only increases in the value of $H$ but also increases in the value of $V$ if marginal increases in the payment rate are also accompanied by higher levels of deadweight loss associated with $\epsilon$.

The value of $R$ appears rather constant with increases in the value of $v$. This is consistent with the DIA theoretical framework: the larger the value of $v$ the more weight that is given to "the reranking resentment of the poorest" (Duclos et al. 2003 p. 19) But giving more weight to the resentment of the poorest need not necessarily increase the reranking. Therefore, in spite of the increase in weight given to the reranking resentment of poorest, the reranking index $R$ changes but very little. Irrespective of this sluggish movement in $R$ in response to increases in $v$, it is clear from the results in both Tables 9.1 and 9.2 that reranking is a major effect that follows direct healthcare payment. In fact at the lower values of $v$ in Table 9.2, reranking has higher values than $H$. The question of which of these is a more serious problem to a society may be depend on social values. It is clear, nevertheless, that reranking is a major problem in healthcare financing through direct payment in Nigeria.

While reranking may be a problem in a strong fiscal system, it has devastating effects in direct healthcare financing. Healthcare financing could become such a shock to a household that it could fall dramatically from the top income level to the bottom. For example, Narayan et al (2000), in a story whose significance was also noted by Wagstaff and van Doorslaer (2001), reported the voice of a young man complaining how the cost of treating his daughter’s severe illness plunged him from the top to the bottom of the income distribution in his community. This story finds an echo in many reported cases in our data of households that fell from the 9th decile at the prepayment income distribution to the 3rd and 4th deciles at the post-payment distribution (see Figs 7.3 and 7.4 above). Several households reported having to borrow money or sell household assets in order to
finance the illness of a household member. Such dramatic changes in a household’s relative standing in the community clearly indicate that reranking is a major problem in direct healthcare financing.

9.6 Comparing Decomposition Results from AJL and DJA Frameworks

Table 9.3 Results from AJL and DJA Decompositions Based on Per Capita

<table>
<thead>
<tr>
<th>Parameters</th>
<th>AJL</th>
<th>DJA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>$G_x$</td>
<td>0.4474**</td>
<td>0.4474**</td>
</tr>
<tr>
<td>Se</td>
<td>(0.1349)</td>
<td>(0.1349)</td>
</tr>
<tr>
<td>$G_{x-T}$</td>
<td>0.4482**</td>
<td>0.4482**</td>
</tr>
<tr>
<td>Se</td>
<td>(0.1349)</td>
<td>(0.1349)</td>
</tr>
<tr>
<td>$RE$</td>
<td>-0.0008</td>
<td>-0.0008</td>
</tr>
<tr>
<td>$V$</td>
<td>0.0126</td>
<td>0.0136</td>
</tr>
<tr>
<td>$H$</td>
<td>0.0086</td>
<td>0.0070</td>
</tr>
<tr>
<td>$R$</td>
<td>0.0044</td>
<td>0.0074</td>
</tr>
</tbody>
</table>

NB. ** Statistically significant at 5% level

Table 9.3 contains the comparable parameters from the AJL and DJA frameworks. It was indicated in the preceding section that in the AJL framework assumes implicitly that the value of $\nu = 2$. This is the value of $\nu$ inbuilt in the Gini index. The DJA model generalizes the value such that $\nu$ on values ranging from 1 to infinity. This implies that a meaningful comparison can be made between the frameworks only under a common value of $\nu$, and this is when $\nu = 2$. Thus, it is immediately evident that the DJA framework allows the policy-maker some latitude to reflect a wide range of social attitudes to inequality into policy rather than being restricted to that predetermined by measurement limitations. This is consistent with democratic principles and public choice theory rather than a rigid environment suggestive of policy dictatorship. On this score, the DJA framework represents a significant improvement on the AJL framework.
Further more, it is noted above that $\epsilon = 0$ implies the null hypothesis $H = 0$. Under the AJL framework, the result $H = 0$ can only occur when the income bandwidths approach zero. Thus, $H_{AJL}$ and $H_{DJA}$ can only be comparable when $\epsilon = 0$ in the DJA framework. Columns 2-5 in Table 9.3 show the results from the various definitions of income band under the AJL: 10%, 5%, 3.3% and 1% bandwidths respectively. While the values estimated from these bandwidths do not necessarily tell us what the values of the parameters are when the widths are zero, it is assumed that the trends in the values of the parameters as the bandwidths get smaller will provide sufficient information as to the possible values of the parameters when the bandwidths approach zero.

Leaving aside the estimated prepayment and net payment income, and thus the redistributive effect, since these have the same values under both frameworks, we consider the differences in the estimated components of the decomposition. In general it is observed that under the AJL framework $V_{AJL}$ tends to increase as the bandwidth becomes smaller, although it increases at a very slow rate. This is explained by the fact that the averaging process in larger bandwidths tend to conceal some differences in vertical distribution than in small bandwidths so that we expect $V_{AJL}$ to increase as bandwidths approach zero. A similar increase is observed in the value of $R_{AJL}$ which increases even faster than $V_{AJL}$ as the bandwidth approaches zero. $H_{AJL}$ moves in the opposite direction to $V_{AJL}$ and $R_{AJL}$. Obviously, as the bandwidth shrinks to zero, every incidence of $H_{AJL}$ turns out to be an incidence of $R_{AJL}$ so that $H_{AJL}$ disappears altogether. It is, therefore not surprising to observe that $R_{DJA} \approx R_{AJL} + H_{AJL}$ as the bandwidth of the AJL framework approaches zero. Notice for example that in Table 9.3 at 1% bandwidth the values of $R_{AJL}$ and $R_{DJA}$ tend to converge because the value of $H_{AJL}$ tends to zero. This clearly suggests that unless the bandwidth in the AJL framework approaches zero the $R_{DJA}$ will seem to be a better estimator of the true $R$, while $R_{AJL}$ will tend to underestimate the true $R$. 

221
With respect to H, the two frameworks actually estimate different conceptual issues. To get the issues involved here well fixed, in a sample distribution of prepayment income it is difficult to see two income units with exactly the same amount of income. This implies that every vertical alignment of the income of two or more units in the post-payment income (as in Fig 7.3) is counted as \( R_{DJA} \). But this can only be counted as \( R_{AJL} \) if the observations move from the prepayment income band to another income band in the post-payment distribution. If every observation stays in its prepayment class at the post-payment distribution, then there is no reranking at all even though there may have been permutations of rank within the income bands in the transition from the pre- to the post-income distribution. Every such rank permutations within class in the transition is recorded as \( H_{AJL} \) not \( R_{AJL} \). But for the DJA, the two types of movements constitute \( R_{DJA} \).

Under the DJA framework, \( H_{DJA} \) is an index of the risk premium an economic agent in a given prepayment income class is willing to pay to prevent the uncertainty that accompanies the transition from pre- to post-payment income distribution that arises from the intervention of the fiscal system. The more arbitrary (or random) the fiscal behaviour, the higher the value of \( \varepsilon \) and so the higher the risk premium an agent is willing to pay to obtain the certainty equivalent in the post-payment income. As Auerbach and Hassett (2001: 49-50) note "Large differences among similar individuals might be viewed as intrinsically arbitrary (regardless of whether they resulted from intentionally "abusive" government behaviour) and therefore more costly to the social fabric, or they might simply be viewed as more costly because individuals compare themselves with those with similar characteristics." This same line of argument is also advanced by Musgrave (1959 cited in Musgrave 1990). Thus, it seems that on this score, the concept of horizontal equity embodied in \( H_{DJA} \) differs fundamentally from what is measured by \( H_{AJL} \). In as much as \( H_{AJL} \) and \( H_{DJA} \) represent fundamentally different ideas, their respective influences on \( V_{AJL} \) and \( V_{DJA} \) will not be the same with the result that the values of \( V_{AJL} \) and \( V_{DJA} \) under the two framework will also differ markedly.
Furthermore, it is noted that the weighting schemes under both frameworks are also different. For the AJL framework, $G_{f(c)}$, the index of inequality in the $i$th band is weighted by $a_z$ which is the product of the populations share of the band and the post-payment income share of the band. This is summed to obtain $\sum a_z G_{f(c)} = H_{AJL}$. It is important to note that $a_z$ is not 'pure weight' as is found in the weighting schemes of Lambert and Ramos (1997), Auerbach and Hasset (2001), and Duclos et al (2003). This is because $a_z$ is contaminated by vertical equity consideration through its dependence on the post-payment income share of the $i$th income band. In other words, the magnitude of $a_z$ apart from being a function of population size of the $i$th band (pure weight), is also a function of post-payment income share of the group (a vertical consideration) This is again a significant difference between the frameworks and also accounts for the differences in the estimates of the two frameworks. It is therefore not surprising to find that $V_{DJA}$ is greater than $V_{AJL}$. In actual fact, Duclos et al (2003) suggest that $V_{DJA}$ under the assumptions: $\varepsilon = 0, \nu = 2.0$, is equivalent to the vertical equity indices of and Kakwani (1977) and Reynolds-Smolensky (1977) as discussed above in sections 4.6 and 4.7 respectively.

9.7 The Robustness of Decomposition Results

Are the decomposition results stable under different assumptions about the measure of economic well-being or utility in both frameworks? In other words, are the results of the models affected by assumptions about equivalence scales and households’ economy of scale? This is the question that is addressed in this last subsection. To address these questions, we present the results of both models under alternative assumptions about equivalence scales and economy of scale.

The utility scales are based on the general equation for intra-household allocations and economies of scale specified in (5.7). In this particular application, the parameters specified in that equation assume the following values: In the first place, for intra-
household resource allocation, each adult in the household is ascribed a unit of adult equivalence so that the parameter \( a \) in (5.7) assumes the value of the number of adults in the household. Each child in the household below age 16 is ascribed 0.5 of adult equivalence (see Buhmann et al 1988). The scale elasticity \( e \) in the formula is non-linear. It takes the value 1.0 under the assumption that there is no economy of scale. For \( e = 0.7 \), it is assumed that it costs less to maintain an additional member of the household. For \( e = 0.3 \) it costs still less to maintain an additional member of the household than for \( e = 0.7 \). These scales are denoted as: Eqvs:1-1 \( e=1 \); Eqvs:1-0.5 \( e=0.7 \); Eqvs:1-0.5 \( e=0.3 \).

