

**STRUCTURAL REFORM OF THE KENYAN HEALTH CARE
SYSTEM**

by

Benjamin Musembi Nganda

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy of the Department of Economics and Related Studies, University of York, England.

November, 1994

— CONTENTS —

LIST OF FIGURES	viii
LIST OF TABLES	x
ACKNOWLEDGEMENTS	xii
ABSTRACT	xiii
CHAPTER ONE	1
1. INTRODUCTION: SCOPE AND PURPOSE OF THE STUDY	1
1.0 Introduction	1
1.1 Pressures for Reform of the Health Care System	4
1.2 The Initial Response: Reforms in the Financing Arrangements	5
1.3 The Need for 'Broader Structural Reforms'	6
1.4 Purpose and Scope of the Present Study	6
1.5 Organisation of the Dissertation	7
CHAPTER TWO	9
2. THE SETTING OF THE PROBLEM—I: THE HEALTH STATUS SITUATION IN KENYA AND THE DISTRIBUTION OF HEALTH RESOURCES	9
2.0 Introduction	9
2.1 Demographic Characteristics and Implications for the Health Services	10
2.1.1 General demographic trends	10
2.1.2 Structure of the population	11
2.1.3 Life expectancy	12
2.1.4 Distribution of the total population and urbanisation	16
2.1.5 Conclusion	17
2.2 Trends in public health expenditures in the 1980's and regional resource distribution	17
2.2.1 Financial resources	17
<i>Trends in public health expenditures in the 1980's</i>	17
<i>The regional allocation of financial resources in relation regional population shares</i>	20
2.2.2 Physical resources	21
<i>Trends in physical resources</i>	21
<i>Facility distribution and access measures</i>	23
2.2.3 Trends in and distribution of human health resources	26
<i>Trends in medical and health personnel</i>	26

<i>The regional distribution of human health resources</i>	31
2.3 Morbidity and Mortality Patterns	33
2.3.0 Overall patterns	33
2.3.1 Regional out-patient morbidity patterns	36
2.3.2 Regional mortality patterns according to hospital based information ..	37
2.3.3 The distribution of health resources and health status—implications for policy	39
2.3.4 Conclusion	41
2.4 Some International Comparisons	43
2.4.0 Some caveats about international comparisons	43
2.4.1 Performance of Kenya’s health sector viewed in relation to other developing countries	44
2.4.2 Conclusion	50
2.5 Conclusions	52
 CHAPTER THREE	 55
3. THE SETTING OF THE PROBLEM—II: THE PRESENT HEALTH CARE SYSTEM IN KENYA—ITS ORGANISATION AND BACKGROUND TO PRESENT PROBLEMS	55
3.0 Introduction	55
3.1 Organisation of Health Services in Kenya	55
3.1.1 Organisation of the public health sector	56
3.1.2 The private sector of the health services in Kenya	58
3.1.3 Relationship between the public and private sub-sectors	59
3.2 A Synoptic Diagnosis of Problems and Their Causes in the Current System	60
3.2.0 Introduction	60
3.2.1 The resource allocation process in Kenya’s public health services system	60
<i>The budget setting and resource allocation process in the public sector</i>	61
<i>Some weaknesses in the budget setting and resource allocation process</i>	63
<i>The need for change</i>	66
3.2.2 The private sector	67
3.3 The Overall Picture and Implications for Policy	69
3.4 A Review of Related Previous Works on the Kenyan Health Care System	71
<i>Assessment and Conclusions</i>	81
3.5 Summary of Key Problems Faced by the Current System that Need Redressing	85
3.5.1 Problems within the public sector	85

3.5.2 Problems within the private sector	85
3.5.3 Conclusions and policy implications	86
3.6 Conclusions	86
CHAPTER FOUR	88
4. ASPECTS OF INEFFICIENCY IN THE KENYAN HEALTH CARE SYSTEM: DEFINITIONS, SOURCES AND IMPLICATIONS FOR HEALTH POLICY	88
4.0 Introduction	88
4.1 The Objective Function in the Kenyan Health Sector and Efficiency	88
4.2 Efficiency in Health Care: Some Theoretical Considerations	94
4.2.1 What should efficiency in health care mean?	95
<i>Providing only medically effective services</i>	97
<i>Providing services that are cost-effective</i>	100
<i>Efficiency also demands concentrating on health services that offer the highest pay-offs in terms of health</i>	100
<i>Efficiency also means providing appropriate scales of the medically- and cost-effective services</i>	102
4.2.2 Quality of care	104
4.2.3 Consumer choice	104
4.2.4 The private sector	105
4.2.5 The question of incentives	106
4.3 Potential for Efficiency Improvement in the Kenyan Health Sector	106
4.3.1 Some evidence of resource mis-use in the Kenyan public health care system	107
4.3.2 The MOH resource allocation is not geared to the country's health problems	110
4.3.3 An alternative 'closer' look at the distribution of recurrent curative expenditures	114
4.4 A Framework for Analysing Health Sector Problems in Kenya	116
4.5 Summary and Conclusions	125
CHAPTER FIVE	127
5. THE EQUITY OBJECTIVE IN KENYAN HEALTH POLICY	127
5.0 Introduction	127
5.1 Equity in Kenyan Health Policy and Health Care	128
5.1.1 The basis for equity concern—health care as an individual's 'right' ..	128
5.1.2 The approach adopted by government to realise the equity objective .	132
<i>Development of health services before independence</i>	132
<i>Prevailing equity conceptions and changes implemented since independence</i>	134

(i) <i>Equity and the public health services after Independence</i>	135
<i>Why the waiver system is ineffective</i>	138
(ii) <i>Equity in the private sector</i>	139
5.1.3 Conclusions	139
5.2 Philosophical and Health Economics Bases of Equity and Their Implications for Kenyan Health Policy	143
5.2.1 Utilitarianism	144
5.2.2 The difference principle	145
5.2.3 Equality of opportunity	147
5.2.4 Sen's capabilities approach	148
5.2.5 Theory of entitlements	150
5.2.6 The decent minimum	151
5.2.7 The non-envy criterion	152
5.2.8 The health maximisation principle (extra-welfarism)	153
5.2.9 Conclusions	159
5.3 Equity and Access to Health Care: Towards Operational Measures	159
5.3.1 What constitutes equity in health and health care?	160
5.3.2 'Need' and equity in health care	161
5.3.3 Need, equity and efficiency	164
5.3.4 Equality of what?—Possible 'candidates' for equitable distribution ..	166
5.4 Policy Implications for Kenya	170
5.4.1 Underlying policy guidelines	171
5.4.2 Targeting resources to specific groups	171
5.4.3 Implications for informational requirements	172
5.4.4 The need for further research	172
CHAPTER SIX	174
6. THE AGENCY RELATIONSHIP AND ITS RELEVANCE TO MODELLING A REFORM STRATEGY FOR THE KENYAN HEALTH CARE SYSTEM	174
6.0 Introduction	174
6.1 Theoretical Approaches to Explaining Variable Efficiency in the Health Sector	175
6.1.1 Non-allocative inefficiencies that may sustain X-inefficiency in the health sector	175
6.1.2 Budget restraint as tool for dealing with non-allocative inefficiencies in the health sector and its deficiencies	177
6.1.3. Budget restraint and the 'proximate' behaviour of providers	180
6.1.4 A model of hospital behaviour suited to Kenyan public hospitals	181
6.2 Towards an Alternative Policy Model for the Kenyan Health Sector	186
6.2.1 The appropriate setting—Some factors to consider	186

6.2.2 Funding and performance incentives under the agency relationship ...	189
<i>Optimal reward structure in a single agency relationship</i>	190
6.2.3 An application to centrally-funded health care facilities	192
6.2.4 Implications for improving incentive compatibility in the health sector	195
6.3 Costs, Outputs and Competition	198
6.3.1 Competitive reforms, regulation and managed care	203
<i>Competitive reforms and managed care</i>	203
<i>Prospective payment systems</i>	204
<i>Competitive contracting</i>	206
<i>Global budgeting</i>	208
6.3.2 Incentives and system performance under competition: What empirical evidence reveals	209
6.3.3 Conclusions	214
6.4 Summary and Conclusions: Implications for Health Policy Formulation in Kenya	215
 CHAPTER SEVEN	 217
7. PROPOSALS FOR RESTRUCTURING THE KENYAN PUBLIC HEALTH CARE SYSTEM	217
7.0 Introduction	217
7.1 The Criteria for Option Evaluation	218
7.1.1 Equity	218
7.1.2 Efficiency	219
7.1.3 Ease/difficulty of implementation	219
7.1.4 Adequacy	220
7.2 Proposals for Restructuring the Delivery of Public Health Care Services in Kenya	220
7.2.1 Option 1: The current system with internal changes	221
<i>Assessment in terms of the evaluation criteria</i>	222
7.2.2 Option 2: Introduction of reimbursement using 'global budgets'	223
A. <i>Introduction of reimbursement using a resource allocation formula</i>	223
(a) <i>The setting of institutional recurrent budgetary needs</i>	223
(b) <i>Setting capital budgets</i>	226
(c) <i>Some difficulties in the use of a formula to allocate resources</i>	227
(d) <i>Assessment in terms of the evaluation criteria</i>	229
B. <i>An alternative methodology for estimating institutional budgetary requirements</i>	230
(a) <i>Setting recurrent budgets</i>	230
(b) <i>Capital budgets</i>	234
(c) <i>Assessment according to the evaluation criteria</i>	234

7.2.3 Option 3: Introducing publicly financed competition in the health services through prospective reimbursement budgets	235
(a) <i>Use of DRG information to set prospective budgets</i>	235
(b) <i>Use of population and utilisation based standards</i>	236
(c) <i>Assessment in light of the set criteria</i>	239
7.2.4 Option 4: Introducing publicly financed competition in the health services through the separation of the financing and provider roles	239
(a) <i>Restructuring the system through giving prominence to primary health care</i>	241
(b) <i>Other necessary changes in the organisational structure of public hospital services</i>	242
(i) <i>Changes in the status of public hospitals</i>	242
(ii) <i>Changes in the way the National Hospital Insurance Fund operates</i>	242
(c) <i>How competition can be injected into the new system</i>	245
(i) <i>Tackling equity and efficiency in the new system</i>	245
<i>How the overall budget would be split</i>	245
<i>Districts as budget holders: How they'll allocate resources to ensure efficiency</i>	246
<i>Allocating the curative services budget resources</i>	246
<i>Allocating resources for primary care at the district level</i>	247
(ii) <i>The process of contracting</i>	249
(iii) <i>Implementing the change—the time framework</i>	250
(d) <i>Some issues to be addressed in the future</i>	252
7.2.5 An assessment of the proposals—a summary	253
7.3 Some 'Observations' about the Private Sector	254
7.4 Summary and Conclusions	256
CHAPTER EIGHT	259
8. CONCLUSIONS	259
8.0 Summary	259
8.1 Conclusions/Recommendation	264
8.2 Weaknesses and Strengths of the Thesis: An Overall Assessment	265
8.3 Research Implications of the Suggested Structural Reform	266
REFERENCES	268

— LIST OF FIGURES —

FIGURE 2.1: VITAL POPULATION STATISTICS: 1950—1990	12
FIGURE 2.2: PUBLIC EXPENDITURE ON HEALTH, EDUCATION, DEFENCE AND OTHER SOCIAL SERVICES, 1980/81—1990/91	18
FIGURE 2.3: RELATIONSHIP BETWEEN GNP, GOVERNMENT BUDGET AND VARIOUS FORMS OF GOVERNMENT EXPENDITURES, 1980/81—1990/91	19
FIGURE 2.4: PER CAPITA PUBLIC SPENDING ON SELECTED SERVICES TRENDS: 1980/81-1990/91	19
FIGURE 2.5: TRENDS IN HEALTH FACILITY STOCK BY TYPE, (1978-1990):	22
FIGURE 2.6: TRENDS IN HOSPITAL BEDS BY TYPE OF OPERATING AGENCY, 1964-1989	23
FIGURE 2.7: TRENDS IN SELECTED CADRES OF HUMAN RESOURCES EMPLOYED IN THE HEALTH SECTOR OF KENYA: 1960—1990	27
FIGURE 2.8: SURPLUSES/DEFICIENCIES IN VARIOUS CATEGORIES OF PERSONNEL: 1991/91	30
FIGURE 2.9: PERCENTUAL DEVIATIONS FROM THE NATIONAL MEAN POPULATION (KENYA) PER UNIT OF SELECTED HEALTH PERSONNEL, 1991/92	32
FIGURE 2.10: PER CENT OF GOVERNMENT SPENDING ON HEALTH AND GNP PER CAPITA IN DEVELOPING COUNTRIES: 1987—1990	45
FIGURE 2.11: PER CAPITA HEALTH SPENDING AND GNP PER CAPITA IN DEVELOPING COUNTRIES: 1986—1990	46
FIGURE 2.12: LIFE EXPECTANCY AT BIRTH (YEARS) AND GNP PER CAPITA IN DEVELOPING COUNTRIES: 1989	48
FIGURE 2.13: UNDER 5 MORTALITY RATE AND GNP PER CAPITA IN DEVELOPING COUNTRIES, 1989	49
FIGURE 2.14: PER CAPITA PUBLIC SPENDING ON HEALTH AND LIFE EXPECTANCY IN DEVELOPING COUNTRIES, 1989	51
FIGURE 4.1: THE OBJECTIVE OF HEALTH POLICY AND POLICY ALTERNATIVES	91
FIGURE 4.2: HEALTH IMPROVEMENT BETWEEN T0 AND T1 DUE TO 'EFFECTIVE CARE'	98
FIGURE 4.3: DETERMINING THE OPTIMAL SIZE OF THE HEALTH SECTOR	103
FIGURE 4.4: COMPARISON OF COSTS AND QUALITY AMONG MOH HOSPITALS, HEALTH CENTRE AND DISPENSARY AND AN NGO HOSPITAL, 1988/89	108
FIGURE 4.5: FLOW OF FUNDS AND INTER-RELATIONSHIPS AMONG PARTICIPANTS IN THE KENYAN HEALTH CARE SYSTEM	117
FIGURE 4.6: A MATRIX OF HEALTH CARE FINANCING AND PROVISION IN KENYA	121

FIGURE 5.1: THE CHAIN FROM GOODS TO UTILITY	149
FIGURE 5.2: DIFFERENCES IN ACCESS COSTS/DEMANDS CAUSE DIFFERENCES IN UTILISATION	167
FIGURE 6.1: THE OUTPUT-MAXIMISING HOSPITAL UNDER DIFFERENT CONDITIONS .	200
FIGURE 7.1: TOWARDS A RESOURCE DISTRIBUTION FORMULA: A DIAGRAMMATIC REPRESENTATION	225
FIGURE 7.2: THE DETERMINATION OF FACILITY COST PER CAPITA	237
FIGURE 7.3: CALCULATION OF HOSPITAL COST—COMPONENTS TO CONSIDER	238

— LIST OF TABLES —

TABLE 1.1: TRENDS IN SOME ECONOMIC INDICATORS, 1970—1990.	3
TABLE 2.1 POPULATION AGE STRUCTURE, 1962—1979 (FIGURES IN '000'S AND PER CENT)	11
TABLE 2.2 LIFE EXPECTANCY AT BIRTH BY SEX: 1950—1990	13
TABLE 2.3 POPULATION DISTRIBUTION BY PROVINCE, 1989: PERCENT OF POPULATION BY SEX, SEX RATIO AND PROPORTION OF POPULATION UNDER 15 YEARS BY PROVINCE.	14
TABLE 2.4 PROPORTION OF POPULATION UNDER 15 IN RELATION TO FERTILITY RATE, INFANT, UNDER FIVE AND CHILDHOOD MORTALITY RATES	15
TABLE 2.5: PERCENTAGE DISTRIBUTION OF URBAN AND RURAL POPULATION BY PROVINCE, 1969 AND 1979.	16
TABLE 2.6: REGIONAL DISTRIBUTION OF RECURRENT PUBLIC HEALTH EXPENDITURE BY THE MINISTRY OF HEALTH: 1989/90 (PERCENTAGES)	21
TABLE 2.7: PHYSICAL RESOURCES AND ASSOCIATED MEASURES OF ACCESS BY PROVINCE, 1989	25
TABLE 2.8: FUNCTIONAL DISTRIBUTION OF MINISTRY OF HEALTH PERSONNEL, 1991/92	28
TABLE 2.9: POPULATION PER UNIT OF SOME SELECTED CATEGORIES OF HEALTH PERSONNEL BY PROVINCE—1991/92	32
TABLE 2.10: PATTERNS IN OPD MORBIDITY BY TYPE OF DISEASE CATEGORY: VARIOUS YEARS BETWEEN 1964 AND 1991	34
TABLE 2.11: DISTRIBUTION BY ICD CLASSIFICATION OF THE TOP 20 LEADING CAUSES OF ADMISSIONS AND DEATHS IN MOH HOSPITALS IN 1990.	35
TABLE 2.12: OPD MORBIDITY PATTERNS BY DISEASE GROUP AND BY PROVINCE, KENYA: 1984 AND 1991—PER CENT AND NUMBERS	36
TABLE 2.13: ADMISSIONS AND DEATHS BY SEX IN REPORTING GOVERNMENT HOSPITALS 1968 AND 1990 BY PROVINCE	38
TABLE 2.14: SIMPLE CORRELATION BETWEEN IMR, U5M, CMR AND DIFFERENT CATEGORIES OF HEALTH PERSONNEL IN KENYA AROUND THE PERIOD 1991/2	40
TABLE 2.15: DISTRIBUTION OF A SAMPLE OF DEVELOPING COUNTRIES BY PER CENT OF GOVERNMENT EXPENDITURE ON SOCIAL SERVICES	44
TABLE 2.16: DATA USED FOR INTERNATIONAL COMPARISONS	54

TABLE 3.1: COMPARATIVE CLINICAL EFFICIENCY INDICATORS, KNH AND AKH, 1986	73
TABLE 3.2: ANNUAL FINANCING GAPS BY RESOURCE CATEGORY, P/PHC SERVICES, 1990	75
TABLE 3.3: GAPS IN RURAL HEALTH FACILITIES AND HOSPITAL OUTPATIENT DEPARTMENTS	76
TABLE 4.1: FUNCTIONAL DISTRIBUTION OF RECURRENT HEALTH CARE EXPENDITURES: 1972/3— 1990/91 (VARIOUS YEARS—PER CENT)	110
TABLE 4.2: DISTRIBUTION OF CURATIVE SERVICES EXPENDITURES, 1991/92	111
TABLE 4.3: BURDEN OF DISEASE BY SEX, CAUSE AND TYPE OF LOSS, 1990.	114
TABLE 4.4: RECURRENT EXPENDITURE ALLOCATIONS BY INPUTS, 1982/3—91/92: (PERCENTAGES)	115
TABLE 4.5 ESTIMATES OF FINANCING OF RECURRENT HEALTH EXPENDITURE BY SOURCE OF FINANCE, 1983/84 (KS MN)	118
TABLE 4.6: TOTAL GROSS RECURRENT EXPENDITURE BY TYPE OF HEALTH CARE PROVIDER, 1983/84 (KS MILLIONS)	122
TABLE 7.1: SETTING HOSPITAL BUDGETS USING DRG INFORMATION: AN ILLUSTRATION	235
TABLE 7.2: PUBLIC HEALTH SECTOR REFORM PROPOSALS ASSESSMENT MATRIX ...	253

— ACKNOWLEDGEMENTS —

I am indebted to Professor A. J. Culyer, under whose wisdom, tutelage and unreserved guidance I was able to write this dissertation at the University of York. I thank him not only for the academic nourishment he imparted through the critiques of all my drafts (some of which I often look at and wonder why I ever passed them to him!), but also for the patience and the speed at which he read all of them. I also wish to thank the other members of my thesis advisory group—Mr J. Suckling and Mr M. Wheeler—for the many useful comments they made of my drafts. I also have many others to thank, all of whom cannot be listed by name. I came to York without any background on health economics. I am grateful to all my MSc. (Health Economics) tutors who introduced me to the subject, particularly Dr A. Jones who supervised my MSc. dissertation and Dr J. Posnett, the Director and tutor of the health economics course in 1990/91.

Financial support from the International Development Research Centre (IDRC) of Canada made it possible for me to come to York, and not only paid all my fees, but also gave me other allowances. The University of Nairobi granted me study leave to pursue this course. The East African Research Consortium provided supplementary field-work financial support. The Department of Economics and Related Studies at York helped with computer memory chips when I got stuck with masses of data that had to be reduced to meaningful statistics. The Leche Trust made a contribution to my upkeep at York when I had financial difficulties. I am indeed grateful to all.

I wish to thank James Musinga and Mercy Reuben who were my principal field-support staff for patiently coping with (my) arduous work for about five months; Messrs S. Muchiri and D. Nzoya who assisted with most information sources at the Ministry of Health; and Ben Obonyo of Sheria House who provided (with permission) most of the unpublished information compiled by the Registrar-General's department.

I am also grateful to George Masila Musyoka and family and Joses Kirigia and other Kenyans—both at York and elsewhere in the UK who made my stay at York less lonely at one point or another. Non-Kenyans also contributed similarly—Professor Culyer and family (UK), John Kudjoe (Ghana), Dr James Flolu (Ghana), Dr (and Mrs Meena) J. Pokharel (Nepal), Hoffman Aipira (Malawi), and R. Mangalore (India).

Finally *but not least*, I wish to record my gratitude to Elizabeth, who has not only patiently endured my long absence but also has taken care of the family, given me hope, moral support and encouragement through constant mail and occasional phone calls. This work is dedicated to her, our children—Esther Mwithi, Mercy Ndinda and Nancy Mwendu for their perseverance and understanding; and my parents, Mr Julius Nganda and Mrs Lydia Nganda, for the sacrifices they made in order to educate me.

Last, the usual disclaimer applies. None of the individuals or institutions named above bear any responsibility, by association or any other way, for any errors that may remain in this work. All such errors are solely mine. A similar disclaimer applies to the ideas expressed herein—except where works of others are used, in which case the source is duly cited.

— ABSTRACT —

Practically all health care systems are in flux. Regardless of whether they are public or private, health systems have begun to re-examine their resource utilisation, cost-efficiency and the effectiveness of services, service provision, adequacy and accessibility, and equity. Kenya is no exception to this world-wide trend. In 1989 the reform process in the health care system was started by implementing some changes in the public health services sector. These appear to have been geared more to increasing the amount of resources available to the public health services sector than to tackling the afore-mentioned. Out-patient user-fees were re-introduced for the first time since their abolition about 25 year earlier, while contributions to the National Hospital Insurance Fund were revised into a progressively sliding-scale according to personal incomes.

This thesis views these changes as inadequate to transform the health system into one that can serve the rising population efficiently and equitably now and in the future (and shows how this is so), and proceeds to develop a framework of what efficiency and equity in health and health care mean. On efficiency, it is suggested that the health care system should aim to provide only services that are medically effective, cost-effective, offer the highest payoffs in terms of health gains, and are provided at appropriate scales. On equity, after reviewing the basis for equity concern in the system, and the approach used by the government to realise this objective, a review of various philosophical and health economics approaches to equity indicates that equity is more than the mere provision of 'equal opportunity of access', which appears to be the underlying conception of equity in Kenya. Specifically, and particularly if efficiency is also taken into account, equity in health care demands that patients who are 'alike' in 'relevant respects' ought to be 'treated in a 'like' fashion, and those who are 'unlike' in relevant respects treated in appropriately 'unlike' ways in a manner similar to the notions of 'horizontal' and 'vertical' equity in public finance economics. A model that can be used as the basis for proposals for health care reform of the public health care system is then developed.

Following this framework, several reform proposals are presented—(i). the current system with 'modest' internal changes, (ii). the adoption of 'global budgets', (iii). the implementation of prospective reimbursement, and, (iv). the separation of the purchasing and provision roles through the implementation of provider markets. These vary both in terms of content and effect. Following their assessment in terms of set criteria (efficiency, equity, ease of implementation/pragmatism, and adequacy—meeting the government's long-term health policy objectives) it has been found that in the long-run the separation of purchaser/provider roles is the most promising proposal for restructuring the system. This would separate organisationally the funding and provision of health care. The provision of health care would largely be privatised, but not the demand for it. The advantage of this separation is that competition and efficiency gains would be created combined with the intrinsic equity of a 'central allocation' system.

1. INTRODUCTION: SCOPE AND PURPOSE OF THE STUDY

1.0 Introduction

Kenya operates a 'mixed economy of health care' that is largely an outgrowth of the pre-independence patterns of provision. It consists of a heterogeneous mixture of government and non-government activities. Service providers consist of both modern and traditional practitioners. The government services are more varied in terms of population coverage and the comprehensiveness of services. The private sector activities are largely biased to curative services. At independence in 1963 the country inherited a health care system that was considered (by the new government) to have many defects (see chapter two) including an 'unacceptable' distribution of health (e.g., in terms of crude measures such mortality and morbidity—see chapter two), and differentials in access to health care services. To 'redress' these defects, the new government adopted a policy of providing free or low cost curative and preventive care services to increase the population's access to health care services¹. The system adopted was based on a technical or rational planning model designed to provide uniform services in a 'Stalinist-type' framework—a centralised command and control decision making framework. The decision-making process was vested in politicians and civil servants at national, regional (provincial and district), and municipal (or sometimes county) levels, while the day-to-day operating authority was the responsibility of government appointed administrators and 'suitably qualified' medical personnel at those levels. This top-down planning model was deemed a publicly accountable arrangement that would ensure the provision of necessary services in a universal and cost effective fashion.

Available aggregate statistics since independence indicate that the system has generally been able to partly achieve the objectives it was designed for.² At independence there were 148 hospitals. By 1989, the number was 183 (MOH/GOK, 1993). Increased training and career development opportunities for health personnel have substantially lowered the population-health personnel ratios. Census statistics show that Kenya's population growth rate rose from 2.5% in 1948 to 3.34% in 1989, having peaked at 3.8% in 1979. This was mainly due to declines in infant and child mortality, higher fertility among women in their child bearing ages and improved medical services

¹ Medical services were declared free only for outpatients and all children, though. Inpatients continued to pay, albeit low, various charges.

² It would not be entirely correct to attribute all the improvements described hereafter to the health services alone. Other factors played their role too, e.g. education, rise in the standard of living, and so on, as indicated.

that helped reduce still births and generally lowered morbidity and mortality. The crude death rate declined steadily from 25 per thousand in 1948, to 12 per thousand in 1990. The crude birth rate on the other hand remained almost unchanged at around 50 per thousand (see Figure 2.1 and the explanation thereof in chapter 2). Infant deaths (per thousand live births) declined steadily from an estimated 184 in 1948 to around 70 by 1990 (although results from a demographic household survey indicate that these 'impressive' overall figures conceal severe regional disparities, with infant mortality varying from 107 per thousand in Coast province to as little as 35 per cent in Rift Valley province (Kenya, 1989)).

Life expectancy at birth has also increased. The 1991 figure of 59 years compares favourably with 50 years for 'least developed countries' in general, 51 years for sub-Saharan Africa, but unfavourably with 61 and 74 years at birth for developing and developed countries respectively (see UNICEF, 1993, Table 1, p. 68). Total fertility—defined as the number of live births a woman gets by the end of her productive life—has declined slightly from 6.7 in 1948 to 6.4 by 1991, although in 1979 it was 7.9. Despite this decline, the total fertility rate in Kenya is still higher than the 6.0 and 3.7 for developing and developed countries respectively, although lower than that of other sub-Saharan countries (6.5). The proportion of infants with low birth weight in 1990 was estimated at 16%, which was lower than the 24% and 19% for least developed and developing countries respectively, and at par with that of sub-Saharan Africa.

But although food availability has grown considerably and has almost kept pace with population growth, acute malnutrition and other nutrition deficiency disorders prevalent at independence—though no longer considered a national problem—still persist (Kenya, 1992, chapter 6). In the period 1980-91, the proportion of children of ages 0-4 years suffering from moderate and severe underweight was 14% and 3% respectively. That of children aged 12-23 months suffering from wasting was 5% while the proportion of children aged 24-59 months suffering from stunting was 32%. Although these proportions are lower than the corresponding proportions in the least developed, developing or even sub-Saharan countries, there are wide variations between provinces and districts. For example, the 1987/88 rural child nutrition survey (Kenya, 1992, *loc. cit.*) found that all the districts in Coast and Nyanza provinces had stunting rates higher than the national prevalence rate (which was 19.6%). In the Coast province, Kilifi and Kwale districts had the highest proportions of stunting, while Siaya district in Nyanza province held this ignoble trait. Other districts with higher than average prevalence rates were Machakos and Meru in Eastern, Nyanjarua in Central, and Narok, Kericho and Nakuru in Rift Valley. The survey also observed that besides food intake, nutritional stunting was closely associated with access to clean water and

sanitation, and literacy rates. In 1990 the proportion of population with access to safe drinking water was only 21% in the rural areas and 61% in the urban areas, and the national average was 30%. The proportion of population with access to adequate sanitation was 19%, 89%, and 34% respectively for rural, urban and the country as a whole.

Overall, Kenya has performed comparatively well (particularly in terms of health status) in relation to the sub-Saharan and other 'least developed' countries (see chapter 2). Substantial achievements in economic and social development enabled the country to make good progress in spite of various constraints—both internal and external. The economy grew at an average rate of 5.2% per annum in real terms between 1964 and 1990 although the growth was not uniform.

	Growth of GDP (%)	Growth of GDP per capita (%)	Balance of Payments Deficit (K£m)	Public debt (K£m)	Inflation rate (%)	External debt service charges as a % of exports	Terms of trade (1982=100)		Government budget deficit as a percent of GDP
							All items	Non oil items	
1970	6.4	4	16.8	—	3	3.5 ¹	—	—	—
1975	3.3	0.3	(16.9)	311.6	18	2.8	—	—	(8.1)
1980	4.0	(0.8)	(72.2)	941.9	13	5.6	122	121	(7.3)
1985	4.8	1.4	94.2	2,565.9	11	15.3	92	87	(4.3)
1986	5.5	1.9	(73.0)	2,846.5	6	15.6	103	93	(8.6)
1987	4.8	1.2	104.4	3,209.6	7	18.5	85	75	(4.0)
1988	5.2	1.6	67.7	3,353.8	11	17.1	88	79	(4.5)
1989	5.0	1.4	80.5	4,101.6	11	18.5	79	70	(6.3)
1990	4.5	1.0	(168.9)	4,758.1	13	14.7 ²	71	62	(6.8)

¹ 1972 figure
² Provisional.

Source: Compiled from Statistical Abstracts and Economic Surveys, various years, 1970—1990.

Table 1.1: Trends in some economic indicators, 1970—1990.

Several factors combined to account for fluctuations in the rate growth, including oil shocks, deteriorating terms of trade, adverse fluctuations in exchange rates of currencies of Kenya's major trading partners, especially the (British) pound and the (US) dollar, increased interest rates on international loans, increased debt burden, general global economic recessions in the late 1970s and early 1980s, droughts, structural rigidities in some sectors of the economy, and the effects of expansionary fiscal and monetary policies in the post-oil crises periods, also external security problems. Besides, the 1970s and early 1980s saw a growing participation of government in

the economy in the provision of social services and in other economic activities through parastatals.³ This participation increased government spending and resulted in unsustainable budgetary deficits. The effect of these crises was reduced growth rates of gross domestic product (GDP) and income per capita, increased balance of payments deficits, increased public debt and high inflationary pressures (see Table 1.1 above). Overall economic performance faltered.

Government—with backing from the World Bank and the International Monetary Fund (IMF)—responded to these crises by implementing policy initiatives designed to restructure the economy. These policies (respectively associated with the institution that backed them) were structural adjustment and economic stabilisation programmes. Structural adjustment policies were meant to correct economic imbalances through (social and political) institutional and economic reforms so as to achieve sustainable and balanced growth and were of a medium to long term nature. Stabilisation policies on the other hand were of short-term nature, primarily aimed at correcting balance of payments dis-equilibria and internal (or sectoral) imbalances. Devaluation, reduction of direct participation of government in economic activities, decontrol of prices, liberalisation of the economy, creation of export incentives, and budget rationalisation were some of the conditionalities imposed for the loan facilities offered by the two institutions. The programmes placed stringent controls on the economy, resulting in a gradual run-down of state subsidies and price inflation. Personal incomes failed to keep up. The public health care system—which in 1991/92 employed some 45,000 people or about six per cent of the total civil service workforce—was particularly adversely affected. A striking fallout has been a gradual decline in nominal and real terms of the proportion of the government budget allocated to the health sector. The public health budget nominally fell from 8% to 5% of total government expenditure between 1964 and 1992, the decline having accelerated under the programmes. In 1991/92, the total health sector (nominal) expenditure amounted to K£213.05 millions, equivalent to Ks. 185 per head and 4.9% of the GDP.

1.1 Pressures for Reform of the Health Care System

Despite this 'good' performance at the overall system level, due to the problems alluded to above, in the eighties the public health services came under critical stress. The government revenue

³ This term refers to a business entity in which the government has a share-holding, i.e. of joint ownership between the government and private enterprise. This is the equivalent of 'quangos' in the UK. The government need not be directly involved in the management of such a corporation, although it has the prerogative to appoint its own board of management should it desire to do so.

base shrunk, leading to cuts in public revenues allocated to various components of government spending. The health sector, being in a rather weak position in the government's list of priority spending (e.g. as compared with education and defence—see the appendix) among others, was particularly adversely affected. As a consequence, service provision failed to expand as anticipated. Equipment to deliver services, personnel to staff facilities, and other inputs (financial and physical) became increasingly inadequate. At the same time the population was increasing and patient expectations were rising. Other contributory factors also played their part—though they were not explicitly recognised, as we show in this dissertation—such as the consequences of the fixed (and almost guaranteed) institutional budgets and 'assured' personnel salaries, both with strong disincentives for higher productivity, creating additional disincentives for efficient utilisation of resources. Against this background, it became imperative to review the structure and organisation of the public health services, with a view to redefining the role of government in the health sector. In line with the thrust of the then World Bank policy towards developing countries health care systems (Akin *et al.*, 1987)⁴, which was still supporting structural reform in the economy, the government, in the Sessional Paper number 1 of 1986 (Kenya, 1986) contemplated a health policy reform that focused on four main areas:

- ◆ finding alternative sources of finance, particularly making the 'beneficiaries' of these services bear a part or all) of the cost of provision, whilst simultaneously making the poor the major beneficiaries of the expanded resources for and improved efficiency in the government sector through appropriate subsidy schemes;
- ◆ Increased use of insurance or other risk coverage schemes;
- ◆ Supplementation of government effort by the non-government sector;
- ◆ Decentralisation of the public services—in areas of planning, budgeting and purchasing to increasingly make use of market incentives to guide decisions where appropriate.

1.2 The Initial Response: Reforms in the Financing Arrangements

Following the above, in 1989, outpatient user-fees were reintroduced in public health facilities at all levels above the dispensary (since their abolition almost 25 years earlier) and most

⁴ whose main contentions were that there was:

- Insufficient allocations of government spending on cost-effective health activities;
- Internal inefficiency of public programs due to under-funding in critical areas, adversely affecting the performance of the sector—reduced effectiveness of personnel, under-utilisation of lower level facilities co-existing with overcrowding in central outpatient clinics and hospitals;
- Inequity in the distribution of benefits from health services due to a bias in investment in the expensive modern technologies that serve a few, mainly urban based clients (who also have better access to non-governmental services), while low cost interventions for the (mainly rural based) masses are under-funded.
- need to mobilise additional resources from other sources, including encouragement of the private sector.

inpatient charges were revised upwards, while the contributions to the National Hospital Insurance Fund (NHIF)—previously pegged at Ks 20 for all qualifying contributors (those earning incomes of Ks. 1,000 and over per month)—were revised into a more progressive schedule in which contributions varied with income.

1.3 The Need for ‘Broader Structural Reforms’

But these changes have not brought substantial revenues to the public health system (as anticipated—see chapter two) mainly because the user charges are only nominal, and do not cover a substantial cost of the actual cost of public health services, while the greatest proportion of the increase in the NHIF contributions gets siphoned to the private sector facilities where most beneficiaries of this Fund seek care. Therefore, the public sector will continue to have inadequate and possibly even greater declines in the real resources available to fund public health care activities. Given the government’s long term objective of ensuring all Kenyans (in need of health care services) have access to such services—and also that the government itself is the single largest financier and provider of health care services in the country—this trend will make it even more difficult to achieve the objectives of the public health care sector (see chapter five) unless corrective action is taken swiftly. If the situation gets worse, it is conceivable that some of the gains achieved so far may be compromised. In this changed environment, health sector activities can no longer be justified on their ability to provide universal services and to enhance social justice only. Productivity and efficiency issues have to be important criteria for judging the performance of health services.

1.4 Purpose and Scope of the Present Study

This thesis is concerned with the design of a basic institutional framework/system for the delivery and financing of personal services as well as preventive medicine in Kenya in this changed environment. The framework developed encompasses hospital and clinic based services, training institutions for doctors, nurses and other paramedical professionals, the role of research in supporting health services, and the role of various public agencies concerned with preventive programs. The aim of the thesis is to provide a sound framework on the basis of which issues such as the following can be addressed: What kinds of health services should exist in Kenya? Who will get them and on what basis? Who will deliver them? How are the burdens of financing them to be distributed? and, How is the power for the control of these services (between the public and the pri-

vate sectors) to be distributed? The way these questions are resolved in the Kenyan health care system will determine the level and distribution of health risks in the population, the actions taken to deal with them, and the degree to which people are helped to regain good health. Macro as well as micro decisions should be made on the basis of principles—moral (about justice) and economic (efficiency)—that serve as a public and final basis about how the health care system should be designed/operate. In this thesis a framework within which health planners and legislators can make more specific and informed policy decisions is developed. The framework provides principled mechanisms to resolve the often conflicting claims advanced by different groups on the health system—conflicts that reflect the fundamental differences that exist between providers and consumers of health care, between different groups of providers or consumers, or between different economic classes, who bear the benefits and burdens of policy decisions differently.

Specifically, the thesis proposes (among others) a reform strategy based on the use of increased competition through contracting in the public sector. The questions raised above are answered indirectly by searching answers for five related questions:

- ◆ What services should be provided, now and in the future?—what criteria should be used to determine what should be provided? and what monitoring/control/regulatory mechanisms should be adopted?
- ◆ The system does not generate the relevant information for the most part, so what system would do this over time if established now?—i.e. what type of information and support would the ‘purchasers’ and ‘providers’ need?
- ◆ For whom should these services be provided, and why?
- ◆ Who should pay for these services, and how? In other words, what funding arrangements should be adopted?
- ◆ What incentives for efficient behaviour—for providers, patients, and funders—should be used/introduced? Related to this are other questions: What would be the role of competition in the ‘new’ system? What role would the central government play? What about the private sector? And finally, what would be the role of training, research and education in the system?

1.5 Organisation of the Dissertation

The rest of the dissertation is organised as follows. The next two chapters provide some background information on the Kenyan health care system, giving a snap-shot of the

setting/environment under which inefficiency has been cultured. Chapter two describes the demographic characteristics of the Kenyan population and the implications for health services demand in the future. The trends and distribution of health resources are then discussed and related to the distribution of the population. The epidemiological situation is then discussed and also related to the distribution of the health resources. The chapter ends with a discussion of the Kenyan health situation in relation to other developing countries. Chapter three provides a description of the administrative and organisational structure of the system followed by a review of some studies done prior to, and during the two financing changes mentioned above. Chapters four and five are about the theoretical developments and literature reviews. In chapter four we review the concept of efficiency in the health care system. In chapter five, theoretical issues of equity—focusing on the various approaches—both philosophical and non-philosophical, and contributions from economics, proposed to shed light on the meaning of equity, are reviewed briefly, and it is shown how efficiency and equity can be incorporated in the health policy. Chapter six considers the issue of incentives in the health care system. Theoretical approaches to improving incentive structure (as well as raising incentive compatibility) are discussed and a model for incorporating incentive structures in the system is developed. Chapter seven discusses four proposals for health care reform in the public health sector and looks at their implications. Some suggestions concerning the private sector are also given. Given the objectives of government policy and reality, the chapter outlines what can be done to make the Kenyan health care delivery system more efficient and equitable, given the existing constraints. Chapter eight concludes the thesis.

2. THE SETTING OF THE PROBLEM—I: THE HEALTH STATUS SITUATION IN KENYA AND THE DISTRIBUTION OF HEALTH RESOURCES

2.0 Introduction

At independence in 1963, Kenya inherited a health care system that was uneven—largely biased towards the hospital sector (curative services), and favouring certain groups (the middle and higher income) and the urban areas. For example, Nairobi was better provided in terms of medical services/facilities because this was where most people with high incomes and the head offices of most philanthropic organisations were located, and also because most doctors and other trained medical personnel preferred to work there since most either were trained there or in similar environments and did not want to leave. Thus the environment from the onset provided no incentive for physicians or facilities to relocate towards the geographical areas with greatest medical need—mainly the rural areas. Therefore the distribution and location of facilities were largely influenced by the benefactors, beneficiaries and providers. Moreover, the practice of case selection that emphasised the development of techniques for curable or potentially curable patients led to a health care system that favoured curative services. Thus the service mix was also biased in favour of curative medicine. As a result, as we show below, health care resources were unevenly distributed—between services, areas, and social groups, and generally ineffectively used, given the country's health problems. This trend has persisted for over 30 years now. The survival to a great extent of the unequal pattern of regional provision per head of crude population (see below) after more than 30 years of public direction by a government which recognised this problem at the time it ascended to power—among others—is the issue addressed in the present and the next chapter.

This chapter provides an overview of the epidemiological situation and looks at the distribution of health resources in relation to the epidemiological profile. The chapter is organised as follows. Section 2.1 provides a summary of the major demographic characteristics and their implications for the health services. Section 2.2 discusses trends in various types of health resources, both over time as well as their distribution between regions. Section 2.3 then discusses the epidemiological situation, both over time as well as across geographic regions within the country and attempts a link of this with the distribution of health resources. Section 2.4 provides some international comparisons. Section 2.5 concludes the chapter.

2.1 Demographic Characteristics and Implications for the Health Services

To study the patterns of morbidity and mortality it is essential to know the demographic background of the country under study, because, among other factors, the frequency of disease and death is related to the age and the sex structure of that population, its geographical distribution and the migration processes, and the level and distribution of health resources. This section examines the major demographic characteristics of the Kenyan population. The focus is on the broader features of Kenya's demographic situation that may provide information about the quality of life in relation to the available health resources.

2.1.1 General demographic trends

The demographic data used in this study comes from the 1979 and 1989 population census reports, supplemented with information from other sources such as the earlier census reports, statistical abstracts and various fertility and health surveys conducted by the Central Bureau of Statistics (CBS) and the Kenya Population Council (KPC).

The recorded population of Kenya at the time of the first national census in 1948 was 5.4 millions. The 1969 census returns showed the population had grown to 10.943 millions, indicating the population had doubled in approximately 21 years. The 1979 census recorded a total population of 15.327 millions, showing a trebling of the population in approximately 30 years. The 1989 census on the other hand indicated that the total population in that year was 21.469 millions—an intercensal growth rate of approximately 3.34 per cent between 1979 and 1989⁵—indicating the population had nearly quadrupled in approximately 40 years. Such a high rate of population growth (as compared to the growth rates in developed countries for example) has grave implications on the ability of the government to provide essential services such as health care, education, and other social services over the long term, besides the retarding effects it has on the capital formation capacity of the economy as a whole. Let us examine some aspects of the demographic situation that have a bearing on 'health needs', such as the structure and the distribution of the population.

⁵ Due to organisational inadequacies in the census exercise, the CBS considers this figure an underestimate and estimated the actual population to have been about 23 millions [Kenya, (1991), CBS, p. 33].

2.1.2 Structure of the population

The structure of the population is similar to that found in other developing countries, in that a large proportion of the population is relatively young as Table 2.1 below indicates.

Year	Age group					Total ²
	0—14	15—29	30—59	60+	N.S ¹	
1962	3,975.5 (46.0%)	2,211.8 (25.6%)	1,974.5 (22.9%)	418.3 (4.8%)	56.1 (0.6%)	8,636.3 (100%)
1969	5,293.0 (48.4%)	2,743.9 (25.1%)	2,317.6 (21.2)	587.9 (5.4%)	0.3 (0.2%)	10,942.7 (100%)
1979	7,410.0 (48.3%)	4,125.0 (26.9%)	3,059.0 (20.0%)	703.6 (4.6%)	29.5 (0.2%)	15,327.1 (100%)
1989	10,258.7 (47.8%)	5,911.4 (27.5%)	4,221.7 (19.7%)	1,159.9 (5.4%)	25.2 (0.1%)	21,468.9 (100%)

¹ Age not specified
² The per centage totals may not add-up exactly to 100 due to rounding.

Source: CBS, Census data, various years.

Table 2.1 Population age structure, 1962—1979 (Figures in '000's and per cent)

The number in the youngest age band (0—14 years) has increased, both in absolute and proportionate terms, between 1962 when it represented 46 per cent of the population, to about 48 per cent by the time the 1989 census took place. Also, the proportion of old people (those aged 60 and over) increased—from 4.8 to 5.4% in 1989. These changes have meant increase dependency burdens. Given the large proportion of those under 15 years of age, the population of Kenya thus generally youthful.

This structure of population can be explained in terms of figure 2.1 overleaf, which shows that although there has been a significant decline in the infant mortality rate, the birth rate has however remained more or less constant. The visual impression of the figure is that there was a slight decline in the birth rate up to around the early seventies after which a slight increase is noticeable. Specifically, in the 1985—90 period, the birth rate per 1,000 inhabitants was 7 per cent lower than for 1950—55. Why it again increased (albeit slightly) is rather perplexing since we would generally expect a continued decline with advances in economic conditions, if the models of demographic trends in the developed countries are useful guides. We can only conjecture this phenomenon occurred due to improvements in health services availability (discussed anon), and the generally good economic conditions in the preceding decade (although the relative and absolute increase in the

proportion of females in the population might also be an important contributory factor—see the next sub-section). Infant mortality rate over the same period fell by just over one half. The death rate, by comparison, also declined, albeit slightly over this period. The decline in mortality rate can be attributed to improved health services situation in the country and the increased emphasis during the late 1960s, the 1970s and 1980s on immunisation, particularly immunisation for childhood diseases. The impact of these is particularly evident in the decline in infant mortality. The unchanged birth rate might be attributed to the ‘delayed response effect’ whereby, given the previous high infant mortality levels, it takes a while for people to adjust to the new lower levels and adjust their family sizes accordingly. The overall result has been an increase in the population in the lower age brackets—notice (from table 2.1) the proportion of those under 30 has increased during the period.

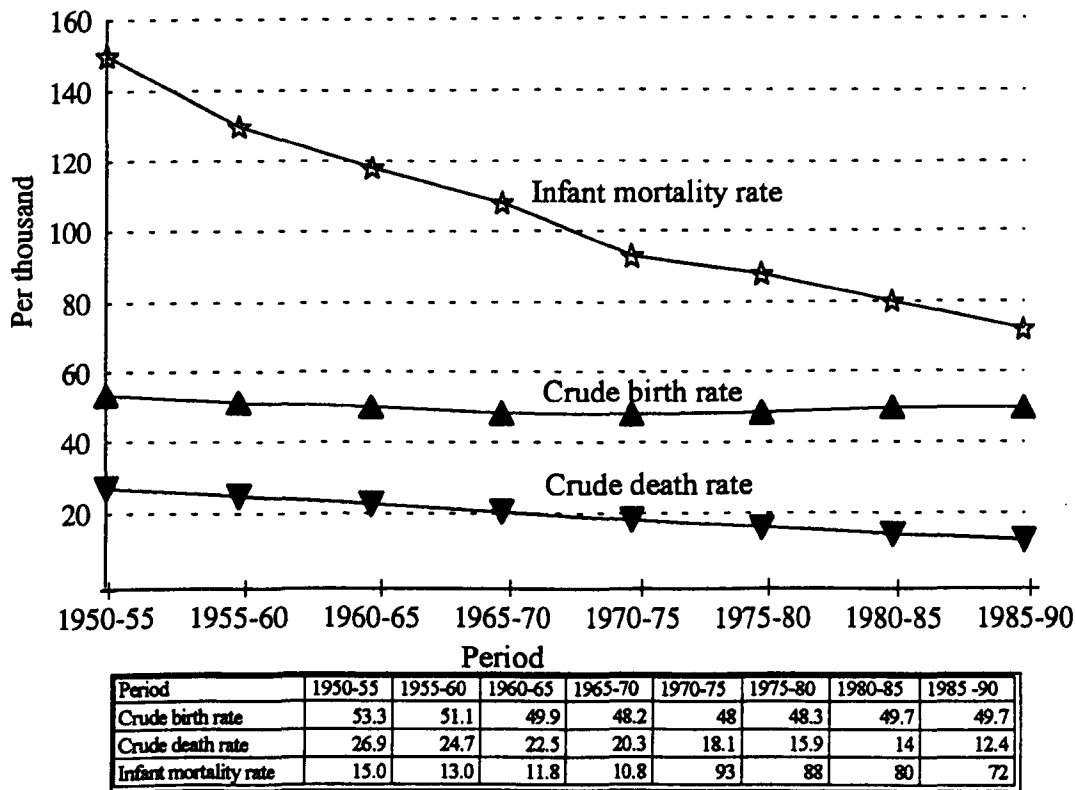


Figure 2.1: Vital population statistics: 1950—1990

2.1.3 Life expectancy

Life expectancy at birth improved from 39 years in the 1950—55 period to just over 55 years in the 1985—90 period. The life expectancy for females has been higher than that of men

(Table 2.2). The result was an increase in the proportion of females in the population. The proportion of females in the 1989 census was 50.4 per cent.

	Period							
	1950—55	1955—60	1960—65	1965—70	1970—75	1975—80	1980—85	1985—90
	5	0	5	0		0	5	0
	Life expectancy at birth in years							
Males	37.1	39.1	41.4	43.8	46.3	48.9	51.2	53.5
Females	40.2	42.3	44.6	47	49.6	52.3	54.7	57.1
All	38.6	40.7	43	45.4	47.5	50.5	52.9	55.3

Table 2.2 Life expectancy at birth by sex: 1950—1990

The population age and sex structure are dominated by the (almost non-discriminating) effect of mortality as it affects different cohorts, and cannot reveal the significant difference that exist between the urban and rural populations. The urban type of age and sex structure is illustrated by the Nairobi pattern that differs widely from that of the rest of the provinces, as shown by the census data, shown for two census years in Table 2.3 below. The pattern shown in the table arises out of in- and out-migration of people that results in the influx of people (mostly those over 15 years) into the urban areas, and this trend seems to become more evident over time as the intercensal changes indicate. In the 1989 census, the percentage of under age 15 population was 47.8 per cent. A similar situation existed in the 1979 and 1969 population censuses.

In general, whereas in the rural areas the proportion of under-age population (under 15 years) for both sexes is well over 48 per cent, in the main urban areas it is lower. Provinces that have a tendency of gaining migrants—such as Nairobi, Coast and the Coast province have a low proportion of children, and a higher proportion of adults (particularly ages 15—44), whereas those losing migrants, e.g., Central, Eastern and Western have a high proportion of children. The Nairobi figure of 31 per cent in 1989 is less than the national average due to this influence of in- and out-migration. Also, there is a tendency for out-migration areas (mainly rural districts) to show a deficiency of the male population. In the urban areas, therefore, there is a tendency for the males to dominate the migration streams. However, in some of the smaller urban centres, the proportion of the under-age population and sex ratios are closer to those of the rural areas.

1979									
Province	Sex composition (per cent) ¹			Population under 15 yrs			Sex Ratios		
	Males	Females	Total	Males	Females	Total	Per cent under age (both sexes)	All : Males per 100 females	Adults: males per 100 females
Nairobi	6.3 (58)	4.5 (42)	5.4 (100)	3.7 (16)	3.9 (18)	3.8 (34)	34	138	169
Central	15.1 (49)	15.5 (51)	15.3 (100)	16.3 (26)	16.2 (26)	16.3 (52)	51.3	96	91
Coast	8.9 (50)	8.6 (50)	8.8 (100)	8 (22)	7.9 (22)	8 (44)	44	102	102
Eastern	17.3 (48)	18.2 (52)	17.8 (100)	18.1 (25)	18.1 (25)	18.1 (50)	49.4	94	87
N. Eastern	2.6 (53)	2.3 (47)	2.4 (100)	2.4 (24)	2.2 (22)	2.3 (46)	45.8	111	111
Nyanza	16.7 (48)	17.8 (52)	17.3 (100)	17.5 (25)	17.5 (24)	17.5 (49)	49	93	85
Rift Valley	21.6 (51)	20.7 (49)	21.1 (100)	21.5 (25)	21.4 (24)	21.4 (49)	49	103	104
Western	11.5 (48)	12.4 (52)	12 (100)	12.6 (25)	12.7 (26)	12.6 (51)	51.1	92	84
All Kenya	100 (50)	100 (50)	100	100 (50)	100 (24)	100 (24)	48.4	99	96
1989									
Nairobi	7.1 (57)	5.3 (43)	6.2 (100)	4 (15)	4.2 (16)	4.1 (31)	31.4	132	152
Central	14.4 (49)	14.7 (51)	14.5 (100)	14.3 (24)	14.2 (23)	14.2 (47)	46.8	97	92
Coast	8.7 (50)	8.4 (50)	8.5 (100)	8 (22)	8 (22)	8 (45)	44.7	101	102
Eastern	17.2 (48)	18 (52)	17.6 (100)	18.2 (25)	18.1 (24)	18.1 (49)	49.3	94	87
N. Eastern	1.8 (52)	1.6 (48)	1.7 (100)	1.8 (25)	1.6 (23)	1.7 (48)	47.6	108	106
Nyanza	15.8 (48)	16.9 (52)	16.4 (100)	17.1 (25)	17.1 (25)	17.1 (50)	50	92	84
Rift Valley	23.6 (50)	22.8 (50)	23.2 (100)	24.2 (25)	24 (25)	24.1 (50)	49.6	102	102
Western	11.4 (48)	12.3 (52)	11.9 (100)	12.6 (26)	12.8 (26)	12.7 (51)	51.2	91	83
All Kenya	100 (50)	100 (50)	100	100 (24)	100 (24)	100 (24)	47.8	98	96

¹ For sex composition the unbracketed figures indicate the share of the province in the national population while the bracketed figures show the distribution within the province or Kenya for the total population. For population under age 15, the bracketed figures show the share of the province's population (by sex) in the total (national) population. Due to rounding biases, there may appear discrepancies between the sex ratios as implied by the sex column compared to the sex-ratio column.

Table 2.3 Population distribution by province, 1989: Percent of population by sex, sex ratio and proportion of population under 15 years by province.

Besides the effects of general mortality and improvements in life expectancy on the population composition within various provinces, there are other factors that might explain the structural differences. Provinces with low fertility (e.g., Nairobi and the Coast provinces) have a low proportion of children, and vice-versa (Table 2.4 below).

Province ¹	Population Under 15 years (1989)	Fertility Rate ²	Infant mortality rate ²	Under five mortality rate ²	Childhood mortality rate ²
Nairobi	31.4	4.6	46.3	80.4	35.7
Central	46.8	6	37.4	47	10
Coast	44.7	5.5	107.3	156	54.5
Eastern	49.3	7	43.1	64.3	22.2
Nyanza	47.6	7.1	94.2	148.5	60
Rift Valley	50	7	34.6	50.9	16.9
Western	49.6	8.1	74.6	132.8	62.9
Kenya	51.2	6.7	58.6	90.9	34.3

¹ Data on mortality and fertility is not available for North Eastern province.

² According to 1989 Kenya demographic and Health survey. Mortality rates are for the period 1979-1989.

Table 2.4 Proportion of population under 15 in relation to fertility rate, infant, under five and childhood mortality rates

Except for Coast and Central provinces, areas with high infant mortality (i.e., under age 1 mortality), childhood mortality (i.e., under age 15 mortality) and under-five mortality portray high proportions of children—possibly due to replacement and precautionary childbearing behaviour. It is difficult to explain why Central province has the highest proportion of children using this approach, since it not only has a lower than the national average rate of fertility, but also has the lowest rates of childhood and under five mortality. A possible explanation is that Central province, being close to Nairobi, besides having a well educated (and therefore mobile population), loses most of its adult population to the City and other urban areas. On the other hand, the Coast province is different in that it not only has a low proportion of children to adults, it also has the highest rate of infant and under five mortality. It is difficult to reconcile these two⁶ but it should be noted that the province contains Kenya's second largest city—Mombasa, and we may be getting the effect of that city's population structure reflected in the overall provincial population structure. The Eastern, Western and Nyanza provinces have high proportions of young populations and correspondingly high fertility and child mortality rates. Western province, with the highest overall childhood mortality also has the highest fertility rate. Similarly, Eastern province, with the lowest (among these three only) childhood mortality also has the lowest fertility (again, among the three only).

⁶ That is, unless we assume the infant and under five mortality is so high as to contribute to this outcome—a questionable assumption since we would expect a decline in the population over time—a result not supported by the facts.

2.1.4 Distribution of the total population and urbanisation

The total population of Kenya is unevenly distributed, with a nation-wide density of 37 persons per square kilometre in 1989. The distribution of the rural population closely follows the pattern of rainfall in the country. Most of the rural population is concentrated in three areas in the country. The first cluster is around the lake region, with a broad base on the Lake Victoria basin. This cluster stretches unevenly into some parts of the Rift Valley, and across onto the eastern side in the Central and Eastern Provinces. These areas experience precipitation levels of more than 500 mm. annually. The second major cluster lies to the East of the central Rift Valley and extends roughly from north of Nairobi to Mount Kenya and down to the Mua hills in Machakos District. The third major population cluster is along the coastal region, besides the major urban centres there. The rest of the country is sparsely populated.

Environmental conditions play a vital role in the health of the population of any country. The distribution of the population between the urban and rural environment is one important environmental consideration. Table 2.5 shows the urban and rural population by province for the 1969 and 1979 censuses respectively⁷.

Province	1969				1979			
	Urban	Rural	Total ¹	% of total ¹ population	Urban	Rural	Total ¹	% of total ¹ population
Nairobi	100	0	100	4.7	100	0	100	5.4
Central	2.7	97.3	100	15.3	5.5	94.5	100	15.3
Coast	30.1	70	100	8.6	30.3	69.7	100	8.8
Eastern	2	98	100	17.4	8.6	91.4	100	17.8
N. Eastern	0	100	100	2.3	17	83	100	2.4
Nyanza	20.7	79.3	100	19.4	7.9	92.1	100	17.3
Rift Valley	6.7	93.3	100	20.2	10.5	89.5	100	21.1
Western	0.8	99.2	100	12.1	5.8	94.2	100	12
Kenya	13.5	86.5	100	100	15.1	84.9	100	100

¹ Some totals may not add up to 100 due to rounding.

Table 2.5: Percentage distribution of urban and rural population by province, 1969 and 1979.

By international standards, the level of urbanisation in Kenya is low. The urban population is concentrated in the two major cities—Nairobi and Mombasa and other upcoming urban ar-

⁷ Similar information from the 1989 population census was not available at the time of writing.

east—Kisumu, Nakuru, and Eldoret, etc. Although in the inter-census years some urban boundaries changed, the number of urban centres with populations more than 2,000 persons increased. Nairobi's share of the total urban population declined with the advent of other urban centres. The rate of increase of the urban population is generally well above that of the rest of the country. This trend continued between 1979 and 1989 censuses. Initial analytical results of the 1989 census data revealed the urban population had increased to about 3.8 million by 1989, giving an intercensal growth rate of 4.8 per cent per annum and a 19 per cent proportionate share of the total population compared to 15 per cent in 1979.

2.1.5 Conclusion

In conclusion, it is therefore notable that the Kenyan population is not only generally youthful, but its sex composition shows there are more females on average. Besides, the population is growing at a fairly fast rate. Moreover, most of this population resides in the rural areas where the provision of health services is a demanding task, particularly given its uneven distribution and the differences in ecological conditions that lead to differential disease patterns across regions. The concerns for the future health of such a population and the ensuing demands upon the health services are undoubtedly real ones.

2.2 Trends in public health expenditures in the 1980's and regional resource distribution

2.2.1 Financial resources

Trends in public health expenditures in the 1980's

Figure 2.2 below shows public expenditure on health, education, defence and other social services in the 1980's in constant (1982) prices. The nominal figures have been deflated using a simple index of government spending which is derived by dividing nominal government spending with constant price government spending as given in the national income accounting manuals. Ideally, sectoral indices should be used but these are not available for Kenya. The figure shows both health and other social services took a small (and almost constant) proportion of the government budget. Education and defence, on the other hand, took comparatively large and increasing proportions. During the period, the economy as a whole did not perform well and health and other social services bore the brunt of this poor economic performance.

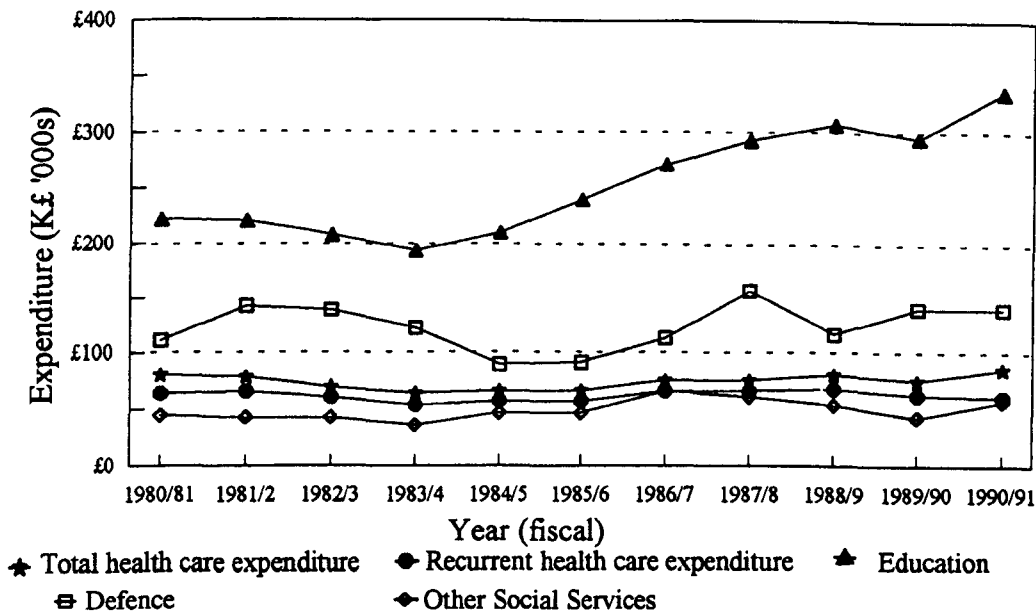


Figure 2.2: Public expenditure on health, education, defence and other social services, 1980/81—1990/91

There is a general relationship, though not very strong, between real GNP changes and various forms of government spending in real terms (see Figure 2.3, below). The government budget appears to follow the trend of GNP with roughly a one year lag. The expenditures themselves in turn appear to relate to the government budget with a short time lag. But education and health expenditure appear to rise and fall more or less in line with trends in real GNP than with the government budget. But the expenditures show greater variance than GNP. The biggest variance is displayed by defence expenditures which on average take a while longer to respond to the changes in real GNP. The government budget swings are less pronounced compared to specific expenditure changes.

Figure 2.4 below shows there has been a general decline in the per capita spending on health, but the rate decelerated in the second half of the 1980s. Per head (public) expenditure on education continued to increase. In many countries, the shares of resources devoted education and health tend to be nearly the same, with health taking a lightly lower share (see below). The share of the GNP devoted to health has fallen over time. In terms of the overall financing of the health services, between 1980/1 and 1990/1 fiscal years⁸ the public expenditure on health as a percentage of Gross National Product (GNP) declined throughout.

⁸ The fiscal year of the Kenya Government runs from July 1 through June 30.

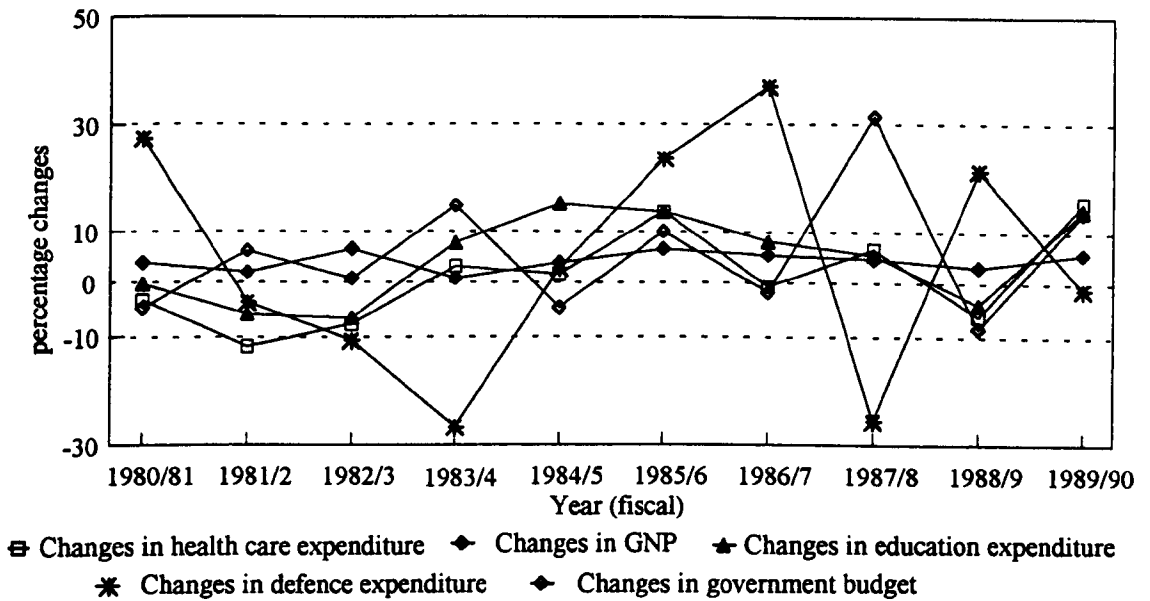


Figure 2.3: Relationship between GNP, government budget and various forms of government expenditures, 1980/81—1990/91

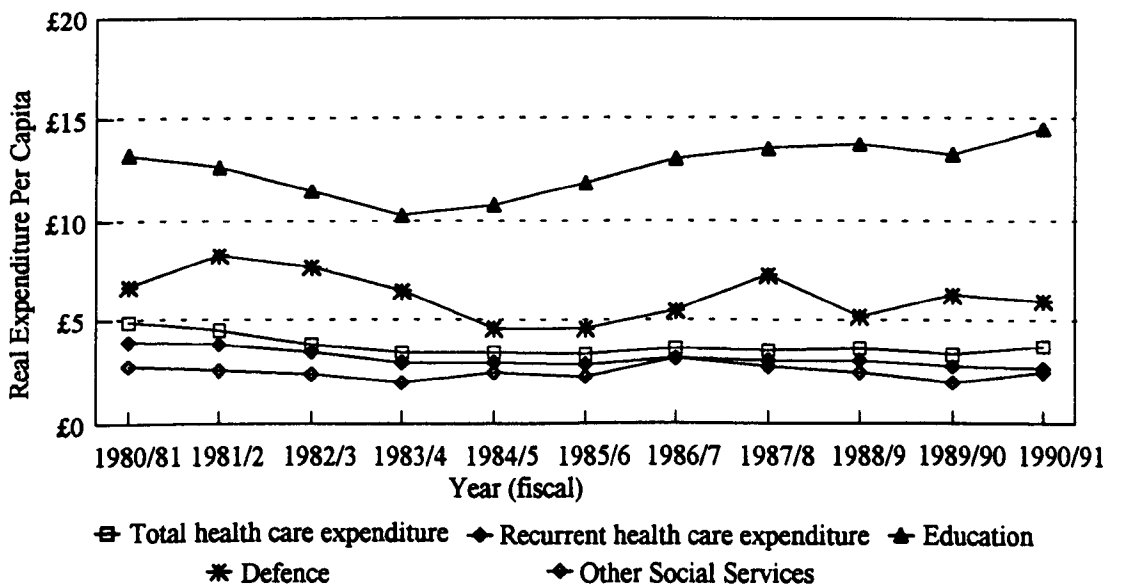


Figure 2.4: Per capita public spending on selected services trends: 1980/81-1990/91

This analysis (of trends in the levels of public health expenditures) reveals two major problems relating to them. First is the low level of GNP which leads to low levels of spending on health care. Second, in recent times there has been a decline in the proportion of GNP devoted to health care. Since we have no reason to suppose that there will be any significant changes in the proportion of GNP resources being allocated to the health sector, and in the worst scenario, this propor-

tion will continue to decline, both in nominal and real terms, any major improvements in the sector must in the main come from better use of the currently available resources—mainly through improved efficiency in the use of resources.

The regional allocation of financial resources in relation regional population shares

The regional distribution of the Ministry of Health resources by provinces is shown in Table 2.6 (below) for the year 1989. The table shows there are disparities in the allocation of Ministry of Health recurrent financial resources between provinces. Column 1 shows the overall share of each province in the MOH budget. Column 9 shows the share of each province's total population. The Coast province, which accommodated only 9 per cent of the total population, received about 15 per cent of the MOH budget. The North Eastern province, which housed less than 2 per cent of the total population received over 4 per cent of the MOH resources, and so on. Column 10 shows the overall implications of such disparities in the allocation—differences in per capita spending on health. This disparity ranges from K£11⁹ per capita in North Eastern province to less than K£4 per capita in Rift Valley and Nyanza Provinces (K£3.52 and K£3.60 respectively). The national per capita spending on health in 1989/90 was K£4.77. Four out of eight provinces had per capita expenditures lower than this average (i.e., Nairobi, Eastern, Nyanza, and the Rift Valley). North Eastern province had more than double the national average. Whether there is any justification for such discrepancies will be assessed— in section 2.3 below—after we have discussed the patterns of morbidity and mortality. Other columns (2—8) show there are variations in allocations of specific budgets.

Thus there are wide per capita regional variations in the allocation of public health care expenditures among the regions in the country, variations that at this point we cannot tell whether they are justifiable until we have examined the distribution of the ill-health burden in the country. This will be necessary since to achieve a balanced and rational distribution of health resources, it is necessary to relate the inputs (resources) to the outcomes (health status). But before we attempt such a comparison, let us further examine the trends in other health related resources—the stocks of physical resources.

⁹ K£1=Ks20.

	All MoH ¹ Exp ² %	Admin ³ Exp ² %	Prov. ⁴ Hosp. ⁵ %	Dist. ⁶ Hosp. ⁵ %	Prom ⁷ Prev ⁸ Med ⁹ %	Rural Health Servs ¹⁰ %	Train. ¹¹ Servs ¹⁰ %	Shared facilit ¹² Servs ¹⁰ %	Pop ¹³ %	Per Capita RHCE ¹⁴ K£
Column→	1	2	3	4	5	6	7	8	9	10
Nairobi	6	11	7	—	9	—	44	7	6	4.34
Central	15	11	11	19	10	19	10	11	15	4.85
Coast	15	13	16	16	19	10	9	16	9	8.42
Eastern	16	13	14	19	13	19	9	14	17	4.38
N. Eastern	4	12	6	3	5	2	2	6	2	10.79
Nyanza	13	13	14	9	14	17	12	14	17	3.60
Rift Valley	17	16	18	18	17	14	10	18	23	3.52
Western	15	11	13	17	13	20	4	13	12	5.98
All Kenya	100	100	100	100	100	100	100	100	100	4.77

¹ Ministry of Health; ² Expenditure; ³ Administrative; ⁴ Provincial; ⁵ Hospitals; ⁶ District; ⁷ Promotive; ⁸ Preventive; ⁹ Medicine; ¹⁰ Services; ¹¹ Training; ¹² Facilities; ¹³ Population; ¹⁴ RHCE: Recurrent health care expenditure.

Table 2.6: Regional distribution of recurrent public health expenditure by the Ministry of Health: 1989/90 (percentages)

2.2.2 Physical resources

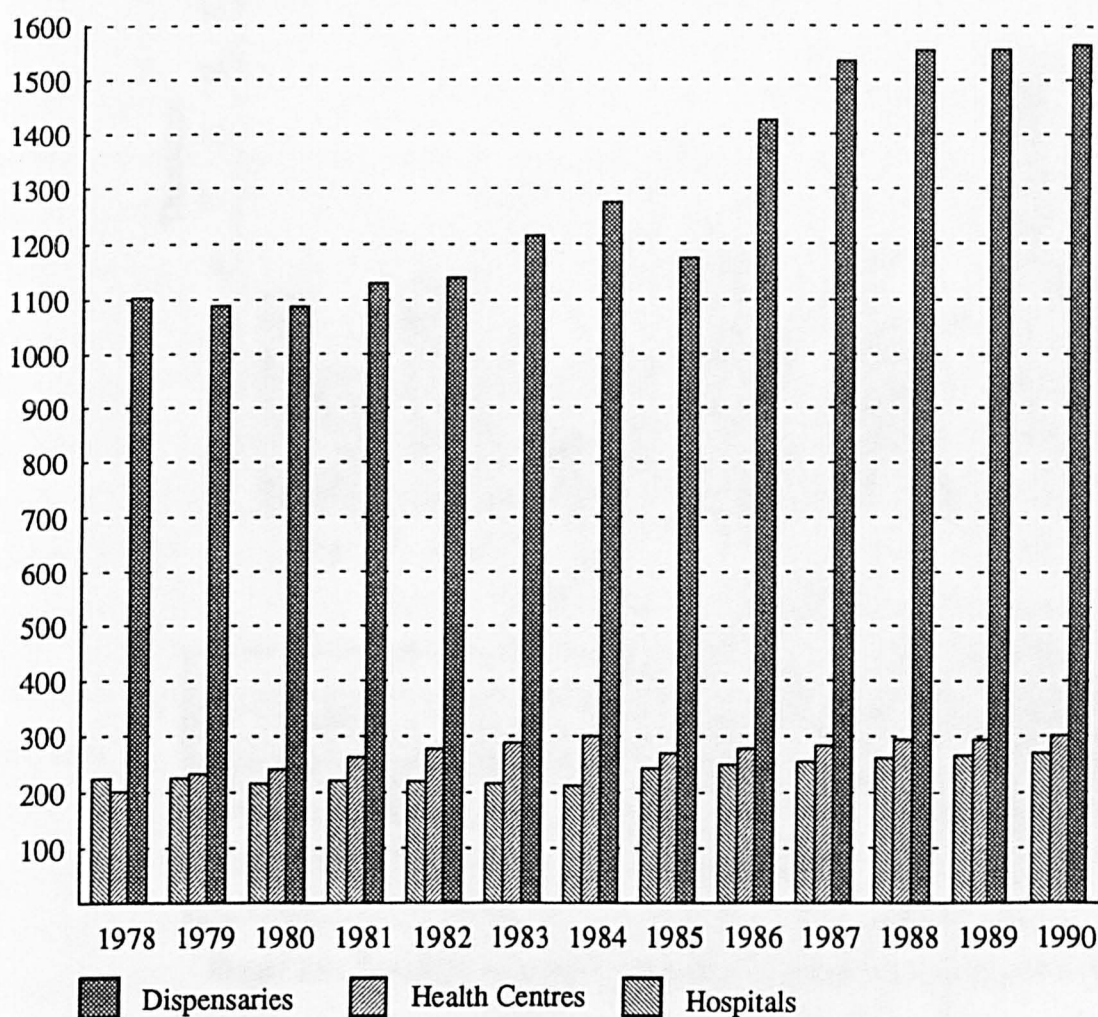
Trends in physical resources

The stock of health facilities in the country grew substantially since the 1960's. Available data shows in 1967 there were 199 hospitals in the Republic, 76 of them (38 per cent) being public hospitals, catering for a total population of about 9 million. In 1970, the numbers for health centres and dispensaries were 195 and 603 respectively. Of the 195 health centres, 173 were operated by the central government, 18 by municipalities and only 4 by missions. Of the 603 dispensaries in the country then (1970), 378 were run by the central government, 152 by missions, 59 by private companies, and the rest (14) by municipalities. By 1978, the proportion of government hospitals (118) had risen to 52 per cent. Figure 2.5 below shows the trends in health facilities in the country for the period 1978—1990. The largest increase during the period occurred in the stock of health centres (49 per cent) while dispensaries increased by 42 per cent. Hospitals increased by about 20 per cent during the period.

The public hospitals vary considerably in size and facilities¹⁰. The national referral hospital, Kenyatta National hospital, and most provincial hospitals are categorised as grade 1 while most district hospitals are graded 2. There are also what are called grade 3 hospitals or cottage hospitals

¹⁰ This is also true of the private and mission hospitals.

—usually without a permanent resident doctor. Grade 1 hospitals are usually well staffed and equipped. Available figures on government hospital size (by bed capacity) show that in 1989, Kenyatta National Hospital was the largest of the grade 1 hospitals, with a total of 1,662 beds. The second largest was Mathare Mental Hospital (1,138 beds). The average size of the provincial hospitals was 460 beds. The largest maternity hospital, Pumwani, had 329 beds. District and sub-district (or cottage) hospitals varied in size from as few beds as 9 (Othaya hospital) to 352 beds (Eldoret Hospital) with an average of 134 beds.

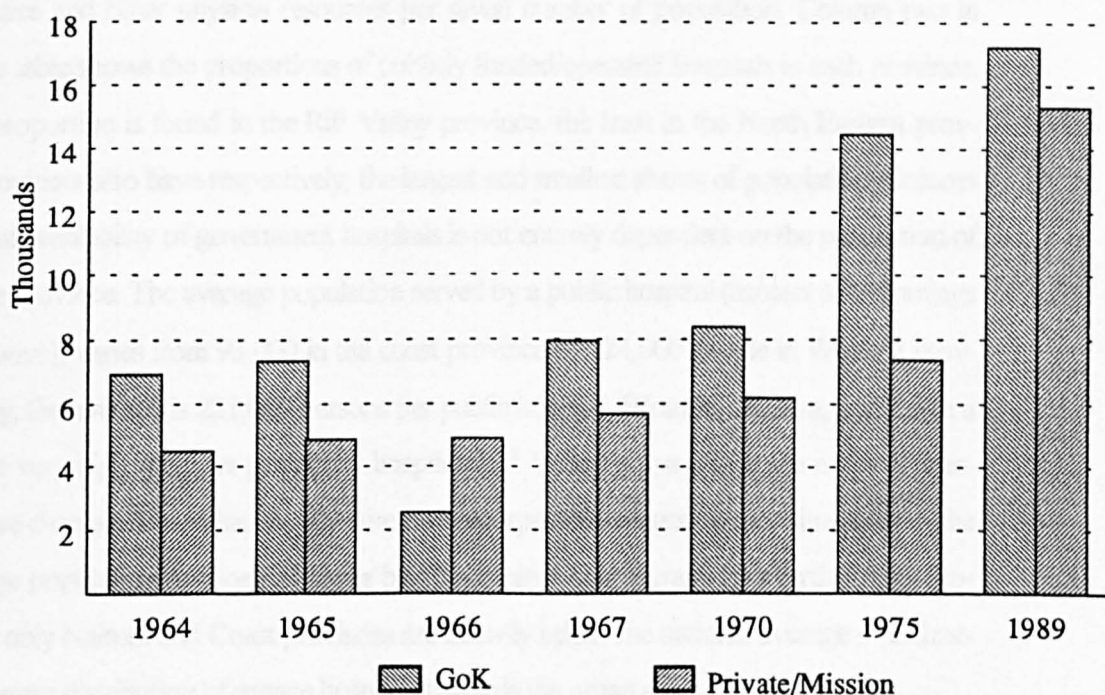


	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Hospitals	225	226	216	221	220	216	213	243	249	254	260	264	268
H/Centres	201	233	241	262	276	288	299	269	276	282	294	294	299
Dispensaries ¹	1,103	1,088	1,087	1,130	1,135	1,213	1,273	1,173	1,424	1,535	1,553	1,555	1,564

¹ Includes sub-health centres; Data for each type of facility are for both private and public sector.