To make the results within each model comparable under different assumptions about well-being, in the case of the AJL model the results of the decomposition using 5% income band are compared across the three utility scales. For the DJA framework, the value of \( v \) is fixed at 0.4 and the value of \( v \) is fixed at 2.0. The estimated parameters are then compared across the alternative scales.

### 9.8 Stability of AJL Results

#### Table 9.4A Stability of AJL Results under Alternative Scales

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( Eqvs : 1 - 1, e = 1 )</th>
<th>( Eqvs : 1 - 0.5, e = 0.7 )</th>
<th>( Eqvs : 1 - 0.5, e = 0.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_X )</td>
<td>0.4474</td>
<td>0.4159</td>
<td>0.4216</td>
</tr>
<tr>
<td>( S_e )</td>
<td>(0.0196)</td>
<td>(0.0190)</td>
<td>(0.0194)</td>
</tr>
<tr>
<td>( G_{X-T} )</td>
<td>0.4482</td>
<td>0.4169</td>
<td>0.4220</td>
</tr>
<tr>
<td>( S_e )</td>
<td>(0.0218)</td>
<td>(0.0207)</td>
<td>(0.0208)</td>
</tr>
<tr>
<td>( R_E )</td>
<td>-0.0008</td>
<td>-0.0010</td>
<td>-0.0004</td>
</tr>
<tr>
<td>( G )</td>
<td>0.0883</td>
<td>0.0913</td>
<td>0.0950</td>
</tr>
<tr>
<td>( C_T )</td>
<td>0.5875</td>
<td>0.5577</td>
<td>0.5970</td>
</tr>
<tr>
<td>( K )</td>
<td>0.1401</td>
<td>0.1418</td>
<td>0.1755</td>
</tr>
<tr>
<td>( V )</td>
<td>0.0136</td>
<td>0.0143</td>
<td>0.0184</td>
</tr>
<tr>
<td>( H )</td>
<td>0.0070</td>
<td>0.0124</td>
<td>0.0085</td>
</tr>
<tr>
<td>( R )</td>
<td>0.0074</td>
<td>0.0039</td>
<td>0.0103</td>
</tr>
</tbody>
</table>
Table 9.4 shows that under the AJL framework the estimated results are not only sensitive to alternative definitions of the income band but also to assumptions about equivalence scales and economies of size in the household. We observe for instance that there are significant differences between the estimated Gini coefficients for per capita and the estimated Gini index for equivalence scales and size elasticities of 0.7 and 0.3. The estimated Gini index is 7.6% higher under the per capita assumption (column A) than under the Eqv:.1_0.5 e=0.7 assumption (column B). But the Gini index is not a monotonically decreasing function of consumption elasticity as the Gini index under the Eqv:.1_0.5 e=0.3 (column C) is over a percentage point higher than under Eqv:.1_0.5 e=0.7. The differences in the main parameters of the decomposition under the AJL framework are summarized in Table 9.5.

Table 9.4B Changes in AJL Parameters under Alternative Scales

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( [(A - B) / B]% )</th>
<th>( [(B - C) / C]% )</th>
<th>( [(C - A) / A]% )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_x )</td>
<td>7.60</td>
<td>-1.30</td>
<td>-5.80</td>
</tr>
<tr>
<td>( G_{x-T} )</td>
<td>7.51</td>
<td>-1.21</td>
<td>-5.85</td>
</tr>
<tr>
<td>RE</td>
<td>-20</td>
<td>150</td>
<td>-50</td>
</tr>
<tr>
<td>V</td>
<td>-4.90</td>
<td>-22.28</td>
<td>35.29</td>
</tr>
<tr>
<td>H</td>
<td>-43.55</td>
<td>45.88</td>
<td>21.43</td>
</tr>
<tr>
<td>R</td>
<td>89.74</td>
<td>-62.14</td>
<td>39.19</td>
</tr>
</tbody>
</table>

It is further observed that the estimated values of the RE under the scale assumptions are important, with Eqv:.1_0.5 e= 0.7 having the highest redistributive effect value in absolute terms. The Eqv:.1_0.5 e= 0.3 has the highest value of estimated Kakwani index of progressivity, and this is significantly different from the values estimated for the other two assumptions. The estimated values of V, H and R are also different from one another under the three assumptions with per capita having the lowest values of estimated V and H while Eqv:.1_0.5 e=0.7 has the lowest estimated re-ranking index value.

The general conclusion that can be made from Table 9.5 is that under the AJL framework adjustments for intra household distribution of resources and size economies of scale are important in the estimation of inequalities. Secondly, larger households and households
with many dependents (children) tend to be poorer. Thus, adjustments for household demographic composition and size economies of scale within limits tend to narrow the counterfactual level of inequality.

### 9.9 Stability of the DJA Results

#### Table 9.5A Stability of the DJA Results under Alternative Scales

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Equs : $1 - 1, e = 1$</th>
<th>Equs : $1 - 0.5, e = 0.7$</th>
<th>Equs : $1 - 0.5, e = 0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_X$</td>
<td>0.4966**</td>
<td>0.4627**</td>
<td>0.4686**</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.1498)</td>
<td>(0.1414)</td>
<td>(0.1423)</td>
</tr>
<tr>
<td>$G_{X-T}$</td>
<td>0.4999**</td>
<td>0.4661**</td>
<td>0.4709**</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.1503)</td>
<td>(0.1414)</td>
<td>(0.1417)</td>
</tr>
<tr>
<td>$RE$</td>
<td>-0.0033</td>
<td>-0.0034</td>
<td>-0.0023</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.0040)</td>
<td>(0.0090)</td>
<td>(0.0080)</td>
</tr>
<tr>
<td>$V$</td>
<td>0.0334**</td>
<td>0.0347**</td>
<td>0.0371**</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.0120)</td>
<td>(0.0149)</td>
<td>(0.0158)</td>
</tr>
<tr>
<td>$H$</td>
<td>0.0253**</td>
<td>0.0250**</td>
<td>0.0263**</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.0076)</td>
<td>(0.0072)</td>
<td>(0.0077)</td>
</tr>
<tr>
<td>$R$</td>
<td>0.0114**</td>
<td>0.0131**</td>
<td>0.0130**</td>
</tr>
<tr>
<td>$Se$</td>
<td>(0.0036)</td>
<td>(0.0043)</td>
<td>(0.0060)</td>
</tr>
</tbody>
</table>

**NB: ** Statistically significant at 5% level

#### Table 9.5B: Changes in DJA Parameters under Alternative Scales

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$[(A - B) / B] %$</th>
<th>$[(B - C) / C] %$</th>
<th>$[(C - A) / A] %$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_X$</td>
<td>7.33</td>
<td>-1.25</td>
<td>-5.64</td>
</tr>
<tr>
<td>$G_{X-T}$</td>
<td>7.25</td>
<td>-1.02</td>
<td>-5.80</td>
</tr>
<tr>
<td>$RE$</td>
<td>-2.94</td>
<td>47.83</td>
<td>-30.30</td>
</tr>
<tr>
<td>$V$</td>
<td>-3.75</td>
<td>-6.47</td>
<td>11.08</td>
</tr>
<tr>
<td>$H$</td>
<td>1.2</td>
<td>-4.94</td>
<td>3.95</td>
</tr>
<tr>
<td>$R$</td>
<td>-12.98</td>
<td>0.77</td>
<td>14.04</td>
</tr>
</tbody>
</table>
Tables 9.5A and 9.5B summarize the effects of changes in the alternative assumptions about intra-household resource distribution and economy of scale on estimated parameters under the DJA framework. If the values of $\varepsilon$ and $\nu$ are kept constant, the levels of inequality in $G_X$ and $G_{X-T}$ change in response to changing utility scales. Relatively, the parameters most susceptible to transformations in utility seem to be $RE$ and $R$. Redistributive effect for instances decreases by as much as 49% when utility is transformed from $Evus : 1 - 0.5, e = 0.7$ to $Evus : 1 - 0.5, e = 0.3$.

In general, therefore we can conclude that both the AJL and DJA models are sensitive to assumptions about household composition and economies of scales. But the core parameters of their decompositions do not show any consistent patterns of differences in respect of assumptions about equivalence scales and scale economies.

9.10 Summary

This chapter has tried to assess an alternative framework for the decomposition of redistributive effect to the AJL decomposition framework which has dominated the empirical literature. The DJA framework while building upon basic principles underlying the AJL framework improves and advances both the theoretical and empirical basis for decomposition of redistributive effects of taxes and social transfers. Three important aspects of this improvement are worth emphasizing. In the first place, the DJA framework generalizes the inequality aversion implicit in the AJL framework and therefore, provides the social decision maker the latitude to choose the value of aversion parameter that reflect the norms and social attitude towards inequality. In this sense, it can be said that the DJA framework is more consistent with democratic principles.

Secondly, the weighting variable in the estimation of the horizontal inequity is purified of the influences of vertical equity in the DJA. This consideration is important given the warning by Musgrave (1990:118) that measures of HE must not “reflect a mix of HE and VE and confuse the two basic issues”. By weighting the HI index by a product of population and post-payment income, the AJL framework seems to have failed to heed
this warning. Thus it seems that in this respect, the DIA framework by weighting the population weight only which serves ‘pure’ weight, the HI index is uncontaminated by VE considerations.

Perhaps the most defining feature of the DIA framework is how it deals with the problem of defining income equals that serves to dog previous attempts at measuring horizontal inequity. The basic command of HE is that “equals should be treated equally” Assuming that preferences and tastes are alike, defining income equals seems arbitrary since in empirical distributions exact equals are difficult to find. It is easier to find ‘close equal’ (Lambert and Ramos 1997) than to find ‘equals’. It is, however, clear that in the limit the bandwidths should approach zero for equals to be really equal. Driving this line argument to its logical end implies that in empirical distributions each observation constitutes its own band which poses the problem of implementability in empirical estimations. By proposing the use of non-parametric kernel regression with known statistical properties to estimate the bandwidth, the DIA framework seems to have transferred this normative task to a statistical exercise. Although the choice of bandwidths under the kernel regression is still an arbitrary decision, the regression model suggests the possibility of choosing a bandwidth that provides the optimal trade-off between bias and integrated mean square error. In this case the choice of bandwidth ceases to be an arbitrary exercise.