Figure 2.5: Trends in health facility stock by type, (1978-1990):

The figure below (2.6) shows that between 1964 (at independence) and 1989, the total stock of beds in the country increased almost three-fold and the proportion of public beds has been generally greater. During the same period, the population just more than doubled. The proportion of public hospital beds has declined slightly in recent times, but the government remains the largest provider of facilities and services in the country. On average about 10 per cent of the total beds are reserved for children (cots).



	1964	1965	1966	1967	1970	1975	1989
GoK ¹ hospital beds	6,931	7,334	2,550	7,928	8,359	14,438	17,219
Private and mission hospital beds	4,499	4,849	4,888	5,680	6,178	7,349	15,315
Per cent of GoK ¹ beds	61	60	34	58	58	66	53
Per cent of private/mission beds	39	40	66	42	42	34	47

¹ GoK: Government of Kenya.

Figure 2.6: Trends in hospital beds by type of operating agency, 1964-1989

Facility distribution and access measures

In this section, simple access measures are used to indicate the extent of variations in availability of physical resources to populations living in different provinces in Kenya, in 1989. The measurement of the volume of work carried out by various medical institutions and how well those institutions serve the target populations for comparative purposes cannot be adequately examined without the development and application of standardisation procedures that take care of differ-

ences in the populations being served—composition by sex, age, incomes, education, etc., and, amongst others, the complexity of patient conditions that the institutions receive. However, important insights can still be discovered without the use of such complicated procedures as we show below. Inaccuracies are inevitable, because of cross-boundary flows of patients, for example, and the differences mentioned above. Consequently, the results must be interpreted with caution.

Table 2.7 provides some statistics which show the extent of variations in physical health resources availability between provinces. The data show that there are great variations in the availability of facilities and other physical resources per given number of population. Column two in part one of the table shows the proportions of publicly funded/operated hospitals in each province. The greatest proportion is found in the Rift Valley province, the least in the North Eastern province. These provinces also have respectively, the largest and smallest shares of population. Column three shows that availability of government hospitals is not entirely dependent on the proportion of hospitals in the province. The average population served by a public hospital (subject shortcomings adumbrated above), varies from 93,000 in the coast province to 424,000 people in Western province. Nationally, the average is 221,000 persons per public hospital. Western, Nyanza, and Eastern provinces have very high numbers per public hospital. Rift Valley is just above the national average. The picture changes somewhat slightly when we incorporate non-government hospitals—the national average population per hospital drops by 50 per cent. This average is nevertheless deceptive since now only Nairobi and Coast provinces are actually below the national average—indicating the very uneven distribution of private hospitals towards the urban areas¹¹.

Rift Valley province contains the largest (absolute and relative) number of health centres and dispensaries, but the population per public health centre is highest for Eastern province, followed by Central province and Nyanza—the only provinces with above average numbers per health centre. The best served region is North Eastern province—with 41,000 people per public health centre. Including non-government health centres lowers the national average from 56,000 to 44,000 persons per health centre, and Nairobi becomes the best served region nationally. Eastern province still remains the worst served region. The accessibility (by numbers) to government dispensaries shows Western province to be disadvantaged, with over six times (or 134,000 persons per public dispensary) above the national average (21,000). The picture remains much the same even after inclusion of private sector dispensaries—showing this to be one province where the private sector operations have had least impact on accessibility, although the national average is now lower 8,000

¹¹ The average for Coast province, excluding Mombasa—the second largest city in Kenya, is 92,000 persons per public hospital. This effect is present in almost all of the other measures of availability given for this province.

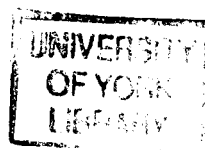
persons per dispensary, and Western province's access figure improves substantially (now 46,000 only)—but still remains about 6 times above the national average! Nevertheless, there has been an improvement since the 1970—75 Development Plan, when the (public) ratio varied from 1:100,000 to 1:50,000, with an average of 1:65,000. In 1989, the dispersion was between 1:41,000 and 1:70,000, with an average of 1:56,000. (Remember, on paper health centres are designed for catchment populations of around 30,000)

PART ONE							
1	2	3	4	5	6	7	8
	Share of GOK Hospitals in the Province %	Population per GOK Hospital ('000's)	Population per Hospital - All ('000's)	Share of GOK H/centres in the Province %	Population per GOK H/centres ('000's)	Population per H/centre - All ('000's)	Share of GOK Dispensaries in the Province %
Nairobi	7	192	64	8	45	33	4
Central	15	207	115	11	71	57	16
Coast	21	93	69	10	47	43	15
Eastern	15	248	133	12	81	59	17
N. Eastern	3	124	124	2	41	41	3
Nyanza	9	395	148	15	61	46	12
Rift Valley	23	222	117	25	51	37	32
Western	6	424	121	16	42	37	2
Kenya	100	221	111	100	56	44	100
PART TWO							
	Population per GOK Dispensary ('000's)	Population per Dispensary - All types ('000's)	Share of Other GOK facilities in the Province %	Population per Other facility type in the Province ('000's)	Share of GOK Beds in the Province %	Population per GOK Bed in the Province (number)	Population per Bed All - Beds (Number)
Nairobi	35	18	87	30	21	379	236
Central	19	15	2	3,110	16	1,149	949
Coast	12	9	6	617	11	956	390
Eastern	21	12	0	n/a ¹	15	1,474	785
N. Eastern	12	11	0	n/a	5	448	448
Nyanza	29	23	0	n/a	8	2,679	835
Rift Valley	15	9	2	4,894	20	1,419	773
Western	134	46	4	1,272	5	2,870	913
Kenya	21	13	100	411	100	1,243	658

¹ Not applicable—this means the province has no 'other facility type'.

Source: Compiled from data provided by the MOH Information Services and the 1989 Population Census Reports.

Table 2.7: Physical resources and associated measures of access by province, 1989

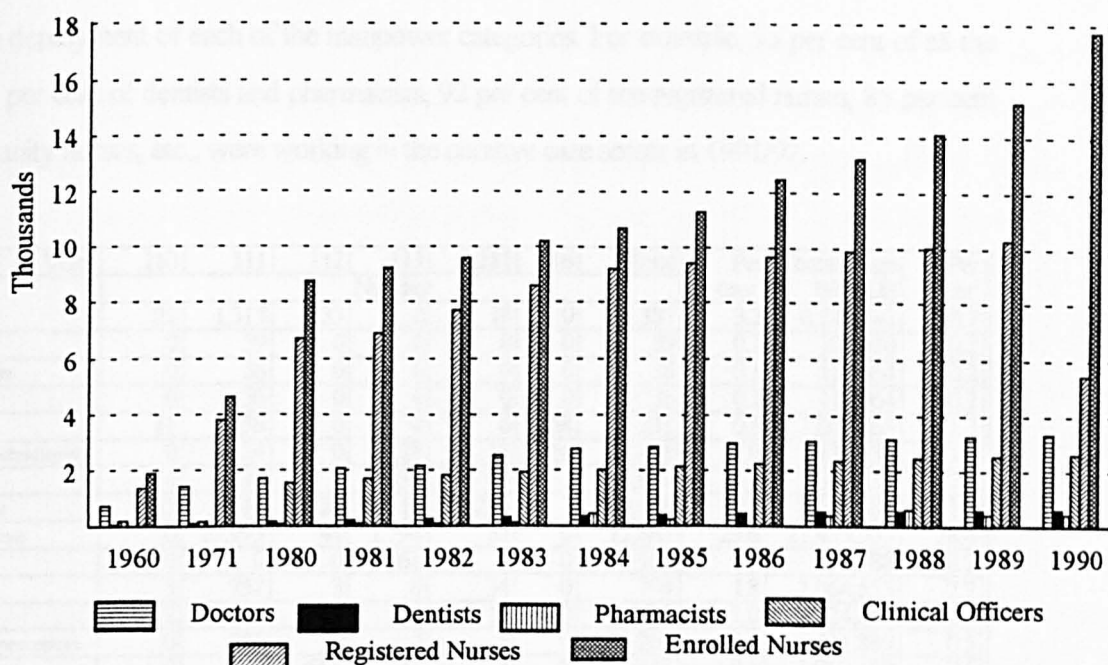


The distribution of 'other health care facilities' included in the table shows a big bias towards Nairobi province. This is because these others include mainly specialised clinics—most of them located at the national referral hospital (Kenyatta) in Nairobi, medical centres and nursing homes (private and public)—also concentrated in major urban areas—especially in Nairobi and Mombasa (Coast). Finally, the table also shows the distribution and availability of hospital beds to populations by regions in 1989, both for public sector and combined. Nairobi had the highest number, although this figure includes beds at Kenyatta National Hospital—which are available nation-wide (this being a national referral hospital). Nairobi also had the least number of people per public hospital bed (376), followed by North Eastern (448), and Central (956) and Coast province (1,149). The other provinces had bed-ratios that were above the national average (1,243), the worst being Western province (2,870 persons per public hospital bed), Nyanza (2,679) and Eastern (1,476).

Overall, using a simple average of all the rankings according to the measures given in Table 2.7, the emerging conclusion is that the Coast province comes out best overall, followed by Nairobi, then the Rift Valley, Central, N. Eastern, Eastern, Western and Nyanza provinces in that order. We soon shall examine whether this distribution of physical resources is correlated with the health situation in the regions. However, before that, something about trends and the distribution of another important parameter of the health system—health manpower.

2.2.3 Trends in and distribution of human health resources *Trends in medical and health personnel*

There also occurred substantial increases in medical and health personnel in the country. Figure 2.7 shows the changes in the stocks of some selected trained manpower over the period 1960—1990. In 1960, there were only 713 registered doctors in the country. By 1990, the figure was nearly 3,400, an increase of nearly 500 per cent. A more dramatic increase occurred in the number of dentists, though in absolute numbers it was small (see the data accompanying the figure). But these statistics should be interpreted with caution as there is no annual licensing of registered medical personnel in the country. The data therefore may overstate the numbers of those actually practising in the country. Some may be dead (or may have migrated out of the country).



	1960	1971	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Doctors	713	1,405	1,691	2,057	2,151	2,514	2,752	2,847	2,980	3,071	3,176	3,266	3,357
Dentists	52	87	162	197	239	289	331	384	441	492	527	561	596
Pharmacists ¹	170	148	60	84	86	113	427	231	—	362	588	413	443
Clinical officers ¹	—	—	1,534	1,681	1,821	1,921	2,001	2,107	2,224	2,355	2,464	2,534	2,630
Registered nurses ²	1,354	3,789	6,692	6,892	7,675	8,547	9,165	9,377	9,627	9,862	10,009	10,289	5,441
Enrolled nurses	1,865	4,612	8,722	9,191	9,518	10,168	10,650	11,248	12,452	13,202	14,078	15,200	17,724

¹ Blanks indicate missing data; ² 1990 data doubtful.

Figure 2.7: Trends in selected cadres of human resources employed in the health sector of Kenya: 1960—1990

The growth of the various categories of medical personnel has been uneven. Notice the phenomenal growth in nursing personnel as compared to other categories. Even within the nursing category, the increase in the stock of community nurses is far greater than that of registered nurses. However, it is not clear why the later declined after 1989.

The effectiveness of health care system's manpower also depends to a large extent on how it is deployed across the various functions performed by the system. The distribution of Kenya's health sector manpower by function in 1991/92 is shown in Table 2.8 (below). In 1991/92, the Ministry of health had a labour force of 43,522 personnel of diverse cadres as shown, with a total wage bill of K£75,510,120. From the last row in the table, 69.2 per cent of all the MOH personnel were deployed in the curative services sector, and only 11.7 and 11 per cent in the promotive, and preventive, and rural health services respectively. This bias towards the curative care services is re-

flected in the deployment of each of the manpower categories. For example, 95 per cent of all the doctors, 100 per cent of dentists and pharmacists, 92 per cent of the registered nurses, 85 per cent of the community nurses, etc., were working in the curative care sector in 1991/92.

Vote ¹	110	111	112	113	114	116	Total	Per cent ³	Total wage bill (K£)	Per cent ³
Category:	Number									
Doctors	21	1,315	33	0	18	0	1,387	3.2	6,548,141	8.7
Dentists	0	39	0	0	0	0	39	0.1	193,630	0.3
Dental technologists	0	38	0	0	0	0	38	0.1	113,364	0.2
Dental Trainees	0	36	0	0	0	0	36	0.1	113,964	0.2
Pharmacists	11	156	0	0	6	58	231	0.5	1,003,863	1.3
Pharmaceutical Technologist	0	324	0	98	13	16	451	1.0	1,126,135	1.5
Clinical officers	4	1,339	1	431	63	1	1,839	4.2	5,349,834	7.1
Kenya Reg. Nurses	35	3,844	28	14	256	1	4,178	9.6	8,625,888	11.4
Kenya Comm. Nurses	0	10,993	44	1,794	52	0	12,883	29.6	21,435,503	28.4
Ungraded Nurses	0	0	0	161	0	0	161	0.4	166,788	0.2
'Therapists' ²	5	752	3	0	34	0	794	1.8	2,064,553	2.7
Radiographers	2	401	3	0	15	0	421	1.0	1,224,854	1.6
Radiographic film processors	0	255	0	0	5	0	260	0.6	411,889	0.5
Health educators	0	0	613	731	1	0	1,345	3.1	1,960,468	2.6
Laboratory Technologists	0	965	70	26	44	0	1,105	2.5	2,062,958	2.7
Public health officers	0	0	170	0	2	0	172	0.4	509,904	0.7
Public health technicians	0	0	2,792	0	7	0	2,799	6.4	4,460,737	5.9
All other ministry personnel	1,386	9,664	1,316	1,345	1,478	194	15,383	35.3	18,137,747	24.0
of which subordinate	447	6,228	385	721	1,043	60	8,884	20.4	7,791,069	10.3
Total, all Ministry	1,464	30,121	5,073	4,600	1,994	270	43,522	100.0	75,510,120	100.0
Per cent ³	3.4	69.2	11.7	11.0	4.6	0.6	100.0	—	—	—

¹ The functional vote categories are:

- 110: General administration and planning;
- 111: Curative services;
- 112: Preventive and promotive services;
- 113: Rural health services;
- 114: Training;
- 116: Central medical stores co-ordinating unit.

² This category ('Therapists') subsumes physiotherapists, occupational therapists, and orthopaedic technologists.

³ Totals may not add up to 100 due to rounding errors.

Table 2.8: Functional distribution of Ministry of Health personnel, 1991/92

The third last column in the table shows the per cent distribution of all the ministry's personnel by type. Doctors constituted 3.2 per cent, dentists 0.1 per cent, community nurses 29.6 per cent, etc. The shares of the lower cadre personnel and other subordinate staff are rather disproportionately large. The last column in the table gives possible clues for this type of distribution. This column shows the shares of each category in the total wage bill. The more specialised manpower (doctors, pharmacists, etc.) generally take a more than proportionate share of the wages while the opposite is true of the more general categories (e.g., community nurses). It is possible that a substitution (by default) has led to the employment of more of the relatively cheaper cadres. Of course, the inadequacy of the higher skill manpower in the country cannot be ruled out as a possible rea-

son. However, the ministry does not have a clear manpower deployment policy to guide recruitment¹². Some guidelines on deployment within the ministry exist, but are not designed to determine appropriate staffing ratios. They are rather arbitrary and deficient for this purpose. Though they specify the *authorised* numbers of personnel of different cadres that *should* be deployed in each functional service or facility, but it appears they are not strictly adhered to, nor do they seem to be related to workloads in various types of institutions. That two hospitals are classified in the same grade, that they have say, equal floor area, number of beds, or even equipment does not mean they should similarly be staffed. The staffing needs, and indeed for many other 'inputs, the quantities deployed there should take into account the workloads that each facility handles. This means it does not make much sense to insist on constructing 'standard' types of hospitals or other health care facilities. Moreover, even with currently established 'staffing norms' appear to change from year to year, not because of changes in the parameters that ought to be used to set the *norms*, but purely to facilitate the retention of the existing number of personnel in a particular service or facility. Consequently, analysing the manpower deployment in the ministry, one gets the impression these guidelines are cosmetic. Facilities would appear in year 0, for example, as over-staffed (according to the *norms*), only to appear the following year as having the *required* personnel—*without any apparent change in the status quo*—since the Ministry officials in charge of staffing decide that that be the 'required staff' numbers for institution.

These shortcomings notwithstanding, we have analysed the data using the 'established guidelines' to determine which categories of manpower are in short supply and which are not, *according to the Ministry of Health's norms*. We have only done this at the macro level since at the micro level there are the constant changes described previously which abrogate their intended effects. The analysis shows the ministry was generally understaffed in 1991/92 by only 2 per cent, and the problem is more pronounced for the higher skill categories, for example, dentists, and less pronounced for the lower skill cadres as shown in the figure below. The figure reveals that the ministry needed only 2 per cent more doctors in that year (relative to the establishment then) to achieve its *required staffing norm* for doctors in the system. This appears to be an under-estimation of manpower requirements by the ministry, traceable to the nature of guidelines in use. The same is true of most other categories of skilled cadres. It appears inconsistent that while the national plans advocate for increased doctor-population ratio, while the ministry's planning indicates the targets are nearly achieved.

¹² Like all other ministries in the country, most manpower recruitment is done on its behalf by the Public Service Commission.

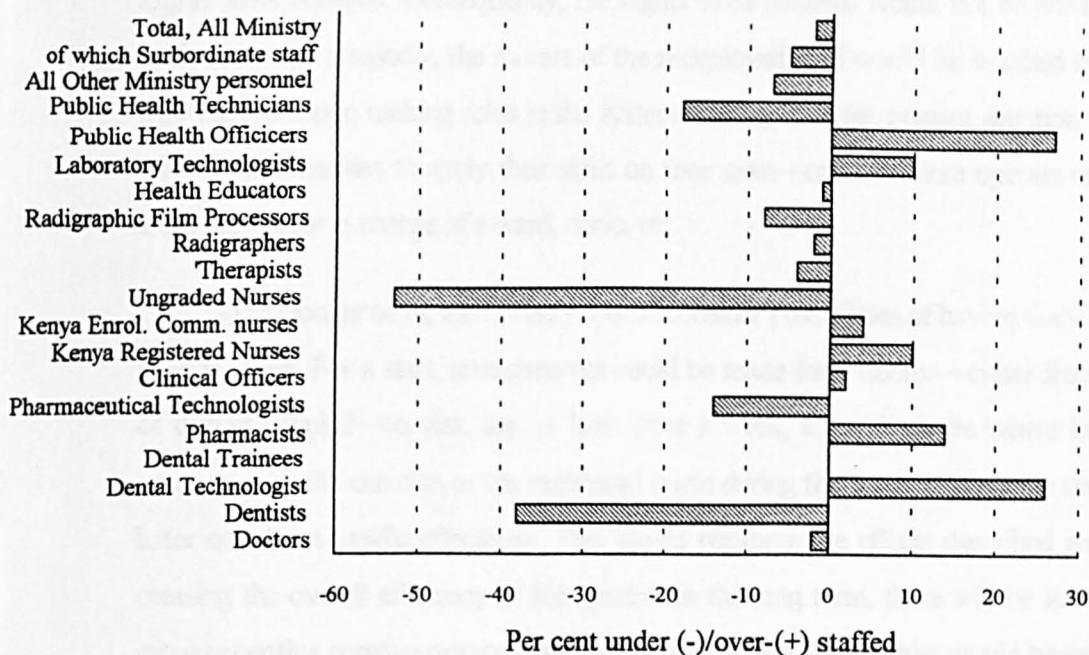


Figure 2.8: Surpluses/deficiencies in various categories of personnel: 1991/91

There is therefore need to establish proper manpower deployment guidelines based on facility and service needs. The ministry has recognised this deficiency [personal communication—MOH officers], but has yet to device and implement them. Besides establishing such guidelines, there is the question of their application. There are structural and practical difficulties in redeploying manpower (due to various political and social impediments), but, if guidelines are to be useful, it is inevitable that some sections will be adversely affected. To minimise the obstacles to implementation, factors likely to cause adverse effects should be carefully considered and where possible minimised if not eliminated. One possible starting point for such a process should be a change in the ministry’s priorities, from curative based care towards preventive, promotive and rural health services. The country has got its priorities mixed up. The bulk of the health system’s workload is handled by the lower level facilities, especially dispensaries and health centres, yet these account for only a small proportion of the total labour force. The staffing requirements of these facilities should be reconsidered. Besides the current staffing norms, the ministry should consider the possibility of gradually deploying more qualified personnel to these facilities—such as registered nurses, community nurses and more clinical officers, to be in-charge of their operations. And this should start with the more experienced of these categories so that the need for supervision by higher skill personnel is minimised. Such a move would have two desirable effects. First, it would lead to improved services in the lower level facilities and better utilisation of the labour that is so expensive to be under-utilised. Improved services in the lower level facilities would mean less referrals to the

higher level facilities. Consequently, the higher level facilities would not be adversely affected by such a change. Secondly, the morale of the redeployed staff would be boosted by the feeling that they have decision making roles in the system—rather than the present situation where they have limited opportunities to apply their skills on their own—most of them operate under instructions from the doctor in charge of a ward, clinic, etc.

In the longer term, the ministry should consider possibilities of having doctors in these lower level facilities. For a start, arrangements could be made for a doctor—either from the sub-district or district hospital—to visit, say, at least once a week, a health centre where he reviews all the cases seen by the clinician or the registered nurse during the week, and decide on those cases the latter could not handle effectively. This would reinforce the effects described above, besides increasing the overall efficiency of the system. In the long term, there will be less demand for the more expensive curative oriented facilities, and those that would exist would be capable of delivering high quality and efficient care since the cases they receive would already have been adequately screened and diagnosed at the lower levels of the system. These reflections call for a well thought manpower training and deployment scheme, as will be discussed in the main text under various proposals in chapter seven.

The regional distribution of human health resources

The distribution of health personnel is usually reflected either in the number of a specified category per unit (thousand or 100,000) population, or alternatively, the population per unit (of that category). The table below (Table 2.9) shows the population per unit of some selected categories of health personnel by region.

The table shows the availability of health personnel is closely correlated with that of the facilities, as would be expected. Coast and North Eastern are the only provinces having population per unit of the selected categories of health personnel that exceed the national averages. Although the North Eastern province is vast and sparsely populated so that the figures may not necessarily represent an accurate picture of availability, the population of this province is increasingly urbanised. The urbanisation clusters are located near main amenity centres and on this account we can accept the hunch that the province has become better served in terms of health resources than many others in recent times.

Health Personnel Category:	REGION								
	Nairobi	Central	Coast	Eastern	North Eastern	Nyanza	Rift Valley	Western	Kenya
	Population per unit of respective health personnel category								
Doctors	4,748	17,415	10,608	21,463	12,352	29,180	20,484	27,696	16,475
Dentists	151,869	703,058	420,977	964,806	266,594	883,845	825,057	731,218	585,902
Pharmacists	33,078	101,455	69,876	129,698	46,423	168,710	118,395	135,707	98,919
Clinical Officers	22,510	12,231	6,984	11,444	5,638	20,098	10,207	26,285	12,425
Registered Nurses	6,908	4,386	3,009	4,625	2,775	10,065	5,581	11,108	5,469
Community Nurses	4,074	1,528	1,111	1,554	874	2,961	1,554	2,903	1,774

Source: Compiled from information printed in the MOH Recurrent Expenditure Estimates 1991/92 (the section on manpower budgets), and population projections based on the 1989 Census district population growth rates.

Table 2.9: Population per unit of some selected categories of health personnel by province—1991/92

A better perspective of the overall distribution of manpower resources between the provinces can be gleaned from figure 2.9 below which shows the percentual deviations of population per unit of the selected categories of the health personnel from the (national) mean.

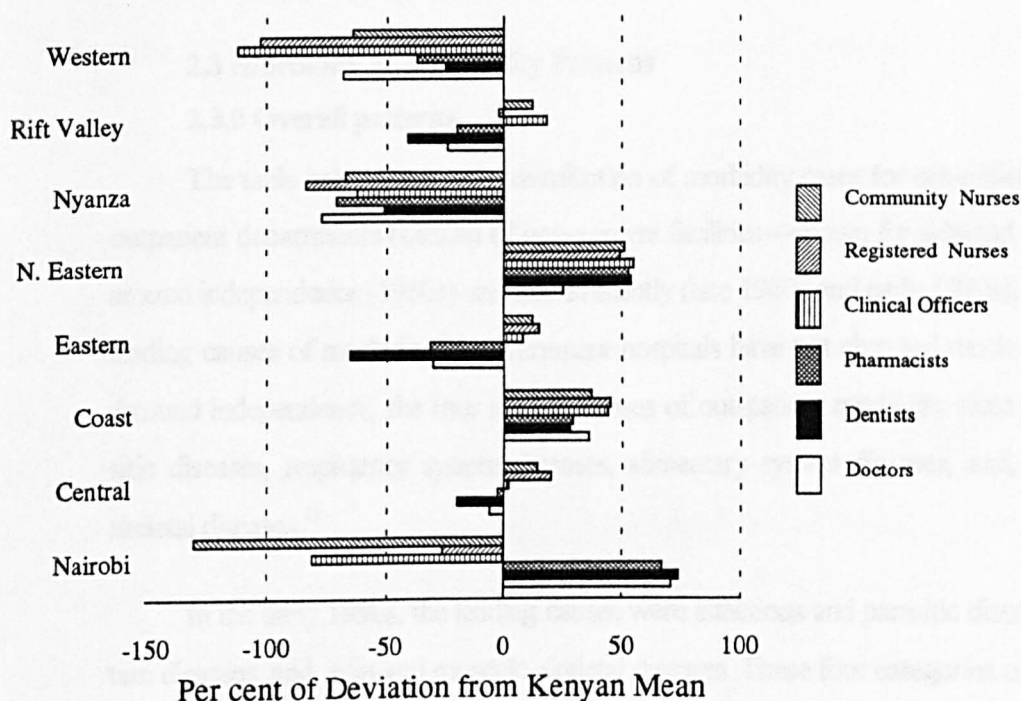


Figure 2.9: Percentual deviations from the national mean population (Kenya) per unit of selected health personnel, 1991/92

Nairobi shows a mixed picture, but is generally well served in terms of the higher skill categories of health personnel (doctors, dentists and pharmacists). Central, Eastern and Rift Valley provinces have below-average availability for the higher skill categories and generally above-average availability for the lower cadres of health personnel. However, the Central province is better off compared to the Rift Valley, which itself is in turn better off than Eastern province. The other two provinces, that is, Nyanza and Western, appear to be equally disadvantaged in all respects and their availability measures are below the national average for all categories of health personnel.

It is also clear from the distribution of human health resources that the two provinces, i.e. Nyanza and Western, are also disadvantaged in terms of the quantity of the health personnel deployed there. The important question at this point is: is there a corresponding disadvantage in terms of the health health in these provinces? Let us now look at the health situation in the country to see whether we can answer this question—i.e., to see whether the regional health status is any way related to the health resource distribution.

2.3 Morbidity and Mortality Patterns

2.3.0 Overall patterns

The table below shows the distribution of morbidity cases for out-patients presenting in the outpatient departments (OPDs) of government facilities—shown for selected years for the period around independence (1960s) and more recently (late 1980s and early 1990s). The data shows the leading causes of morbidity in government hospitals have not changed much since independence. Around independence, the four leading causes of out-patient morbidity were infectious and parasitic diseases, respiratory system diseases, alimentary system diseases, and, skin and musculo-skeletal diseases.¹³

In the early 1990s, the leading causes were infectious and parasitic diseases, respiratory system diseases, and, skin and musculo-skeletal diseases. These four categories of diseases accounted for over three quarters of all OPD cases. In the parasitic and infectious diseases group, the leading causes of morbidity have been malaria, diarrhoeal diseases and early childhood diseases such as measles, whooping cough, etc. In the respiratory tract diseases group, pneumonia is the single largest cause of morbidity. In 1991 about 88 per cent of all OPD visits were caused by infectious and

¹³ We exclude diseases of the nervous system and sensory organs whose prominence waned off almost immediately on the assumption that this might have been a result of traumas of the fight for independence. Also excluded is the general category of ill-defined diseases and accidents.

parasitic diseases; respiratory system diseases; skin and musculo-skeletal diseases, and, ill-defined diseases, fractures and injuries. Thus close to 90 per cent of the health problems that are handled by curative services are problems that are largely preventable and possibly at low cost (see chapter 4, section 4.3.2 for a discussion of this aspect of the health services).

	1963	1964	1984	1985	1986	1987	1991
Infectious and parasitic diseases	17.4	26	42.1	40.5	36.2	37.8	38.5
Malignant growths/neoplasms	0.1	0.2	0	0	0.1	0	0
Allergic, endocrine system and metabolic diseases	1.7	2.2	1.6	1.7	1.6	1.4	5.7
Diseases of the nervous system/sensory organs	24.2	5.5	5	4.9	5.2	4.5	4.1
Circulatory system diseases	0.2	0.4	0.3	0.4	1	0.3	0.7
Respiratory system diseases	16.5	20.5	20.3	23.4	23.1	22.4	23.3
Alimentary system diseases	12.4	13.4	0.4	0.9	1.3	0.9	1.1
Genito-urinary system diseases	2.3	3.1	0.3	0.4	0.3	0.3	0.2
Skin and musculo-skeletal diseases	9.4	1.6	8.1	7.7	8.3	9.4	10.3
Ill defined diseases, accidents, fractures and injuries	15.8	20	21.9	20	22.8	23.2	16
Total	100	100	100	100	100	100	100

Source: Compiled from various MOH Annual Reports

Table 2.10: Patterns in OPD morbidity by type of disease category: Various years between 1964 and 1991

Data on hospital admissions and deaths also reveal a pattern of largely avoidable problems that is no less disconcerting than outpatient data. The table below shows the 20 leading causes of admissions and deaths in MOH hospitals in 1990. Between them, the top 20 diseases accounted for 63 per cent of all admissions, and 54 per cent of all hospital deaths. Concerning admissions, of the 63 per cent, infectious and parasitic diseases; and diseases of the respiratory system, respectively, accounted for 25 and 13—i.e. 38 per cent of all the admissions due to these top 20 diseases (which translates to 40 and 20 per cent respectively for admissions due to all causes).

In a 1990 Provincial and District Health Services Study (REACH/MOH, 1990), a sample of patient records reviewed in a sample of facilities revealed that in 1989, burns, injuries and wounds, fractures and malaria accounted for 26 per cent of all inpatient admissions into the Nakuru Provincial General Hospital. At Naivasha District hospital, bronchial pneumonia, abortion, malaria, and burns accounted for 34 per cent of the inpatient admissions. In the private sector facility in the sample (Mercy Hospital), malaria, pneumonia, abortion, and measles accounted for 47 per cent of ad-

missions. In Naivasha and Mercy hospitals, malaria and abortion accounted for 37 and 42 per cent of admissions to the female wards, respectively; malaria, wounds, injuries and traffic accidents accounted for 52 and 27 per cent of the admissions into the male wards. In the paediatric wards, bronchial pneumonia and malaria accounted for 34 and 44 per cent of admissions respectively.

ICD ¹ Group	Description	Admissions (per cent) ²	Deaths (per cent) ³
1	Infectious and Parasitic diseases	24.94	15.91
2	Neoplasms	0	8.05
3	Endocrine, nutritional, metabolic diseases.	1.55	7.89
4	Diseases of blood and blood forming organs	6.46	6.18
5	Mental disorders	0.22	3.13
6	Diseases of the nervous system/sense organs	0.48	2.2
7	Diseases of the circulatory system	0.42	1.7
8	Diseases of the respiratory system	12.45	1.45
9	Diseases of the digestive system	0.17	1.31
10	Diseases of the genitourinary system	0.32	1.24
11	Complications of pregnancy, childbirth and puerperium	13.73	1.21
12	Diseases of skin and subcutaneous tissue	0.16	1.06
13	Diseases of musculo-skeletal system and connective tissue	0	0.73
14	Congenital anomalies	0	0.49
15	Certain conditions originating around perinatal	0.01	0.39
16	Symptoms, signs and ill defined conditions	0.54	0.36
17	Injury and poisoning	1.5	0.23
	Total	62.96	53.54

¹ International Classification of Diseases. Entries with zeroes should be interpreted to mean that no disease(s) from the group featured among the top 20 causes of admissions or deaths in 1990.

² As per cent of total (nation-wide admissions due to all causes).

³ As per cent of total deaths due to all causes.

Source: Compiled from data supplied by the Health Information System of the Ministry of Health.

Table 2.11: Distribution by ICD classification of the top 20 leading causes of admissions and deaths in MOH hospitals in 1990.

These data suggest that a large fraction of inpatient cases—by far the most expensive level of treatment in the system—are preventable through other interventions. Concerning mortality, the table also shows the single largest cause of mortality was infectious and parasitic diseases—these are also preventable. If we add together all causes of mortality that are preventable (e.g. infectious and parasitic diseases, diseases of the respiratory system, and (most) diseases of the skin, musculo-

skeletal system and connective tissue), preventable deaths rise to at least 25 per cent. It is our intention to propose a reform strategy that will create effective mechanisms for tackling these essentially 'public goods' type health problems (as well as the less public ones that today gobble up the lion's share of health resources).

2.3.1 Regional out-patient morbidity patterns

Data on morbidity by region (provinces) for 1984 and 1991 (the only two years where an indication of the response rates of surveyed institutions is given) is shown in Table 2.12 below¹⁴.

1984									
	Nairobi	Central	Coast	Eastern	North Eastern	Nyanza	Rift Valley	Western	Kenya
Infectious and parasitic diseases	26	24	45	39	48	54	40	50	38
Malignant growths/neoplasms	0	0	0	0	0	0	0	0	0
Allergic, endocrine system & metabolic diseases	2	1	2	1	3	2	3	1	6
Diseases of the nervous system/sensory organs	5	5	4	5	5	4	4	3	4
Circulatory system diseases	0	0	0	0	0	0	0	0	1
Respiratory system diseases	22	28	17	21	24	18	25	20	23
Alimentary system diseases	1	0	0	1	1	1	1	1	1
Genito-urinary system diseases	0	0	0	0	0	0	0	0	0
Skin and musculo-skeletal diseases	7	8	10	8	8	10	9	9	10
Ill defined diseases, accidents, fractures & injuries	37	33	20	24	10	11	17	15	16
Total	100	100	100	100	100	100	100	100	100
Population (Thousands)	1,049	2,697	1,573	3,176	372	3,061	3,965	2,153	18,046
Share of total population (per cent)	6	15	9	18	2	17	22	12	100
Share of total morbidity (per cent)	1	14	10	15	1	28	26	17	100
1991									
	Nairobi	Central	Coast	Eastern	North Eastern	Nyanza	Rift Valley	Western	Kenya
Infectious and parasitic diseases	35	27	46	22	48	54	40	50	38
Malignant growths/neoplasms	0	0	0	0	0	0	0	0	0
Allergic, endocrine system & metabolic diseases	1	0	2	43	3	2	3	1	6
Diseases of the nervous system/sensory organs	5	4	5	3	5	4	4	3	4
Circulatory system diseases	0	2	0	0	0	0	0	0	1
Respiratory system diseases	28	30	21	14	24	18	25	20	23
Alimentary system diseases	1	2	1	1	1	1	1	1	1
Genito-urinary system diseases	0	0	0	0	0	0	0	0	0
Skin and musculo-skeletal diseases	10	13	11	7	8	10	9	9	10
Ill defined diseases, accidents, fractures & injuries	19	22	13	10	10	11	17	15	16
Total	100	100	100	100	100	100	100	100	100
Population (Thousands)	1,480	3,289	1,971	3,962	372	3,772	5,306	2,712	22,863
Share of total population (per cent)	6	14	9	17	2	17	23	12	100
Share of total morbidity (per cent)	2	22	12	8	1	19	26	9	100

Source: Compiled from data published in Statistical Abstracts, various years.

Table 2.12: OPD morbidity patterns by disease group and by province, Kenya: 1984 and 1991—per cent and numbers

¹⁴ The data for 1987 give similar information but the total for Nyanza province appears inconsistent with the published figures. The 1991 data are provisional.

The table shows in 1984 the outpatient department (OPD) regional pattern was similar to the national average as discussed above, with the exception of Central province where diseases of the respiratory tract are predominant. This difference is explained by the high altitude of the province, which makes the climate different from that of the rest of the country. Nairobi, which is also high altitude, has above the national average cases for respiratory diseases too. Over 50 per cent of all the OPD cases in each of Nyanza and Western provinces in that year were due to infectious and parasitic diseases, reflecting another ecological variation—low lying mosquito infested areas, where malaria accounted for a substantial proportion of these cases.

The table also shows the proportions of population residing in each province and the corresponding burden of total morbidity as a per cent of the national totals. Nyanza, Western and the Coast provinces had a higher share of morbidity than their shares of the population in 1984, although the situation was more pronounced in the first two provinces. The analysis of section 2.2 above showed these two provinces had below average shares of most categories of health resources. It therefore appears that the relatively high share of morbidity in them (the provinces) owes its origin partly to this distribution. At least this is true for Nyanza province. After all, neither province has more of the private sector resources than others.

2.3.2 Regional mortality patterns according to hospital based information

A comparison of the pattern of hospital mortality between regions is given in Table 2.13 below for 1968 and 1990. The 1968 data is extracted from Bonte (1974). The following points should be taken into account when interpreting the data. First, the 1968 data combines Nairobi and Central provinces. Second, the 1990 data is only for the top twenty leading fatality diseases in the country as at then. Third, the disease classification procedure used differs between the periods—the former is more aggregative as compared to the latter. Despite these differences however, it is possible to get an idea about the trend pattern between the two years.

The table shows that in 1968 at the national level although only 86 males got admitted for every 100 females, there was a higher casualty for the males—124 males died for every 100 women who died. Only Nyanza and the North Eastern provinces had death-sex ratios of less than 100, possibly on account of the small absolute numbers. On the other hand in 1990, the sex ratio of deaths increased slightly to 127, and Nyanza and Western provinces this time have lower ratios

than the rest of the country. The Western province ratio is very low—again possibly due to the low absolute numbers. The sex ratio for Central province, including Nairobi in 1968 is low probably due to the low sex ratio of Nairobi, which is evident in the 1990 figures.

Region	1968										
	Admissions ¹			Deaths ¹			Sex ratios ²		Clinical mortality ³		
	Males	Females	All	Males	Females	All	Adm.	Deaths	Males	Females	All
Central, including Nairobi	22,143	26,040	48,183	1,818	1,563	3,381	85	116	8	6	7
Coast	10,552	8,573	19,125	602	345	947	123	175	6	4	5
Eastern	7,884	12,089	19,973	749	658	1,407	65	114	10	5	7
North Eastern	1,489	993	2,482	24	35	59	150	69	2	4	2
Nyanza	5,266	5,447	10,713	194	235	429	97	83	4	4	4
Rift valley	20,415	25,060	45,475	1,570	1,163	2,733	82	135	8	5	6
Western	5,137	6,430	11,567	290	247	537	80	117	6	4	5
Kenya	72,886	84,632	157,518	5,248	4,246	9,494	86	124	7	5	6
Region	1990										
	Admissions ¹			Deaths ¹			Sex ratios ²		Clinical mortality ³		
	Males	Females	All	Males	Females	All	Adm.	Deaths	Males	Females	All
Nairobi	86	535	621	13	91	104	16	14	15	17	17
Central	13,520	25,930	39,450	926	546	1,472	52	170	7	2	4
Coast	17,818	18,734	36,552	858	528	1,386	95	163	5	3	4
Eastern	11,385	14,143	25,528	662	357	1,019	80	185	6	3	4
North Eastern	1,528	1,381	2,909	128	106	234	111	121	8	8	8
Nyanza	10,247	13,070	23,317	386	624	1,010	78	62	4	5	4
Rift valley	12,627	22,072	34,699	896	611	1,507	57	147	7	3	4
Western	1,281	9,098	10,379	93	257	350	14	36	7	3	3
Kenya	68,492	104,963	173,455	3,962	3,120	7,082	35	127	6	3	4

¹ Absolute numbers; ² (Males/Females)x100; ³ (Deaths/Inpatients)x100;

Compiled from data by Bonte (1974) and other data from the MOH Information Services.

Table 2.13: Admissions and deaths by sex in reporting government hospitals 1968 and 1990 by province

There is no discernible pattern between the two periods since some areas such as Coast, Nyanza and Western provinces have registered declines, while Eastern, North Eastern and the Rift Valley recorded rises in the inpatient sex ratios. It is not clear whether such changes are linked to the health situation with respect the sexes in the respective areas, and there is need for further investigation to determine what factors most likely contributed to these trends. But it is also notable that the change in inpatient death sex-ratios is closely related to the changes in admission sex-ratios, which have similarly changed.

In 1968, as in 1990, males had a higher clinical mortality than women¹⁵, that is, for a given number of both sexes each admitted, males suffered greater fatal casualties than females, as depicted in the last 3 columns of table. This pattern exists in all regions except in Nyanza province in the two years. Regionally, the worst clinical mortality ratios were recorded in the central province, including the Nairobi area and the Eastern province in 1968—whose average exceeded the national figure. The Central province figure during that year is probably exacerbated by the adverse Nairobi figure that is evident from the 1990 figure. There occurred an improvement in all regions between the two years except in North Eastern province (excluding Nairobi area). The high Nairobi figure for 1990 can possibly be explained by the complexity of the cases received by the national referral hospital (Kenyatta National Hospital) and other specialised institutions there—cases whose survival chances are low. Possibly there are other explanations—such as the presence of a large teaching hospital there, or the lack of a district hospital, to help ease the burden of admission for Nairobi residents, etc. But there does not appear to exist any discernible correlation between the distribution of health resources (as discussed previously) and the pattern of hospital inpatient mortality. Western and Nyanza provinces, which were shown to be relatively disadvantaged in terms of health resources' distribution, do not possess significantly worse records of mortality in this case. Similarly, areas shown to be favoured by the resource distribution (e.g., Coast province) have more or less the same patterns of mortality as the disadvantaged areas, some are even worse (e.g., North Eastern).

2.3.3 The distribution of health resources and health status—implications for policy

Thus far we have analysed the distributions of health resources and 'health' in the country. The available data, though scanty and probably unreliable, generally suggest that regions with more health resources are better-off in terms of health status. This is as would be expected. Theory predicts a strong relationship between the health inputs and the health status, i.e., areas with more health inputs (resources) will tend to have better health status (e.g., lower mortality or morbidity, and so on), and vice-versa. This type of analysis is best approached through the health production function. However, the absence of reliable data on the health situation at district levels (the availability of which would have given us adequate degrees of freedom to estimate a health production function for the country or even regions) precludes this type of analysis. But it is still possible to

¹⁵ Why this is so is hard to rationalise, but it is possible that women, being from the very beginning the target of most health awareness campaigns, consult early enough before their health situations deteriorate too much. May be men engage in more 'risky' jobs.

show the same relationship (between inputs and outcomes) holds for Kenya, albeit using crude and rather highly aggregated (provincial) data.

	Doctors	Clinical officers	Registered nurses	Enrolled nurses
Infant mortality rate	-.5448 (.103)	-.3376 (.229)	-.4020 (.186)	-.4030 (.185)
Under 5 mortality rate	-.5659 (.093)	-.4588 (.150)	-.5510 (.100)	-.5211 (.115)
Childhood mortality rate	-.5576 (.097)	-.5845 (.084)	-.7083 (.037)	-.6375 (.062)

Table 2.14: Simple correlation between IMR, U5M, CMR and different categories of health personnel in Kenya around the period 1991/2

The table above shows the simple correlation between various definitions of under age 15 mortality (usually a better measure of overall health status in a given region than total mortality) and various categories of health personnel, calculated using data on mortality and health personnel for Kenya's provinces around 1991/2. The correlations are negative as would be expected. The figures below the coefficients (and in brackets) show the calculated (probability) levels at which, for a one tailed test they can be accepted as significant. Although these (significance) probabilities are rather high, the important thing to note here is that there exists an inverse relationship between mortality and these health care inputs. This is true for most health inputs and has implications for the redistribution of health resources in the country. First, if all the regions were resourced on the same per capita basis as the best resourced province (e.g., the Coast or North Eastern provinces), we can expect an improvement in the overall health situation in the country, although this conclusion is subject to whether those areas receiving the extra resources are capable of converting them into increased health (see chapter four for more on this). But this situation may be highly hypothetical since we have seen the actual situation is one of declining resources to the health sector. Second, if instead we were to redistribute the existing resources 'equally' across the provinces, two outcomes may be the case. One, if the losing provinces were already having an 'excess' of resources over their actual needs (which is highly doubtful), we can expect an overall increase in the overall health status. If otherwise, the overall outcome depends on a comparison between the gains and losses. If the gaining areas are able to generate more 'health' out of those resources than the losing ones, then on balance the overall health situation would improve. If not, there will be a net loss of 'health'. Indeed, given the heterogeneity of the epidemiological situation across regions, it is

not possible to tell *a priori*, the outcome of such a redistribution of resources. It would be necessary to develop a needs based formula for redistributing resources, one which makes use of all the available information on inputs and outputs for the different areas, which leads to the same conclusion arrived at above—we need to know the relationship between the inputs and outputs in the health sector.

2.3.4 Conclusion

The analysis of morbidity and mortality data from various sources given above gives a general perspective on the “demand” for health services in Kenya and helps us to approach the problem of those who are most at risk in terms of the socio-economic characteristics of the population. But it is not very revealing on some important questions that are important for the formulation of health policies, such as the kinds of risks that are unequally distributed. The only truly satisfactory way of approaching these matters is by means of exhaustive health and morbidity surveys such as those undertaken in the developed countries (and some developing countries). Nevertheless, the analysis done so far helps us at least to begin to approximate the extent of inequality of risk, both by types (groups) of individuals or by region. Unfortunately, most of the available (published) statistics are not really very helpful for an analysis concerned with the distribution of health and the differential impact of sickness and death, at least not in the form they are currently published. We have in mind data such as those on mortality through notification of deaths, or the type published in the statistical abstracts on outpatient morbidity, for example. Unless one is able to do a considerable amount of work with the semi-processed data available at the Registrar-Generals and/or the Ministry of Health ‘archives’, there is no alternative but to use broad classifications - e.g., by regions, to approximate the more meaningful socio-economic distinctions such as urban/rural, high/low income, etc. Such an approach is obviously fraught with biases/errors. The data by provinces are of course indirectly related to such distinctions, as the provinces have unequal rates of urbanisation or education, and vary in gross products per capita, as well as in disease ecology. But the comparison of average data per province leaves us ignorant of these differences within each province (and these may paint a totally different picture)—despite favourable values for health indicators, extreme inequalities may exist (as is apparent from the implications of the data on the Coast province in this and the previous chapter). With these caveats in mind, the following main conclusions follow.

Mortality, and infant mortality in particular, which is still high by international standards, has been declining over the past 40 years or so, and appears to be on a downward trend yet, although the rate of decline has somewhat slowed since the early seventies—the fastest rate of decline occurred in the period immediately after independence, possibly because at the time medical services were availed to groups hitherto then missing that “right” under the colonial administration. Infant mortality is much influenced by ‘exogenous’ causes deriving mainly from the environment—infectious diseases for example, which can be related to medical services availability, the nature of housing sanitation, etc. It has not been possible to demonstrate changes in mortality through time due to data unavailability, particularly on mortality itself, but the scanty data that is available indicate there are substantial differences in infant mortality rates as between provinces. Without data on provincial GNPs, it would not be possible to explain adequately why any variations occur over time. Nevertheless, it pays to point that the more economically advanced provinces (e.g., Nairobi, Central and some parts of the Rift valley province) have generally lower infant mortality rates. The same is true of fertility rates. There does not appear to exist a correlation between urbanisation and infant mortality - Nairobi, wholly urban, has a higher infant mortality than Central, Eastern, and Rift valley provinces. However, it must be realised random fluctuations occur from year to year and reliable comparisons can only be made on the basis of a long time series data (using moving averages for example).

The data on overall mortality might give a better picture since it is available at more than one point in time, is more broadly based and includes cause of death, but is deficient in accuracy. This shortfall notwithstanding, it reveals substantial differentials in the distribution of different kinds of health risks in Kenya. Deaths due respiratory and infectious/parasitic diseases are widespread in the country. These deaths (especially those due to infectious/parasitic diseases) are usually associated with low health standards—the result of poor environmental conditions and/or poor health facilities. But the former vary more with geographic and climatic condition.

The evidence of the unequal spread of health risks between Kenya’s provinces discussed above is in addition to other, probably great inequalities that can be expected to exist within each province—e.g., between rural and urban dwellers, higher and lower income groups, etc. This unequal spread will obviously be related to broad factors of socio-economic nature and to inequalities and in the availability and efficiency of the health services to the different areas and/or groups. The next chapter will review the organisation and development of the health services sector and assess policies that have been implemented to rectify the unequal distribution of health risks (if any) in

Kenya. But before then, let us first look at how Kenya's health situation compares with that of other developing countries.

2.4 Some International Comparisons

2.4.0 Some caveats about international comparisons

Data bases for different countries are seldom truly comparable, but cross-country comparisons can sometimes be indicative of deficiencies in one country's practices as compared with others. Some studies have examined whether countries with the same per capita spend the same fraction of their health care resources on hospital services. The evidence, subject to (possibly large) data imperfections, suggests the allocation varies greatly from country to another (see for example Mills, 1990a, b). The general hypothesis emerging is that most countries currently allocate health resources sub-optimally. However, cross country analyses can be highly deceptive because of a number of factors (De Ferranti, 1983a). First, the data rarely reflect real differences across countries due to disparities in the quality and definitional conventions used. Second, in some cases the published data may cover only the expenditure by the ministry of health, in others, other public and private providers are included. The way the health sector is defined also is important—in some instances it may simply mean medical care to individuals, mass immunisations and control of vector borne diseases; or it may be more comprehensive—also encompassing environmental sanitation, family planning, public health education and promotion, nutrition programmes, solid waste disposal, etc. Differences also occur depending on whether only recurrent expenditures are given, or development expenditures are also included. The conventions used to determine what counts as recurrent/development expenditure also differ.

The data can also be misleading since it may fail to account for differences in purchasing power of the various countries' currencies (i.e. purchasing power parity). Also, figures on current (average) spending elsewhere are not necessarily a reasonable guide to how much any particular country ought to spend. That average may be too high or too low. Moreover, cross-country statistics do not reflect complementarity between the health and other sectors, and in other cases, they simply do not allow for differences in population structure and composition, the country's health profiles, and the costs of appropriate technologies for prevention and cure. While some studies have attempted to rectify for some of these deficiencies¹⁶, in general, their approaches do not re-

¹⁶ For example, the studies by Tait and Heller (1981), and Golladay and Liese (1980).

solve the problem since if the current allocation is sub-optimal, any information based on such data will serve to perpetuate the existing inefficiencies and inequalities.

What follows adds some perspective to the development of health expenditure in Kenya set in an international perspective, subject to the foregoing caveats.

2.4.1 Performance of Kenya's health sector viewed in relation to other developing countries

By invoking exchange rates, it is possible to relate health expenditure per head for different countries to the per capita incomes in those countries. This has been done in the present work. Table 2.15 shows the distribution of a sample of 60 developing countries for which data on government expenditure on health, education, defence and other social services and GNP per capita was available.

Per cent of government expenditure spent on various services	Health ¹	Education ²	Defence ³	Other Social Services ⁴
Less than 5 %	25	5	7	21
5% but <10%	26	10	21	12
15% but <20%	6	35	20	15
20% and more	2	9	8	9
Total Number of countries	59	59	56	57

¹ Mode: 6 per cent; average: 7 percent.
² Mode: 19—20 per cent; average: 15 per cent.
³ Mode: spread almost evenly between 4 and 14 per cent, with a hunch around 8—9 per cent.
⁴ There is no clear distribution pattern across countries but the mean was 12 per cent.

Source: Compiled from information contained in Table 2.16 at the end of this chapter.

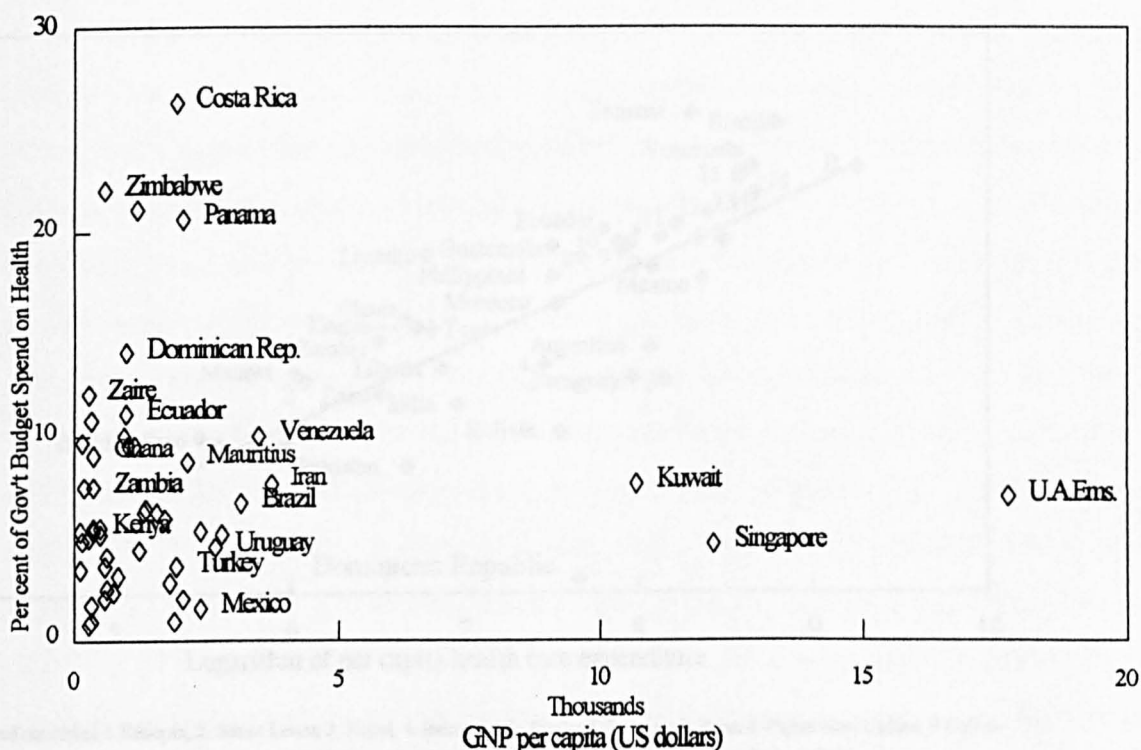
Table 2.15: Distribution of a sample of developing countries by per cent of government expenditure on social services¹⁷

The table shows most developing countries allocate about half as much of government resources to health as to education or defence. In general, the health sector receives less than 10 per cent of the budget (the sample average was 7 per cent) while education, defence and other social services¹⁸ receive about 15, 13, and 12 per cent respectively (with sample modes of 19—20 per

¹⁷ The data used in the table are for the period 1987—1990, derived from two main various sources—UN Government Finance Statistics Yearbooks and UNICEF, 1993.

¹⁸ Other social service expenditures are those on social security and housing, among others.

cent for education, 8—9 per cent for defence. Kenya allocates only about a quarter as much of the government budget to health as to education and defence. the figure below depicts this information pictorially using the GNP per capita against government spending on these services. Even without discounting outliers e.g., Colombia, Costa Rica, Panama, etc., (all with low GNP per capita but with high allocations of government spending on health) and Kuwait, Singapore, United Arab Emirates, etc., (these with high GNP per capita but low government budget allocations to health), there is no apparent correlation between the amount that governments of developing countries allocate to health and their GNPs per capita—some with low per capita incomes allocate comparatively higher proportions of their government’s budget to health and vice-versa.



Note: For most levels of government spending on health, only the country with the highest GNP per capita is named, and only where it is conveniently placed for labelling; U.A.Ems: United Arab Emirates

Figure 2.10: Per cent of government spending on health and GNP per capita in developing countries: 1987—1990

However, as is the case in developed countries, there exists a positive relationship between GNP per capita and per capita spending on health as the figure below shows (the relationship is approximated by the fitted [regression] line AB in the figure).¹⁹ Both per capita health spending and per capita GNP are in constant 1987 prices and have been adjusted for purchasing power parity.

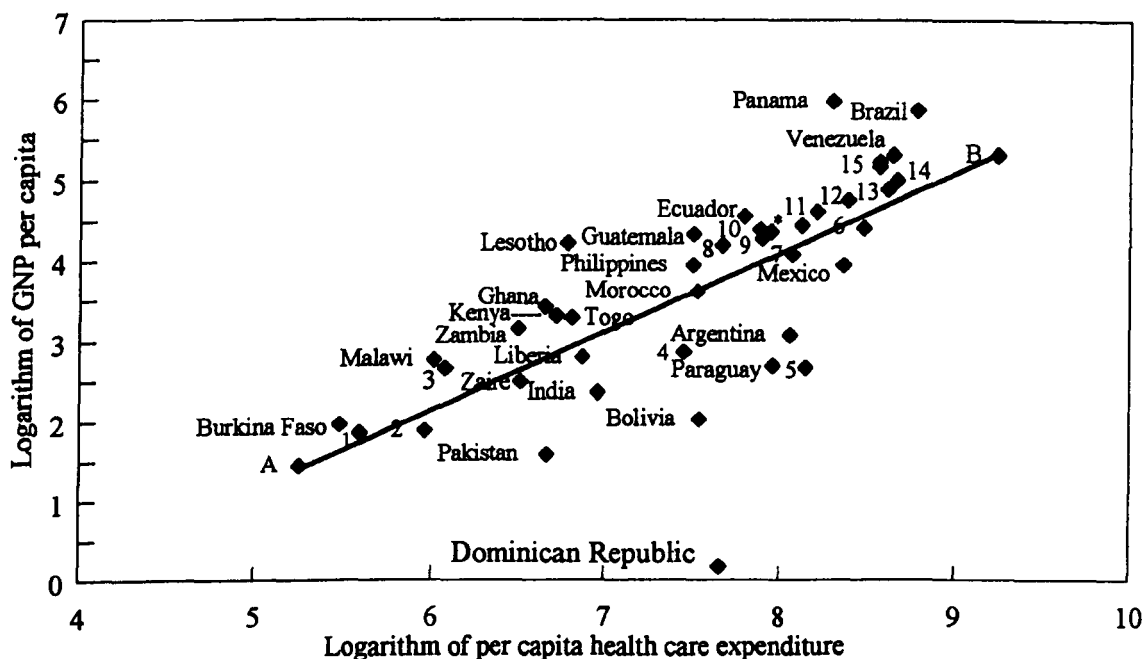
¹⁹ This result is similar to what has been found in the developed countries—that the amount spent on health care in any country has little to do with the degree of state involvement in finance, but has a great deal to do with the level of national income (Abel-Smith, 1967; OECD, 1977; Newhouse, 1977).

The data are for different years between 1986 and 1990. Therefore the relationship is only approximate. The fitted (double-logarithmic) line (AB) takes the form

$$\ln HCE_{pc} = \alpha + \beta \ln GNP_{pc}, \beta \geq 0 \dots\dots\dots (2.1)$$

where \ln is the natural logarithm, HCE_{pc} is health care expenditure per capita, GNP_{pc} is GNP per capita, α and β are parameters. The specific equation is²⁰

$$\begin{aligned} \ln HCE_{pc} = & -3.6568 + 0.9699 \ln GNP_{pc} \dots\dots\dots (2.1a) \\ & (6.0498) \\ & R^2 = 0.484, \\ & n = 41. \end{aligned}$$



Numbered countries: 1 Ethiopia, 2. Sierra Leone, 3. Nepal, 4. Indonesia, 5. Thailand, 6. Egypt, 7. Peru, 8. Papua New Guinea, 9 Colombia, 10. El Salvador and Jordan, 11. Tunisia, 12. Turkey, 13. Malaysia, 14. Mauritius, 15. Uruguay and Syria, * Chile.

The line AB is fitted on the basis of equation 2.1a in the main text.

The mark for most countries is immediately to the right, except for Togo, and Mauritius (14). Where two or more countries are named on the same line, the first one is represented by the first mark on the right, etc.

Figure 2.11: Per capita health spending and GNP per capita in developing countries: 1986—1990

The bracketed term is the *t-ratio*, and is highly significant. The coefficient of $\ln GNP_{pc}$ can be interpreted as the elasticity of HCE_{pc} with respect GNP_{pc} . Though we can reject the hypothesis $\beta=0$, we cannot reject the alternative that $\beta=1$, even at the 1% level of significance, hence health is

²⁰ The sample of usable observations dropped due to gaps in some countries' data.

a normal good in developing countries (if only for those in the current sample, which represent all levels of development in the developing world).

The observed variations in health care expenditure per capita (public and private) in developing countries shows a very wide variation—a factor of over 30 in conservative estimates (see the column on health care expenditure per capita in Table 2.16 at the end of this chapter) unlike in the developed countries where the variation factor is only about five (see e.g., Culyer, 1987, 1988b). Another difference is that whereas variations in the developed countries' national income almost entirely accounts for all the variations in HCE_{pc} , in the developing countries it only explains just over 50 per cent of the variations.²¹

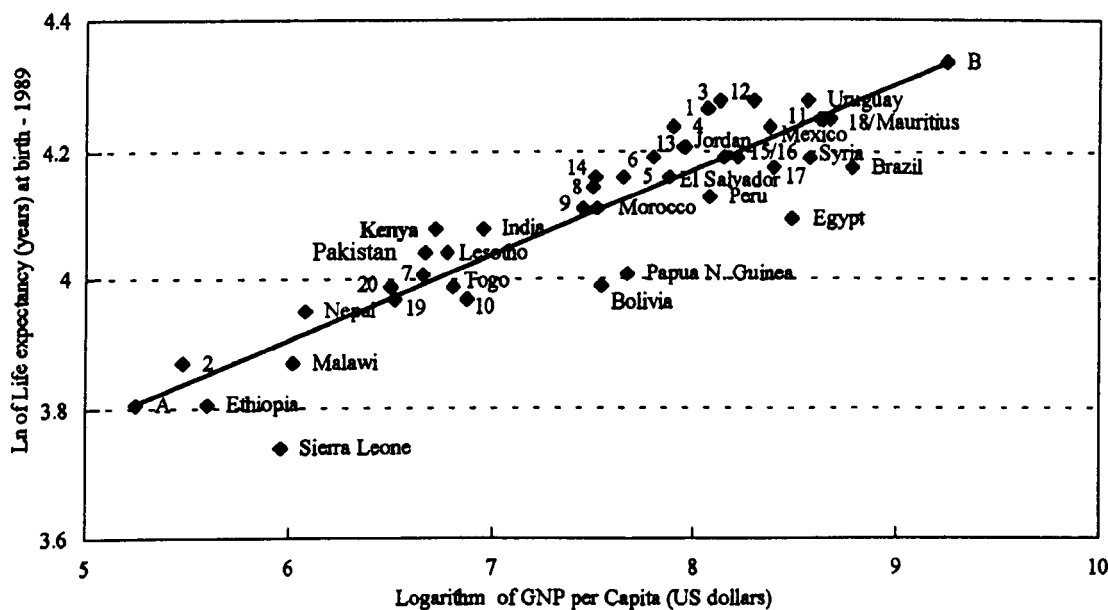
The above result has important implications for health expenditure policy in developing countries. Since the income elasticity is positive, it means the consumption of health care services, made possible by available disposable income, will be directly or indirectly influenced by the distribution of income—any reduction in disposable income will reduce people's (or a country's) ability to consume health care services. Specifically, any increase in the inequity of the distribution of disposable income will increase inequalities in the access to health care services, thereby aggravating any existing inequalities in health. For Kenya, this has the following implications. Available income distribution information shows that Kenya, a low income country according to UN classifications, has an inequitable income distribution, favouring the high income groups (UN, World Development Report, 1986)—the lowest 20 per cent were earning a mere 2.6 per cent of the total income in 1976 (the most recent year on which Kenya's income distribution data is available), while the richest top 20 per cent earned 60.4 per cent of the income. The adoption of a regressive payment (for health services) structure e.g., flat fee schedules, under these circumstances will adversely affect the poor and merely serve to worsen health inequalities between the rich and the poor—and it will be shown such inequalities already exist.

Thus although on average the proportion of government budget allocated to health in Kenya is less than (the sample) average, the per capita spending (when we take into account pri-

²¹ A simple linear function shows
 $HCE_{pc} = -15.8899 + 0.0349GNP_{pc}$
(6.704)
 $R^2=0.535,$
 $n=41.$

meaning that an increase of GNP_{pc} by \$1 will increase health spending by 3 cents only. This conforms with the general trends in these countries—where as a whole the average spending on health care is about 3% of their GDP. In the developed countries, the corresponding figure is about eight cents or 8 per cent of the GDP (Culyer, 1988b). This as shown by the equation 2.1a implies the demand for health care is income elastic—a rise in income always generates a rise in health expenditure. [What factors account for the remaining 50% of the variations in health care spending per capita in the above equations?]

vate health expenditures and purchasing power parity), given her capita income, is above the average—the line AB in Figure 2.11 above. But this does not tell us whether or not the health of the population is on average better. The per capita income is an inadequate measure of the welfare of the population as it ignores distributional aspects and says very little about the living conditions of the people. Per capita income is only but one aspect of welfare (some others being security, freedom, longevity, health status, literacy and nutrition. Health status is the most frequently used of the non-monetary indicators of a country's development performance. It can be approximated by various measures such as life expectancy at birth, or infant mortality. These measures show how the country has made use of the available resources to provide the basic necessities of life for its population (e.g., food, safe drinking water, sanitation, housing, etc.), including health care services.



Numbered countries: 1. Argentina, 2. Burkina Faso, 3. Chile, 4. Colombia, 5. Dominican Republic, 6. Ecuador, 7. Ghana, 8. Guatemala, 9. Indonesia, 10. Liberia, 11. Malaysia, 12. Panama, 13. Paraguay, 14. Philippines, 15. Thailand, 16. Tunisia, 17. Turkey, 18. Venezuela, 19. Zaire, 20. Zambia.

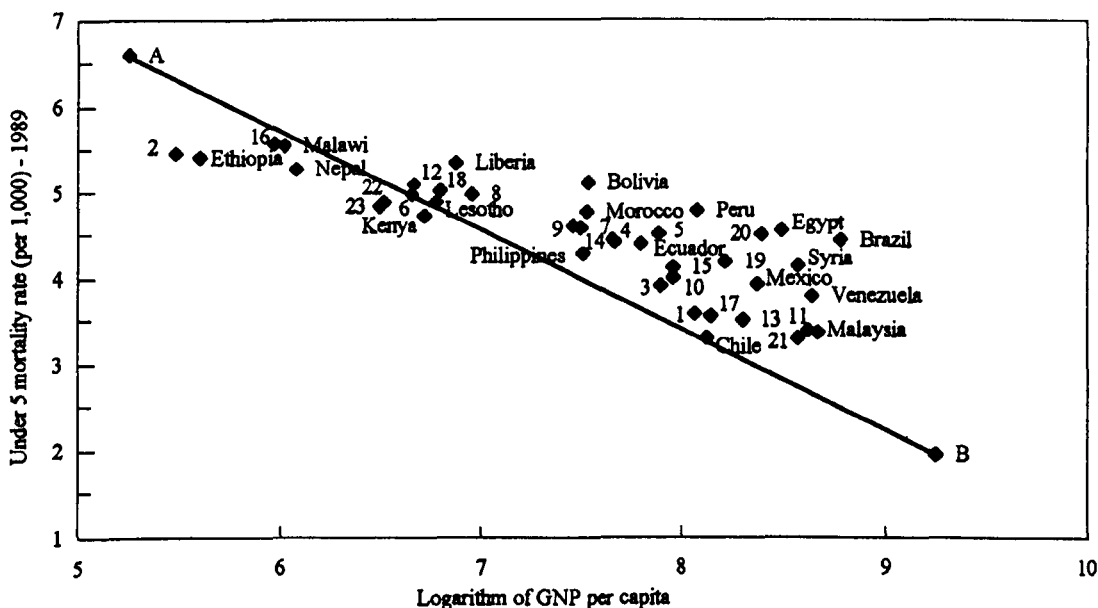
Figure 2.12: Life expectancy at birth (years) and GNP per capita in developing countries: 1989

In Figure 2.12, each country's GNP per capita is plotted against the corresponding life expectancy at birth. It is evident that countries with a high GNP per capita also are associated with a high life expectancy (for their citizens of course) at birth.

The line AB in the figure above is drawn on the same principles as equation 2.1 above and is based on the equation

$$\begin{aligned} \ln LEY_{89} = 3.1179 - 0.1314 \ln GNP_{pc} \dots\dots\dots(2.1b) \\ (12.240) \\ R^2=0.793, \\ n=41. \end{aligned}$$

where LEY_{89} is life expectancy in years at birth in 1989. The equation indicates that for every percentage increase in GNP_{pc} , life expectancy rises by 0.13%. An alternative comparative approach with similar conclusions is also evident from an examination of the relationship between the GNP per capita and the infant mortality rate—shown in Figure 2.13 and equation 2.1c below.



Numbered countries: 1 Argentina 2 Burkina Faso, 3. Colombia, 4. Dominican republic, 5. El Salvador, 6. Ghana, 7. Guatemala, 8. India, 9. Indonesia, 10. Jordan, 11. Mauritius, 12. Pakistan, 13. Panama, 14. Papua New Guinea, 15. Paraguay, 16. Sierra Leone, 17. Thailand, 18. Togo, 19. Tunisia, 20. Turkey, 21. Uruguay, 22. Zambia, 23. Zaire.

Figure 2.13: Under 5 mortality rate and GNP per capita in developing countries, 1989

The line AB is also fitted on the basis of a double logarithmic function of the form of equation 2.1. The equation used is

$$\begin{aligned} \ln USMR_{89} = 12.7237 - 1.1657 \ln GNP_{pc} \dots\dots\dots(2.1c) \\ (-9.341) \\ R^2=0.691, \\ n=41, \end{aligned}$$

where $USMR_{89}$ is the 1989 under-five (years) mortality rate. Both equations 2.1b and 2.1c and the corresponding figures indicate there is a strong correlation between longevity and mortality on the one hand and the GNP per capita, on the other. Moreover, the health status improves with improved GNP per capita. Equation 2.1c shows that a unitary increase in GNP per capita causes a more than proportionate decline in under five mortality rates, which perhaps explains why some of these countries experienced unprecedented rates of population growth. Figure 2.13 shows Kenya lies below the line AB which means given her GNP per capita, she has a lower than average under five mortality rate. The figure shows most countries with high incomes per capita lying above the average, which might indicate that it becomes increasingly difficult to lower the mortality rate as income per capita increases in these countries—but lower levels have been achieved in the developed countries—which either suggests there is something the developing countries are not getting right. But the diagrams also show there are differences even between countries with similar GNP per capita, e.g., between the Philippines and Indonesia.

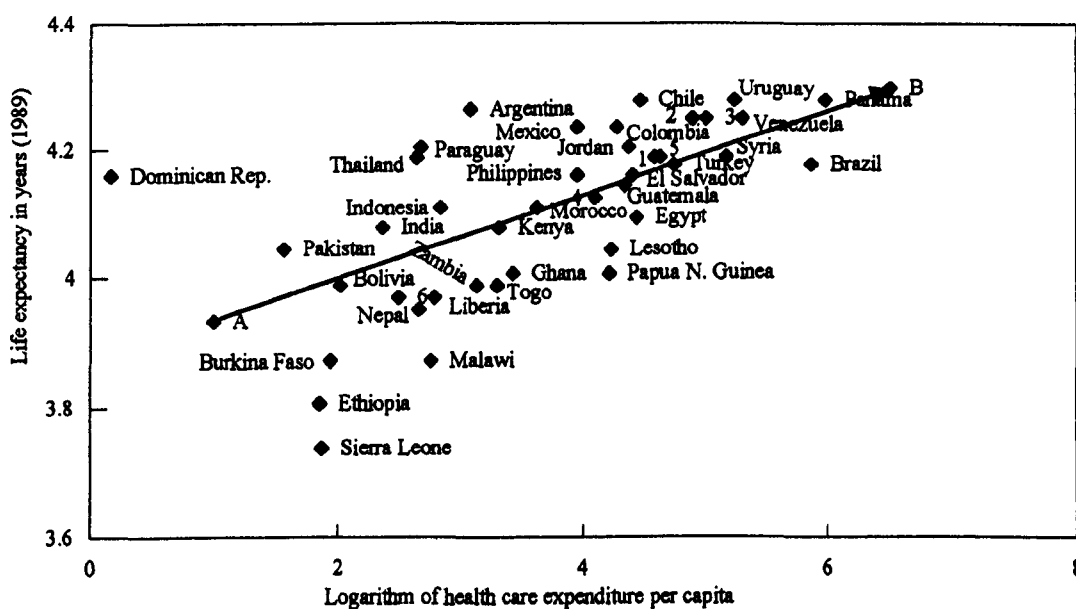
The foregoing shows there is no simple/single relationship between income growth and improved health status. The latter largely depends on the policies (directly or indirectly) affecting health pursued by individual countries. The importance of health status as a measure of a country's health policies drives most governments to play a prominent role in providing health care services and/or in the financing it. In a half of the countries in the sample used above, the government is the largest participant in the health care sector. This is because of a desire to provide a health care system that can ensure there is equal access for the citizens.

2.4.2 Conclusion

The general conclusion from the analysis presented in this sub-section is that whereas the proportion of state governments' budgets allocated to the health sector are not correlated with the level of GNP per capita, the amounts of health spending per capita themselves are correlated with health. In terms of the proportion of public resources allocated to health by the government, Kenya is not significantly different from most other developing countries, but in terms of the expenditure per capita, it is above average. The data shows a positive relation between the per capita health spending and the general health of the population (reflected in Figures 2.12 and 2.13, and explicitly depicted in Figure 2.14 below by the line labelled AB). This figure shows the countries with a higher GNP per capita at each level generally perform better—most of them lie above the line AB. In other words, high per capita health expenditure per se does not necessarily result in higher gains

in longevity—some countries spending less than Kenya have higher life expectancy for their populations, because they have a higher GNP. This suggests that the increase in GNP is a necessary (but not sufficient) condition for improvements in the health sector outcomes, as it is likely to lead to increased public (and private) health expenditures. However, as adumbrated above, the increase in the amount of resources going to the health sector is uncertain, given the low level of GNP and the poor performance of the economy in recent times.

It therefore appears that although Kenya has performed relatively well in relation to other developing countries in similar stages of development, the implementation of structural reforms (in addition to other adverse external factors) and the generally poor economic performance are likely to lead to a constant or declining real resources to the public health sector. Therefore, future increases in the output of the health sector have to come, in the main, from efficiency gains in the health sector itself. Consequently, it is imperative that means be sought of increasing the efficient use of the available public sector resources in order to attain the same objective, that is, improved health situation for the population. How this is to be done is the substance of this Thesis.



Numbered Countries: 1. El Salvador 2. Malaysia, 3. Mauritius, 4. Peru, 5. Tunisia, 6. Zaire.

Figure 2.14: Per capita public spending on health and life expectancy in developing countries, 1989

2.5 Conclusions

This chapter has analysed the current Kenyan health care system in terms of trends and the distribution of resources in relation to the country's health problems. A number of problems have been identified in relation to the public health care expenditures. Most notable among them are (i) the low level of GNP which results in low allocations to the health sector, and (ii) variations in the regional per capita allocations of public health care resources. Two tentative conclusions can be drawn from the analysis of this chapter. First, although Kenya has performed relatively well in relation to other developing countries in similar stages of development, the implementation of structural reforms (in addition to other adverse external factors) and the generally poor economic performance are likely to lead to a constant or declining real resources to the public health sector. In this environment, future increases in the 'output' of the health sector have to come, in the main, from efficiency gains in the health sector itself rather than from freshly injected resources. Second, as the analysis of the resource distribution in relation to regional health status shows, there is need to study the impact of public health expenditures upon health (in more detail than has been done in this work) as one issue for policy evaluation (this should form a priority area for future research in the Kenyan health sector). For in order to allocate public money among different regions or facilities (hospitals, health centres, and so on), it is not only the knowledge of output which may prove useful, but also knowledge regarding the productivity of different inputs as well as the interrelationships between them (e.g., substitutability). This would enable policy makers and resource allocators to identify the optimal combination of inputs and enable them to evaluate the extent to which increases in these inputs (through the expenditures on them) would increase the health status. Thus an analysis of the interplay between the health status (outcome) and health resources is an important link that must guide the allocation of public health expenditures among/between different programs. An alternative but essentially similar approach to achieving similar results will be presented in chapter four after we have looked at the factors that have possibly shaped the present distribution of health resources in the next chapter.

The following table presents the data used in section 2.4 for international comparisons.