Furthermore, the DIA approach leads to a fundamental shift in the substance of horizontal equity. What it measures is no more the aggregate loss or gain in rank between individuals in different income bands in transiting between pre- and post-payment distributions (as measured by $R_{A/L}$). Not even the aggregate changes in individual ranks in-between transition (as measured by $R_{DIA}$). It now measures the risk or uncertainty that in the post-payment distribution one may end up in one or the other of the extreme positions among one’s prepayment equals in the post fiscal intervention distribution. This is an experiment that is akin to Rawls (1972) “veil of ignorance”. Being risk averse, individuals would prefer to pay risk premiums to obtain a level of income that is certain in the post-payment period than take a gamble with best and worst scenarios. In this sense, the theory builds upon fundamental human psychology. In this sense also, it is
possible to compute the cost of inequality by knowing how much individuals are willing to pay as risk premiums. However, the problem with this psychological construct is that the source of inequality among prepayment equals is not just the intervention of the fiscal system, (in this case the healthcare payment system), but their inner psychological disposition towards risk. From purely utilitarian perspective, it would seem that the propensity to take risk would differ according to prepayment income levels with those at the lower end of the distribution more prone to taking the gamble, implying that HI will be higher at the top than at the bottom of the prepayment distribution.

Judging from the empirical result, it seems that the AJL model underestimates the components of redistributive effects: V, H and R principally because of their different conceptual and methodological assumptions. Furthermore, expectedly the estimations show that the AJL model is sensitive to the definitions of income band. The smaller the income band the smaller the level of H and higher the levels of V and R. This confirms the trade-off alluded to by Musgrave (1990) that a fiscal policy that addresses V may end up worsening H and vice versa. It therefore becomes a normative choice as to which of these should command higher priority. However, under the DIA model increasing the efficiency of the fiscal system and removing dead weight loss would reduce uncertainty, improve horizontal equity, and possibly vertical equity as well if dead weight loss increases with level of income.

Finally, the estimations suggest that the results from both frameworks are to a significant extent determined by assumptions about the intra-hold distribution of resources and economies of scale. If poorer households tend to have larger sizes and also have greater number of children, this would tend to reduce the level of inequality. But, as noted above, this is only within limits. This is because improvement in well-being in this particular case is not a monotonic function of the economy scale elasticity.
Chapter 10

Summary of Findings and Conclusion

10.1 General Summary

The primary policy concern that motivated this research was to investigate the possible distributional impact of the method of financing healthcare that relies on direct payment for healthcare services. To do this, however, it was considered expedient to analyze the prior distribution of health itself and socioeconomic gradients in health distribution which accentuates the need for policies that address the redistributive effect of healthcare financing. Nigeria provides a good case study in this respect. Government subsidy on healthcare is minimal. More than 70 percent of health financing is done through out-of-pocket payment and the intervention of third parties is almost non-existent. Households therefore demand healthcare services to the limit of their income. Further more, the supply of healthcare services is largely deregulated. There are large numbers of small formal and informal healthcare suppliers of healthcare products and services which presents a picture of the ideal of perfect competition. Thus, a deregulated demand and deregulated supply of healthcare services implies a healthcare market that relies more on competition and the price system determine consumption. Such a healthcare financing arrangement has potential to breach a fundamental principle of equity in healthcare financing and in the presence of large information asymmetries lead to sub-optimal social welfare outcomes.

The general objective of this research has been to analyze the possible redistributive effects of financing healthcare out-of-pocket given the market environment for the demand and supply of such healthcare services. Doing this requires a fair and more complete picture of the distribution of health and income. That is, a clear picture of the
prior distribution of health itself is important in order appreciate the direction of burden of financing healthcare. The realization of the broad objective of the study requires more detailed specification of the objective and the development of instruments for realizing them. Appropriate instruments for the achievement of the specified objectives are therefore developed both under a general framework and at more specific levels of empirical applications.

On account of severe limitations of the existing data, a multistage sampling design for household survey was used to generate appropriate, population representative, sample data for the study. The empirical methods of estimation and result generation are designed to as much as practicable capture the data-generating process which is the goal of any econometric analysis. In all, every of the objectives listed in section 1.4 of the study is considered to have been realized with appropriate instrument.

10.2 Summary of Main Findings

In general, the results show that health as measured by self assessed health (SAH) and the health utility index (HSI), has a distribution that unlike income distribution. Health distribution is skewed to the left. This implies that majority of people report from good to excellent health under SAH and similarly, majority indicated HSI of above 70 on a 100 point scale. The non-parametric kernel regression and normal distribution function are used to explore the shape of health distribution. These are illustrated in Figs 6.1 to 6.3

The growing general policy worries in developed countries about socially generated inequities in health that tend to favour those that are better placed in society: those with better education, more income, being male, being younger in age etc are found to be even more justified for developing countries like Nigeria. Evidence from this study point to significant inequalities not only in pure health (Chapter 6) but more disturbingly in socially generated inequalities. Health distribution among the Nigerian population follows the socio-economic gradient suggesting that those at the lower rungs of the socio-economic distribution are less healthy than those at the top. While some of these differences may be justified and therefore considered natural inequalities (for example,
health inequalities based on age), others might be considered as inequities resulting from social patterning of behaviour. For example, the results show there are significant differences in health inequalities based on gender and level of education. There are no \textit{a priori} reasons to expect that women should report poorer health than men.

Using the extended concentration index, the study also finds that the extent these socioeconomic factors contribute to health inequalities depends also on the social decision-maker's aversion to inequalities in health (see Table 6.4 and Fig 6.6). The Ordered Probit model was used to analyze the determinants of health status. The determinants of health status are considered to be different from the determinants of health inequalities. While a variable may contribute to health status it need not necessarily contribute to health inequality if it is evenly distributed in the population. However, the study finds that many of the variables that determine health status also contribute significantly to health inequalities. Some of these variables include income, age, level of education, sex, and disability status; though income was not as significant as \textit{a priori} expectation would suggest.

From the methodological perspective, it's gratifying to note the near equivalence of results obtained using two different measures of health: self-assessed health (SAH) and the health utility index using the Visual Analogue Scale of the EQ-5D. The study finds that for most parts, the results from the two instruments seem to converge. However, this convergence needs to be further validated with other populations and data.

Before analyzing the core redistributive effects of healthcare financing, the study explores the statistical relationship between income (proxied gross household expenditure), healthcare payment and net payment income distributions. This investigation revealed that healthcare financing also follows the income gradient (Table 7.1). Households at the lowest quintiles of the income distribution also pay relatively less for health. In other words, mean households' healthcare purchases vary according to their level of mean income. A further disturbing fact is that, possibly driven by high prices of healthcare, more than 55% of the population seek healthcare from what could be regarded
as informal sources and self medication or home treatment. Only 21% of the people reported seeking health from a public hospital or clinic. The remaining 79% sought health from one form of private sector healthcare provider or the other, or resort to self-treatment. Average price per visit to a hospital is about N230 which is about the 35% of mean income of the lowest income quintile in the distribution. Furthermore, price for medical care varies depending on which healthcare provider is sought. On the average, private hospitals charge the highest amount of money, followed by the public hospitals, mission hospitals and traditional medicine practitioner, in that order. The cheapest source of healthcare is the pharmacy/drugstores alternative. More than half the population patronizes this last group of healthcare suppliers. The burden of healthcare financing is not evenly distributed. While the poorest deciles in the income distribution spend about 8% of their income on healthcare, the top income quintile spends nearly 14% of its income on healthcare indicating perhaps that the low income households are unable to afford health expenses and so are excluded from healthcare utilization. This indicates the need for policy intervention in order to help these lower income households have access to healthcare. The proportion of the income of the top household devoted to healthcare as out-of-pocket payment compares with the total per capita healthcare expenditure financed from other sources in some other countries.

The analysis of the incidence of catastrophic health financing is summarized in Table 7.2. It shows that the incidence of catastrophic financing of healthcare varies depending on the catastrophic threshold adopted and the assumptions about aversion to catastrophic financing. However, the table shows large incidence of catastrophic financing in the population whether measured as catastrophic headcount, as catastrophic gap index, or even on assumption of no aversion to catastrophic financing. For example, the result shows that even without any aversion considerations, if catastrophic financing is defined as 5% of income, then the incidence of headcount catastrophic financing is 29%. It decreases to 22% if it is defined as 10% of income. The situation gradually gets worse as the aversion parameter \( \nu \) is increased to 5.
Similarly, Table 7.3 shows that health financing contributes to a very large extent in worsening the poverty situation in the country. For all the three poverty measures used in the study, and for all utility transformations on income, it is found that healthcare financing worsens income poverty. For example, it contributes about 7% to headcount index under per capita income. It accounts for as much as 18% and 22% of total poverty gap and square poverty gap indices, respectively even under a moderate assumption of scale elasticity of 0.7. Thus, in all cases, healthcare financing is shown by this study as a major source of impoverishment among the Nigerian population.

Turning now to the part of the study that analyzed the redistributive effect of healthcare financing, two issues are rated as primary importance here. The first is to unpack the components of redistributive effect. The second is the evaluation of the frameworks that are used to for this purpose in literature since no such attempts have been made before. Yet it is clearly recognized that methods and assumptions underlying them are critical to outcomes and policy relevance of results. In this respect the study finds that, in general, there are significant inequities involved in direct healthcare financing in Nigeria and there also significant differences in results obtained under different methodologies and underlying assumptions.