Country	Year data refers to	Population: mid-year figures (million)	GDP ¹ per capita US\$	Health care expenditure as per cent of GE ³	Education as a per cent of government expenditure	Defence expenditure as per cent of GE	Other social services as % of GE	Per capita HCE ² \$	Life expectancy at birth 1,989	Under five mortality rate 1,989
Argentina	1989	31.93	1,887	1	10	10	40	2	71	36
Bhutan	1990	1.52	154	5	10	na	9	3	49	193
Bolivia	1990	7.4	748	2	18	14	18	3	54	165
Botswana	1990	1.3	2,803	5	21	13	18	50	64	87
Brazil	1990	150.37	3,150	7	3	4	26	72	65	85
Burkina Faso	1987	7.89	132	5	14	18	2	1	48	232
Cameroon	1989	11.54	na	3	12	7	10	7	53	150
Chile	1988	12.75	1,732	6	10	8	35	30	72	27
Colombia	1989	29.19	1,197	21	42	na ⁴	5	10	69	50
Costa Rica	1990	2.99	1,907	26	14	na	16	128	75	22
Dominican Rep.	1990	7.17	990	14	10	5	21	1	64	85
Ecuador	1990	10.78	986	11	18	13	3	16	66	80
Egypt	1989	51.8	1,809	3	13	13	26	18	60	94
El Salvador	1990	5.25	1,142	10	18	25	7	11	64	90
Ethiopia	1988	48.59	119	3	11	37	12	1	45	226
Gabon	1985	0.99	3,700	na	na	na	na	na	52	167
Ghana	1988	14.13	368	9	26	3	12	5	55	143
Guatemala	1989	8.94	939	10	20	13	8	11	63	97
Guinea-Bissau	1989	na	na	5	5	4	17	na	na	na
India	1990	827.1	338	2	2	17	7	1	59	145
Indonesia	1990	179.3	610	2	9	8	2	3	61	100
Iran	1990	14.61	3,751	8	22	10	20	na	67	64
Jordan	1990	4	1,091	na	15	21	19	20	67	55
Kenya	1990	24.03	351	5	20	10	6	5	59	111
Korea (Dem Rep)	1990	42.87	569	2	19	23	11	19	70	31
Kuwait	1988	1.96	10,696	8	14	14	22	na	73	20
Lesotho	1990	1.77	329	11	18	7	7	19	57	132
Liberia	1988	2.43	483	5	11	9	3	6	53	209
Madagascar	1990	11.2	na	8	20	9	2	3	54	179
Malawi	1989	8.02	205	7	9	5	3	4	48	258
Malaysia	1990	17.76	2,390	5	19	9	11	37	70	30
Mali	1988	7.83	na	2	9	8	3	2	45	287
Mauritius	1990	1.07	2,135	9	14	1	20	47	70	29
Mexico	1989	84.49	2,376	2	12	2	10	9	69	51
Morocco	1987	23.2	816	3	17	15	8	7	61	116

Table 2.16 (continued)

Country	Year data refers to	Population: mid-year figures (million)	GDP per capita US\$	Health care expenditure as per cent of GE	Education as a per cent of government expenditure	Defence expenditure as per cent of GE	Other social services as % of GE	Per capita HCE \$	Life expectancy at birth 1,989	Under five mortality rate 1,989
Myanmar	1988	39.97	275	5	14	19	15	2	61	91
Pakistan	1986	99.16	322	1	3	29	10	1	57	162
Panama	1990	2.42	2,045	21	17	5	24	116	72	33
Papua N. Guinea	1988	3.56	1,023	9	15	5	5	27	55	83
Paraguay	1990	4.28	1,230	4	13	13	15	5	67	61
Peru	1986	19.84	1,299	6	21	18	0	13	62	119
Philippines	1990	61.48	603	4	16	11	3	6	64	72
Sierra Leone	1990	4.15	175	10	13	10	3	2	42	261
Singapore	1990	3	12,177	5	20	24	9	na	74	12
Sri Lanka	1990	16.99	473	5	10	7	15	109	71	36
Syria	1990	12.12	2,048	2	7	32	5	8	66	62
Thailand	1990	56.08	1,372	6	19	18	6	1	66	35
Togo	1987	3.22	384	5	20	11	11	6	54	150
Trinidad & Tobago	1989	1.21	3,540	na	na	na	na	na	71	18
Tunisia	1990	8.07	1,548	6	17	6	18	32	66	66
Turkey	1990	56.1	1,935	4	19	12	4	17	65	90
U.A.Emirates	1989	1.55	17,755	7	15	44	4	na	70	31
Uruguay	1990	3.1	2,671	4	7	9	51	31	72	27
Venezuela	1990	17.53	3,474	10	20	6	13	74	70	44
Yemen	1987	7.31	576	3	19	23	2	6	51	192
Zaire	1988	33.46	265	12	25	56	19	2	53	132
Zambia	1988	7.53	361	7	7	na	4	10	54	125
Zimbabwe	1987	8.64	583	22	3	3	na	3	59	90

¹ Gross Domestic Product

² Health care expenditure

³ Government expenditure

⁴ n.a. means data not available.

Table 2.16: Data used for international comparisons

3. THE SETTING OF THE PROBLEM—II: THE PRESENT HEALTH CARE SYSTEM IN KENYA—ITS ORGANISATION AND BACKGROUND TO PRESENT PROBLEMS

3.0 Introduction

This chapter describes the administrative and organisational structure of the health care system and shows that a logical outcome of it is inefficiency and inequity in mobilisation and use of available health resources. The rest of the chapter is organised as follows. Section 3.1 outlines the administrative and organisational set-up of the public health services and its relationship with the private sector. Some stylised factors that explain the observed variations in performance within the system are then discussed in section 3.2. Since our interest is in the reform of the public health care system, we focus on the effects of the financial budget practice on efficiency motivation and its differential effect on institutional performance, according to some basic behavioural postulates because we believe (and we will illustrate this) this is one of the biggest sources of problems in the present system. Section 3.3 summarises the main implications for policy. Section 3.4 reviews previous (related) works on the Kenyan health sector with a view to identifying the main concerns identified, and assessing whether the (identified) concerns and the recommendations proffered have really been the fundamental ones and or whether they have been incentive compatible. Section 3.5 summarises the key problems faced by the current system that need redressing. Section 3.6 summarises the chapter.

3.1 Organisation of Health Services in Kenya

Kenya is divided into eight (civil/political) administrative units called provinces—Nairobi (an urban-extra provincial district), Central, Coast, Eastern, North Eastern, Nyanza, Rift Valley and Western. Each province is further subdivided into units called districts. Some provinces have more districts than others. In 1990, there were a total of 42 districts. Each district is divided into smaller political and administrative units called divisions (sometimes sub-districts). The divisions constitute the smallest political units (constituencies). For administrative purposes, however, the divisions are further subdivided into locations, the locations themselves in turn divided into the smallest administrative units called sub-locations.

The health sector has a fairly large public sector that is funded and run by the government, and a small but *not* insignificant private sector. The Ministry of Health (MOH) is responsible for designing and implementing national health policy. It has a supervisory role over the conduct of the private sector, besides its own general responsibility for health matters in Kenya.

3.1.1 Organisation of the public health sector

The activities of the MOH are organised on a central basis with four levels according to the political/administrative set-up described above. The four levels are central, provincial, district and rural. However, with the advent of District Focus for Rural Development, the district is increasingly becoming the focal point for co-ordinating health activities within each district.

The central level is the headquarters from which political, administrative and professional matters are directed and co-ordinated and policies and decisions made. The minister for health is the political head. The permanent secretary in the MOH is the head of the administrative activities of the ministry. The professional head bears the label of the Director of Medical Services (DMS).

The central government, through the MOH, is responsible for organising and administering government hospitals, which together with the health centres and dispensaries form the core of clinical facilities in Kenya. The hospital service is based on an integrated system of sub-district and district hospitals, provincial hospitals and the central consultative hospital in Nairobi—Kenyatta National Hospital—with a full range of modern facilities, including a cardiology clinic, a kidney transplant unit, a cancer research centre, a medical and a dental training school (parts of the University of Nairobi) for training medical doctors and dentists respectively. A medical college attached to the hospital trains high rank nurses. In Nairobi (the capital city), besides the headquarters of the Ministry of Health and the national referral hospital, there are other specialist institutions including the main mental hospital, an infectious diseases hospital with poliomyelitis and respiratory wings, an orthopaedic rehabilitation centre, medical research laboratories, an insect-borne diseases research institute and radiological department. The expertise focused in Nairobi is available to receive difficult or infectious cases referred there by supporting hospitals—both nationally and internationally in the eastern and central African region. The Central Medical Stores are also located in Nairobi. These supply pharmaceutical and laboratory equipment to the whole country, including a range of vaccines, some of which are exported.

There are eight provincial hospitals at Embu, Garissa, Kakamega, Kisumu, Machakos, Mombasa, Nakuru and Nyeri. These are administered by a Provincial Medical Officer of Health, and have bed capacities of between 300 and 800. These institutions are the general hospitals for their districts but also serve some specialist needs of the district and mission hospitals and private clinics. They have fully qualified medical officers of health on their staffs, and provide specialist as well general services, clinical laboratory and diagnostic X-ray services.

Below the provincial level are district and sub-district hospitals. The District Medical Officer of Health is responsible for the management of the district hospital, besides their other functions concerning the local health services. The sub-district hospitals are under the charge of a hospital assistant. They provide simple medical care.

Below the district and sub-district hospitals are health centres and dispensaries. Much of ambulatory health services are provided through a network of health centres. These are under the direct supervision of a District Medical Officer of Health, supported by a Health Inspector, a Public Health Nurse and a Clinical Officer. Their aim is to bring medical aid and forces of preventive hygiene to local people, especially those in rural villages. The health centres are mostly located in large villages or small towns. They are ordinarily designed for catchment populations of around 30,000. Health centres generally offer various services—preventive and curative—mostly adapted to local needs. Primary emphasis is directed toward the preventive and promotive aspects of health, but now there are pressures on curative services in these facilities. This pressure has been temporarily met in most areas through the construction of dispensaries.

The dispensaries, besides relieving the pressure off the health centres, fulfil an important role in providing a wide coverage for preventive health measures, which is a primary goal of the health policy. They are run by qualified nurses, although some are still under 'graded dressers' who can make elementary diagnosis and give treatments. Sometimes, mobile health units, operated by health centre staff pay regular visits to areas beyond the health centre and dispensaries' immediate sphere. Such mobile units are especially suitable for serving the nomadic peoples—the Maasai, the Turkana and the West Suk, for example. Mobile clinics are also used to take services to the people where the local population, though sedentary, is widely scattered, making it difficult for some people, given their low incomes and the poor transport network, to reach the site facilities.

In addition, Local Government Authorities, especially the Nairobi City Council and several municipal councils operate facilities at the health centre and dispensary equivalents in their respec-

tive areas of jurisdiction, to supplement the MOH facilities. They are also in charge of sanitation, public and environmental health in their localities.

The public health system, as is evident from the above, is a hierarchical referral health care system. The dispensaries are the system's line of first contact with patients (although in some areas health centres and even hospital outpatient departments are first points of such contact). A pervasive characteristic of a hierarchical referral health care system is a pyramid-like structure of health institutions, through which basic and tertiary health services are provided, in principle, to everyone. The apex of the system in Kenya is the Kenyatta National Hospital. The system permits movement of patients from the base of the national health system to its apex and vice-versa. The movement of the patients in the referral system is intended to be initiated by the health professionals who manage the public health care system.

3.1.2 The private sector of the health services in Kenya

The government health service is supplemented by privately owned/operated hospitals and clinics, mission hospitals and clinics (operated mainly by the Protestant and Catholic Missions, but a few by the Islamic community and other religious organisations). Except for the mission facilities and a few scattered private clinics, the bulk of the private sector health facilities are located in urban areas (for reasons already discussed). These consist mainly of private hospitals, private dental services, private medical and nursing services and private pharmaceutical services. Some employers also operate 'on site' clinics for their employees' needs. Occasionally, Kenyans go abroad for specialised treatments such as heart transplants (although successful transplants have been done in the country—in both private and public facilities) and brain tumour surgeries.

In the main towns there are private general hospitals that vary in size. Nairobi alone has six large private hospitals and several private maternity homes. In the rural areas, hospitals have been built by various agricultural and commercial enterprises for their employees. There were 50 private hospitals in Kenya in 1989. In addition there are other smaller institutions that provide in patient medical facilities

There were 34 mission hospitals in Kenya in 1989. They have qualified personnel—although most have been 'poached' from public sector, where salaries are generally low. Missions played an important part in introducing medical services to rural populations. But much of their services are mainly curative oriented. The standards of services in these facilities are usually high, sometimes

comparable to those of the purely for profit hospitals (often reputed for their high quality of care—although see chapter six why this interpretation need not be correct).

The number of medical doctors and other medical personnel in the private sector is not well known. It is difficult to estimate the number of doctors working in the private sector in Kenya because doctors working in the government sector engage in private practice on a considerable scale as consultants or are employed again in the private sector as consultants. Similar problems exist for other medical personnel. Long serving clinicians with considerable experience, previously barred from private practice were allowed to enter practice in 1990. This must make it even more difficult to determine the personnel in the private sector.

3.1.3 Relationship between the public and private sub-sectors

As shown above, the public and private sectors co-exist in the Kenyan health care system. The public service sector is strongly centralised, but is not supported by an efficient hierarchical network for its management—rules are executed at provincial, district and institutional levels by politically responsible bureaucrats, under a hierarchical set of formal authorisations and innocuous controls. Funding and provision are not separated (more of which below). On the other hand, the private sector is generally independent and non-integrated—with diverse providers operating almost independently, subject to (some) control by the Ministry of Health. One remarkable behavioural feature of this structure of the health system has been the ‘dual practice’ phenomenon—where some Kenyan physicians hold appointments simultaneously in the public and private sectors—especially those in urban areas. For example, a 1988 study on the health services situation in the Nairobi Area (REACH/KNH, 1988) found that about 22 per cent of doctors in private practice in Nairobi had employment outside their clinics. This phenomenon has never been fully challenged by the public health authorities and whether it is an outcome of the relative importance of the private sector or a consequence of its existence is open to debate. However, using simple economic concepts, its implications in the system can be shown to work to the disadvantage of the public health services, given the present structure. In addition to this behavioural aspect of the health care system, some other aspects of the organisation and funding of the different sectors of the system need to be considered when assessing the incentives structure in the system as a whole, before general postulates about the system incentive structure can be drawn. Focusing mainly on the funding arrangements, and taking into account the organisational structures and funding arrangements in the different sub-sectors of the system, the next section examines the incentives prevalent in the

current system, and concludes there is need for a more co-ordinated effort if the objectives of the system are to be attained.

3.2 A Synoptic Diagnosis of Problems and Their Causes in the Current System

3.2.0 Introduction

In addition to the behavioural factors alluded to above, the nature of funding of the public sector has generally not encouraged a more cost-conscious management of the resources in Kenya's public health sector. In fact, as we show below, the method presently used by the Ministry of Health (and the National Hospital Insurance Fund (NHIF)) to allocate resources to the provider units (to reimburse providers in case of the NHIF) has not provided an explicit and objective financial constraint within which questions of operational efficiency in the sector can in principle be supported.

3.2.1 The resource allocation process in Kenya's public health services system

In countries with a certain degree of decentralisation, an initial distribution in favour of regional health authorities usually precedes the final allocation although the criteria for allocation is determined in principle in a fashion consistent with the implementation of globally-set planning policies. Demographic variables—such as population in the region, morbidity and mortality factors, and so on, are mostly considered at this intermediate stage with account taken of some territorial or social inequality factors (e.g. see DHSS, 1976). However, in practice, other implicit criteria may also be present, such as those based on pure incrementalistic patterns, political lobbies, or some other amorphous criteria such as those offering a generous one-time treatment to some areas for one year and none the next.

Whenever no explicit criteria are set for the latter redistribution (i.e. allocation to regions and facilities), the previous year's expenditure, or what has already been actually spent in the current year takes on a systematic nature—one which in reality leads to some form of cost-based reimbursement. This may fail to translate the financial environment of resource limitation (scarcity) into the internal behaviour of the health providers in order to induce incentives for improving performance over time. Most of these features are found in the system of budget coverage for expenses in the Kenyan public and private segments of the health sector, particularly in the former.

The budget setting and resource allocation process in the public sector

In the Kenyan public health services sector, the roles of purchasing agent and provider are combined under one agency—the Ministry of Health (MOH). The central government provides funds from general tax revenues. The Ministry of Health, through the organisational framework already described, is responsible for allocating funds to the various functions and facilities. The budget is usually the main instrument by which decision makers plan the activities of the public health sector. The budgets of various public providers have in the past not been drawn up in a way that related resources to specific objectives—or outputs—that are to be achieved. Instead, they are designed for administrative purposes: to ensure appropriations are used in carrying out specific tasks. The current budgets are ‘item’ budgets—appropriations are allocated according to categories or items of expenditure. There are columns for actual amounts spent on major categories and the corresponding items of expenditure within each, for the previous year (called approved appropriations), and what is ‘estimated’ to be the requirements for the present year. The only innovation introduced in the planning process recently was that of a ‘forward rolling’ budget where an attempt to ‘look forward into the future’ is made. In spite of this, the impression one gets is that it merely compares the present year with the previous because even the forecast budget is based on the previous year’s budget out-turn. Thus the budget is one of budget by ‘items’ of expenditure by categories rather than by objectives.

The items of expenditure are broken down according to various functional divisions of the Ministry of Health. The main functional categories are: General administration and planning; Curative health, Preventive medicine and promotive health; Rural health services; Health training; National health insurance; and Medical supplies co-ordinating unit. These are further broken down into sub-divisions. For example, Administrative and planning expenditure is broken down into various components: Ministry of Health headquarters administrative services; Ministry of health headquarters professional administrative services; Provincial professional administrative services (further broken down by province); Medical legal services; and Planning and feasibility studies. Similarly, the Curative Health category is broken down into Kenyatta National Hospital, Provincial hospitals (further broken down by province); District hospitals (up to 1989, this appeared as a single expenditure category but presently it is broken down by district and by hospital); Psychiatric services (by institution); Non-Governmental Organisations (grants to); Spinal injury hospital (only one); Biomedical and hospital engineering; and Dental services. Within these sub-categories of expenditure the budget lists the particular items of expenditure, for example, the amount allocated for personal emoluments, different types of utilities, maintenance of buildings, purchase of equipment,

and so on. Apart from expenditures, there are entries under each category and or sub-category showing 'appropriations in aid'—income receipts that arise as a result of service 'sales', and so on. For example, under each curative health facility—say a district or a provincial hospital—appropriations in aid (in the form of 'sales and fees for services rendered', 'boarding fees', and 'other miscellaneous receipts') are indicated. This type of budget by function and expenditure item categorisation is used for both the recurrent and capital development budgets.

The available resources are not shared between regions and or facilities according to any systematic basis. Although the Ministry of Health has recognised the need for allocating resources on the basis of factors such as size, demography, sex, mortality and morbidity factors (MOH, 1992), it has yet to incorporate such considerations into the budgeting process. The revenue allocation process is usually on the basis of expected workloads implied by facilities' work plans, the sum of which determines the overall allocation to any region—not the factors mentioned above, i.e., the Kenyan public sector operates under a system of revenue allocation based on a retrospective spending. The mechanism of revenue allocation to the decentralised health care units proceeds through three main stages:

- ◆ determination of the global sum of expenditure available to the health sector. This decision is faced at the central government level and involves negotiation between the MOH and The Treasury.
- ◆ the distribution of the funds to the health care sub-sectors (such as curative versus preventive medicine, hospital versus ambulatory care, inpatient versus outpatient treatment, etc.). This type of decision is faced at the MOH and to a certain extent reflects government objectives although political interference is not unusual.
- ◆ finally, in relation to the distribution amongst the specific units charged with the provision of that particular type of health care. This type of decision is also faced at the MOH level.

In general, there are no decentralised stages for planning and allocation of health resources regionally, although in recent times, with introduction of District Focus for Rural Development, the MOH is also contemplating decentralisation. The decentralised facility network is managed under a type of bureaucratic network in a hierarchical fashion. In general, there are no explicit criteria set for the distribution of resources between hospitals. There has been reliance on the previous year's expenditure—or what has actually been spend—as the basis for current allocation. Such an allocation procedure may fail to instil the financial discipline necessary to provide incentives for improving efficiency over time into the internal behaviour of hospitals. This is because such an allocation, based on retrospective spending as it is, is nothing more than a cost-based reimbursement—where

the reimbursement is based on actual (self) reported expenses. Thus, one can argue, it ultimately results in already incurred expenditure being reimbursed.

The process of resource allocation of revenues generally begins with requests put forward by each department or facility, e.g. a hospital, in advance to the MOH. These budget forecasts are statutorily based on the probable expenditure in the current year, adjusted for cost increases anticipated from expected facility activity—in reality they are little more than ‘guesstimates.’ For each item, the forecast figure has to be accompanied by the probable expenditure in the present year. Thus, the current amounts are obviously used as the basis for an incrementalistic pattern (of the budget) in lieu of a full expenditure appraisal. In general, the forecasting exercise consists of two parts—one for the current expenditure, the other for capital expenditure. Standardised procedures for presentation of information are used, constructed according to certain administrative formalities, rather than from a specific functional design for budget evaluation principles. A set of guidelines is provided annually for this purpose, providing codes, definitions, and delimiting the scope of items included, often with an indication of priority areas, particularly in investments and equipment acquisition. Manpower expenses are determined according to staffing levels as approved by the Ministry of Health. The rest of the budgetary items—medical and non-medical inputs—are often left to the institution’s discretion. In the hospital sector, there is no formal separation of outpatient and inpatient services—this is also left to the institutions to establish the relative sizes. Consequently, little is hardly known about the use of resources in the respective activities since there is a lot of resource sharing. Besides, budgets are often not linked to global growth of the economy and are rarely adjusted for inflation (see section 3.4 below). The resulting ‘bids’ form the basis for incrementalistic pattern of allocation instead of a full expenditure reappraisal. The rural health services’ budgets are residually determined—i.e., receive whatever is left after curative, administrative budgets, etc. have been completed.

Some weaknesses in the budget setting and resource allocation process

The type of budgeting process described above, as a tool of planning, cannot in its present form assist the ministry officials in deciding how to allocate scarce resources efficiently and equitably (these concepts are discussed in the following chapters) among the various goals and or activities. The budgeting process has various shortcomings. First, it is not output oriented. There is hardly any economic or epidemiological appraisal. Each sub-category (of allocation) is limited to pointing out the differences between, say, this year’s requests as compared with the previous year’s

amount. This shows the budget requests are done mainly on an incremental basis—last year plus a bit more. This has the implicit assumption that the current level of expenditure (say, for a particular facility or even on an input) is ‘the correct’ level of expenditure and what is needed are only marginal changes to incorporate cost factors that are likely to affect either the demand for or supply of health care services, such as expected increase in caseloads, inflation, or medical advances and technological changes that have cost implications, and so on. Moreover, the allocation is generally on the basis of the existing health sector infrastructure. This may also fail to account for inequality in the current distribution of resources. While the total requirements of, say, the hospitals are given, there is no indication of why a particular level of doctors, nurses, and so on should be employed in a particular institution—because the technical relationship between inputs and outputs is hardly known. The setting of staffing norms is therefore an arbitrary exercise, more based on convenience than technical considerations.

Second, under the current budgeting procedures, institutions are simply reimbursed all expenses at full cost, irrespective of their contribution or otherwise to the objectives of the health care system. Moreover, since reimbursement is on the basis of past costs, and given that any savings made by an institution during the budget period revert to The Treasury—any unspent allocations at the end of the accounting period are automatically returned—institutions (facilities) have incentives to spend all the allocated appropriations. To do this may involve encouraging activity as much as possible just to ensure all allocations are fully accounted for. This may have adverse effects on efficiency since most activities undertaken have not been evaluated for their worth—neither to the patient nor to the objectives of the system as a whole. Thus the present budget practices penalise facilities for being efficient and ‘reward’ inefficient facilities.

Another confounding factor is that doctors and other health personnel are paid a salary that is unrelated to performance, one instead determined mainly by grade, length of service, and so on. Such wages are in principle ‘attached to jobs rather than workers’ without account to the external earnings levels in the private sector²², or the worker’s actual level of performance. A complex set of administrative procedures—such as ‘job-groups’ that are mostly dependent on seniority and experience in the bureaucratic hierarchy, together with some consideration of formal qualifications—determine the level of wages. If we assume significant job-skills are acquired simply by experience on the job, this (wage hierarchy set-up) in theory reduces potential bargaining costs, particularly in

²² *The Weekly Review* (June 3, 1994), reporting on a dispute between the government employed doctors and their employer, states that in the private sector doctors earn four times as much as their counterparts in the public sector (*Weekly Review*, June 3, 1994, *The Medical Profession: Doctors Issue Strike Threat*, pp. 19–20).

sectors with high informational asymmetries (as in the health sector). But this provides a constant perverse incentive environment, since it is actually based on job promotion for life—which can conceivably be achieved simply by giving a ‘modest effort’ while the passage of time guarantees promotion into the next rung up the wage ladder. In such cases, problems of ‘shirking’ could be expected to abound, and thus inefficiency. This perhaps explains the general lack of incentives for efficient performance over time, and the continued haemorrhage of the public sector health personnel—aspects much decried by the authorities in the Kenyan public health care sector. The government has not recognised the need to provide incentives (monetary and non-monetary) for doctors—and other health personnel—to be efficient in their resource allocation. The financial and non-financial incentives that exist are not meant to encourage efficient behaviour (but are given as inducements to instead serve in isolated, usually small urban towns in arid areas—but the distribution of health personnel still shows these are the areas with the greatest number of unfilled vacancies according to the Ministry of Health’s own staffing norms).

Another important outcome of the allocation process discussed above is that instead of correcting inefficiencies in resource distribution, it serves to entrench them since the (present) allocation is tied to existing institutional capital infrastructure. Consequently, much reliance has been made of the capital budget to improve health infrastructure capital distribution, but it appears the revenue budget dominates the process, since capital projects can be postponed. As noted elsewhere, even this approach has been prone to errors, e.g., when the government constructed a facility in an area already served either by purely for profit or non-governmental facilities.

Finally, the absence of a more complete budget appraisal and expenditure evaluation policy on which to link those restraints has led to problems of inefficiency and inequity in the health sector since often the allocations may depend on irrelevant factors such as the negotiating abilities of facility managers. Personal connections may also play a role in determining the outcomes of budget allocations. Moreover, the approval of the budget is not necessarily the final constraint according to which activities are to be carried out. Budget overruns are common²³. No specific pattern can be demonstrated to exist in the sequential process of budget submission, approval and execution, with respect to their initial and final outcomes, in terms of the level of reduction, degree of overrun, and the future use of any of the former budget bases for new planned expenditure. It appears there is an

²³ Sometimes the MOH resorts to arbitrary ‘across-the-board’ sweeping cuts in the budget proposals. This seems an exercise in futility because before the year is over, most hospital managers (or heads of other public health departments/facilities) will be back in the MOH corridors presenting requests for supplementary funding to defray deficits and, unfortunately or fortunately, most requests are fulfilled—even if only partially. Thus, in general, the approval of the budget is not, strictly speaking, the final expression of the financial constraint according to which activity has to be carried out.

implicit assumption that financial restraint policies can encourage efficiency. This appears to stem from the assumption that the resulting imposed cost-savings imply parallel improvements in resource utilisation. Whether such policies can achieve improvements in efficiency is a debatable matter. First, what is to be understood by efficiency in this context is not clear. If it translates into 'doing the same but cheaper', it needs to be proven that the same outcome is actually achieved and it would be necessary to account for variations in circumstances from which this result can plausibly be expected. If the same refers to what someone else does, it needs to be proven that similar circumstances prevail in the comparison, especially if from the alternative providers we expect identical performance. Second, the expenditure restraint just serves to validate past levels of expenditures, since they (the past levels) are used as bases for future projections. We have seen that those bases are the result of the factors such as the budget revenue allocation grounded in the current retrospective system. Therefore, this is a weak basis for expenditure rationalisation, since it incorporates the consequences of a perverse set of institutional arrangements. Past levels of expenditure cannot provide a sound basis for an expenditure rationalisation projected into the future. Third, the implementation of general cash limits with across-the-board increases/cuts tend to deal with potential 'unequals' as 'equals'. In this case, for example, the haphazard increases penalise previously cost-conscious spenders but rewards past profligacy—leaving the latter (higher spenders) relatively better-off after implementation of the policy, given the indifferent basis of application of these general increases. As a result, the policy introduces questions of comparative justice and certain dislocation effects that encourage some gaming strategic behaviour (on the part of providers) in budgeting in order that the initially prospective financial constraints can be evaded.

Thus, neither the reimbursement nor the expenditure restraint policies provide any effective incentive motivation for cost-conscious management of health resources. 'Across-the-board' increases based on self-reported costs in a context of inadequate monitoring and expenditure evaluation cannot intrinsically contribute to the expenditure rationalisation targets required if a serious and constant effort for improvements in the allocation of public resources is to be pursued in the health sector. Moreover, this policy introduces a set of distortion against a more rational basis for development of health sector activity.

The need for change

The short list of some of the major problems inherent within the mechanism of resource allocation in the public health services discussed above shows the system, in its present form, cannot

substantially assist officials in deciding how to allocate scarce resources to meet the objectives of efficiency and equity in the health care system—there is a lack of appropriate incentives. In short, there is need for establishing a more solid basis on which a consistent and rational evaluation of policy should be permanently grounded. In the absence of a complete reform of the budgeting methods, system budget overruns are common, and under the influence of facility administrators (particularly hospital administrators) current levels of expenditure are often used to adjust budgets where such overruns have occurred. In theory, in order to prevent such outcomes, facility administrators are ordered to follow budget execution rules strictly (with resources being released in tranches periodically). But if, under ‘exceptional circumstances’ problems of over-spending arise, facilities are allowed to request for additional funding. Though theoretically no presumption exists for automatic approval, the process is nearly a formality for dealing with budget overruns without appearing to make the exercise simple. Lobbies and political pressures may be exerted for the full coverage of the overruns, but the final outcome often depends on subjective elements such as personal abilities of facility administrators to achieve additional funds, and so on.

In conclusion, the system operates in practice as a retrospective reimbursement. In this context, we cannot expect any general motivation for a cost-conscious management other than that which is *ad hoc*. The question now is: what alternative system or framework or modifications of the present system can be put in place to ensure that health care resources are used efficiently without at the same time abandoning that cherished concept of equity? The efficiency implications of the process of resource allocation, and the environment of their application, will be analysed in the next chapter. According to the theory to be developed in chapter six, there are a number of alternatives that might be tried/adopted. These alternatives will be discussed in chapter seven.

3.2.2 The private sector

Growth of the private sector has both positive and negative consequences for the health system. It is for example claimed that the better-off patients will use private hospitals and thus free public resources to serve the poor (Akin *et al.*, 1987). On the other hand, it widens the gap of inequality. Given that the state aims towards equitable health care services provision, but the private sector providers are often driven by the desire to maximise profits, there may a divergence between the objectives of the state and those of the private sector providers. Private physicians may for example over-prescribe care in order to generate higher incomes for themselves or their organisations, which is costly both to the individual and society, and may have adverse effects on state

objective of increased access to care. The concern with profits may lead the private sector private sector providers to behave in ways considered unethical. In 1992, there was a general public uproar when one private hospital was reported in the local media to be refusing treatment to accident victims because payment was not assured. There has also been report of another hospital that refused to put a renal patient on dialysis because 'enough' deposit had not been paid. Private providers may also promote expensive treatments and high cost procedures or high technology equipment even in the absence of evidence about their effectiveness. Besides an absence of concern for public health services, the private sector providers, particularly those in poorer areas, may cut costs by providing low quality care, or by relying on unqualified, less costly personnel. In other cases, the motivation for profits also may mean that the private sector providers are unlikely to engage in public health care activities (as shown in chapter four). This means they fail to promote preventive practices which reduce morbidity and the resultant number of consultations. Private sector providers could play an increased role in the early detection of disease and prevention *if encouraged* to use screening procedures for groups at risk of specific diseases (where efficacy has been demonstrated), and to engage in immunisations.

The growth of the private sector also attracts trained personnel from the public sector, often at a considerable social cost and may thus undermine the public services. This problem is particularly acute in Kenya. The large differences of earnings that exist between the public and private sector encourages trained to leave the public sector to work full time in the private sector, so creating skilled personnel shortages in the government facilities. For example, the Nairobi Area Study (REACH/KNH, 1988) survey data showed that about 68 per cent of privately practising doctors in Nairobi were previously public sector employees. In particular, 52 per cent of them were previously working for the Ministry of Health, 10 per cent from the University and 7 per cent from the Nairobi City Council. Some private sector doctors also work in public institutions (mainly as consultants). It is conceivable that their private practice receives priority over government sector work (again due to remuneration differences), in addition to giving them an opportunity to solicit for patients from the public sector. The study quoted above concluded that

There is need to determine the effect that the rapid expansion of the private sector in Nairobi since independence has had on the structure of the sector. It is important to determine, for example, whether expansion has made the private sector more competitive and what has happened to the costs. Twenty years from now a large proportion of the private doctors will be nearing retirement. Government should have a policy of replacing retired doctors in the private sector to ensure that medical care in this sector is not disrupted by the natural retirement of private sector doctors (page 56).

The wisdom of this advice is questionable. While it is a stated government objective to encourage private sector expansion in order to ease the burden on the public services, it is not entirely up to the government to provide training for all private sector requirements. Besides, if the competition mentioned is a reality, the problem of skilled personnel haemorrhage to the private sector will continue and there would be no need for such a policy. The important issue is how this affects the effectiveness of the public health services.

As the state has a role in ensuring safe and appropriate health service provision to the population, it must have mechanisms through which to liaise with private sector providers. In considering ways to improve service provision to the whole population, there is need to analyse the interrelationships between the government and the private sector, in particular, the actual and potential mechanisms for improving this relationship. A conceptual framework to assist towards this end will be developed in chapter six, where an attempt will be made to outline the key areas of interaction, the concerns and relationships between them, after development of a theoretical framework that will place this relationship in context.

3.3 The Overall Picture and Implications for Policy

The analysis up to this point was an overview of the Kenyan health care system—its organisation, administration, and the distribution of financial and physical resources and personnel—and the general health situation. It is now apparent that Kenya inherited an inherently inefficient and inequitable system. The survival, to a great extent, of the unequal pattern of regional provision per head of crude population—and the persistent unequal burden of health risk—after more than 30 years of public intervention can however be heuristically explained.

The separate development of the different segments of the health care system that had occurred before independence persisted after independence largely because of an absence of coordinated or explicit national health services planning policy. Two factors were largely responsible for the apparent lack of any coherent planning even in the face of a government that was committed to elimination of inequalities as early as 1965. The first is the quality and the distribution of existing capital stock at independence, and the second is the revenue costs of the previous systems.

The hospital capital existing at independence was largely unevenly distributed, some of it aged. Despite the effort by the government to alter the distribution, substantial capital was required—both for improvement and construction. Taken together, these would have required a large

capital budget to remedy them. But even more noteworthy, there was no explicit criteria for provision and this led to duplication of effort. It has not been unusual to find a government, private or mission facility being constructed in an area where another facility of a similar type exists due to this uncoordination in the system.

Moreover, this uncoordinated planning also extended to the financing side. While private health insurance was not widespread and that which existed catered only for specific groups of the population, the introduction of public sector insurance (the National Hospital Insurance Fund) put a check to its growth. But the legal imposition of a flat contribution rate restricted the development of this particular public institution and no attempt was made to cater for those in the rural areas—largely because the scheme, like the private sector insurance—focused on the working group, most of them based in urban areas. Consequently, capital development on any broad scale was precluded by unavailability of adequate funds. Moreover, for a newly independent country, there were many other equally pressing areas requiring urgent capital infrastructure—e.g. industrial and social capital was necessary—and the development of the health care infrastructure had to compete with other sectors. Those sectors politically favoured, e.g. education, naturally won. Besides inefficiencies in resource allocation, a number of other factors further impeded the development of the health sector. Notable among them is the efficiency in resource use in the public sector. Usually, the allocation is not related to expected health status improvements—because as shown in the description of hospital sector resource allocation, no systematic methodology is used. The productivity of resources is therefore hardly known. There is an absence of specific norms for provision of various types of services and consequently, the concentration of capital continues to be urban biased. Thus there is need to study the impact of these public expenditures upon health as one basic issue for policy evaluation. The failure of the ministry of health to drive any substantial inroads into the pattern of inequality existing at independence largely derives from the method of funding used to allocate budgets. Facilities and regions (districts and provinces) receive a budget based on the previous annual budget with allowances for new capital or services anticipated in the coming year and for inflation and growth—when cuts were necessary, frequently the ‘victims’ have been capital projects. The other main problem, as hinted above, has been poor economic performance.

Another problem has emanated from reliance on donors—who mostly constructed facilities but did not equip or staff them. In most cases the MOH failed to incorporate the effects of such facility developments in its budget—so the budget provision for these new facilities was inadequate—this further constrained their efficiency and the effectiveness of the health care services they provided. The reliance on ‘harambee’²⁴ effort had similar efforts—facilities were constructed under

communal efforts that could not be operated effectively/efficiently. Lack of policy guidelines on the donor community/harambee meant there was failure to apply new capital wholly to the benefit of the worst-off regions. This continued to perpetuate inequalities, besides compromising efficiency. In order to allocate budgets for public money among different regions and or facilities (hospitals, health centres and so on), it is not only the knowledge of the output that may prove useful, but also the knowledge about the productivity of various inputs as well as the interrelationships between them (substitutability) enables us to identify the optimal combination of inputs and enables us to evaluate the extent to which increases in these inputs would increase the health status. There has to be criteria for allocating resources that takes into account the inter-relationships between the health status, health resources and technology if we are to achieve an effective and equally efficient health care system. This is the main task this thesis undertakes—the development of mainly theoretic constructs for optimal resource allocation. We subsequently examine the theoretically appropriate patterns of regional provision and policy attempts to redistribute resources. However, the fact that there is unequal per capita provision does not of itself demonstrate an inappropriate distribution—such an assumption would beg the question of what comprises an appropriate distribution. These issues are taken up in Chapter five. But we first review some works on the Kenyan health sector in order to outline what are currently thought as the problems that need redressing.

3.4 A Review of Related Previous Works on the Kenyan Health Care System

Studies done prior to (and during) the implementation of the initial financing reforms looked at various aspects of the Kenyan health care system, including (i) issues of efficiency, principally from a managerial point of view, i.e., identifying how more output could be achieved from the resources available to the Ministry of Health, (ii) estimating the level of additional resources needed to supplement the allocations to the public services from The Treasury, (iii) identifying options for mobilising resources in the private sector towards this end, and (iv) simulating the likely impact of introducing cost-sharing in the public sector. The purpose of this review is to collate these studies so as to identify the main concerns raised, the recommendations proffered and assess whether those issues were really the fundamental ones and whether the changes/reforms proffered were incentive compatible.

²⁴ 'Harambee' is a Kiswahili word for an informal 'communal effort' directed at pooling resources for a common goal—where members contribute according to their self-assessed abilities. The system has successfully been used to raise funds for construction of health facilities, particularly dispensaries and health centres; to finance medical expenses for catastrophic illnesses for some individuals; assist bereaved families; to sponsor students for studies; and so on.

The studies carried out on Kenyatta National Hospital (REACH/KNH, 1988), the Nairobi Area Study (REACH/MOH, 1988) and the Provincial and District Health Services (PADS) Study (REACH/MOH, 1990) focused on management efficiency improvements. In 1987, Kenyatta National Hospital (KNH), through a presidential order of April 1987 was granted independent status, governed by a Board of Directors. This order and the empowering act of parliament gave KNH the freedom to adjust its operations in order to achieve greater efficiency in service provision and raise funds from sources other than the Ministry of Health (MOH). The hospital initiated a consultancy study (REACH/KNH, 1988) to identify ways to improve management and efficiency, and consider cost-sharing for selected services. The study analysed virtually every aspect of the hospital's operations in order to prepare it for the new status. One of the main findings of this report was that management was too over-centralised to achieve effective control, with most departmental managers reporting directly to the Director, although there existed a layer of senior management to whom these managers should have been reporting. The established lines of authority were thus not being followed. The resulting organisational structure—where appropriate lines of authority are not followed—'short-circuits' the 'established' decision making framework, often overloading the top layer of management (the Director) with 'trivial issues' rather than concentrating on 'strategic issues'. Other major problems identified included lack of proper financial and physical planning in many spheres of operations that often led to increased costs, lack of service co-ordination resulting in shortages of essential supplies and under-utilisation of costly services of operations, lack of adequate information to carry out the above, and general under-funding. Some selected 'low level' inefficiencies identified based on comparisons are summarised in the Table 3.1 below.

The study made several recommendations, including charging for drugs and diagnostic services, introducing fees for inpatient and outpatient services, and strategies for generating additional revenues besides fees, such as seeking contracts with employers or making arrangements with insurers and other third party payers to provide services to their clients, and to raise additional revenues through establishment of an 'amenity' ward where doctors would pay 'rent' for using the hospital's highly specialised facilities to treat their private patients. Following the findings of this study, the World Bank proposed and co-funded a project designed to introduce a series of changes covering strategic planning, changes in organisational management, including changes accounting, financial, materials, pharmaceuticals and maintenance protocols that would increase efficiency and quality of services provided in every department of the hospital.