The study highlights the fact that the healthcare financing system in Nigeria is pro-rich in its redistributive effect. This is despite the apparent progressivity of the financing system, implying that this progressivity is not sufficient to translate into a pro-poor redistributive effect. However, the statistical significance of the redistributive effect is not decisive enough to warrant clear conclusions about this regressive effect. This holds true under both the AJL and DJA frameworks, and under all assumptions made in the study about the level social aversion to inequality. Part of the reason for this ambiguity in result seems to be the fact that, the average proportion of household income devoted to healthcare financing (about 9%) is not sufficiently large to have a decisive influence on redistributive effect. Yet, the proportion is considered very large in relation to the source of financing: households’ budgets which are generally poor.
Notwithstanding, this is a remarkable result. It shows that out-of-pocket payment can also exhibit evidence of progressivity even if it is in this case more an indication of proportionality. In other words, private financing of healthcare could also lead to pro-poor equity. However, this pro-poor equity could as well be evidence that the poor under-utilize healthcare services because they are unable to afford the cost. This then indicates that there is unequal access to healthcare in Nigeria based on income since household financing while not equal to household consumption (particularly if there are exemption policies in place) reflects more household consumption health consumption as opposed to public financing.

For the purpose of estimation, vertical equity is defined as "appropriate differentiation among unequals" (Musgrave 1990) or unequal treatment of unequals. Similarly, horizontal equity is defined as the equal treatment of equals while reranking is defined as the gain or loss in rank in the transition from pre- to post-payment income. In this respect the study finds that there are significant vertical and horizontal inequities as well as reranking effects in the out-of-pocket payment mechanism. This is true under both frameworks but particularly so under the DIA framework. The estimated components of the redistributive effect are found to be sensitive to the definitions of income band under the AJL approach. It is also confirmed that under the AJL approach there is a policy trade-off involved in the desirable levels of vertical and horizontal inequities. A similar trade-off exists under the same framework for reranking and horizontal inequity. Higher levels of horizontal inequity imply lower levels of re-ranking, and vice versa.

Furthermore, the study also establishes that under the DIA framework, horizontal inequity is sensitive to assumptions made both in respect of the level of social aversion to inequity (v) and the level of aversion to uncertainty in net incomes (ε). Higher values of v are associated with increasing income inequality both in the pre and post-payment distributions, increasing regressivity in redistributive effect, increasing vertical and horizontal inequities but decreasing reraking effect. Higher values of ε are associated with increasing inequality in pre- and post-payment income distribution, higher redistributive effect in absolute terms, higher vertical and horizontal inequities, and a
declining re-ranking effect. Unlike in the AJL model, the horizontal and vertical equity indices seem to move in the same direction under the DJA framework.

In particular, the study finds large differences in the values of vertical and horizontal inequities, as well as reranking effect under the two frameworks. In fact, the analysis reveals that reranking effect estimated from the DJA framework is approximately equal to the size of the reranking effect and the horizontal inequity combined, estimated from the AJL framework. This results from the differences in their conception of reranking effect. While the AJL sees reranking only when there is an overlap between two distinct prepayment income bands in the post-payment distribution, the DJA framework sees reranking in every instance of rank mutation in the transition from pre- to post-payment distribution. It is, therefore, not surprising that the reranking effect from the DJA model is equal to the reranking effect from the AJL model only when the horizontal inequity estimated from the AJL model approaches zero, implying that the income bandwidth of the AJL model must also approach zero.

There is also a fundamental conceptual difference in what the frameworks estimate as horizontal inequity. In other words, horizontal inequity in the AJL framework is not conceptually the same with that in the DJA framework. While the horizontal inequity in the AJL model \( H_{AJL} \) estimates the aggregate level of inequity induced by the payment system among pre-payment equals in the post-payment distribution, the horizontal inequity in the DJA model \( H_{DJA} \) estimates the aggregate level of risk premiums engendered by uncertainty in the transition from the pre- to post-payment income distribution, given the level of aversion to uncertainty.

**International Comparison of Redistributive Effect of Direct Healthcare Financing**

Table 10.1 compares some results from the redistributive effects of healthcare financing found in Nigeria with those of some other countries. The results for Vietnam are from Wagstaff and van Doorslaer (2001) while the results for the OECD countries are from
van Doorslaer et al. (1999). The comparison is restricted to only the comparable aspects of the results. For all the countries, the decomposition framework is the AJL model.

**Table 10.1 Redistributive Effects of HCF in Fourteen Countries**

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>RE</th>
<th>V</th>
<th>g</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nigeria</td>
<td>2004</td>
<td>-0.008</td>
<td>0.0138</td>
<td>0.0883</td>
<td>0.1419</td>
</tr>
<tr>
<td>Denmark</td>
<td>1987</td>
<td>-0.0021</td>
<td>-0.0021</td>
<td>0.0077</td>
<td>-0.02654</td>
</tr>
<tr>
<td>Finland</td>
<td>1990</td>
<td>-0.0039</td>
<td>-0.0037</td>
<td>0.0150</td>
<td>-0.2419</td>
</tr>
<tr>
<td>France</td>
<td>1989</td>
<td>-0.0033</td>
<td>-0.0031</td>
<td>0.0091</td>
<td>-0.3396</td>
</tr>
<tr>
<td>Germany</td>
<td>1988</td>
<td>-0.0014</td>
<td>-0.0010</td>
<td>0.0098</td>
<td>-0.093</td>
</tr>
<tr>
<td>Ireland</td>
<td>1987</td>
<td>-0.0015</td>
<td>-0.0014</td>
<td>0.0093</td>
<td>-0.1472</td>
</tr>
<tr>
<td>Italy</td>
<td>1991</td>
<td>-0.0023</td>
<td>-0.0017</td>
<td>0.0201</td>
<td>-0.0807</td>
</tr>
<tr>
<td>Netherlands</td>
<td>1992</td>
<td>-0.0005</td>
<td>-0.0004</td>
<td>0.0095</td>
<td>-0.0377</td>
</tr>
<tr>
<td>Portugal</td>
<td>1990</td>
<td>-0.0111</td>
<td>-0.0084</td>
<td>0.0336</td>
<td>0.2424</td>
</tr>
<tr>
<td>Sweden</td>
<td>1990</td>
<td>-0.0005</td>
<td>-0.0004</td>
<td>0.0018</td>
<td>-0.2402</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1992</td>
<td>-0.0011</td>
<td>-0.0011</td>
<td>0.0030</td>
<td>-0.3619</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1992</td>
<td>-0.0021</td>
<td>-0.0017</td>
<td>0.0079</td>
<td>-0.2229</td>
</tr>
<tr>
<td>United States</td>
<td>1987</td>
<td>-0.0128</td>
<td>-0.0109</td>
<td>0.0271</td>
<td>-0.4603</td>
</tr>
<tr>
<td>Vietnam</td>
<td>1993</td>
<td>-0.0277</td>
<td>-0.0116</td>
<td>0.059</td>
<td>-0.0421</td>
</tr>
<tr>
<td>Vietnam</td>
<td>1998</td>
<td>-0.0174</td>
<td>-0.0109</td>
<td>0.054</td>
<td>-0.0025</td>
</tr>
</tbody>
</table>

The point to note in all the countries is that the overall redistributive effect of direct healthcare financing is negative. That is, direct healthcare financing tends to be pro-rich in redistributive effect. However, the magnitudes of this pro-rich effect differ from country to country. Vietnam (1993 and 1998) and United States seem to have larger redistributive effect (in absolute terms) than most other countries including Nigeria. However, the value of vertical equity for Nigeria is positive as against the negative value obtained for the rest of the countries including Vietnam (for both years). Thus, this result is striking because as we observed above, direct healthcare finance can also potentially be progressive in vertical equity. This progressivity of direct healthcare financing is clearly
also shown by the positive value of the Kakwani index of progressivity (K) for Nigeria as against its regressivity in other countries included in the table. It is possible that this progressivity is driven by the large proportion of household income absorbed in Nigeria by direct payment. For example, the value of g for Nigeria is more than 11 times the value for Denmark, and more 29 times the value for Switzerland. It seems, therefore, reasonable to conclude that it is possible for direct healthcare payment to be progressive on household income as the Nigerian case has shown, but it is also to be noted that such progressivity is achieved at a cost: the possible exclusion of households who cannot afford the high value of g from utilizing healthcare. Thus, progressivity of direct healthcare financing may imply regressivity in utilization of healthcare.

**Robustness of Estimated Models**

With respect to the robustness of the estimated parameters to assumptions about household consumption scale elasticity, it is found that in both frameworks, the estimated parameters are sensitive to assumption made about household economies of scale and composition. However, it is noted that there is less inequity in pre- and post-payment income distributions in both models when the per capita unit scale is transformed into utility scales of adult equivalence, and consumption elasticities of 0.7 and 0.3. This is, perhaps, the consequence of the presence of large number of large households and households with preponderance of children in the study area. However, there are no corresponding monotonic decreases in the components of the redistributive effects. For example, it is found that controlling for values of ε and ν in the DJA the redistributive effect increases marginally in absolute terms from 0.0033 in per capita to 0.0034 in scale economy of 0.7 before declining steeply to 0.0023 in economy scale of 0.3. Similarly, it is found that keeping the definition of the income band constant, the value of V varies from 0.0136 under per capita to 0.0143 and 0.0184 for economy scale of 0.7 and 0.3 respectively while \( R_{A,k} \) declines from 0.0070 under per capita to 0.0039 for economy scale of 0.7 and rises to 0.0103 for the scale of 0.3. Thus, the effects of household composition and scale elasticities of consumption on the components of the redistributive effect do not seem to be a priori determined.
In general, this study identifies fundamental technical differences between the AJL and the DJA frameworks for decomposing redistributive effects. To highlight some of the differences, it notes that while the AJL framework is implemented under a rigid political/ethical environment, the DJA framework is implemented under a more ethically flexible and politically democratic environment. The DJA model does this by proposing an ethical/political parameter of aversion to inequality. Thus instead of assuming the implied ethical value in the Gini index (neutrality in respect of whether inequality happens at the top or bottom of the distribution) as is the case with the AJL model, the DJA model generalizes the ethical parameter to accommodate any level of political or ethical aversion to inequality. This is considered more democratic and consistent with the basic postulates of the public choice theory which emphasizes that social decision-making is a process involving trade-offs between those whose welfare is improved by a policy and those that are hurt by it.

Secondly, the DJA framework purifies the weighting scheme used in the AJL framework \( (a_x) \) for computing horizontal equity, of its contaminations with vertical equity influences. This it achieves by disassociating the population shares of prepayment equals (which is income independent) from the proportion of the post-payment income of prepayment equals (which is income dependent). Therefore, in the DJA, unlike in the AJL, the weighting scheme \( a_x \) becomes 'pure' weight.

Thirdly, in order to solve the problem of arbitrary definition of income equals, the DJA framework proposes the use of non-parametric kernel regression function which provides an optimal trade-off between bias and mean square error in the estimation of the joint distribution of the pre-and post-payment incomes. In so doing it removes the element of arbitrariness in the definition of income equals that weakens the theoretical grounds of the AJL model and transforms it into a statistical process in which the distribution has well known statistical properties. The result of this is a more efficient and consistent estimation of the components of the redistributive effects.
10.3 Policy Discussion and Conclusions

At different points in the course of this presentation, some policies have been suggested to deal with specific issues arising from the study. At this point, we now consider some general conclusions and policy directions that are suggested by these findings. This study has shown that there is a combination of high inequality in income and inequities in health distributions in the Nigerian population using Enugu state as an example. Both forms of imbalances are created and sustained by the social structure. Those at the lower rungs of the social strata are disadvantaged on both counts. On the other hand, the very nature of ill-health and its attendant financial demands as a random shock to the household requires that healthcare financing arrangement should be such as to mitigate the effects of this shock and enhance the capacity of households especially to cope with its negative effects. The rule of equity and fairness requires that public policy should be so designed as to offer the most protection to those who are most socially vulnerable and equal protection for those in equal need from the adverse financial effects of ill-health. Unfortunately, this does not seem to be the case under the present financing arrangement for healthcare financing in Nigeria. What shows up rather is a situation where little or no protection is available to poor households with the result that people finance healthcare according to their ability to pay. Those who are unable are implicitly excluded from healthcare utilization in time of need. Financing healthcare becomes for most households an exercise in managing catastrophe.

The redistributive out-comes from this arrangement are predictable: at best healthcare financing is proportional to income implying zero redistributive effect, at worst it is regressive, implying that the poor are implicitly subsidizing the rich. The results of this study lie within the purview of a priori prediction. It induces wide gaps in post-payment income of prepayment equals. It not only reranks but also impoverishes non-poor households as much as it sinks prepayment poor households into deeper poverty. In spite of the existence of huge inequities and inequalities in health and income distributions, and in the financing of healthcare that favour the well off, the study also finds that healthcare financing is not only a burden to the poor, it is also a burden to the rich and
therefore seems to limit the scope for cross-subsidization and redistribution of income through the existing health financing arrangement.

What policy directions are suggested by these general findings? The findings do not suggest mere alterations in levels of application of existing policy instruments in order to address perceived shortcomings of inequitable health financing arrangement. For example, it would not be sufficient to merely increase the level of control and surveillance over the ethical conduct of healthcare suppliers. They rather point to fundamental imbalances in the social structure in general which manifests itself in a particular way in the health sector. Policy solutions would, therefore, seem to require more fundamental approach to the particular manifestation of structural imbalances in the health sector and its financing arrangement. In this respect, this study points to three levels of structural approach to reform the health sector and address the problem of inequities in health and income redistribution in Nigeria. Firstly, there is need for wider economic reforms aimed at narrowing the huge inequalities in resource distribution and which would seek to give every individual and household access to basic needs that are promotive of health and well-being. Secondly, the findings suggest a policy reform of the supply side of the healthcare market. Thirdly, the study suggests a reform of the healthcare financing system in the country. The next few paragraphs attempt to summarize the rationale for these policy suggestions.

Why the need for a broad based socioeconomic structural reform in Nigeria? Section 2.4 of this study cites a quotation from the editorial of one of Nigeria's leading newspapers, to the effect that a few powerful and influential groups seems to have captured the state. While this few live in opulence comparable to that in any part of the world the large majority live in poverty. The results from this study seem to confirm that observation (see Table 7.1). Furthermore, throughout the report, but particularly in Chapter 2, this study emphasizes the close links between the socioeconomic environment and health. Several studies in the past have also identified this close connection between socioeconomic structures and health of the population. If, the diseases of the population do indeed arise from the structural defects of the society, to paraphrase Rudolf Virchow, an important
angle from which to view the Nigerian health profile is the social defect symptomized in huge resources imbalances that give the "top 1%" free access to affluence while the "bottom 75%" live in penury, where the top richest households that finance healthcare spends 3300% of the amount spent by the bottom 60% of households for the same purpose. This level of resource imbalance gives a few access to affluence and denies the majority the very basic requirements for human life.

Thus, if the Nigerians are to enjoy "a state of complete physical, mental and social well-being" as defined by WHO and if the large gaps in the health of Nigerians is to be meaningfully addressed, it seems that the right starting point is for policy to address the provision of basic human needs that give rise to poor population health. The Alma-Ata Declaration (1978), recognizing the indispensability of social structures in provision of good health for all emphasized the comprehensive approach that provide both health services and also addresses the defects of the social structure that give rise to poor public health. These basic social, economic and political needs, the absence of which causes poor health include food, water, shelter, education, electricity, employment, clean environment, transportation, access to healthcare services. The basic needs also include the protection of rights of women and children, community empowerment, and addressing imbalances in social and economic power relations etc. When these basic needs are equalized for every individual in the population, then those 'defects in society' to which "the diseases of the population" are traceable would, to a large extent disappear. Then also would the huge inequalities in health that are determined by the social structures be largely eliminated. In summary, therefore, inequities in health distribution that give rise to inequities in healthcare financing could be broadly addressed by reforming the social structures that generate inequities in health.

How can policy achieve such socioeconomic structural reforms? There are multiple policy instruments at the disposal of government that can bring this about in Nigeria. These instruments include new commitments to redistributive justice through fiscal reforms in terms of greater accountability and transparency in the use of available resources, greater drive towards compliance with national tax laws and plugging of
leakages in the national resources, directing public spending to critical human areas such as health, education, nutrition, water resources, roads, electricity, sanitation, women empowerment, and making access to these basic needs the right of every citizen. This requires, not necessarily an increase in government fiscal outlay (though as we shall argue later, this is also required) but rather a rearrangement of priorities. For example, the huge amounts of public expenditure on opening and maintaining foreign missions and embassies in every country of the world, the large amounts used in purchase of military equipments, government administration, and in flying top government officials abroad for treatments that could be obtained at home and so on, could easily be used to provide some of the basic needs of the large population of the country.

There are several examples of countries that have achieved high level of public health at low cost which Nigerian policy makers can emulate. A good number of studies have shown that attaining population good health is not necessarily a function of income, at least not of income alone. Indeed, the relationship between income and public health has quite easily been shown to be a contingent relationship implying that there is no causal necessity between the two variables. While high GDP, for instance, may tend to facilitate improvements in public health, such effects have been shown to be generally mediated through other channels such as social capital and spending on social facilities including health and education.

It is, therefore, not surprising that several communities and countries have been able to achieve longevity of life, avoidable suffering and premature death for their populations with minimum income. A number of studies have also shown the contrary evidence of communities and countries with abundance of resources and income levels that are unable to achieve decent healthy life and well-being for their populations. But there are also a number of other countries such as Costa Rica and Sri-Lanka, China, Indian state of Kerala that have achieved drastic reductions in mortality and morbidity without much of economic growth.
The foregoing is only to show the contingent nature of the relationship between high incomes and population good health. What policy lessons could be drawn from this? It is simply that while more resources are certainly required, more economic growth, greater foreign investment, reductions in resource-outflow etc, far more can be achieved in public health, in reductions in inequalities in health distribution, reductions in child and infant mortality, reductions in unnecessary pains and sufferings, and improvements in life expectancy and wellbeing of the Nigerian population given the present level of resource constraints in the country. Thus, resource constraints alone is not the sole reason for the deplorable state of public health and impoverishment in Nigeria but how the available resources are managed and on what they are spent. What is required is to rearrange national priorities in order to direct available resources to areas that have critical positive multiplier effects on human welfare, areas that balance costs against benefits. For example, a policy that shifts resources from over-bloated government administrative paraphernalia (some government officials have up to 15 official cars) that often takes more than 30% of national budget and some states budgets to provide meals for school children will not only boost the literacy rate but also contribute significantly to reductions in undernourishment, ill-health and low performance among children and other positive multiplier effects.

Furthermore, the anticipated new policy direction requires sustained effort at curbing corruption and patronage in government which give the rich and the influential limitless access to the public till while depriving basic needs to the powerless majority. For example, in a recent interview on Federal Radio Corporation of Nigeria (FRCN) 'Tuesday Night' program about October 2004, a senior special assistant to the president in the Due Process Office and Price Intelligence Unit, Dr Obi Ezekwesili, announced that the country now gets 45 kobo worth of services for every one naira it spends on public projects (that is, 45 cents for every $1), and this is up from previous rate of 25 kobo for every N1. She also announced that in less than three years of its establishment her office saved the country N102 billion (an equivalent of nearly $1billion) from inflated and unexecuted contracts. This amount is more than the entire national budget for education in 2000. Add to this, illegal siphoning away of public funds by public office holders, then
a picture of the level of leakage and waste in government that denies most Nigerian access to basic needs begins to emerge.

The foregoing is meant to underscore the fact that inequalities in health and healthcare financing can only be meaningfully addressed in the context of general improved conditions of living in Nigeria. It is not to say that improved funding of health is not important. Rather, it emphasizes the fact that improved funding becomes more productive in the context of synergies between improved social conditions of living and health. In other words, the socioeconomic conditions constitute important inputs into the public health production function. There is definitely a critical need to scale up the health share of the national budget in line with the Abuja 2001 Declaration and other international standards. A situation where health services take less than 3% of national budget as against 15%, $2 per capita as against $34 benchmark for Low Income Developing Countries, 0.2% of GDP as against 1% is not only inconsistent with the country’s avowed commitment to observe these international standards but also widens the gaps in health of the poor and the well-off and implicitly widens the income gaps as well.