	Abdominal Surgery Discharges		Malaria Discharges		Pneumonia Discharges		Normal Deliveries	
	KNH ¹	AKH ¹	KNH	AKH	KNH	AKH	KNH	AKH
Average length of stay (days)	8.7	7	3.9	3.4	5.2	4.6	2.4	2.9
Range of stay (days)	1-47	1-47	1-15	1-8	1-41	2-12	1-26	1-5
Laboratory requests (per patient)	1.8	5	2.9	6.5	1.2	3.2	0.3	1.9
Per cent of completed laboratory requests	—	—	68	100	67	100	19	98
Radiology requests (per patient)	0.4	0.6	0.04	—	0.3	1	0.3	2.1
Completed radiology request (per patient)	85	94	0	—	100	93	—	—
Other procedures requested (per patient) ²	2.2	2.5	0.7	—	0.3	—	0.6	0.9
Percent completed of other procedures requested (per patient)	98	100	86	—	67	100	71	82
medications prescribed (total) ³	31	69	39	32	27	20	13	31
Medications (per patient)	4.9	10.3	3.4	4.8	3	4	1.4	4

¹ KNH: Kenyatta national Hospital; AKH: Aga Khan Hospital. For each condition, random samples were selected from the hospital records. These varied in size between hospitals and conditions.

² These vary from blood transfusions, parasitic cultures, and so on.

³ Number of different medications prescribed.

Source: Compiled from REACH/KNH, 1988, Exhibits III.B.1-1 — III.B.1-4.

Table 3.1: Comparative clinical efficiency indicators, KNH and AKH, 1986

The Nairobi Area Study (REACH/MOH, 1988) was launched mainly to examine ways of strengthening services provision in the City of Nairobi in view of the changed status of the Kenyatta National Hospital which was previously available to most City residents, and partly to examine efficiency of services provision within the city public facilities in order to recommend ways of reducing the costs of producing them and devise ways of expanding both the public and private sector capacity to cope with workload diverted from Kenyatta as well as with future levels of demand. The study analysed the Nairobi health delivery system, the cost and utilisation of services' both in private and public facilities, the scope for efficiency improvements and general policy, planning and management issues. The Study concentrated on three areas (i) outpatient services, (ii) preventive and promotive services, and (iii) maternity and mental health services. The study identifies twelve problem areas which were classified into priority levels:

(i) First priority group:

- uneven distribution of utilisation,
- inappropriate/inadequate financing mechanisms,
- inefficient use of resources,
- inadequate supply of resources,
- inadequate policy guidelines for operation and co-operation,

- lack of mission clarity for facilities and systems,
- inadequate communication/coordination,

(ii) Second priority group:

- inappropriate/inequitable distribution of resources,
- lack of management skills,
- lack of evaluation of inputs and outputs,
- limited information for planning,
- non-operational information reporting systems.

The study used a random sampling design and the sample (154 facilities) indicated that roughly 64 per cent of the facilities in 1987 belonged to the public sector—Nairobi City Council, the MOH and parastatals, while 13 per cent were mission facilities and the rest private. This high proportion of public facilities is indicative of the general situation in the country, but, in Nairobi it should have been lower—had private clinics and surgeries been included. The study found considerable unexploited potential for increasing efficiency and the effectiveness of curative services. The most acute problems related to inefficient use of existing physical capacity, ineffective deployment of staff, unco-ordinated flow of patients between service levels, inadequate administrative and management systems, and lack of accountability to organise and co-ordinate City-wide health services. Among some notable findings was the low inpatient bed occupancy rates for all types of facilities—public as well as private—generally under 90 per cent, except for two public facilities, the spinal injury hospital, which had a 92 per cent occupancy rate, and the main maternity hospital, with a 107 per cent occupancy rate. The private facilities had generally higher occupancy rates. However, for maternity beds, the opposite was true, with occupancy rates generally higher for public facilities. The public sector operated over 80 per cent of all beds in the city but handled approximately 70 per cent of all admissions, while. For maternity beds, the rates varied from 26 to 129 per cent. Most of the public facilities had rates around 100 per cent or more—probably indicative of a pricing structure which makes private sector facilities unavailable for maternity cases. But this could mainly be attributed to the fact that maternity cases are not covered by private insurance, as most private sector patients use this means of reimbursement. In terms of outpatient workload distribution between the sectors, the public sector handled respectively 84 per cent of all maternity cases and outpatient services. In terms of service delivery efficiency, large variations existed not only between the sectors, but also between various units within each sector. There were more broken equipment in the public service facilities also. Personnel appeared more under-utilised in the public sector, and so on. The Ministry's response to these problems was to develop a planning process that would result in a strategic plan for providing health services in the Nairobi area. Later it was decided to produce such a plan for the whole country. Consequently, the Provincial and District

Health Services Study, PADS, (REACH/MOH, 1990), a series of sub-studies, was launched, to examine preventive and primary health care, user fees implementation and district health services.

Government considers primary and preventive services an efficient and cost-effective means of providing health care to the population. Shortages of drugs, supplies, transport, insufficient staff, and other key resources are the main factors constraining the facilities providing these services. The objective of the PADS preventive and primary health care (P/PHC) gap study (MOH, 1990) was estimate the additional financial resources required to operate the P/PHC system at full capacity. The study examined the efficiency of existing operations, assuming location, mix of P/PHC services to be optimal. This is a major weakness of the study since access factors differ significantly between regions (see below), indicating that the current locational matrix of facilities is not optimal. There is need to take into account differences in supply and demand factors. The study surveyed approximately five per cent of the MOH facilities and identified two major constraining gaps—one in recurrent expenditures needed to provide services at full capacity, and the other a one-time investment expenditure required to bring the already existing facilities to full capacity. A summary of the annual financing gap by resource category is shown in Table 3.2 below.

Category	Current expenditure (Ks m)	Expenditure gap (Ks m)	Gap as a per cent of expenditure
Drugs	275.5	79.1	29
Equipment	6.3	7.2	114
Transport	19.3	17.3	90
Training	42.2	5.1	12
Supplies	41.1	61.3	149
Patient food	1.7	6.0	352
Maintenance	3.3	83.3	2,524
Staff	765.5	170.5	22
Total	1,156.8	429.8	37

Source: Compiled from data extracted from various tables, in MOH, 1990.

Table 3.2: Annual financing gaps by resource category, P/PHC Services, 1990

The total annual recurrent expenditure gap was estimated at approximately Ks 430 million (then Ks 23=US \$1) or 37 per cent of the recurrent expenditures for P/PHC. In terms of categories, the largest (absolute) gaps were for personnel and drugs, though they are relatively small in percentage terms. Building maintenance and patient food expenditure gaps, though small in mone-

tary terms, were the most under-funded items, as shown by the large percentage gaps. These gaps cannot be viewed in isolation as one item acts as a constraint on one or more others. For example, transport deficiency constraints personnel movement, food, drug and patient transport, and so on.

The one-time investment gap, showing the level of investment expenditure required to upgrade equipment, vehicles, staff and space to the full capacity requirements was estimated at Ks 375m (US \$16.3 m), which was an under-estimate as it excluded building and equipment rehabilitation costs.

Table 3.3 below indicates that rural health facilities had larger gaps in most resource categories, except drugs. These facilities therefore experienced relatively larger expenditure deficits than hospital out-patient departments (OPD). If account is taken of volumes of workloads, the rural health facilities are in a worse situation since, since although hospital OPDs provided about 10,000 consultations per month, while the health centres' and dispensaries' averages were 3,000 and 2,000 respectively, the large number of health centres and dispensaries means dispensaries and health centres provided more than 60 and 30 per cent of all out-patient consultations, respectively. Hospital OPDs provide less than 10 per cent of the consultations.

Type of facility/category of expenditure	Current expenditure (Ks m)	Expenditure gap (Ks. m)
RURAL HEALTH FACILITIES		
Drugs	196.7	43.9
maintenance	2.1	73
Personnel	309	81
Supplies	13.3	37
Total	521.1	234.9
Gap as per cent of current spending	—	44
HOSPITAL OUT-PATIENT DEPARTMENTS		
Drugs	79	35
maintenance	1.2	10.3
Personnel	456.5	89.5
Supplies	14.6	23.7
Total	551.3	158.5
Gap as per cent of current spending	—	29

Source: Compiled from various tables in MOH, 1990.

Table 3.3: Gaps in rural health facilities and hospital outpatient departments

Thus there is evidence here, though based on a study whose methodology is shaky, that the MOH should reform its budget allocation from hospital based curative services in favour of rural based health services. However, there is need for better a designed study to establish marginal productivity of resources in order to put resources in those areas with greatest returns, taking into account factors such as local population structure, health needs, as well as other sources of service provision in the localities. Other implications of the study are that Government should not approach the problem piece-wise. This will not solve the underlying (complex) relationships between the various categories.

Another PADS study, the PADS Nakuru District Health Services Study (REACH/MOH, 1990), attempted a comparison of private and public providers in a sample of District facilities. The Study does not explain why or how Nakuru District was chosen as the study area and it is not clear whether the study results could be generalised. Moreover, there is no explanation of the sampling design used for identifying facilities and patients included in samples. These are major flaws which should be borne in mind as we discuss the study.

The study identified four problem areas in the delivery of public health services: planning and budgeting inadequacies, production inefficiencies, lack of cost containment measures, and inadequate attention to quality of care. In public facilities, there were inadequate allocations of financial and physical resources to areas with chronic shortages such as drugs, supplies, transport, equipment and building maintenance, and so on. This study also found the public facilities had a worse record of equipment maintenance, which lead to increased service costs as it led to referrals to higher cost facilities, or increased lengths of stay. Patients were often admitted several days prior to operations, leading to longer stays, and inefficient use of staff time, in addition to other resources. In addition, there were no standard treatment protocols for common conditions, which results in large variations in costs for similar conditions in comparable facilities. The study also suggested charges would induce patients to leave hospital sooner than later—cases were documented of patients remaining in public hospitals for several days after being officially discharged and the need for cost-monitoring. It also suggested the MOH should invest in more capital, labour and medical and other supplies in lower level facilities in order to ease the burden on the higher level (that are also higher cost) facilities. Concerning quality of care, it was noted that most patients were dissatisfied. In addition, it was found that the volume of out-patient services as well as the number of new curative cases had declined over the previous two years, prompting the study to conclude that a smaller proportion of the population was receiving MOH services than two years previously, a phenome-

non interpreted to mean that unless the overall health of the people had improved substantially (unlikely), people were receiving services from other providers, probably paying to obtain them.

The decision to implement cost sharing in the health sector was a logical reaction to pressures from many concerns as noted above. From government's policy statements, the main objective of cost sharing was to generate additional resources, although other desirable objectives were also in sight. The introduction of cost sharing was based on an agreement between the Kenya Government and the World Bank that the revenues so generated were in addition to, and not substitution for government funding to the MOH. Such revenue, if unspent at the end of the year was not to be returned to The Treasury as Government accounting procedures required. Thus cost-sharing revenue was not to feature in the MOH Appropriation-in-Aid account. Now, whether cost-sharing would lead to higher revenues depended on several factors, including the response of consumers to the charges (prices), the cost of the collection system, and incentives for collection. In addition to all the studies reviewed above which included surveys of patient willingness to pay and perceived quality of services in public facilities, there were also other studies on the likely effects of fees (Mwabu, *et al.*, 1989; Kirigia, *et al.* 1989; Ellis *et al.*, 1990). These studies indicated patients were generally willing to pay for services if that would improve their quality.

An *ad-hoc* study by ministry officials (MOH, 1991) evaluated the problems likely to confront cost-sharing implementation—fees administration, revenue collection procedures, incentive structure and impact on utilisation. It turned out that institutional and management systems were not quickly restructured to accommodate the implementation of user fees. The existing accounting, security and receipting mechanisms were inadequate to handle the volume of anticipated revenues. Moreover, many aspects of the fee program were not fully understood by patients and staff—including the waiver fees, which services fees were to be paid for, and so on. The process of introducing the user-fees itself was fraught with difficulties and (political) uncertainties, and consequently it was a stop-go introduction—but despite the negative attitude towards cost sharing within the administration, Kenya introduced charges for public health services beginning November 1989, although the take-off was never smooth (Dahlgren, 1990). There was evidence after user-fees implementation that attendance declined and the revenues collected were not as much as previously anticipated (Mwabu, *et al.*, 1991; MOH, 1991). Data indicated there was about 38 per cent drop in the utilisation of services. Mwabu *et al.*, (*op. cit.*) summarise the outcome for the period between November 1, 1989 and September 1990 when fees were operational²⁵, thus:

²⁵ Concerns about the effect of the new fees on specific patient groups first led to reductions in maternity fees on January 23, 1990, exemption of civil servants on April 17, 1990, and finally to suspension of outpatient consultation fee on September 1, 1990, but other fees, e.g., inpatient fees, maternity

The outpatient services are highly inelastic with respect to user charges. Nonetheless, cost-sharing led to a reduction of up to 38 per cent in demand for outpatient services as a result of the large percentage increase by which outpatient fees were adjusted. Consistent with this finding, suspension of registration fees is associated with a remarkable increase in service use. This increase occurred despite the retention of some other fees like the laboratory and X-ray fees. By March 1991, eight months after the removal of registration fees, service demand had recovered substantially and was only 8.5 per cent below its original level. Patients are more sensitive to laboratory fees than to registration fees; laboratory fee elasticity of demand is about 5 times larger than that of the registration fees. User charges compress health care demand. Their implementation or upward adjustment therefore should be preceded by investments that compensate for this shortcoming.

The result, taken in context of the pre-user fee situation shows what patients perceive as possible benefits of fees *changes* when confronted with actual fee demands. As a result of this unanticipated behaviour, the revenues generated were far less than previously projected. Between December 1, 1989 and June 30, 1990, only Ks 92 million had been collected, representing approximately 2 per cent of the recurrent budget (adjusted for the period) of the MOH hospitals and health centres, and only one per cent of the total recurrent budget of the MOH. This was far less than the expected figure of 5 per cent. The Ministry had anticipated collections would amount to 20 per cent of the total recurrent budget by the year 2000. This now looks an ambitious goal. To ensure revenue collection was maximised, fee expenditure guidelines were prepared according to which 100 per cent of all the funds collected are retained in the collecting district, 75 per cent to be used by the collecting facility while the 25 per cent is redistributed for community preventive health care in the district. The amount remaining in the facility would be used for facility and service rehabilitation, improvement of management and service delivery, and improvements in quality of care. Public hospitals, were encouraged to claim reimbursement from the National Hospital Insurance Fund (NHIF). It turned out that they had only managed to recoup 15 to 40 per cent of the potentially claimable revenue from NHIF.

Some of the problems listed above are largely because in the past there did not exist a clear health policy that defined national health care goals and the strategies of realising them, delineating the specific roles of various providers and institutional framework to synchronise the activities of the different participants in the sector. In March 1993 the Ministry of Health produced the 'Strategic Plan of Action for Financing Health Care in Kenya', (MOH/GOK, 1993) whose objectives, besides evaluating the then 'current' situation facing the health sector, were to

- ◆ develop a public policy role of the Ministry of health.
- ◆ provide a sharper focus of the Ministry of Health activities in the health sector.
- ◆ find ways of targeting public subsidies for curative care to the poor.

fees, X-ray and laboratory fees continued to be in force.

- ◆ explore means of encouraging increased role of the private sector in the financing and provision of curative care services.
- ◆ explore how alternatives to tax and out-of-pocket spending on health care through a diversification of social financing mechanisms could be strengthened.

This plan specified for the first time some important elements that a national health policy should contain, including

- ◆ some strategies for achieving the 'specified' goals, e.g., by
 - improving resource allocation within the public sector and mobilisation of resources through cost sharing and better use of private sector resources.
 - increasing efficiency, effectiveness and equity.
 - improving quality and quality assurance.
- ◆ delineation of the expected roles of the private sector in relation to the provision and financing of services—to assume increasing importance in the provision and financing of curative care services (services that have 'private good' characteristics).
- ◆ specifying the role of the Ministry of Health in terms of the extent of involvement in the provision of care, by devolving itself from the provision of curative care services and concentrating only on services that have 'public good' characteristics, mainly health preventive and promotive care services, and in extreme circumstances, providing all care in certain disadvantaged areas or Districts where the operation of the private market is adjudged insufficient.

Among (many) other recommendations (much of the plan is nothing else but a summary of all the recommendations proffered in the studies summarised above), the strategic plan recommended the following strategic goals and options for action in the health care sector as policy initiatives likely to have greatest impact (these have been selected on account of their relevance to the objectives of this thesis):

- ◆ directing public spending on curative care to target groups with specific health needs, including under-served areas, poor populations, communicable disease patients, and patients with chronic illnesses.
- ◆ Improving quality of services at MOH facilities, and encouraging the use of nearby low-cost facilities.
- ◆ creating an enabling environment and reducing government imposed costs of providing health care services through the private sector, particularly with respect to target populations, using appropriate incentives such as reducing unnecessary burdens through reduced bureaucracy, simplified licensing fees (e.g., by abolishing double license fees as often required of a provider who sets up activity in a Municipality or Local County Council area, where they pay a license fee to the central government and another to the municipality or local council authority) and tax structures, encouraging rural location and relocation through selected reduction of duties on medical equipment, setting aside of free land for private health facilities, provision of targeted subsidies for those operating in underserved areas, and so on.
- ◆ enhancing private sector role in publicly supported curative care by engaging in gradual privatisation of selected MOH curative services, departments, and facilities through pilot projects.
- ◆ strengthening the role of the National Hospital Insurance Fund (NHIF) as a social insurance scheme by, among others, developing a role of NHIF as a financier of new health development.

- ◆ Improving efficiency in allocation of scarce budgetary resources, through: allocation of resources using a formula built on the need for resources under consideration; redeploying from over-staffed facilities to under-served facilities that are in need of similar skills; allocating physical resources across administrative boundaries and facilities on the basis of assessed needs; and developing capital budgets budgeting plan based on target group priorities and actual demand for services.
- ◆ lowering the cost per unit of output of the MOH health services (i.e. achieving acceptable quality of services with minimum cost) through: increased use of budgeting not only as a tool for planning but also to pursue efficiency goals at different decision-making levels; giving different level decision makers budgets and responsibility to manage them; allowing decision makers to retain the whole or part of cost-savings within the facility at the end of the financial year; encouraging the use of economic efficiency/evaluation criteria in the choice of least cost intervention policies; and training decision makers at different levels on economic appraisal methods.
- ◆ managing MOH staff to achieve maximum productivity from available staff by: improving the working environment; reducing bureaucracy involved in making decisions at facility levels; improving the procedures for the procurement and distribution of drugs and medical supplies; and encouraging the use of (effective) treatment protocols.
- ◆ managing resources for drugs and supplies to achieve maximum benefit for patients through: strengthening the medical supplies co-ordinating unit (MSCU) management through the implementation of internationally approved pharmaceutical and equipment procurement protocols and workload based on estimation of need and allocation of supplies; finalisation and implementation of clinical guidelines; and so on.

This thesis provides some of the 'nuts and bolts' required to achieve/implement some of the above objectives/strategies.

Assessment and Conclusions

In collating these studies, we should ask: did they address the fundamental issues? Were they incentive-compatible in their proposed changes/reforms? Were they evidence based—are there reasons for believing them to have the effects hoped for from them? In short: (a) did they ask the right questions? (b) if they did, did they come up with plausible answers?

The administrative and organisational arrangements within which the Kenyan public health services sector providers operate make it inherently inefficient compared to private providers. The budget allocation process is the single most important factor. The PADS study (REACH/MOH, 1990) corroborates this view. Reviewing the budgeting process within a sample of public facilities, this study observed that in 1988/89 fiscal year, the treasury allowed the MOH a general increase of 4 per cent in recurrent expenditure over the previous year (while the rate of inflation was 10 per cent). Most facilities received only a fraction of their requests. (An interesting observation made by this study was that 99% of the increase [in recurrent allocations allowed by The Treasury to the facilities in the sample] went to salaries, leaving only 1% to cover all other categories of expendi-

ture.) As a result, many facilities cannot function efficiently and most have accumulated bills, which are carried over from year to another. This is contrary to official government policy since facilities are not allowed to incur expenses for which no financial commitment has been allowed when funds are insufficient, except in exceptional circumstances, and then only for 'essential services'. This study however found the most under-funded items of expenditure were patients' food; electricity; telephone; drugs and dressings; purchase of stationery; uniforms and clothing; and personal allowance expenses. That facility managers have a discretion in deciding 'what is essential' is apparent from the list. Moreover, some facilities did not have any allocations for buildings and equipment repairs. Thus, as shown above, the worst affected areas from this type of budget process have been non-personnel costs. This, combined with the practice of paying outstanding bills first when new allocations are made, leave many facilities in a perpetual vicious cycle of shortages, leading to a number of problems. This study observed that

It is difficult to establish the precise budgetary gaps of government health facilities due to the methods used to make budgetary estimates and the final allocations by the Ministry of Health . . . The methods used in making budget estimates and actual allocations lead to inappropriate estimates of the actual needs of health facilities. Budget allocations do not take into account changes in utilisation levels of health facilities that result from population growth or changes in demand for services. (p. 41).

The private facility included in the sample did not have as big a problem of financial shortfalls (with pending bills being only 3.2% of total budget allocations). One of the main outcomes of this type of budgeting process has been the non-functioning of the referral system. The Nairobi Area Study observed that many services that could be provided at lower level facilities were being referred to higher level facilities due to lack of equipment and personnel in the lower level facilities (of the public health service). This suggests there has not been proper planning that places emphasis on costs of treatment at appropriate levels. Consequently, upper level facilities become congested with patients who would have been served at a lower level (and at lower cost) facility. This not only increases the cost of providing medical care in the public sector, but also affects the quality of services as perceived by patients since the result is (probably unnecessary) increased travel and waiting times for all patients. Second, although all facilities, public as well as private experience shortages of trained personnel, the situation is worse in the public sector, but the major difference is that comparable facilities within the public and private sector use different personnel ratios. For example, the study cited above found that the doctor facility ratio in private, mission and parastatal facilities was generally smaller than in the public services, yet patients perceived the quality of service in the former to be better (although it was also observed that the workload of the private sector is generally lower). The private sector was observed to make more use of part-time staff as a

method of alleviating staff shortage, a rare occurrence in the public sector. Another significant observation was in the deployment of skilled health personnel in administrative responsibilities—the public services tend to deploy more of the higher level personnel in these activities compared to private and NGO sector providers.

In the Provincial and District Health Services Study considerable variations in the health sector were also found to exist. The average cost (to the hospital) per outpatient visit was found to be lower for private hospitals than for public hospitals, but the cost of the drugs per outpatient curative visit was much higher in the private sector. Cost per inpatient day was lower for the private sector. Similarly, the average length of stay was lower for private providers. Workloads reveal a different pattern. The highest workloads was for the private sector (in terms of volumes of work in relation to staff numbers)—in general, private providers manage more workloads with lesser staff.

The considerable variations which can be observed in the performance of comparable facilities as indicated above between the public and private sector *may* indicate that such variations cannot be explained by the single notion of 'facility uniqueness' and therefore there is need to examine how the performance of facilities in the public sector can be brought to par with that of their private sector counterparts. The main drawback of the studies cited above is that there was *no effort to define the output of facilities*—say in terms of *patients' health states*. However, in general, these studies are indicative of the general situation, although their results are not entirely suitable for comparison purposes because their methodology did not control for factors such as the peculiarities of facilities, or outcomes. More research is needed to clarify the issue of productivity differentials between the public and the private sector.

Moreover, efficiency comparisons in the above sense might not be revealing since we are actually not comparing like with likes. The public sector providers have different targets and not just efficiency, although it is important. Second, the sources of income and the way they are committed to activities also is different. Third, the public providers operate within an environment in which conditions are imposed on them by an outside regulator (the government) who specifies minimum standards, extent of geographical and social coverage, and so on. This commutes the comparison (Lindsay, 1976). The private sector operates more in line with the postulates of the competitive market (and the pareto criterion) while the public providers operate within a bureaucratic public administrative environment. The actual production characteristics, and the internal policies and or regulations that govern their operations may make it hard for them to operate more or less like the private providers. Since the existing data, in spite of inadequacy in terms of comparability, show

the private sector providers to be rather efficient, it is our intention to examine how the public providers in the system can be made to operate in an environment which, while not losing touch of the social obligations, promotes competition. Substantial attention has been devoted to the examination of methods of encouraging public providers to be more competitive. The literature bearing on these aspects will be presented in chapter five after a definitions of efficiency and equity in chapters two and three.

Generally, the studies summarised above (excluding the strategic plan) have used approaches that have facilitated the identification of the many weaknesses of the public health care system and have offered suggestions and proposed methods of how to solve them. The main weaknesses of the studies are that they had a narrow focus and did not offer incentive-compatible recommendations—either on a single facility or a group of them or in some cases on only a single topic—and concentrated a lot on managerial inefficiencies. In general, the issue of efficiency has not been treated as a systematic problem that can only be solved through a combination of technical *and allocative* concerns, simultaneously interacting, either at system-wide or sectoral level. Ignoring allocative efficiency considerations may result in prescriptions that improve the particular area of spending but this does not necessarily lead to optimal use of resources in the whole system. Second, focusing on costs, while appropriate, ignores other aspects in the environment that interact to affect the level of costs, such as the demand for services, marginal cost of alternatives ways of providing the service, the general environment within which activities take place, and so on.

From the above review of the efforts towards a better health care system we can summarise what appears to be the main issues in the Kenyan health system today as follows:

- ◆ resource allocation,
- ◆ cost-sharing,
- ◆ role of the private sector,
- ◆ role of social financing mechanisms,
- ◆ incentives for efficiency.

This thesis departs from the methodology adopted in previous studies on the Kenyan health care system in that we draw on economic principles and lessons from other health care systems in order to develop a reform strategy within whose framework technical and economic efficiency, equity, and incentive compatibility, both at micro and macro levels become focal issues of the health sector activities. The basic framework of analysis is developed in the next two chapters. We start by looking at the concept of efficiency in health care in the next chapter.

3.5 Summary of Key Problems Faced by the Current System that Need Redressing

3.5.1 Problems within the public sector

The major problems identified in this sector include:

- ◆ lack of adequate financial resources to continue/improve on past levels of support.
- ◆ inappropriate distribution of public expenditures between functions—which makes them less effective.
- ◆ inequitable distribution of public expenditures, which leaves most of the rural populations without adequate services.
- ◆ inappropriate use of facilities due to breakdown of the referral system as well absence of appropriate ‘incentives’ for consumers to use levels of facilities commensurate with their ‘need’, due to lack of prices that reflect the true opportunity cost of services in terms of local market conditions.
- ◆ inefficiency in service delivery—occasioned by inadequate supportive resources in most areas due to poor planning, lack of incentives for staff to use resources judiciously and lack of cost containment measures.
- ◆ Poor quality services due to shortages of necessary complementary inputs. Due to poor capital investment planning, the recurrent cost implications of many facilities were not properly budgeted for.

3.5.2 Problems within the private sector

The major problems identified in this sector include:

- ◆ financial crunch affecting mission facilities which have provided services since independence in areas underserved by both the public and purely private or ‘for profit’ providers.
- ◆ geographical disparity with a bias of location towards urban areas.
- ◆ little is known about the range of services provided by this sector, and the efficiency of its operations (other than that the services are mostly curative), although given its ‘competitive’ nature, one may assume away the efficiency concern.
- ◆ lack of co-ordination of activities of this sector and those of the public sector, that often leads to duplication and therefore waste of resources.

3.5.3 Conclusions and policy implications

From these observations, the following immediate broad (policy) proposals appear plausible:

- ◆ there is urgent need to focus explicitly on setting strategic direction and objectives that at the same time still makes the organisation of operating activities the responsibility of professionals;
- ◆ In the public sector, production should be decentralised and decisions delegated;
- ◆ evaluation and feedback should be a regular feature of all production;
- ◆ patients should have the opportunity to choose between alternative types and or providers of care;
- ◆ the responsibility for finance and production in the public sector should be separated, with more frequent use of external purchases where these promise better service than public provision;
- ◆ Need to change the engagement contracts of the public health sector employees if any incentive related gains are to be extorted out of them.

Beyond these (implicit) structural shifts at the system level, there is need for substantial changes in the decision making processes inside individual provider organisations. The major impetus to effectiveness should come from improved social dynamics (incentives and communication). Command and control relationships should be replaced by incentive-driven interactive relationships. Issuing instructions (that are expected to be implemented) should be replaced by dialogue—by a market style exchange between producers and consumers by participatory discussion among employees inside organisations, and an emphatic and respectful discourse with patients. In short, there is need to change the existing structural arrangements within the public health service.

The rest of this thesis is concerned with developing the framework for such a structural change. In the following two chapters we first develop the basic ingredients of the conceptual framework focusing, respectively, on efficiency and equity. In chapter six we shall then present a framework which can be adopted to replace the present command and control relationships.

3.6 Conclusions

The account of the organisation and structure of the Kenyan health care given in this chapter reinforces the conclusions of the previous chapter and identifies further problems concerning the public health care expenditures. Significant among these is that usually allocation is not related to expected health status improvements—because no systematic methodology is used. Second, the

productivity of resources is hardly known or even evaluated. The general conclusion that may be drawn from the previous two chapters' analyses is that, whatever proposals for policy and reorganisation of the institutional and organisational framework for providing health care in Kenya, there are two principal areas upon which research interest should be focused. The first concerns the relationship between the inputs and output. The second concerns the equitable distribution of those resources so that there is maximum impact of them on the health of the population.

The rest of this thesis focuses on these issues. In the next chapter, we develop the theoretical and conceptual framework for assessing the relationship between the inputs and outputs which also takes into account the efficiency with which resources are used, either within the regions or in services.

4. ASPECTS OF INEFFICIENCY IN THE KENYAN HEALTH CARE SYSTEM: DEFINITIONS, SOURCES AND IMPLICATIONS FOR HEALTH POLICY

4.0 Introduction

This chapter discusses efficiency in health care, presents some prima facie evidence of inefficiency in the Kenyan health care system, and provides a framework for examining aspects of efficiency in the system, pointing areas where some reforms are possible/necessary. The rest of the chapter is organised as follows. The next section establishes the objective function that should form the basis for efficiency in the Kenyan health care sector, viz, the quality of care, consumer choice, how the private sector is to be treated in the scheme of health services, and the notion of incentives needed to encourage both efficient and equitable use of the available resources. What constitutes the efficient use of resource resources is then discussed in section 4.2, together with other related evaluative concepts of the health care sector. Section 4.3 then presents some evidence which suggests there is inefficiency in the system. In section 4.4 a snap-shot framework for analysing the issue of inefficiency in the health sector is outlined. Section 4.5 summarises the main arguments of the chapter.

4.1 The Objective Function in the Kenyan Health Sector and Efficiency

In the last chapter we outlined the general setting of inefficiency in the Kenyan health care system. To discuss efficiency in the health care system, we need some basis for evaluating it. This is provided by the system's objective. The current national development plan (Kenya, 1994a) simply states that 'The long-term objective of the health sector is the achievement of *Health for All by the year 2000*' (p. 229). This is a rather amorphous statement for which it is difficult to give content and interpretation. The first national development plan (Kenya, 1966) appears to provide a better statement of the objective of health services, in the context of efficiency:

In drawing up the health programme for the period 1965/70, the Government has kept in mind the fact that an *improvement in the health of the nation is fundamental* . . . Thus, expenditure on health services must be accepted as an essential investment even though the returns may not be easy to calculate . . . Following the precepts of African Socialism as set forth in Sessional Paper No. 10, it is the Government's long-term objective to provide an adequate level of free basic services to all its citizens. *The country does not have the financial resources to provide all these free services at once. Therefore, the long-term objective must be reached through a series of steps as resources permit.* The introduction of free medical services for out-patients and all children in 1965 was an important step in this process. . . . A determined attempt will be made to bring health services

increasingly within reach of all people in order to *promote improvement in the level and standard of national health*. (Kenya, 1966, p. 314, *added emphasis*).

It appears from the passage that government considered 'free' care an 'end' rather than a 'means'—but providing free services is not necessarily an effective means of maximising health gains as we shall demonstrate below. But the context of this passage is best interpreted in light of the Kenyan constitution which guarantees 'equality' before the law, regardless of economic condition, political affiliation, religion, and so on, recapitulated in the Sessional Paper Number 10 mentioned in the passage above (Kenya 1965) to be 'in the sense of equal opportunities to receive certain services and other benefits, among them, health care' (pp. 1—2)—(equity is the aspect implied here and it is discussed in the next chapter as an objective of health services in Kenya). However, the assurance of such a principle has to be interpreted within existing resource constraints since, although it is ethically desirable that every Kenyan achieve the best level of health the available resources can provide, there are resource constraints that impose the need for choice between competing ends. This crucial constraint is recognised by policy makers, as the underlined section in the passage indicates. This immediately suggests that whatever resources are used for the purposes of promoting health have to be used in such a fashion as to maximise their impact on the nation's health (Culyer, 1992)—see the first paragraph on page 94 below. We shall see in the next section this requires an evaluation of the contribution of those resources to health in various uses to which they may be put in the sector and that sometimes hard choices have to be made.

In the health sector, a multitude of choices has to be made with regard to production, provision, and the distribution of the health services such as what should be produced, how much should be produced, how the production of the health services should be organised, i.e. which sector—public, Non-Governmental Organisations (NGOs), or private—should provide what services as well as how they should be distributed regionally and between individuals. It is therefore necessary to identify who, within the system, makes/should be making these choices/decisions. This is important because it enables us to interpret the context in which the objective function is translated into quantifiable/measurable outcomes in the health sector. Knowing the people (or groups of people) making decisions, the environment in which they make them, and so on, can enable us to evaluate these decisions appropriately and where necessary, design appropriate mechanisms to influence these decision makers to make decisions in ways likely to lead to realisation of the health system's objective. In so doing, we can be able to determine how maximum effectiveness and efficiency in use of those resources can be achieved.

A discussion of the health policy framework in Kenya appears to be a good starting point. Figure 4.1 is a schematic depiction of health policy and various levels of resource allocation in Kenya. The objective of health policy is in health *per-se*, placed at the top in the diagram. Health policy, which is directly concerned with the pursuit of this objective is placed next below. The next row shows that there are several alternatives of producing health and public health policy (in the diagram directly related to health policy) is only one of them. Other policies are put at par with public health policy because, though not directly concerned with health *per-se*, do nevertheless have impacts (mostly indirect), on it. Therefore health policy competes for scarce resources with other policies (we can think of this as a decision of determining the sectoral shares in the national cake—at what we call the ‘pre-strategic’ level, and this is outside the scope of our analysis). Health policy is composed of programs. The programs consist of a set of projects that in turn consist of activities, these in turn being broken down to treatments, techniques or services. Each of the programs can be expanded all the way down to activities. In the diagram, we have only expanded two of them, and these are discussed only in relation to the public health services. Moreover, the expansion is not exhaustive. The diagram therefore is not an in depth representation of all elements of health policy, but the expanded branches will serve to illustrate the major issues.

The levels where resource allocation decisions are made are indicated by circled and numbered nodes. The first level of decisions are what we call here ‘strategic level’, the circle numbered 1, determines shares to these programs. In Kenya most decisions affecting the health of the population at this level are made by planners and politicians, although in recent times, with the decentralisation of government administrative activities and the introduction of the District Focus for Rural Development strategy, the planning for the health needs of the population is being done often with collaboration of the local people in order to get a better matching of resources with local needs. This development notwithstanding, strategic decisions of allocating the total funds available to the health sector between its various activities are still made by planners at the MOH. Such decisions determine the potential output mix of the health sector, i.e. how large each of the programs would be in the MOH budget. This level of allocation results in determining allocations that will go to various geographic regions. Decision makers at this level can be expected to be interested in allocating resources so as to maximise their impact on outcomes—i.e. health, or some other proxies of it such as mortality and morbidity. The question is, particularly for the public services, have they been doing so? We show below that it is doubtful, and will argue the resource allocation should be related to the patterns of morbidity and mortality. In later chapters we shall outline the procedure for incorporating such factors when making decisions at this level.

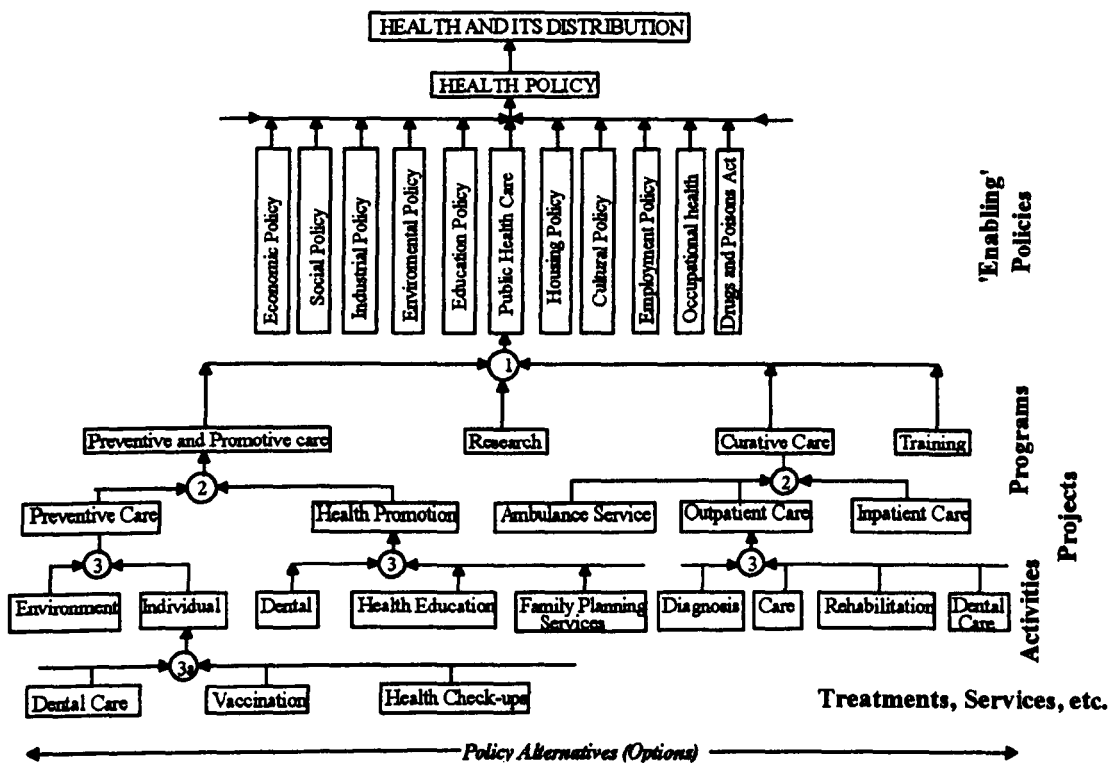


Figure 4.1: The objective of health policy and policy alternatives

The second level of decision making in the system allocates resources between the various projects—such outpatient services, inpatient care, preventive care, and so on. Thus for example, when a hospital (an example of a program—for providing hospital services in a certain area) has received its allocation from the MOH, it decides how much of them it will allocate for outpatient, inpatient, maintenance services, and so on²⁶. We show below that because of lack of focus in the budget, a number of items get underfunded, mainly due to constraints imposed by level 1 decision makers, and propose a simple way of rearranging the accounts in order to be able to detect areas where underfunding is likely to make severe constraints with the program activity.

The third and last decision-making level in the system concerns allocating project resources to consumers i.e. at the case level (e.g., at facility level) by medical specialists working within the budget allocated to them or the facility, These decide how to treat patients—whether to hospitalise or treat them as outpatients, and so on. These decisions are not only about the effectiveness of the treatments they offer, but also their efficiency. Most of the studies reviewed in the previous chapter have concentrated their effort on this level, mostly on efficiency (see chapter 3, section 3.4). From the definitions of efficiency and effectiveness (given in the section below), it will become clear that

²⁶ In reality this decision is made before the funds are allocated to it for program heads or managers produce activity budgets to back their budget requests.

this piece-meal treatment of problems in the health sector goes only a part of the way in solving problems of resource allocation in the system. Another aspect of resource allocation decision, labelled as 3a, is that by individuals. Although in the diagram this comes under preventive care, it may appear under any of the activity branches. This shows individuals affect resource allocations by the decisions they make about their health care seeking behaviour. We show below (section 4.3) that when individuals seek health care at levels not commensurate with their health states, this affects the efficiency with which health resources are used in the system. This aspect and what can be done about it will also be discussed in this study, mainly by examining the methods used in other systems to cause individuals to seek health care at 'appropriate' facility levels and compare this with the course taken in Kenya.

But, although on some occasions we will consider choices made by individuals themselves (node 3a), we shall in the main assume that the 'output' of the health sector is allocated between individuals by doctors and other health personnel, both administrative and professional, who work in various facilities within the ministry (as well as others in the private sector). For ease of exposition, we shall refer to such an individual (or groups of same) as the 'policy maker'. The questions we must consider are what criteria should these (policy makers) use and to what effect? How can they make decisions that ensure the output from the 'limited' resources is maximised, i.e. in a fashion so as to maximise the impact on the nation's health of whatever resources are availed to this end? We start on the premise that, *since budgets (synonymous with resources) are, and will always be limited, it is desirable to attempt to get the most out of them.* That involves making formal, well-considered judgements about how best to allocate those budgets and the services they will be used to produce. What should be the basis of such considered judgements?

The pareto optimality condition appears to provide the most comprehensive theoretic construct for assessing a system's efficiency since it takes into account the way resources are allocated, the way production is organised and the distribution of the goods and services to consumers—all current issues in the Kenyan health care system. The principle says that a change is desirable if it makes some individuals better-off without making any others worse-off. This is the main value judgement on which the pareto optimality condition is based. The impact of changes is usually analysed by use of the social welfare function (SWF), which relates different allocations of resources to different levels of social welfare. In this work, the objective of health policy as identified above can be recast as the maximisation of health, subject to the condition that it also is equitably distributed. This rider means health policy in Kenya should take distributional issues directly into account. Specifically, resources benefiting the poor are given more weight than those benefiting the rich.²⁷

The 'welfare' or the maximand is interpreted as the individual's health status or some proxy of it, which depends on the inputs of health care, and other health enhancing inputs (including intra- and extra-personal factors), available to the individuals, adjusted to take into account the factors such as externalities, uncertainty (about the timing and effects of illness on the side of consumers, and about the outcomes of various diagnostic regimens on the side of the providers), merit good argument and monopoly nature of health services production. Only then can the available options (of allocating health resources between individuals) be ranked on a scale of better or worse, depending on their overall impact on the SWF.

The use of the pareto improvement criterion to choose among ways of committing scarce resources presupposes that the securing of potential pareto improvements is the social objective. There may be difficulties of interpretation of who makes such decisions in society. But we want all players—from top policy makers through managers, professionals, to patients—who commit resources to various uses at various levels in the system to behave appropriately. In the literature, two broad interpretations (of the decision maker) can be distinguished. These are the *decision making approach* and the *paretian approach* (Sugden and Williams, 1978). The decision making approach presumes that the social objective is an objective pursued by a social decision maker who makes decisions in the interest of the public. The decision makers in this case will normally be expected to choose an objective corresponding to the pareto improvement because they know that, with the backing of the government, with its powers to convert potential pareto improvements into actual pareto improvements, it is possible to translate the theoretical potential pareto improvements into a practical policy. Hence, projects satisfying the potential pareto improvement criterion can be implemented, with appropriate tax changes, etc., to make everyone better-off. In this approach therefore the selection of the objective function is at the discretion of the decision maker. If the decision maker errs in forming judgements about the objectives, it may turn out that a better alternative is omitted. Second, the decision maker may, by design or error, choose a course of action and stick to it without looking at alternatives, because they believe 'it is the correct' one.

²⁷ But it need not be so. Given the potential enormous costs (e.g., informational and administrative) of the application of distributional weights (Harberger, 1978), and following Ng, 1990, it is only necessary to concentrate on efficiency, with an understanding that the result is subject to distributional qualifications. Such an approach may be based on a premise such as

For any alternative (A) using a system (a) of purely equality-type preferential treatment between the rich and the poor, there exists an alternative, B, which does not use preferential treatment, that makes no one worse off and achieves the same degree of equality (of, say, the health status, access to health care, or welfare, etc.) and raises more government revenue which could be used to make everyone better-off. (Cf. Ng, 1990, p. 248).

This means the pursuit of equity in health policy be limited to consideration of incentives, on both the producer and provider sides. In this formulation, irrespective of the recipient, a shilling spent on health would be treated as a shilling, if efficiency is to be achieved. Then, any desired level of equality can be achieved by use of redistributive mechanisms.

The paretian approach, on the other hand, presumes that the objectives of social decision makers are 'distilled from a consensus of value judgements of the individuals who make up the society, and they should be propositions which command universal, or, at least very wide assent' (Sugden and Williams, 1978, p. 91). The value judgements derive from welfare economics and are concerned with economic efficiency and distributive justice. In the health domain, economic efficiency would be about the overall health of the community. It also could be viewed from the input side—as pertaining to the size of the total care resources available to the community, since the pareto optimal distribution of inputs should serve the ultimate objective (outcomes). The distributive justice issues are about the way this (health or health resources) is/are shared amongst the people. The potential pareto improvement criterion identifies changes in economic efficiency: a change that produces a potential pareto improvement is one that increases economic efficiency. But potential pareto improvement and increases in economic efficiency mean precisely the same thing. The basic premise of the paretian approach is that, *ceteris paribus*, an increase in economic efficiency is a good thing. If social welfare is measured by the level of economic efficiency and distributive justice, then an increase in economic efficiency is desirable, provided no corresponding decrease in distributive justice occurs. However, if there are other dimensions of social welfare, as some writers have argued there are (Sen, 1982; Culyer, 1990), then the pareto improvement criterion goes only a part of the way to satisfying increases in social welfare. The decision making approach appears closest to the framework in which resource allocation decisions are made in Kenya and will therefore be adopted as the basis for identifying changes with potential pareto improvement.

4.2 Efficiency in Health Care: Some Theoretical Considerations

In order to avoid confusion in interpreting/evaluating the alternative reform strategies proposed for the Kenyan health care system later on in this thesis, we have to first clarify a number of issues at the outset. This includes giving appropriate meanings to evaluative terms like 'efficiency' and 'equity'; laying down the criteria for assessing the desirability of making greater use of one system of provision rather than another (e.g. the NGO private sector as opposed to the private for-profit providers); and the use of professional and financial incentives (such as peer group review or pay-related incentives), and the relative desirability of increasing consumer choice. Alternatives might be judged in terms of how they address the issues of efficiency, equity, quality of care and consumer choice. The alternatives are instrumental to achieving these ends. Except the issue of equity, which is addressed separately in the next chapter, this section discusses these issues.

Efficiency is the maximisation of the size of the total health of all members of the community, achieved by evaluating in terms of opportunity costs, the contribution to it of alternative ways of committing scarce resources and choosing those alternatives that minimise the opportunity costs. One method, which *appears* to have been the basis for analysing efficiency issues by previous studies on the Kenyan health sector, is the estimation of unit costs (or average costs) of delivering services, and *choosing* the method with the lowest unit cost per unit of output. Focus on the average is fundamentally wrong for efficiency considerations. We have used the term '*appears*' deliberately since in previous studies all that was done was to compare the unit costs by different producers, therefore, there was no evaluation of alternatives²⁸. It is usually preferable to do this comparison only for one unit (or similar units) of production, since in such a case, 'other extraneous' factors affecting production costs are readily 'controlled' for. According to this approach, production is technically efficient if the inputs employed produce the maximum possible output. The total cost of a service is the sum of the fixed and variable costs. Dividing the total cost by the total units produced (such as total inpatient maternity cases) results in an estimate of unit costs. On the basis of figures derived this way, it is possible to extrapolate from one case to total expected workload. But this type of extrapolation is not entirely correct since average costs may rise or fall with increasing activity, depending on whether the activity is subject to increasing or decreasing returns to scale. Although the method is simple and straightforward, and probably accurate within narrow ranges of activity, stretched beyond bounds, it may give implausible results, especially where little attention is paid to other determinants of cost such as how they respond to changes in overall resource allocation, marginal adjustments in input mix due to relative price changes, and so on. Unit cost estimates may turn out to be lower because of resource reallocation that changes consumer or producer behaviour, or alters the mix of inputs. Thus, in general, providing a service at the lowest cost per unit of output (in a technical sense) does not necessarily guarantee optimal use of resources.

4.2.1 What should efficiency in health care mean?

The *ideal* procedures for making choices, either between projects in one sector or between sectors are widely documented.²⁹ A similar procedure for the health care sector resource allocation may *similarly* be defined.³⁰ Prior to committing resources to activities that affect the current and future benefits to society, a thorough analysis of the country's health problems, their causes and a

²⁸ Although we can think of public versus private provision of health services as alternatives.

²⁹ Squire and van der Tak (1979), Mishan (1972), Irvin (1978), Little and Mirrlees (1974).

³⁰ The analysis which follows is developed along lines suggested by Culyer, 1992, and de Ferranti (1983a, b).

thorough evaluation of the existing system should be done, taking into account the country's objectives and priorities. Such a sectoral evaluation should include an analysis of the financing sources, and a detailed analysis of the allocation of expenditures with a view to determining the unit costs of each activity and how these are likely to behave over time. This requires an estimate of current and future demand patterns. This type of analysis should be conducted using marginal rather than average costs. Such analysis provides enough information to enable more focused decisions concerning recommendations on objectives for adopting certain procedures, new investments and feasible means of achieving them, to be made, taking into account the technical effectiveness of implementing each decision, their administrative requirements and implications, etc. This means estimates of full costs (both capital and recurrent, pre- and post-implementation, financial and economic)—using, for economic costs, both efficiency and shadow prices—have to be made. Then follows an evaluation of the expected effects of each option on the health status, as reflected in reductions in mortality or morbidity. This analysis will of course take into account the information about the target population, the disease patterns and disease agents and the interventions' effectiveness, etc. Finally, quantifying the estimates of the value to country of the resulting health improvements, the benefits of each option should be quantified, and reduced to a single (monetary) unit. Comparing the costs and benefits of each intervention, based on pre-established criteria [such as the net-present-value (NPV), benefit-cost (B/C) ratio, internal rate of return (IRR), etc.], the activities (or, as usually called, projects) to be funded can be chosen. As a final step, the financing side of *the promising* activities are analysed to ensure that they are affordable. It is advisable at this stage to evaluate the probabilities that those groups, or institutions, required or expected to commit resources for implementing the projects will be able and willing to do so. The net-resource position of each project thus assessed is determined. Also, at this stage the long-run sustainability of recurrent costs after project completion and the replicability of the project(s) (elsewhere or later) should be examined. Of course, the analysis of the planned sources of funds would take into account their implications for efficiency and equity, as well as implications for the public, quasi-public and private mix of services or facilities (if it is a crucial aspect) and other structural considerations (such as the flexibility of administration, etc.). The important ingredients of such a framework are outlined in this section.

The notion of efficiency is much broader than the cost minimisation interpretation given to it in previous studies on the Kenyan health care sector (discussed in chapter three). An efficiency concept that is relevant to a health care system such as Kenya's has four elements which are cumulatively increasingly comprehensive—*effectiveness or efficacy of health care services (which is the*

sister terminology to the concept of 'technical efficiency' in other economics literature); their marginal cost-effectiveness; service-mix payoffs in terms of health; and, scale of efficient operation for different mixes of effective services (see Culyer, 1992). Let us look at each in turn.

Providing only medically effective services

In the health sector, technical efficiency is closely related to 'effectiveness' but the two are not synonymous. Suppose society were faced with the choice between committing resources to health or to defence and 'somehow' this choice-problem was resolved in favour of health. It would soon be realised there are many uses to which those resources can be put in the health sector itself. If we are to maximise the returns from these resources, i.e., use them efficiently, we need to use them 'effectively' in the particular lines of actions we choose to employ them in (in the health sector). Each of these lines of actions can be summarised by three elements—inputs, their transformation and output. The input is the resources we have at our disposal—which can be any combination of personnel, materials, facilities, and so on. The output would normally be the result of an 'input-transformation process'—through the health production function, for example, and expresses the outcome attributable to taking that particular course of action, for example, improvement in the health status. If we commit resources in a particular use—in curative care for example—three possibilities exist. We might record an improvement, a decline or there may be no change in the health status. Effectiveness applies only to the case where the net change in the desired outcome is attributable to the action in question and is positive (see e.g. Cochrane, 1972; Culyer 1976). Suppose the action is taken at time t_0 at which point a health state H_0 prevails and some other state $H_{t_1}^e$ is realised at time t_1 , then the action is effective if (and only if) $H_{t_1}^e > H_0$. If in the absence of the action the health state would have deteriorated to $H_{t_1}^d$, in cross-sectional analysis the effectiveness of the action is measured as $E^t = H_{t_1}^e - H_{t_1}^d$ (see figure 4.2 below).

Over time, effectiveness would be measured by the shaded area in the figure. Thus effective care is one which improves health status. But effective care is not necessarily efficient care as we show below.

Effectiveness of health care thus gives the first notion of efficiency in health care—only services for which there is credible evidence about their ability to alter for better the patients' course of illness should be provided, i.e. provide services that are effective in the sense illustrated above. Clinical epidemiologists refer to this as efficacy. Culyer, 1992, succinctly puts it this way:

Given an objective, such as returning the patient to normal functioning as speedily as possible, one should therefore seek those combinations of diagnostic procedures, physicians' time, medicines, surgical procedures, inpatient and outpatient care, health service and family caring, and the patient's own time, that are most effective. To use more of any of these resources than is necessary for an effective impact is wasteful and inconsistent with the objective of maximising the impact of health resources on the health of the community. For, if more than is necessary is used, the excess could have been used at no cost to the patients in question in order to further the health of some other patient. Thus, overall community health . . . is lower than it need be.

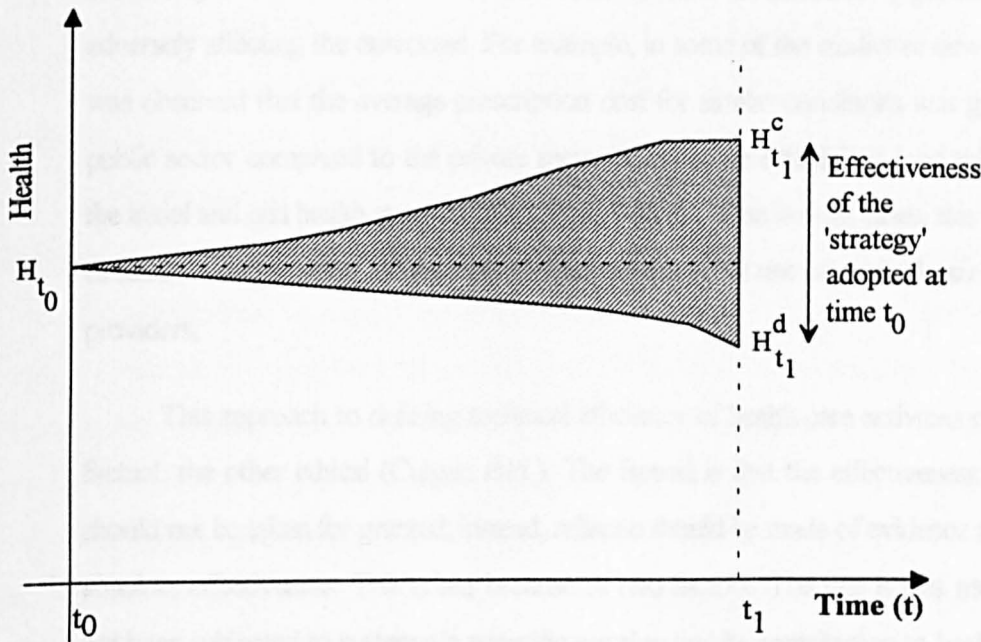


Figure 4.2: Health improvement between t_0 and t_1 due to 'effective care'

Technical efficiency is thus a function of effectiveness and the actual inputs. Technical efficiency, also termed operational efficiency, is about deciding the best way of doing an activity *that has been found to be effective*—that is, selecting from among the alternative means of achieving the same outcome, the option that maximises the output for a given amount of resources (cost). Another way of ensuring the same result is to minimise the cost of producing a given level of output. This is the standard neo-classical rule of operation for the profit maximising producer. For example—to continue the example used above—suppose, as it is, Kenya is interested in reducing the child mortality and morbidity level due to common childhood diseases. The country can either treat those who have ill-health as a result of these diseases, or engage in a mass vaccination campaign to prevent the need for such treatments in the first place. Assume for the time being the end result achieved by either way is the same—equally reduced mortality and morbidity—that is, effectiveness is the same. The question we want to ask ourselves now is, 'which approach is the more

(technically) efficient one?' Since by assumption the effectiveness is the same in either case,³¹ assessing the costs of each determines which method is technically more efficient. The one which uses least resources—i.e. inputs (hence least costly)—is the more operationally (or technically) efficient method. In summary, technical efficiency is about whether it is possible (by using a different combination of inputs) to produce the same flow of throughput (number of children vaccinated) or of output (a reduction of mortality or morbidity to certain levels) by using *fewer* of one or more of the complementary inputs. If this is possible, then the present process of doing the activity is not technically efficient. Resources could be released from the (ineffective) process without necessarily adversely affecting the outcomes. For example, in some of the studies reviewed in chapter three, it was observed that the average prescription cost for similar conditions was generally higher in the public sector compared to the private sector. If it can be established (and this was not done) that the initial and end health states of the patients was the same in both cases, this would be an instance of technical inefficiency (or absence of effectiveness in the use of medical care resources) for public providers.

This approach to defining technical efficiency of health care activities raises two issues: one factual, the other ethical (Culyer, *ibid.*). The factual is that the effectiveness of medical practices should not be taken for granted, instead, reliance should be made of evidence about their relative or absolute effectiveness. This arises because of two factors. The first is that much medical care has not been subjected to systematic scientific scrutiny and its contribution to health is therefore hardly known because, as Culyer (*ibid.*) reasons,

much practice is based on that species of gossip known as the case study, that there is often a strong 'medical signature' in explanations of costly variations in clinical practice, [but] the statistical design of much clinical research in the past has left much to be desired.

The second reason is that *health care is only but one of the many determinants of health*, and often it has only a marginal impact on it (e.g., see Auster, Leveson and Sarachek, 1969; Grossman, 1972; Silver, 1972; Hadley, 1982). In developing countries, for example, mortality is more closely related to levels of poverty, illiteracy, malnutrition, poor housing, poor personal and environmental hygiene, cultural practices, and so on, than to the distribution of health care services (see for example, Kalipen, 1993; Sharif, Huq and Saleheen, 1993). Such information, usually derived by estimation of aggregate health production functions, would be useful for Kenya since it would enlighten health policy decision making, for example, on how health care resources should be allo-

³¹ This is actually the implicit assumption of operational efficiency—that there exists effective options (e.g. effective health care).

cated between geographic areas or among population groups, or whether increasing health care consumption is the best policy for achieving our health objectives.

Providing services that are cost-effective

The second stage of efficiency, *cost-effectiveness*, assumes technical efficiency, but recognises that there might exist several technically efficient ways of performing a specified health activity. It is therefore becomes necessary to evaluate the cost of combining inputs differently, as implied by each of the methods. The interest here is to select from among all the available technically efficient (or medically effective) methods, the one that is least costly. If there exists a less costly method than the one in current use, the present method is not cost-effective, and, by re-allocation, resources could be released from it and used to improve outcomes elsewhere, or, for the same activity, more output could be realised for the same level of cost. In the health sector, there exists several substitution possibilities—between drugs, between surgery and medicine, between primary and curative care, between use of more qualified and less qualified personnel (e.g., doctors versus nurses), and so on, all enabling to achieve a specified objective. Assuming none of them represents inefficient (or ineffective) care, there is need to narrow the definition of efficiency. The requirement that *whatever services are being provided should be provided at least cost* provides the criterion for discriminating between the available technically efficient methods. Again, 'to incur a higher cost than is necessary . . . is wasteful and inconsistent with the objective of maximising the impact of health resources on health in the community' (Culyer, *ibid*).

Efficiency also demands concentrating on health services that offer the highest pay-offs in terms of health

In the health sector, the assumption of processes having similar effectiveness is far fetched. Sometimes (possibly most of the times) one method may be the more effective one but also the more expensive one (Donaldson and Gerard, 1993). In such cases, the marginal cost-effectiveness ratio should be used as a guide to decision making: the lower the ratio, the greater the technical efficiency (see also Sugden and Williams, 1978; Curry and Weiss, 1993, chapter 3; Pearce and Nash, 1981, chapter 4). These ratios can be developed through use of various techniques collectively known as 'economic evaluation'. This involves judging the merit or worth-whileness of some action (or activities, projects or programs). It aids in decision making concerning the choice between actions, or even whether, in the case of a single project, it is worth committing resources to. The programs may be mutually exclusive (substitutes), complementary, or independent. The economic

evaluation can be *ex-ante* or *ex-post*. In the *ex-post* case, the purpose of the evaluation is to decide whether to continue or abandon a program, or modify its implementation. In the *ex-ante* case, the decision at stake is whether to initiate the planned action or not, or, in the case of multiple projects, which one to initiate where they are substitutes or which combination (where they are independent). Economic evaluation assumes that the non-economic evaluation has already been done—that is, the effectiveness of the actions has been established. The economic evaluation then quantifies these (effective) actions in terms of their economic efficiency. The purpose is to ensure that not only ‘worthy’ activities are undertaken, but that only those with the biggest pay-offs in terms of health are the ones that use the scarce resources.

The notions of technical efficiency and cost-effectiveness on their own are thus not discriminating enough to ensure attainment of optimal allocation of resources in the health sector. They do not take into account the people’s valuation of that care as it necessarily implies all individuals should receive ‘standardised’ services. Second, in addition to the form of care provided, we should take into account the resources that people, individually or collectively, are willing to make available to pay for it. This it means it is immaterial whether the quality of the care by a particular activity is of the highest best quality possible in a technical sense, but whether the level and quality of that care represents the best use of the economy’s resources, given the people’s tastes and preferences for it as well as for other forms of medical care *that provide similar outcomes*. This means the ‘benefits’ derived from the alternative medical options of obtaining a given objective are to be compared with the costs of producing them—the opportunity cost in terms of foregone benefits. Higher quality care (in a technical sense) is better only if the benefits people derive from increased quality offset increases in the costs that would be associated with it. Taking these factors into account requires we introduce people’s ‘needs’ for health care into the analysis. ‘Often a situation is regarded as inefficient if people are not getting as much or as high a quality of care as the system could feasibly (though not costlessly) produce at that time’ (Pauly, 1971, p. 4). The problem with this definition of efficiency in health care resource allocation is that it does not take into account the cost of the ‘needed’ services—as medical technology changes, so will the demands. It therefore becomes necessary to evaluate whether, given the people’s preferences and tastes for particular health care services, health resources are being allocated to those services (in the sector) that secure maximum benefits to the community. In the theory of production, this problem is resolved by the application of cost minimisation techniques for a specified level of output, or what is the same, maximise the output for a given set of resources (assuming a given desired outcome is the objective). If for the moment we assume the output of health care services is ‘health’ this means with a

given amount of resources allocated to the health care sector, health should be maximised. In order to make judgements about efficiency in such cases, it is necessary to have measures of health that are appropriate for the kind of procedure or service that at least enable comparisons to be made. If health is 'appropriately' defined, then it would be possible to apply this criterion by allocating health care resources so as to maximise the excess of 'benefits' or health over costs of producing it (Feldstein, 1963, p. 23). The number of life years gained or saved, as measured by reductions in mortality or life expectancy at a certain age (e.g. at birth, 25, 45, and so on) are two common such measures. The Quality of Adjusted Life Years (QALYs) is an innovation recently developed aid to determining the productivity of different health care interventions (see Torrance, 1976), and is discussed in detail in the next chapter. With explicit assumptions about the relative values of health or ill-health, it enables a complete and comprehensive assessment of the effects of alternative procedures on outcomes, including the option of doing nothing, to be made (Culyer, 1992).

Together, the above 'principles' of efficiency in the health sector define a framework, which involves ethical as well as economic judgements, about the rationing of health care resources.

Efficiency also means providing appropriate scales of the medically- and cost-effective services

Once the low level efficiencies have been established as defined above the next stage is to determine overall economic efficiency. In this stage, it is decided which actions are worth allocating resources for, taking into account the effects of allocating those resources to those actions, rather than others outside the health sector. This is why it is also called allocative efficiency. Allocative efficiency is useful at the planning stage where decisions are made regarding which activities are worth doing. At this stage, the scale—i.e. relative sizes—of programs are determined, and it is assumed the question of their technical or operational efficiency has been settled. The aim of allocative efficiency is to ensure that society does not lose by having resources allocated to one activity rather than another. If technical efficiency exists, economic efficiency is synonymous with pareto efficiency. This may be termed as 'full efficiency'.

These notions of efficiency in the health sector imply that resources or costs are not to be considered in isolation from what they are believed to enable to be accomplished. Second, efficiency in health care is a moral notion that entails the maximisation of the impact of health care. There is a correspondence here—that inefficiency necessarily implies patients (prospective or actual) are less well-off than they need be. Third, the notion of full efficiency can be applied in principle to organise thinking in making judgements about the appropriate rate of total resource

deployment to health care as a whole (Culyer, Donaldson and Gerard, 1988). In particular, this 'full efficiency' condition ensures that, in terms of the diagram of figure 4.3 (below), we are at or somewhere close to the point *b*.

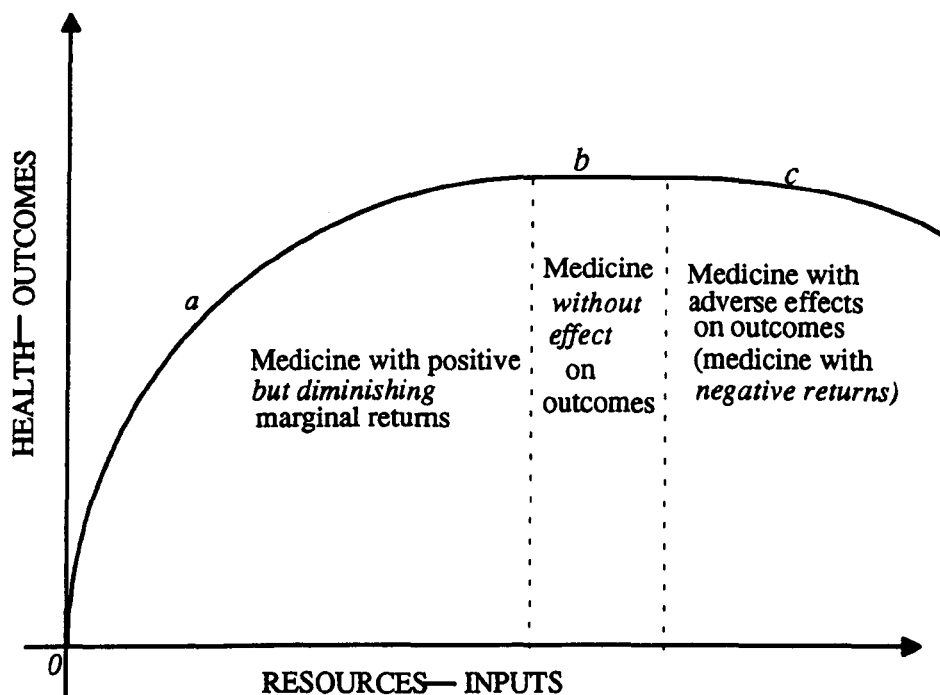


Figure 4.3: Determining the optimal size of the health sector

In terms of the efficiency notions used above, effective medicine is implied by the segment of the total 'productivity' curve where extra real resources to the health sector lead to increased outcomes (points such as *a*). We should stop allocating resources to the health sector at the point *b*, although as the above analysis shows, that point would have to be determined by comparing the benefits and costs—so the outcomes here should be interpreted as 'net of costs'. Allocations in cases such as that depicted by *c* are inefficient.

The framework for efficiency outlined above can now be related to the Kenyan health policy framework described in section 4.1 above. Viewed together with the process of resource allocation as described in chapter three (see section 3.2.1), it is now apparent that, rather than move from top to bottom in figure 4.1 when allocating resources, as is the present practice, the process should actually be the reverse—moving from the bottom upwards. It is only through such a process that we can ensure that the resources allocated to the health sector meet all the efficiency criteria set out above. We saw in chapter two that there is no matching of resources with the 'health needs' of

various regions in the country. Such of course would be the outcome we would expect to in a situation where the allocation is done, in terms of figure 4.1, from top to bottom, rather than the other way round, where at each stage, the relevant efficiency notion is taken into account in order to arrive at an 'optimal' allocation of resources to the health sector.

4.2.2 Quality of care

The quality of care is often interpreted by patients in terms of the inputs used, processes (how the care is provided, i.e. the environment) or outcomes. Viewed in terms of inputs or processes, the above outlined efficiency procedure is sufficient (with qualifications to be outlined in the next chapter, about how we treat factors that have no direct bearing on outcome enhancements, such as hotel services)—the appropriate quality will be that which is cost-effective in achieving the ultimate goals of health care activities. Here we define the quality of care in terms of the patient's well-being in terms of final outcomes, their perception about the inputs and processes notwithstanding. This is important because in Kenya, there is a tendency (among patients, and even within the medical fraternity) to associate good quality with non-health affecting factors. Whilst these factors *may* have some (as yet undetermined beneficial) effect on final outcomes, the quality of care should not be determined in terms of factors other than those help advance the cause of medicine—improvement in health outcomes. This is why we rule out those other factors.

4.2.3 Consumer choice

The choice of initial point of contact or consultation in the system by the consumer (in the Kenyan health sector this inadvertently is closely tied with the choice of the financial intermediary), and the choice of further medical treatments, diagnostics and so on, and the choice of ancillary services are important, as they differ depending on individual circumstances. The first often relates to the consumer's ability (or inability) to exercise choice about utilisation. The fewer the impediments to such choices (e.g. through the availability of more 'points' of contact with as little inhibitive factors—such as charges), the greater the consumer's freedom. Since, apart from the initial decision to consult, consumers generally rely on the advice of doctors as agents, the fewer the impediments to them, the higher the choice. The crucial issue now facing the Kenyan health policy is how to develop the health system without excessive reliance on government revenue. In this context, three main options that are being implemented include:

- ◆ cost recovery in the public service through user fee charges

- ◆ increased reliance on insurance to cover the cost of care in the public sector
- ◆ increased use of the NGO and private sector, through encouragement of reliance on privately financed risk-sharing schemes.

These options have different implications for consumer choice and will be crucial in determining the relative balance between the MOH and the private sector, and there is need to examine the available evidence on their effects on utilisation. There is need to ensure that only approaches that do not deter—or adversely limit the consumers’ freedom or distort their choice—consumers from making the initial contact are encouraged.

The second choice relates to consumption of services after the initial contact has been made by the patient. This essentially depends on the agency relationship that exists between the patients and doctors and to a large extent what happens here is much dependent on the financial intermediary between the patient and the provider. Relationships that distort the incentives of the providers as agents (for example, by encouraging hasty consultations and off-hand referrals, or reward providers on the basis of procedures performed or those that encourage the treatment of consumers as ‘objects’ in a process rather than as the focus of activity) are to be avoided. In this context, there is need to sift through the literature to determine (with a view to encouraging the use of) methods that cultivate an environment in which providers of all kinds have incentives that enhance the ‘agency’ role.

The third choice relates to those aspects of the health care system that may have no direct bearing on medical outcomes and which may therefore be less subject to scrutiny in terms of equity—things like hotel services, and so on. These, to a certain extent, depend on personal circumstances and often influence the choice between providers—public, private or NGO.

These choices are to a large extent influenced by the terms under which health services are provided under different settings, particularly taking into account the financial intermediary between the provider and the patient. This affects the range of (clinical) services available to consumers and there is need to encourage only those that entitle consumers to a comprehensive set of effective services that are provided efficiently, and equitably (this is taken up in the next chapter).

4.2.4 The private sector

How do we treat the private sector in our scheme of health services, particularly the proprietary providers? The approach adopted here is that these be judged in terms of what they enable to

be accomplished for patients. That is, we simply view them as instruments to furthering the objective of improved health status of the patients—that in the process of doing so they make profits is neither here nor there. If in pursuit of profits—particularly in a competitive environment—their activities can produce *outcomes that are both efficient and equitable*, then their activities should be judged on these counts only. The test question is thus: are the private for-profit providers any better at accomplishing *efficient and equitable* care? If the answer is an unqualified ‘yes’, then there is *prima facie* case for using the private sector proprietary providers. That is our stand, and its test is empirical rather than ideological.

4.2.5 The question of incentives

This thesis is largely concerned with questions of incentives that will enhance the patient-agent relationship in an efficient and equitable manner for the system as a whole. The view adopted is that if one incentive structure accomplishes the objectives of the health sector better than the alternatives, there is *prima facie* case for its use (provided the administrative and management costs of implementing the scheme do not outweigh the anticipated benefits).

In the next section we present some evidence which is indicative of inefficiency in the Kenyan public health system, and raise a number of questions about the alternatives available to thwart such inefficiencies, most of which are answered in later chapters of the dissertation.

4.3 Potential for Efficiency Improvement in the Kenyan Health Sector

The aim of this section is to provide some *prima facie* evidence to demonstrate there exists inefficiency or waste in the allocation and/or use of existing resources, particularly in the public health services, arising mainly from the way resources are allocated/used among/in various health activities: curative services, community and primary health services. In terms of the health service budgeting and accounting system these allocations are determined by the allocations of inputs used in their production—personnel, materials, drugs and supplies, maintenance of buildings, equipment and transport and so on. Following the framework of efficiency discussed above, economic efficiency would require that resources be allocated to such health sector activities such that their effect on health is maximised. But there are difficulties in determining whether this is the case for the health sector as a whole since its output is not easily defined and or measured, even where one’s concern is only for a disease-specific or for a narrowly defined health activity. Hospitals, for exam-

ple, provide a complex mix of services (to tackle different disease problems) whose impact on various outcome measures such as morbidity, mortality, disability, and quantity and quality of life vary considerably. Elsewhere, research has tended to concentrate on limited but measurable aspects of the health sector activities (Mills, 1990a), including the analysis of the cost-effectiveness of particular facility-based treatment patterns, or looking at the characteristics of patients attending particular types of facilities to see whether they are being treated at the 'right' levels—in terms of medical standards, (e.g. Heller, 1978). In this section, both these approaches, supplemented with international comparisons (where relevant), are used to examine whether the current patterns of allocation and resource use in the Kenyan health sector are economically efficient.

4.3.1 Some evidence of resource mis-use in the Kenyan public health care system

In chapter two, we showed that the health problems precipitating use of curative services are dominated by morbidity and mortality from largely preventable causes such as infectious and parasitic diseases, respiratory system diseases, accidents, injuries and burns. In terms of the framework for efficiency sketched above, there is a wastage of resources, both at micro and macro levels, because diseases that can be treated at low cost levels in the system continue to be attended to at higher cost levels in the system.³² The following example illustrates there is a potential for efficiency improvements in the present system. The PADS (1990) study found that in the 1988/89 fiscal year, the Government of Kenya spent Ks. 68.90 and Ks 17.41, respectively per outpatient visit in the Nakuru Provincial General Hospital (PGH) and the Naivasha District Hospital³³. On the other hand, the lowest MOH Health Centre and MOH Dispensary were found to have per outpatient unit costs of Ks 16.76 and Ks 11.36, respectively, but patients probably received lower quality care at the PGH (and the District Hospital) than in other levels of the same type of care. Figure 4.4 shows how.

Unit costs per outpatient visit are plotted along the left vertical axis. On the horizontal axis are the names of three hospitals (two MOH, one NGO), the least expensive MOH health centre, and the least expensive MOH dispensary. Along the right vertical axis is plotted the percent of a sample of outpatient diagnoses cross-checked to verify the extent to which the symptoms had been

³² Notice that we are talking of the relative or comparative cost of facilities at which these diseases are being attended to. Theoretically, it is possible that the cost per case prevented may exceed the cost per case treated. This is a different angle to the analysis which we suggest must also be considered but is not of immediate import for the present argument.

³³ It should be noted, however, that the study did not conduct the analysis for a specific disease (which would have given a more impeccable case for the observed cost differentials). Nevertheless, it is instructive.

correctly identified (were based on a review by an independent physician). The NGO facility had the lowest unit cost per outpatient (Ks 7.00), but this might not be a good guide to costs in NGO facilities compared to MOH facilities because it did not include volunteer labour costs. Nevertheless, there are two ways in which system inefficiency can be shown to exist. First, if patients could be referred to the private facility and the government meets the cost of treatment there, the NGO hospital could have treated almost nine *more* patients for the cost of one at the PGH³⁴, Naivasha District Hospital and the least cost MOH health centre each could have treated 3 more patients while the least cost dispensary could have treated 5 more patients, probably with more satisfactory outcomes than the PGH.

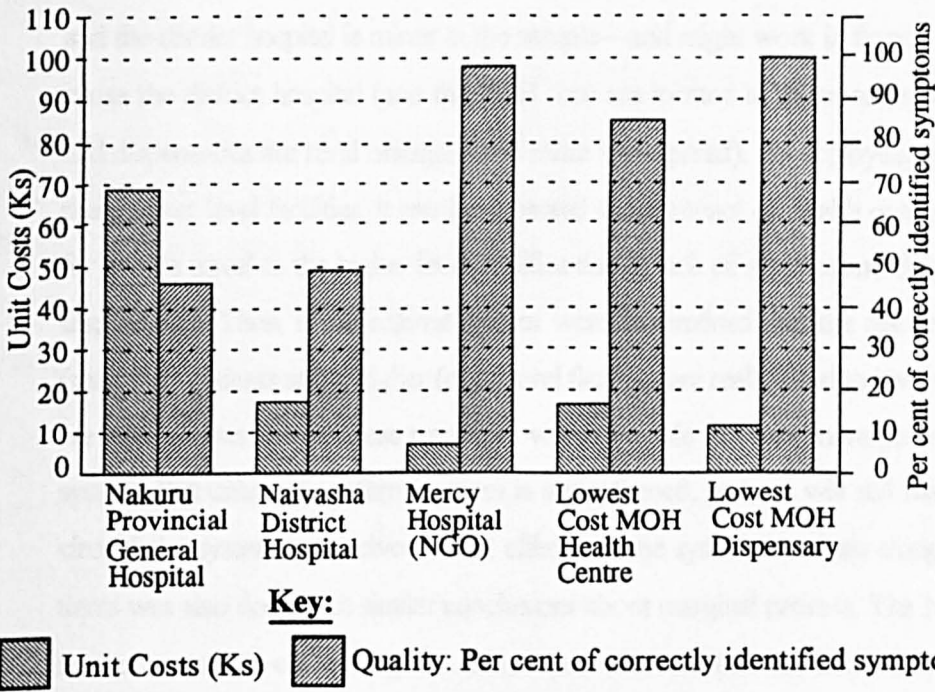


Figure 4.4: Comparison of costs and quality among MOH hospitals, health centre and dispensary and an NGO hospital, 1988/89

The second way of illustrating *prima facie* inefficiency is that of examining whether any of the patients treated at any one of the facilities in the study could have been treated elsewhere 'probably' with equally good results but at a lower cost (to society and perhaps the patients them-

³⁴ Although such a comparison is not perfect as the cost is often also affected by case mix, complexity and other factors. However, the general point is that referring patients to the NGO facility and footing the bills rather than treating them in government facilities may result in more patients getting treated at the same cost, provided two conditions obtain: (1) the quality of care does not deteriorate in the NGO facility as a result, and (2) the cost (both to the government and patients) does not defray any potential gains to be had from the exercise.

selves—travel, waiting, and so on). The PADS study found that out of an average of 300 outpatients treated in the PGH each day, only 20 per cent were referred there. That meant 240 or 80 per cent of patients were ‘self-referred’. Of these self-referred patients, 80 per cent (192) had left a facility nearer their home to come to the PGH, 15 per cent because of lack of drugs at the facility (closer to them), 20 per cent because they considered services there poor. Taking other ‘trivial’ excuses for coming to the PGH into account, the study concluded that about one third of the patients seen at the PGH did not have any genuine reason to be there. A similar analysis done for Naivasha District Hospital and health centres, estimated that one quarter of outpatients in the district hospital could have had their ‘needs’ (see next chapter on more about the meaning of this term) satisfied at a health centre near their home, while between 13 and 17 per cent of patients visiting health centres could have been attended at a dispensary. The difference in unit costs between the health centre and the district hospital is minor in the sample—and might work in favour of the health centre because the district hospital (and the PGH too) are located in urban areas, while the health centres and dispensaries are rural oriented (and more widespread). By deploying more resources towards these lower level facilities, it can be expected more impact on health outcomes would be realised, for people travel to the higher level facilities due to lack of supplies and drugs in health centres and dispensaries. Thus, if the referral system were streamlined and the resources used to treat such (marginal) patients at the higher (cost) level facilities are reallocated to lower level facilities, then, as we have shown above, these resources would be able to cause more persons to be treated in the system. But unless the referral system is strengthened, patients will still have incentives to ‘short-circuit’ the system with adverse cost effects on the system. Analysis along similar lines for inpatients was also done with similar conclusions about marginal patients. The Nairobi Area Study had similar comments concerning the referral procedures within the City health services. (It is one of our objectives to propose a reform that will streamline this referral process).