The point that is being made here is that more effective monitoring and tracking of government expenditure would not only redistribute income but also provide the resources necessary to improve the well-being, including health, of every Nigerian. The obvious anticipation is that those who feed on public treasury and those whose power of patronage and influence would be affected by such policy of resource transfer will attempt to frustrate it. Again the issues boil down to commitment and the political will to do that which needs to be done in order to free resources to be spent on critical areas that translate income into better population health.

It is strongly believed that a commitment to implementation of this policy of reform in public spending will not only lead to significant resource redistributions in the country but also achieve significantly improved health status for every Nigerian. Improved population health results in greater economic productivity as the complementarity between growth and equity discussed in Chapter 3 becomes actualized. When the general
health status is improved, the disease burden is lowered; then, health is equalized to a large extent for everyone, and with it, large disparities among households in financing healthcare resulting from over dependence on curative healthcare are also reduced.

The findings of this study point also to fundamental weaknesses of the healthcare market which result in excessive high costs of seeking healthcare. Healthcare is priced beyond the reach of the poor and the not so poor as shown in section 7.2. There, it is noted that the average charge for treatment in a private hospital is N1869.38 which is 65% of the monthly income of an individual living just on the poverty line. Similarly, the average cost in a government hospital is nearly as high, N1438.24. A visit to a drug store or pharmacy shop costs about N231, and 55% of the population patronize the informal drug dealers than seeking proper treatment when they are sick because they cannot afford the price of formal providers. The high cost of care results not only in vertical and horizontal inequities and reranking effects but also in catastrophic financing and impoverishment of households. But the high cost of obtaining healthcare is only a symptomatic manifestation of the underlying disease in the healthcare market itself. It therefore requires fundamental approach if policy is to effectively address this problem.

There are four major factors that seem to drive high prices of healthcare in Nigeria. One is preponderance attention to curative care that impacts on the proportion of household budget devoted to healthcare (relatively large value of $g$ discussed earlier in Chapter 8). Competition to maximize profit by favourable selection of only those healthcare services that yield profit seems also to be another driving force behind high cost of care. Thirdly, the inadequate attention to primary, preventive and promotive healthcare services and health education among the population is another force dictating the price momentum. The fourth factor is over reliance on imported drugs to meet local needs.

We consider there are basically two key related policy issues that are involved in any policy attempts to reduce the high cost of obtaining healthcare in Nigeria – the issues of the appropriate market structure for pricing of healthcare services and the appropriate method of financing healthcare in a developing country. While there is a growing trend
among the rich nations towards stronger roles for the market in the health sector for efficiency reasons, there must also be an acceptable minimum level of government presence in the delivery of health services for a developing country like Nigeria. Due to large and extensive market failures in the healthcare market as discussed at length in Chapter 3, the prediction is that the outcomes of an unfettered competitive market model in healthcare would be suboptimal. This would be so in developed countries with all forms of consumer rights and protection, and even more so in developing countries like Nigeria where consumer protection, and the intervention of third party are almost non-existent.

The policy response to this healthcare situation must start from an appreciation of the weaknesses of the supply side of the Nigerian healthcare market. In the first place, the organizational structure of the public sector healthcare as noted in Chapter 2 is complex and not easily responsive to the health needs of the population. The command structure which follows the federal structure of the country does not seem suited for a result-oriented healthcare system. FMoH is suppose to have overall supervisory role on health matters over the entire country but it is ineffective in this role because SMoH are not directly responsible to FMoH but to their state governors. SMoH is not obliged to take orders from FMoH because the two bodies are autonomous. This further implies that FMoH cannot monitor directly the activities of the local government Department of Health (LGDH) which are by constitution responsible for vital health programs at the grass-roots, including primary healthcare (PHC). Data and other planning information do not flow from LGDH through SMoH to FMoH. This impedes the stewardship role of the FMoH and severely limits ability to monitor and control epidemiological developments in the country. With severe limitations of capacities at the frontline delivery point of the health system such as the LGDH it becomes difficult for FMoH to access the extent health services such as child immunizations are delivered.

Furthermore, the relationship between the public and the private sectors of the health delivery system is that of competitors not that of co-operators. Again, it is difficult for information to flow from one side of the supply system to the other. This makes planning
difficult both at the federal and state levels. It also makes healthcare delivery expensive. This is only to underscore the fact that allocative inefficiency that results from non-integrated health system also contributes to high cost of treatment. The allocative inefficiencies that characterize the Nigerian healthcare delivery has been sufficiently discussed in section 2.9. These allocative inefficiencies both in the public and private sectors of the supply market are passed on to the consumer in the form of high medical bills.

While the government sector healthcare delivery is characterized by incentive problems ranging from lack of measurable goals and objectives for the evaluation of staff and bureaucracy and lack of commitment a different set of weaknesses characterize the competitive private sector. The most prominent of these include (see section 7.2) unethical practices, importation, distribution of fake and adulterated drugs, hawking of drugs without regard to dosage, and practices whereby patients are required to deposit money before treatment can be commenced and is discontinued when deposit is exhausted. It also suffers from unnecessary fragmentation and duplication of facilities with severe material and human resources necessary for adequate healthcare. With extensive information asymmetries, agency problems and moral hazard, the incentive for healthcare suppliers to be involved in unethical practices is very high indeed.

Furthermore, and as extension of information problem mentioned above, drug importation has become one of the most lucrative businesses in the country. This accounts largely for the preponderance of imported fake and adulterated drugs in the country because importers of drugs want to maximize profit. This result to not only high cost of genuine drugs but also to large number of treatment failures and sometimes drug poisoning.

For policy to address the problem of interrelation between FHoH and SMoH requires constitutional changes because it is embedded in the federal structure of the country. However, there is a lot that can be gained if stakeholders see public health not as arena for power and supremacy but as a cooperative game in which everyone is stands to win.
In this case cooperation between FMoH, SMoHs and LGDHs would go a long way to achieve the goal of healthcare delivery to the people. It is also possible for the FMoH to use incentives to elicit cooperation from SMoH. Such incentives could come in the form of subventions for health programs FMoH wants the SMoHs and LGDHs to implement. The same sort of instrument – incentives could be extended to the private sector.

It is obvious that in a developing country like Nigeria with severe budget constraints and competing development objectives, there are limited resources for every sector including health sector, it is then unrealistic to expect that the government can provide all the health services necessary for the population. There are also serious government failures in delivery of healthcare. The large array of the formal and informal private suppliers and their extensive patronage by consumers is a measure of the relevance of the private health sector. Without strong private healthcare supply to complement public sector supply, it would be difficult to reach many remote villages in the country with any form of healthcare.

What is being suggested here is first, there is need for policy of mergers and acquisitions in the private health sector market to build a responsible private sector supply-side of the healthcare market so that the urge to maximize profit at all costs that characterize the atomistic and individualistic business in the health sector with the associated wastes and inefficiencies that make healthcare expensive could be minimized. There is need for deliberate policies that encourage the emergence of large private healthcare providers and pooling of resources with capacity to attract private sector investment in healthcare services. Such big health organizations would be required to have minimum number of delivery outlets including rural areas just as is the practice in the banking sector.

This implies the need to raise the requirements for registration of private hospitals and clinics especially in the urban areas of the country. There is a lot to gain from having such big private healthcare organizations. It would cut costs, increase efficiency in resource allocation, encourage technological advancements in healthcare, train personnel, make monitoring easier, control the importation of drugs, increase drug production and
distribution channels. In the envisaged system, the informal drug vendors would act as centres for primary healthcare under the supervision of LGDHs. The rural supply by the private sector would be complemented by increased government investment in rural healthcare. Non governmental organizations (NGOs) and international donor agencies should be encourage to supply the rural areas to complement government supplies. But this also necessitates increasing the surveillance and monitoring capacities of the SMoHs and LGDHs to ensure compliance with official regulations and ethical requirements.

In order to ensure compliance with ethical standards there is need to for creation of incentives for the professional associations in the healthcare sector to be more proactively involved in monitoring and enforcing standards of professional conducts among their members. This would include the setting up of professional courts with powers to try erring members and withdraw license to practice of those convicted. Second, the ministries of health and related agencies must be actively involved in the surveillance of all categories of players in the healthcare market with strict enforcement of the laws governing medical practice. The reduction in the unwieldy number of registered private marketers of healthcare would make it easier for monitoring agencies to ensure compliance. The policy would also include the enlightenment and empowerment of consumers to report unethical practices. The inclusion of health education in school curriculum would also be part of this policy.

More fundamentally, there is need for a healthcare financing reform in the country that separates healthcare supply and healthcare financing. The old practice of integrating supply and financing which was the dominant practice in the country before the oil burst of the early 1980s and is still the practice in government supply of healthcare leads to corruption, uneconomic decisions, waste and allocative inefficiencies in the system. The present practice of direct payment leads to inequity. In this respect, the role of the National Social Health Insurance becomes critical. This agency which is yet to be operational would act as the third-party in the healthcare market.
However, in place of a monopoly SHI out-fit, there should be at least four in order to avoid the economic problems associate with monopolies and to encourage competition among the SHI firms as well as avoid collusion. Their role would be to buy healthcare on behalf of the population who in turn make their contributions to the scheme. For those in the formal sector of the economy, it is easy to make this contribution through both employers and employees commitments as the scheme currently envisages. Employers of labour would be required to pay certain percentage of the contribution while employees would be required to pay certain percentage. Those in the informal sector of the economy usually comprise the poorer population. The difficulty then is how to cover this group. In this respect it is possible to use their tax assessment rates to determine their contribution. For those who are unemployed, the poor, the aged, pensioners, school children students etc, the cost of their coverage is borne by the government. However, in order to reduce consumer moral hazard, it is important that consumers pay a fixed minimum amount at the point of consumption. Those who do not want to participate in the SHI scheme should be allowed to register with private healthcare insurers.

While the scheme could be implemented in stages, for example, it may start with those in formal employment, the overall objective should be that within a specified time frame everyone must be health insured. The reimbursement mechanism is then negotiated between the insurance companies and the healthcare providers. With large healthcare suppliers, the difficulties involved in such negotiation would be largely reduced. A national health insurance commission should be created, as is currently proposed in the NHI plan, to supervise the scheme and ensure that SHI firms' compliance with the objectives of equity in access and utilization of healthcare services in the country. The full implementation of the scheme would also ensure effective income redistribution through healthcare financing.

It is acknowledged that there are practical difficulties in implementing such policies. For example, opposition may likely come from professional association and healthcare practitioners whose huge profits may be affected by these policies. It may also be difficult to implement the health insurance scheme in the informal sectors with
difficulties arising, for instance, from identifying the poor. Such difficulties are not insurmountable. What is required in essence is a commitment to improve the welfare of the generality of the people based on equity and cross-subsidization in healthcare financing. The outstanding performance of the National Agency for Food Drug Administration and Control (NAFDAC) in the last three years under the leadership of Dr Dora Akunyili and which has won national and international applause for putting order into the chaotic drug situation of the country has shown that with committed leadership the Nigerian health sector could be reformed and made responsive to equity in financing and income redistribution.

Finally, the results from the methodological applications suggest two points for future research. In the first place, there is need for further investigation into the near-equivalence of the results obtained from the SAH and HSI. It is important, for instance, to establish the external validity of this result. The clear differences in the estimated parameters of the decomposition results using the AJL and DJA frameworks also require other research to establish their external validity. When this is done, researchers may well consider the use of the DJA framework for future research and analyses.
REFERENCES


Arber, S. (1997), 'Comparing inequalities in women's and men's Health in Britain in the 1990s' Social Science & Medicine, 44 (6), 773-787.


Arrow, K. J. (1963), 'Uncertainty and the welfare economics of medical care' Social Science & Medicine, 44 (6), 773-787.


Atkinson, A. B. (1970), 'On the measurement of inequality' Journal of Economic Theory, 2, 244-263.


Burstrom, B. and Fredlund, P. (2001), 'Self-rated health: is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes' *Journal of Epidemiology Community Health*, 55 (11), 836-840.


Center for International Health Information (1999) [www.cihi.ca/-3k](http://www.cihi.ca/-3k).


Franklin, M. F. (1999), 'Comparison of weight and height relations in boys from 4 countries' American Journal of Clinical Nutrition, 70(suppl), 157S-162S.


Gravelle, H., Wildman, J. and Sutton, M. (2002), 'Income inequality and health: what can we learn from aggregate data' Social Science & Medicine,


Haavio-Mannila, E. (1986), 'Inequalities in health and gender' Social Science & Medicine, 22 (2), 141-149.


Hanstrom, O. (1999), 'Does the work environment contribute to excess male mortality?' Social Science & Medicine, 49 (1999), 1349-1356.


Heller, P. S. (1982), 'A model of demand for medical and health services in
peninsular Malaysia Social Science & Medicine, 16,267-284.


Homedes, N. and Ugalde, A. (2005), 'Why neoliberal health reforms have failed in Latin America' Health Policy, 71,83-86.


Hurst, J. (1985), Financing Health Care in the USA, Canada and Britain, Kings Fund Institute, London.


Klavus, J. and Hakkinen, U. (1996), 'Health care and income distribution in Finland' Health Policy, 38,31-43.


Lahelema, E. and Valkonen, T. (1990), 'Health and social inequalities in Finland and elsewhere' Social Science & Medicine, 31 (3), 257-265.


Lambo E., (2003), 'Making good health a real dividend of democracy for Nigerians under the new administration' Notes by the Eyitayo Lambo, Hon. Minster of Health to the members of the National Health Council, Abuja, 2003.


Lewis, A. (1954), 'Economic development with unlimited supplies of labour' Manchester School of Economic and Social Studies, 22 (May), 139-191.

Macintyre, S. (1997), 'The black report and beyond what are the issues' Social Science & Medicine, 44 (6), 723-745.

Macintyre, S., McKay, L. and Ellaway, A (2005) 'Are rich people or poor people more likely to be ill? Lay perceptions, by social class and neighbourhood, of inequalities in health' Social Science & Medicine, 60, 313-317.


Skoufias, E. (1998), 'Determinants of child health during the economic transition in Romania' *World Development*, **26** (11), 2045-2056.


Szalai, J. (1986), 'Inequalities in access to health care in Hungary' Social Science & Medicine, 22 (2), 135-140.


van Doorslaer, E., Koolman, X. and Jones, A. M. (2003), 'Explaining income-


278


Wagstaff A. (2005), 'The bounds of the Concentration Index When the Variable of Interest is Binary, with An Application to Immunization Inequality' Health Economics 14, 429-432.


Wagstaff, A. and van Doorslaer, E. (2000), 'Equity in health care finance and


West, P. (1997), 'Health inequalities in the early years: is there equalization in youth?' *Social Science & Medicine*, 44 (6), 833-858.


Yoder (1989), 'Are people willing and able to pay for health services' *Social Science & Medicine, 29* (1), 35-42.
APPENDIX
QUESTIONNAIRE FOR HEALTH INEQUALITIES AND INCOME REDISTRIBUTIVE EFFECTS OF HEALTH CARE FINANCING

A - GENERAL INFORMATION

A.1 Enumerator's Name
A.2 Enumerator's ID number

A.3 Respondent's Name

A.4 Senatorial District

A.5 Stratum Code

A.6 EA Code

A.7 Enter Today's Date
A.8 Enter Time Now

A.9 Respondent Number (from roster on the next page)

A.10a Senatorial Code
A.10b Stratum Code
A.11 EA Code
A.12 Household Code

Comments

QUESTIONNAIRE INSTRUCTION

STEP 1: BUBBLE IN ALL THE ANSWERS. VERIFY THERE IS ONLY ONE ANSWER PER QUESTION, UNLESS OTHERWISE SPECIFIED.

STEP 2: EXAMINE ALL NUMBER GRIDS AND MAKE SURE THE NUMBERS BUBBLED MATCH THE NUMBERS WRITTEN IN THE BOX ABOVE EACH GRID.

Create a unique reference number above, by combining Stratum Code, EA Code and household Code, EA and household number. Write the number NOW here:

REF. NO.
### B - INFORMATION ON HOUSEHOLD MEMBERS

Write down names of persons who normally live and eat together in this household, starting with the head, EXCLUDE VISITORS.

<table>
<thead>
<tr>
<th>MEMBER NUMBER</th>
<th>MEMBER'S FIRST NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Head)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

#### B.1 ENTER THE SEX OF EACH MEMBER LISTED IN THE BOXES ABOVE

- **Male**: M
- **Female**: F

#### B.2 How long has /NAME/ been away in the past 12 months?

- **Never**: 1
- **Less than 6 months**: 2
- **6 months or more**: 3

#### B.3 Who contributes most to the household income?

Interviewer, check one member

#### B.4 In the absence of the head, who is making decisions for the household?

Skip if one member household

#### B.5 What is /NAME/’s relationship to the head of the household?

- **Head**: Head
- **Spouse**: Spouse
- **Child**: Child
- **Parent**: Parent
- **Other relative**: Other relative
- **House help**: House help
- **Not related**: Not related

#### B.6 AGE at last birthday (in years)

- Fill 00 if under one year old
- Begin with 0 for ages 0-99
- Fill 99 for all ages ≥100

#### B.7 Marital Status

- **Single/Never married**: 1
- **Living together**: 2
- **Married**: 3
  - (Monogamous): 4
  - (Polygamous): 5
- **Divorced/Separated**: 6
- **Widowed**: 7
C - EDUCATION

C.1 Can /NAME/ read and write?  
-YES 
-NO

C.2 Has /NAME/ ever attended school?  
-YES (01-06) 
-NO (07-10)

C.3 What is the highest educational grade /NAME/ completed?  
-00 Primary
-01 Pre-primary
-02 Primary
-03 Intermediate
-04 Secondary
-05 Senior Secondary
-06 Further Education
-07 Technical
-08 Vocational
-09 Other

C.4 How many years did /NAME/ spend in school altogether?

D - OCCUPATION

D.1 Interviewer: Mark here members who are 15 years or older and continue with these members (refer to age on page 2)

D.2 What was /NAME/ main work status during the past 4 weeks?

- Unemployed
- Employed
- Own Account Worker
- Regular Employee
- Casual Employee
- Domestic Service Worker
- House Help
- Student/Apprentice
- Pensioner
- Investor

D.3 What industry was /NAME/ mainly working during the past 4 weeks?

- Agriculture, Forestry and Fishing
- Mining Industries
- Construction Industries
- Manufacturing
- Transport and Communication
- Wholesale Trade
- Retail Trade
- Financial and Insurance Activities
- Service Industries

D.4 What sector was /NAME/ mainly working during the past 4 weeks?

- Public
- Private Formal
- Private Informal
- Pensions/Unemployed-Public
### E - HEALTH

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.1 Is [NAME] physically handicapped or disabled?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Include members only if handicap prevents him/her from maintaining a significant activity or schooling.

| E.2 Is [NAME] sick or injured in the last four weeks? |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |

| E.3 (If answer to E.2 is 'Yes') Was [NAME] able to engage in normal activities (work, school, taking care of family) during the period he/she was sick? |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |

| E.4 If answer to E.3 is 'No. How many days did [NAME] miss work/school due to illness in the past 4 weeks |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |

| E.5 If yes to E.2. What sort of sickness/injury did [NAME] suffer? |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |

Intervener, you may mark more than one answer. The codes are written in parenthesis to the right of each sickness or injury.

| E.6 Has [NAME] consulted a doctor, nurse, pharmacist, health professional, dentist or traditional healer for sickness during the past 4 weeks? |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |

| E.7 What kind of health provider did [NAME] see? |   |   |   |   |   |   |   |   |   |    |    |    |
| Public hospital/Clinic |   |   |   |   |   |   |   |   |   |    |    |    |
| Community Health centre |   |   |   |   |   |   |   |   |   |    |    |    |
| Private Hospital/Clinic |   |   |   |   |   |   |   |   |   |    |    |    |
| Private Doctor / Dentist |   |   |   |   |   |   |   |   |   |    |    |    |
| Missionary hospital |   |   |   |   |   |   |   |   |   |    |    |    |
| Pharmaceutical store/Chemist |   |   |   |   |   |   |   |   |   |    |    |    |
| Traditional healer |   |   |   |   |   |   |   |   |   |    |    |    |
| Treated at home |   |   |   |   |   |   |   |   |   |    |    |    |

| E.8 If [NAME] was treated for injury/illness, how many times did he/she visit the healthcare provider in the past 4 weeks? |   |   |   |   |   |   |   |   |   |    |    |    |
| YES |   |   |   |   |   |   |   |   |   |    |    |    |
| NO |   |   |   |   |   |   |   |   |   |    |    |    |
### E.9 Is [NAME] covered by any form of health insurance scheme?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑</td>
<td></td>
</tr>
</tbody>
</table>

### E.10 If Yes, which of the following health insurance scheme is [NAME] covered by?

<table>
<thead>
<tr>
<th>Public health insurance</th>
<th>Social health insurance</th>
<th>Private health insurance</th>
<th>Community based pre-payment scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### E.11 How much does [NAME] pay per annum as premium for this scheme?

| Amount (€) | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ |

### E.12 Does [NAME] enjoy any form of free healthcare service?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑</td>
<td></td>
</tr>
</tbody>
</table>

### E.13 If yes to E.12, what type of free healthcare does [NAME] enjoy?

- Free treatment of hospital cases
- Free treatment of children
- Free treatment of family members

### E.14 By placing a tick (✓) in at least one box in each group below, please indicate which statements best describe your own health state today.

#### I. Mobility
- I have no problems in walking about
- I have some problems in walking about
- I am confined to bed

#### II. Self-care
- I have no problems with self-care
- I have some problems with self-care
- I am unable to wash or dress myself

#### III. Usual activities
- I have no problems in performing my main activity (e.g. work, study, housework, family or leisure activities)
- I have some problems performing my usual activities
- I am unable to perform my usual activities

#### IV. Pain/discomfort
- I have no pain or discomfort
- I have moderate pain or discomfort
- I have extreme pain or discomfort

#### V. Anxiety/depression
- I am not anxious or depressed
- I am moderately anxious or depressed
- I am extremely anxious or depressed

### E.15 How would you generally rate your state of health now? (Interviewer: Address this question to particular members of the family aged 12 years and above.)

<table>
<thead>
<tr>
<th>Excellent</th>
<th>Very good</th>
<th>Good</th>
<th>Fair</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑</td>
<td>☑</td>
<td>☑</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
E.16 To help people say how good or bad a health state is, we have drawn a scale rather like a thermometer on which the best state you can imagine is marked by 100 and the worst state is marked by 0. We would like you to indicate on this scale how good or bad is your own health today. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is.

*Interviewer: This ruler should be repeated for each member of the household who is 12 years or above.

| Best imaginable health state | Best imaginable health state | Best imaginable health state | Best imaginable health state | Best imaginable health state | Best imaginable health state | Best imaginable health state | Best imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state | Worst imaginable health state |

E.17 What is the distance from your house to the health care provider where /NAME/ was treated? (Distance in kilometers)

| Amount (H) |

E.18 How much did it cost to transport /Name/ to the health care provider where he/she went for treatment?

| Amount (H) |

E.19 Was /Name/ admitted or treated as outpatient?

1 If admitted 2 If not admitted

E.20 If /Name/ was admitted, how many days did he/she stay in the treatment place in the last 4 weeks?
E.21 If /Name/ was admitted, did anybody stay with him/her to look after him/her?

Yes
No

E.22 Who looked after /Name/ during the admission period? 

Interviewer: you may bubble more than one

E.23 Did /Name/ transport to visit the sick person where he/she was admitted?

Yes
No

E.24 If yes to E.23. How many times did /Name/ transport to treatment place either to send food or something else?

No of times

E.25 How much was /Name/ charged for:

| Card in (N) | | | | | | | |
| Diagnosis in (N) | | | | | | | |
| Drugs in (N) | | | | | | | |
| Total in (N) | | | | | | | |

E.26 Did the Household have to sell any of its belongings in order to pay for the treatment?

Yes
No

E.27 If answer to E.26 is "Yes" what was/were the item(s) sold

Amount (N)

E.28 Did this household borrow money to pay for the health care of /NAME/?

Yes
No

E.29 If the answer to E.28 is "Yes" how much was borrowed?

Amount (N)

E.30 From whom did the household borrow to finance the health care service?

Relatives
Friends
Business partner
Social Clubs
Bank
Isusu as soc
Church group
Other
### F - HOUSEHOLD ASSETS

**F.1 Does the household own any of the following? Interviewer: include items only if they are in working order**

<table>
<thead>
<tr>
<th>Item</th>
<th>Owned since more than 1 year ago</th>
<th>Owned since 1 year ago or less</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric iron</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refrigerator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video deck</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stereo system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mattress or bed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watch or clock</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sewing machine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern stove</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bicycle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car or truck</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**F.2 Owned since more than 1 year ago**

**F.3 Owned since 1 year ago or less**

---

### G - HOUSEHOLD ENERGY & WATER EXPENDITURE

**G.1 About how much did the household spend last month on... Interviewer (repeat question for each expenditure category on the following list)**

#### G.1a. ITEM (ENERGY)

<table>
<thead>
<tr>
<th>Item (Energy)</th>
<th>Item Code</th>
<th>Amount Spent in Naira</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wood</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerosene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charcoal/Coal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torch batteries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diesel oil for generators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others (Specify)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### G.1b. ITEM (WATER)

<table>
<thead>
<tr>
<th>Item (Water)</th>
<th>Item Code</th>
<th>Amount Spent in Naira</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water vendors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water rate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NB: Fill in "0" (zero) if nothing is spent on an item.**
**H – FOOD SPENDING AND CONSUMPTION**

The food spending here refers to expenditure on all food that members of the household have bought and eaten. It does not include food that has been bought for resale or used for commercial purposes.

**Interviewer:** Introduce this section by saying: “I have a list of different kinds of food that the members of the household may have eaten in the past week or month. As I call each one, please indicate whether or not it was bought or consumed.

**H.1** Was Name/ bought or consumed by the household in the past two weeks? Repeat this for each item.

<table>
<thead>
<tr>
<th>A Food name</th>
<th>B Eaten/ bought in the last 2 weeks</th>
<th>C What was the amount spent or Quantity purchased in the last two weeks?</th>
<th>D How much was received as a gift</th>
<th>E How much was eaten from own production in the last two weeks?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Quantity (Kg)</td>
<td>Amount (N)</td>
<td>Quantity (Kg)</td>
</tr>
<tr>
<td>Yam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carri</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akpu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abacha</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akidi/Beans</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plantain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bread</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breakfast Cereal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meat</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Onions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tomatoes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetable oil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palm oil</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margarine/Butter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eggs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bananas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruits (e.g.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oranges</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soft drinks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groundnuts</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okra</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akara</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biscuits</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### I - REGULAR NON-FOOD SPENDING

**1.1 Interviewer asks:** "In the past two weeks, about how much was spent on...

<table>
<thead>
<tr>
<th>Category</th>
<th>Amount (₦)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C Igarette/Tobacco</td>
<td></td>
</tr>
<tr>
<td>Beer, wine, Spirit</td>
<td></td>
</tr>
<tr>
<td><strong>ENTERTAINMENT:</strong></td>
<td></td>
</tr>
<tr>
<td>Home Video/VCD</td>
<td></td>
</tr>
<tr>
<td>Sports</td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
</tr>
<tr>
<td><strong>PERSONAL CARE ITEMS:</strong></td>
<td></td>
</tr>
<tr>
<td>Soaps, Shampoos, Haircuts</td>
<td></td>
</tr>
<tr>
<td>Newspapers/Stationeries</td>
<td></td>
</tr>
<tr>
<td>Envelopes/Stamps</td>
<td></td>
</tr>
<tr>
<td>Telephone</td>
<td></td>
</tr>
<tr>
<td><strong>REGULAR TRANSPORT:</strong></td>
<td></td>
</tr>
<tr>
<td>Petrol, oil car service</td>
<td></td>
</tr>
<tr>
<td>Bus, taxi, or train</td>
<td></td>
</tr>
<tr>
<td><strong>MISCELLANEOUS</strong></td>
<td></td>
</tr>
<tr>
<td>Washing powder</td>
<td></td>
</tr>
<tr>
<td>Child care</td>
<td></td>
</tr>
<tr>
<td>Religious fellowships</td>
<td></td>
</tr>
<tr>
<td>Dues of organizations</td>
<td></td>
</tr>
<tr>
<td>Informal taxation</td>
<td></td>
</tr>
<tr>
<td>Donations</td>
<td></td>
</tr>
</tbody>
</table>

### J - OCCASIONAL NON-FOOD SPENDING

**1.1 Interviewer asks the respondent:** "Whether any of the was purchased by the household in the past one year.

<table>
<thead>
<tr>
<th>Category</th>
<th>Amount (₦)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kitchen equipment (E.g. Pots, Pans, Torches etc.)</td>
<td></td>
</tr>
<tr>
<td>Home maintenance/repairs of dwellings</td>
<td></td>
</tr>
<tr>
<td>Bedspreads/Towels</td>
<td></td>
</tr>
<tr>
<td>Furniture and household appliances</td>
<td></td>
</tr>
<tr>
<td><strong>CLOTHING AND SHOES:</strong></td>
<td></td>
</tr>
<tr>
<td>Shoes</td>
<td></td>
</tr>
<tr>
<td>Clothes</td>
<td></td>
</tr>
<tr>
<td>Curtains (Window Blinds)</td>
<td></td>
</tr>
</tbody>
</table>