There is also another ‘hidden’ potential for efficiency. Reallocating resources in the manner described, would cause some of the lower level facilities with high per unit costs due to excess spare capacity to improve on performance since in these facilities manpower and other inputs may presently be under-utilised in many of them, provided other factors preventing people from using them (quality, and so on) are overcome. However, the crucial issue here is determining the desirable mix of facilities that results in an appropriate balance between the higher level facilities (particularly hospitals) and others in the system (and not whether the higher level facilities should be there) because the higher levels have the important supportive role in terms of supervision and referrals (Mills, 1990a). The application of the efficiency principles discussed in the sense discussed

here can help to ensure this balance is got and kept right. Given the scenario (about the pattern of diseases) sketched in chapter two and the evidence above, which is illustrative of the public health services, the next question to ask is 'what has been done concerning the resource allocation to alleviate these problems?'. In the next sub-section we show little has actually been done to realign resource allocation and the country's health problems.

4.3.2 The MOH resource allocation is *not* geared to the country's health problems

Besides the micro issues raised above we can also show that the pattern of resource allocation at overall (public) system level has not been altered in response to the health needs of the country. We concentrate on the public health sector due the paucity of private sector information, but also because the allocation pattern within the public health sector is of more immediate policy relevance since the government can influence the health sector directly and indirectly through the way the health ministry allocates its resources. The categories used in the analysis concentrate mostly on recurrent expenditures. Capital expenditures can reveal little else that will not already be evident from the recurrent allocation. The allocations to the eight main functional categories of expenditure used by the Ministry of Health (MOH) in Kenya are shown in Table 4.1 for selected years between 1973/74 and 1991/92³⁵.

Year	1973/74	1978/79	1983/84	1988/89	1989/90	1990/91	1991/92
General Administration and Planning	5.15	4.57	8.42	4.34	3.18	2.96	3.25
Curative Health	68.34	66.92	67.65	65.68	68.81	68.2	69.75
Preventive Medicine and Promotive Health	6.6	5.85	5.26	11.7	10.72	9.73	7.41
Rural Health Services	13.03	9.63	10.97	11.52	12.08	12.82	13.29
Health Training	6.64	9.25	6.63	6.06	5.07	5.77	5.63
National Health Insurance	0	0	0	0.03	0.59	0	0
Medical Supplies Co-ordinating Unit	0.24	1.91	1.06	0.66	0.72	0.53	0.67
Medical Research	0	1.87	0	0	0	0	0
Total	100	100	100	100	100	100	100

Compiled from data on the MOH Recurrent Expenditure Estimates, various years.

Table 4.1: Functional distribution of recurrent health care expenditures: 1972/3—1990/91 (various years—per cent)

³⁵ The data used in the analysis are based on 'actual' or realised allocations, rather than planned expenditures.

The overall picture can be gleaned from the trends in more recent years. Allocation to preventive and promotive health care programs declined between 1988/89 and 1991/92. Although under the terms of the Health Care Financing Project government was to maintain its funding of the MOH at least at levels that prevailed in the fiscal year 1988/89, and to increase its funding for primary and rural health care, and preventive/promotive health services, none of this appears to have occurred. Although rural primary health services recorded an increase in the share of the budget, *that increase appears to have come completely out of the share devoted to preventive and promotive services*. On the other hand, expenditure on institutional care, during the period actually rose, although we saw in chapter two that Kenya's health problems are mainly public health oriented.

The factor contributing to most of this growth was expenditure on public hospitals—mainly provincial and district hospitals, which alone accounted for 78 per cent of the total allocation for curative services in 1991/92, a representative figure for the period. Expenditure on psychiatric hospitals was only 3 per cent of the curative services budget. The other major component in the curative services bill is the Kenyatta National Hospital, which in 1991/92 had a share of 15 per cent. The latter, when combined with the provincial and district hospitals' share means hospitals alone took about 94 per cent of all the curative care budget (see table 4.2 below).

Purpose of Expenditure:	Percent
Kenyatta National Hospital	15.2
Provincial General Hospitals	20.6
District Hospitals	57.7
Psychiatric Services	3.3
Grants to NGOs	1.2
Spinal Injury Hospital	0.5
Biomedical/Hospital Engineering Services	0.2
Dental Services	1.2
Total	100

Compiled from MOH data on Recurrent Expenditure estimates, 1991/92

Table 4.2: Distribution of curative services expenditures, 1991/92

The other components of curative services allocations are to non governmental organisations; biomedical engineering; very specialised care institutions such as the spinal injury hospital;

and, dental services. These had very small allocations in 1991/92. Such a distribution of hospital expenditure has implications for the accessibility of health care services—these as well as the term ‘access’ itself will be taken up in the next chapter. Another notable observation from the table is the virtual decline to zero of allocations for medical research. Without research into the country’s health problems and how they can be tackled, resources will continue to be allocated to areas where their contribution to society’s health are hardly known.

The development budget does not present any consolation either. Although the share allocated to curative care between 1973/74 and 19991/92 declined from 76 per cent to around 34 per cent, while the share of rural health services increased from 4 per cent to 24 per cent, and that of preventive services from 5 per cent to 8 per cent, these changes have not been sustained in the period 1988/89 to 1991/92 (during which the shares of preventive and promotive services, and rural health services *declined*, whilst that of curative services remained almost unchanged). On the other hand, over the period 1973/4—1991/92, health training allocations declined from 15 per cent to a mere one per cent, while general administrative and planning’s share rose from almost zero to about 32 per cent—clearly there has been some juggling of resources in this budget too.

In terms of international standing, Mills, 1990a, comparing data from various studies of the health sector in developing countries concluded that:

- ◆ hospitals absorb approximately 30-50 per cent of health sector expenditure
- ◆ hospitals absorb approximately 50-60 per cent of current government health sector expenditure
- ◆ hospitals absorb approximately 60-80 of government national health facility expenditure and possibly 70 per cent of district level health expenditure
- ◆ around 60-80 per cent of hospital expenditure can be absorbed by central and general hospitals, the remainder going to district hospitals.

Lack of data on private sector hospital expenditures makes it difficult for us to conclusively argue Kenya’s position as far as the first count is concerned. As for the other counts, in 1991/92:

- ◆ Kenyan hospitals absorbed about 69 per cent of recurrent government health expenditure, which was on the upper side compared to Mills’s findings (taking into account the caveats about the representativeness of her data)
- ◆ about 84 per cent of government national health facility expenditure went to hospitals
- ◆ central and general hospitals absorbed about 40 per cent of the government budget³⁶.

³⁶ This proportion excludes private sector facilities.

Thus, according to available information, Kenya's hospital sector takes a more than proportionate share of the health sector resources.

There are not many studies on cost-effectiveness of various health interventions from the developing countries, but there is credible evidence that in general preventive care can be a more efficient form of intervention than curative care. The 1993 World Bank's World Development Report devoted mainly to health issues asserts:

Spending that reduces the incidence of diseases can produce big savings in treatment costs. For some diseases, the expenditure pays for itself even when all the indirect benefits—such as higher labour productivity and reduced pain and suffering—are ignored. Polio is one example. Calculations for the Americas made prior to the eradication of polio in the region showed that investing \$220 million over fifteen years to eliminate the disease would prevent 220,000 cases and save between \$320 million and \$1.3 billion (depending on the number of people treated) in annual treatment costs. The program's net return, after discounting at even as much as 12 per cent a year, was calculated to be between \$8 million and \$480 million. (World Bank, 1993, pp. 19—20).

This report also argues that the decline in communicable diseases—largely preventable—particularly those of childhood, in the high income countries has led to much of the decline in mortality in those countries (*ibid.*, p. 23). Indeed, world-wide, the burden of disease—measured in terms of disability adjusted life years (DALYs)—largely stems from diseases that are preventable, as the following table shows. Of more immediate relevance to the present study, the report, commenting on childhood diseases concludes

The burden of these largely preventable or inexpensively curable diseases of children is far larger in sub-Saharan Africa (*ibid.*).

In summary, chapter two showed that most of Kenya's health problems are largely preventable. This chapter has shown that the health care resource allocation has not paid much attention to this important fact. The evidence quoted above shows that preventing is likely to be less costly than treating (at least for most diseases in the country). Given the epidemiological situation and the characteristics of the Kenyan population (as discussed in chapter two), it is apparent that a substantial amount of the health care resources should be devoted to preventive services. Yet, the allocation of resources to curative services in Kenya remains high and unchanged, consuming over two-thirds of the public health sector budget. A juggling of resources between the preventive and promotive health services, research and training to increase the funding of rural primary health services appears to have taken place. This will not be helpful at combating the real health issues since rural primary services also include ambulatory clinic-type curative services. The result may well have been a fall in real resources devoted to purely public health activities.

Sex and outcome (DALY)	Disease and category		
	Communicable ^a	Noncommunicable	Injuries
Male			
Premature death	259	152	70
Disability	47	146	39
Female			
Premature death	244	135	33
Disability	74	142	20

Note: DALY, disability adjusted life year.
a. Includes maternal and perinatal causes.

Source: World Bank, 1993, Table 1.1, p. 25.

Table 4.3: Burden of disease by sex, cause and type of loss, 1990.

Thus the persistent disease patterns that can be effectively combated through community health and preventive programs have not yet led to major changes in the budget, and the current trends to do not suggest such a change is likely to occur under the current arrangements. The main reason for this impasse, we conjecture, is that recurrent spending is much tied with the existing distribution of the health capital. Another reason is the contractual arrangements presently in force for public health service personnel—most of them are civil servants. Any change in the system that does not change this contractual relation, which could largely be blamed for absence of efficiency incentives, goes only a part of the way. Mechanisms to implement such changes exist and this thesis will discuss some later. The next sub-section shows some juggling of resources appears to have occurred in favour of personnel costs, the single largest component of the health services budget, in addition to indicating the particular items on which the juggling of resources occurred, in its favour.

4.3.3 An alternative 'closer' look at the distribution of recurrent curative expenditures

The functional allocation described above reveals the allocation of resources in the Kenyan public health sector is a biased towards hospital institutions, but that alone cannot tell us what it is in the system that the resources are used for. To get this, we have re-classified the data by inputs. Although expenditure breakdown by inputs will not give us information about the implications of the present allocation patterns (it does not correlate the expenditure on inputs to the actual services provided or their benefits), it nevertheless can be useful in revealing changes over time in that it may pinpoint allocation problem areas in the system. The (macro) data on expenditure by input category used in this section differs slightly from that used in the previous analysis since in reclassi-

fying the data, we found items recorded under recurrent expenditure which belong to capital expenses—such as construction of buildings and purchase of vehicles. These are omitted from the analysis. The result is shown in Table 4.4 below.

The largest input-cost component, as would be expected of a labour intensive industry such as the health sector, is personnel costs, which increased from K£31 million in 1982/3 to K£36 million in real (1982) prices between 1972/73 and 1990/91 (although in 1986/87 they rose to K£45 million.). The main cost component in personnel costs is salaries, wages and personal allowances which accounted for over 58 per cent (on average) over the period³⁷. On average, the share of personnel costs in the total MOH recurrent cost bill rose from about 53 per cent in 1982/3 to 60 per cent in 1991/2.

	82/3	83/4	84/5	85/6	86/7	87/8	88/9	89/90	90/91	91/92	Avg. ¹
Personnel costs	54.5	59.4	55.3	62.5	67.4	59.3	60.5	54.8	53.5	57.5	58
Building & equip. maintenance	0.3	0.4	0.7	0.5	0.6	0.7	0.8	0.6	1.1	1	1
Transport operating expenses	4	3.8	3.8	3.7	3.1	2.7	2.9	3.3	4.1	3.6	4
Drugs and medical supplies	19.5	14.6	20.2	14.1	12.7	13.1	12.3	14.9	3.8	14.8	15
Utilities	4.3	3.9	3.9	4	3.3	2	2.1	3	2	2.9	3
Subsidy/grants to organisations	8.6	8.4	7.2	6.5	5	15.8	15.3	16.4	17.3	12.9	11
Other expenses	8.8	9.6	8.9	8.7	7.8	6.3	6.1	7	8.2	7.2	8
¹ Period average											

Compiled from the MOH's Recurrent Expenditure Estimates, various years.

Table 4.4: Recurrent expenditure allocations by inputs, 1982/3—91/92: (Percentages)

The second major cost component was drugs and medical supplies, whose share declined from 19 per cent in 1982/3 to just above 12 per cent in 1991/2. Practically all the other items in the cost bill experienced a similar fate during the period, with the exception of subsidies and grants to organisations which increased from less than K£5 m to around K£10 m between 1986 and 1987.³⁸

There are no standard set rules to judge the efficiency of resource allocation by input category, but de Ferranti (1983a) suggested a general rule of thumb³⁹:

³⁷ Other cost components in this category include gratuity and pensions contributions, passage and leave allowances.

³⁸ This item is actually not an input, but is included for completeness. It represents the costs to the MOH of handouts to organisations such as the Flying Doctor Services, mission hospitals and international organisations such as the World Health Organisation, etc.

³⁹ This is based on cross-country comparisons of about a dozen countries' experiences. This does not of course make such 'rules of thumb' authoritative, but they can act as useful guides for 'general' al-

“Investigate further whenever

- salaries account for less than 40% or more than 80 % of total recurrent costs,
- drugs are less than 10% or more than 35%, or,
- transportation is less than 5% or more than 15%,
- maintenance and repair of buildings and equipment are less than 1.5% and 15 %, respectively, of replacement value. ”(de Ferranti, 1983a, p. 50].

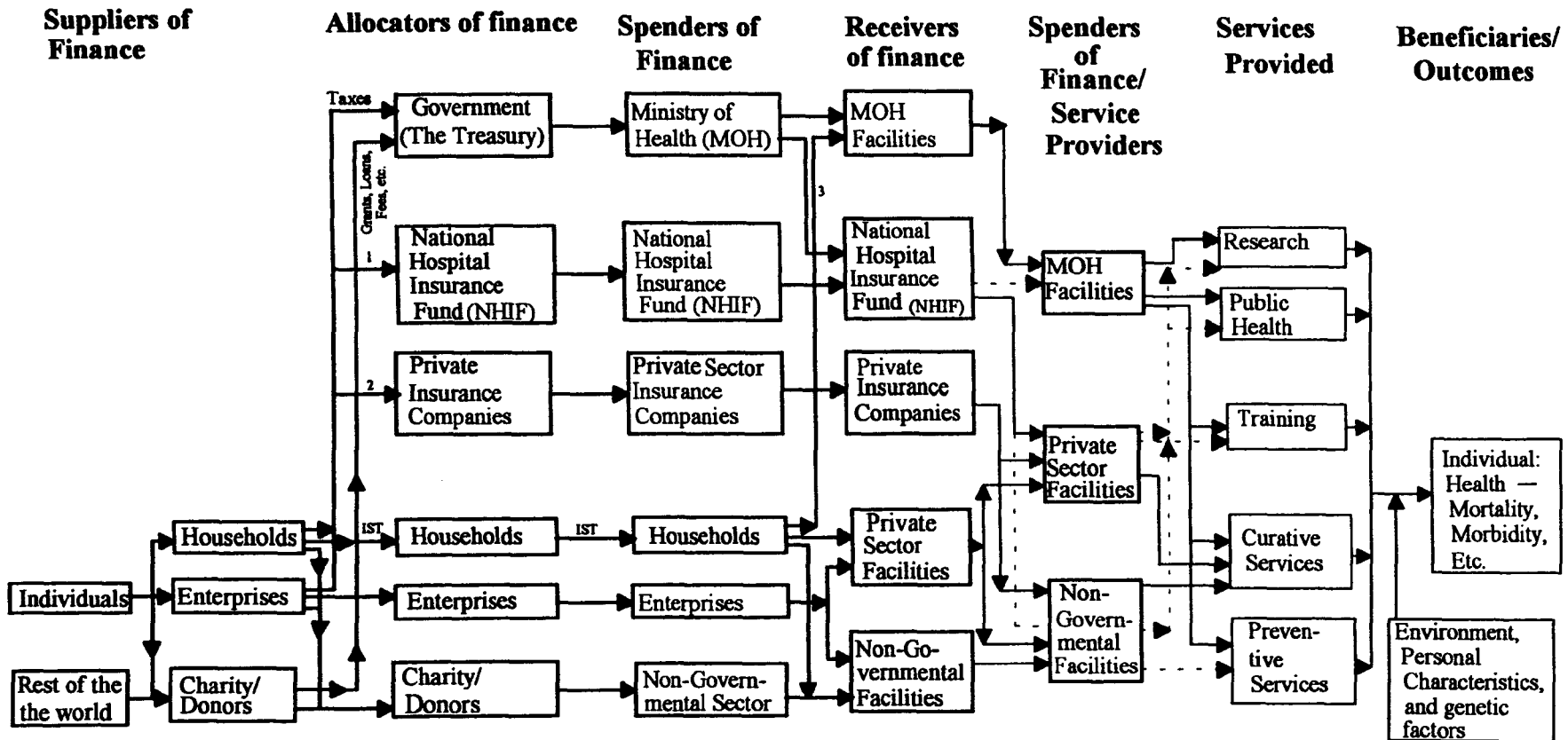
Mills (1990b), reviewing the literature on developing countries' hospital cost structures came to similar conclusions. According to this comparative approach, Kenya is *prima facie* not providing enough resources for building and equipment maintenance and transport. Moreover, judging by the pattern evident in Table 4.4 (of declining share of these two categories in total allocation), the situation worsened over time. Various *ad hoc* studies have pointed out some of these deficiencies (e.g., MOH, 1990). The problem of insufficient allocations for maintenance and transport, and other input categories stems from the large bill incurred in personnel costs, that continues to rise. Because of the limited budget allocation from the central government, some items were squeezed, including drugs and medical supplies. This, as the review of previous studies on the Kenyan health sector has shown, has impaired the efficiency of the government funded health service.

The next section outlines a simple framework that will be used later as the basis to develop some structural reform proposals for the system.

4.4 A Framework for Analysing Health Sector Problems in Kenya

This section presents a schematic model of the interaction between various participants in the Kenyan health system, identifying the suppliers of finance for provision of health services and the users of those funds (see figure 4.5 below) with a view to isolating those areas where inefficiencies (with the types of inefficiency and their implications identified) might be expected to arise.

The expenditure on health services by the government is only a part of the total funds available for 'health production' in the country. This is because the health sector derives its finances from sources other than the public sector. For example, households contribute a sizeable amount through private purchase (mostly in the private sector). Also, some finances to the health sector come from donors, both residents as well as from external sources (what we call here the 'rest of the world').



Legend:

- 1: Employer/Employee Contributions to the National Hospital Insurance Fund
- 2: Private individual and enterprise purchases of health insurance
- 3: User-fees paid directly to Ministry of Health (MOH) facilities by patients
- ▶ Weak linkages in the current system.

Figure 4.5: Flow of funds and inter-relationships among participants in the Kenyan health care system

Reliable data on the expenditure and financing of health sector expenditures from various sources is not collected routinely in Kenya⁴⁰. The only attempt (known to us) at generating this type of data is that by a World Bank *ad hoc* mission on 'Expenditures and financing of the health sector in Kenya' (MOH/World Bank, 1986). According to these estimates, the Ministry of Health (MOH) is the (single) largest financier of health services, as Table 4.5 below shows. Data given are for the fiscal year 1983/84 but are representative of all years, including today. The Table shows the total funds emanating from the various sources in 1983/4.

Source	Finance (K£ m)	Per cent of Total
1 Government:		
Ministry of Health	1,206.8	42
College of Health Sciences	25.4	1
Non-free Appropriations-in-aid of MOH services	3.7	<1
Municipalities	152.6	5
National Hospital Insurance Fund (NHIF)	109	4
2 Missions:		
Foreign contributions	13.1	
Local contributions	9.4	} 1
Non-fees earnings	6.8	
3 Other Non-Governmental Organisations:		
Foreign contributions	11.3	<1
Local contributions	2.1	<1
4 Private companies	53.6	2
5 Private Insurance	35.3	1
6 Out-of-pocket payments	1,175.6	42
7 International donors	71	2
TOTAL	2,876	100

Source: MOH/World Bank Report on 'Expenditure and Financing of the Health Sector in Kenya, Ministry of Health, February, 1986.

Table 4.5 Estimates of financing of recurrent health expenditure by source of finance, 1983/84 (Ks mn)

The two largest sources of finance by far were central government revenue through the Ministry of Health (42 percent) and the out-of-pocket payments (41 percent of total⁴¹). The Na-

⁴⁰ That is besides the public sector information.

⁴¹ Although we have doubts on the representativeness of this figure, especially given the attitude of the mission that estimated it—that patients should pay for the government health services they receive in full through various cost recovery arrangements, as they do in the private sector. The 1982/83 urban household budget survey showed that households were spending only 0.8 % of their income on health related expenses. The low income earners spent only 0.6 % of their income on health, while the middle

tional Hospital Insurance Fund provides only about 4 percent of the recurrent health expenditure, much of it paid to private and mission facilities. Much of the private health insurance is in the form of company group schemes. International donors also support some recurrent expenditures to the voluntary sector and to the government, but their contribution is only small at 1 percent of the total health budget.

Therefore, we can say that in the 'production of health' in Kenya, three groups of 'strategic' decision makers are involved—households (as individuals or groups of them), the private sector, and what we loosely term as 'charitable organisations' or simply 'others'. However, because in the end all finances for funding health services must come from individuals, the above classification is only adopted for convenience (the diagram of the linkages in the health sector reflects this fact). These groups vary in their relative importance in the system. They perform four main functions (some of which are 'delegated' to other third parties) either as *suppliers, allocators, spenders, receivers and users of finance, producers of services and beneficiaries*, in their various capacities/roles as householders, workers (in enterprises) or benevolents.

As *suppliers* of finance, these groups deny themselves, or are denied, the use of resources in other activities in order that these resources are used in the health sector. Here, the chain of transactions may be simple or complex, or, put in other words, horizontal or vertical—simple (horizontal) in that suppliers perform the next stage (*allocation*—discussed next) directly, or complex (vertical) in that the second role is transferred to another party. Thus, in the first case, the suppliers of finance also act as allocators of resources, in the second case, this decision is passed to other parties (e.g. the government). An example of the first (horizontal) case is where households, enterprises and benevolents (these last in their individual capacities) have leeway in deciding how to spend their resources, while in the latter resources are transferred to government (compulsorily as taxes or fees), to the NHIF (compulsorily as mandated contributions of individuals or enterprises), to private insurance companies (as premiums), or to charitable organisations or other suppliers—a case of *intra-sectoral-transfers*, shown as 'IST' in the diagram (e.g. through 'harambee'⁴², for example).

Suppliers of resources (individuals) are the ultimate receivers of the services of the health sector and their welfare is one of major concerns of this thesis. Therefore, according to this framework, individuals finance all health care expenditure by paying insurance premia or by *self-insuring*—otherwise termed out-of-pocket expenses, or through the pooling of resources through third

and upper income households spent 0.8% and 1.1% respectively.

⁴² See footnote 24 above.

party payers such as the government (though tax-collections that finance the services (both from individuals as well as corporations)), the National Hospital Insurance Fund, and through the place of employment. Within this framework and available statistics, the public sector's contribution is by far the largest source of funding for the health care sector,⁴³ providing over 50 per cent of its total recurrent and development expenditures. This contribution has been rising nominally at seven per cent per annum in recent times. Enterprises finance health care by paying insurance premia, in whole or in part, by self-insured arrangements and/or by providing some clinic services themselves. It was estimated in 1988 that about 60,000 persons and their families had some group health insurance cover (Vogel, 1990a, b). The benefits and premiums of group covers vary by company. Most policies operate on a reimbursement basis and include co-payments and are subject to financial ceilings for claims under certain headings, e.g., operation charges, etc. These benefits are besides those obtained under the NHIF. This has often led to too frequent and frivolous use of in-patient services. Most private health insurance schemes favour the middle and upper income classes in Kenya, as elsewhere (Vogel, 1990a, b). A few companies dominate the health insurance industry. This is apparently perplexing given that private motor insurance is already widespread. But the private insurance market for hospital care seems to have been pre-empted by the NHIF, established in 1966. This Fund is similar to Social Security and covers about 2.1 million people (roughly 10 per cent of the total population).

Thus health sector expenditures, regardless of the financing procedures used, are in the end borne by people—either as households, individuals (users) or tax-payers, and there is interest in assessing how this burden is distributed. Data on who bears which types of health expenditures is not available. Consequently, one cannot easily tell how much is borne by the poor, and so on. Nor is data available categorising expenditures in urban/rural classification. Hence it is hard to comment on the distribution of the health care expenditure burden—either between users and non-users, or between the urban/rural dimension. The distribution of the burden by income level is usually a dominant concern. Since the data to allow this type of analysis for Kenya is lacking, the burden implications of the Ministry of Health expenditure is here examined with the aid of data on the composition of central government revenue. According to this data, only about 29% of the health expenditures funded by central government comes from direct taxes. The greatest proportion of health expenditures funded by the central government comes from indirect taxes—whose average contribution to the exchequer between 1963/4 and 1990/91 averaged 58%. Only a small propor-

⁴³ Note: Except for the support given to certain mission/charitable organisation facilities, the government does not fund privately provided care, nor can 'public patients' be treated in private facilities unless they make private payment arrangements.

tion of the health expenditures (less than 15%) comes from other sources. Given the large share of indirect taxes in central government finance, and that it is possible for some corporate tax burden to be shifted to households, one may safely conclude that the tax system is on overall regressive. Therefore the rich may not be contributing a large share of the proportion of government revenue to finance the MOH expenditures. The pooling of resources involves fiduciary relationships in which an institution spends the money on behalf of the consumers. These arrangements have implications for access to health services (or more formally, equity). Chapter five will try to give an interpretation to the equity conception in Kenyan health policy, and, among others, discuss the meaning of equity in financing and equity in service provision.

Those that receive the funds may be termed as the *allocators*. As already evident from the diagram, allocators could be the households, enterprises or some other 'third party' (government, NHIF, private insurance company, or a benefactor). We see from the diagram that allocators of resources are almost synonymous with *spenders of resources*. They allocate resources to different *receivers of these resources*. These are the providers of health care services. In Kenya, as in other countries, there are three types of providers: the MOH, non-governmental, charitable or non-profit organisations (NGOs), and for-profit private practitioners, clinics and hospitals. The figure below shows a matrix of the financing and provision of health care relationships.

FINANCING	PROVISION		
	PUBLIC	PRIVATE	
		FOR PROFIT	NOT FOR PROFIT
TAX	The government health service—hospitals, health centres and dispensaries	Non-existent	Support to mission and other 'charity' facilities
INSURANCE	Limited	Private practitioners and private hospitals	Mission hospitals
USER-CHARGES	Government health service	As in the above cell	Mission hospitals
OTHER: —aid —charity	Some government and other 'charity' facilities	Not known	Mission hospitals

Figure 4.6: A matrix of health care financing and provision in Kenya

Table 4.6 shows the estimated shares of total expenditure handled by various types of providers for the year 1983/84. The issues arising here border on allocative efficiency and these issues

will be discussed in chapter six. (Our main interest there will be to examine how government should allocate resources availed to it from the various sources, and implications thereof).

The users of funds (providers of health services) *perform the important function of allocating the resources*—by deciding from whom to purchase services. There are three sources—the MOH facilities, non-governmental (private) organisations, and the for-profit providers. Receivers also have to decide how to allocate the available resources to various services provided by the health sector. The decisions made at this level are governed both by 'operational' or technical efficiency as well as effectiveness considerations. These are the issues most previous studies on the sector have been concerned with. The users of funds can be seen to allocate resources to five different types of activities—research, training, curative, preventive/promotive, and public (or community) health services. This is done for the benefit of individuals, who are the final beneficiaries of the activities of the health sector. These allocations, combined with other factors, such as personal characteristics, environmental factors and genetic factors determine (or produce) the final output of the health sector—the individual health status (which may be measured by aspects such as mortality and/or morbidity rates, for example).

Provider	Expenditure (K£ m)	Per cent of Total
1 Government:		
Ministry of Health	1,210.4	42
College of Health Sciences	25.4	1
Municipalities	160.6	6
National Hospital Insurance Fund (NHIF)	8.4	<1
2 Mission facilities	68.9	2
3 Other Non-Governmental Organisations:	36.8	1
4 Private companies	11.4	<1
5 Private Market:		
Institutions	264.8	9
Practitioners	216	8
Drugs (non-institutional purchases)	680	24
Other out-of-pocket expenditures	93.3	3
TOTAL	2,876	100

Source: MOH/World Bank Report 'Expenditure and Financing of the Health Sector in Kenya.' Ministry of Health, February 1986.

Table 4.6: Total gross recurrent expenditure by type of health care provider, 1983/84 (Ks millions)

These allocations have economic significance because their financing has different impacts depending on the source of finance. If reliance is on market forces for example, there will be an inadequate financing and therefore consumption of public health or community services unless the government intervenes because it is easy for people to benefit from these services without paying the full cost. Such services include vector control programmes. In this case, one household 'free-rides' on another's spending because the benefits cannot be internalised. An example of this problem is the control of mosquitoes (carriers of malaria causing vector). The control of mosquitoes (e.g. through spraying) by one household without a corresponding effort from other households soon leads to realisation that 'it is not worth the effort' because mosquitoes—the carriers of the malaria causing vector—have no respect for 'property rights'—without organised community-wide control of the vector, one household's efforts are ineffective. Such services therefore have spillover effects (externalities) which are hard to internalise and only community-wide intervention, through government taxation and spending, can ensure enough quantities (of them) are provided. The public intervention solves the *'free rider'* problem while reducing the user cost of service and therefore encourages use (Culyer, 1971). This characteristic, that often distinguishes health care from other goods and services, prevents the efficient allocation of it in a competitive market from being realised. Externalities occur when a third party receives some benefit or suffers some loss without choosing to do so (Culyer, 1971; Cullis and West, 1979). In general, there are externalities when someone's production or utility function is affected by others' acts of production or consumption, or, when someone's utility is affected through feelings of altruism or envy by changes in another's well-being, and these effects are not paid or compensated for. These are the interdependence condition and non-price condition, respectively (Dasgupta and Pearce, 1972, p. 118). Externalities cause the market value (private or internal cost or benefit) not to coincide with the total value (social benefit or cost) of these activities or changes. If the social benefits are greater than the private ones, there are positive externalities (external economies). If the social costs are greater than the private (internal) ones, external diseconomies are said to exist.

There are four possible sources of externalities in health (Culyer, 1978)—altruism, concern for financial or economic impacts of ill-health, concern for equality (individual as well as geographic), and egoistic behaviour, all relevant for Kenya. All call for some form of collective intervention, though not necessarily public provision of it. These give health services some aspects of *public goods* and the competitive market cannot function efficiently to ensure their efficient and equitable distribution⁴⁴. We showed above that many health problems in Kenya are of this nature and

⁴⁴ Note, however that the 'publicness' of a good is not a sufficient condition for government intervention. The pareto efficient allocation may still be attained in this case (see e.g., Cowell, 1986, chapter

this has to be taken into account by any reform strategy that encourages reliance on market forces to determine services offered by service providers. But this does not mean that government should be involved in the production of such services. Instead, mechanisms can be put in place to ensure that the efficient level of such care is produced in the market even in the presence of externalities.

On the other hand, curative services mostly benefit only individuals who consume them. In this case, there is little difficulty in *inducing individuals to adequately finance them*. Government intervention might be necessary because the poor may not have adequate access to such services or because a catastrophic illness may impoverish even a middle income-class household. But the government does not have to foot the full cost of the service. Instead, people might be required to contribute, according to 'their ability to pay', up to a certain maximum (per annum, for example).

Preventive and health promotive services fall between the two polar extremes discussed above. They have characteristics of both public and curative health care services—people are willing to pay for such services but they may not purchase enough of them unless some incentives or penalties are used to induce them to purchase enough quantities. However, in general, inadequate financing of public and preventive/promotive health services eventually leads to high expenditures for curative services. Treatment tends to be more costly in terms of financial cost, and since inadequate prevention/control allows a greater proportion of the population to pass from the less severe (less costly to treat) phases of illnesses, a greater demand for curative services occurs, which further crowds out financing for public health. The system gets 'grid-locked' on a treadmill where more and more is spent on curative services (easily justified because of high demand for services) but less is invested in public health and prevention activities that might actually reduce the need for current levels of spending on curative care. The analysis of the public sector allocations above has shown that the system is currently 'caught in this trap' which suggests mechanisms are needed to correct the resource allocation patterns.

The MOH currently provides, through the public health service delivery system, all five types of health activities discussed above.⁴⁵ NGOs provide preventive and curative services, and in some areas public health services. The private (for-profit) sector providers concentrate on curative services, but also offer preventive services to those who pay for them. These providers cannot be ignored as they provide a substantial proportion of health services in the country. However, their objectives are at times at variance with those of the state. For example, the for-profit providers are

8).

⁴⁵ However, the training of doctors is under the Ministry of Education. Only nurse and other para-medical personnel training is charged to the MOH vote.

geared towards maximising profits, but due to the problems mentioned above, ignore public health. Furthermore, there is no integration with government services, besides attracting professional personnel from the public sector. The NGO providers (and some private providers too, especially those in poor areas⁴⁶) may sometimes provide poor quality services. But both NGO and private (for-profit) providers are capable of delivering public and preventive health services but do not provide these services due to the problems discussed above. Also, at present, these providers also provide limited training of health personnel. Government's decisions about how to finance and deliver these services affects the development and viability of non-government providers because it can choose to purchase them from other providers or deliver them. These decisions affect both the quality and quantity of health services available to the population. Given appropriate incentives, these providers can play an increased role in the provision of all the four types of services and we later show what type of incentive structure needs to be implemented to make this possible, following the criteria discussed in section 4.2 above (see the sections on private sector and incentives). This is one other objective of this dissertation.

4.5 Summary and Conclusions

This chapter was concerned with the development of the basic principles that should govern the efficient provision of health care in Kenya, providing *prima facie* evidence of some aspects of inefficiency, particularly in the public health services and the development of a framework within which efficiency issues can be analysed. The basic efficiency principles developed here stipulate that the system should

- ◆ provide only medically effective,
- ◆ provide such services cost-effectively,
- ◆ concentrate on cost-effective services that offer highest pay-offs in terms of health, and
- ◆ provide such services at appropriate scales, i.e., scales that ensure the share of resources going to health activities secures maximum possible benefits that in any other alternative use of the same.

Following these criteria, and in conjunction with the epidemiological and demographic profile of the Kenyan population, it might turn out to be more rewarding in the long-run to shift resources towards preventive services for which there is evidence elsewhere that the payoffs are far better than those on curative services. The review of the *prima facie* evidence of inefficiency in the

⁴⁶ These, motivated by profit maximisation, may cut-costs by providing low quality care. They may rely on unqualified, less costly personnel, or simply ignore standard hygienic conditions.

system showed that there are various forms of inefficiency, both at micro and macro levels. The main micro inefficiencies concern the inappropriate use of health resources particularly by consumers, while the macro level inefficiencies concern the way resources are allocated to various activities in the system as a whole or, within facilities, allocations to various health problems.

The health policy decision making framework was used relate the above inefficiency aspects to various levels at which decisions are made and a model of the health sector was presented in which various types of problems facing the system were outlined.

The main conclusion arrived at in this chapter is that there are aspects of inefficiency in the Kenyan health care system that need redressing if the system is to maximise the returns from the currently available resources in terms of their impact on the health outcomes for the population. However, besides efficiency, equity is the other cardinal objective of the health system. We turn to this objective in the next chapter.

5. THE EQUITY OBJECTIVE IN KENYAN HEALTH POLICY

5.0 Introduction

The question of equity in health and health care in Kenya has until recently been largely ignored or only received scant attention of policy makers and legislators. Resource scarcity which we saw led to concerted efforts to improve efficiency in the system does not appear to have elicited an equally important question: since there aren't enough resources to provide for all people's health needs, what can be done to ensure that whatever is available goes round in an equitable manner? Achieving efficiency gains within current activity levels—or even increasing the level of spending—is not synonymous with greater equity, as policy makers might (mistakenly) want to believe, unless the concept of equity is a 'threshold' type (e.g. decent minimum) and the efficiency gains can enable more people to receive the threshold level. But even though threshold-type policies (of equity) may shift whole distributions upwards—thereby (probably) pushing lots of people above the threshold level—there would still be inequalities, since it applies only to those at the lower end of the distribution. The recent reforms illustrate this problem well—they brought into sharp focus the need for a clear policy on access to health services. This chapter shows that the current policy does not address the relevant issues and therefore will not lead to significant improvements in the equity situation unless changes to the orientation of health policy are implemented. The historical perspective of the concern for equity in health services shows most 'equity' proclamations in policy documents relate to 'increased access to health care services for the greater proportion of the population'—which need not necessarily translate into 'equitable access'—rather than 'equity in access/utilisation' to/of them, or even equity in health itself. In order to assess the system with respect to equity performance, we first review various conceptions of equity, philosophical and non-philosophical specifications (from economics), in order to distil operational conceptual and empirical definitions of equity and contrast them with those implicit in the Kenyan health policy. We finally review some empirical evidence that indicates there is inequity in Kenya. These concepts and other issues raised in this chapter are used to provide the basis for an interpretation and evaluation of performance of the Kenyan health care policy.

5.1 Equity in Kenyan Health Policy and Health Care

5.1.1 The basis for equity concern—health care as an individual's 'right'

Health policy issues—be they pertaining to equity or efficiency—are invariably intertwined with ethical and other value judgements. To adopt a value-free economic approach would omit (probably) the most interesting and contentious questions of health policy. The distribution of health care resources is one such issue—not amenable to value-free theoretical economic analysis (others may include what the health care system should look like and what kind of health care interventions deserve priority). But although there is ubiquitous interest in what counts (should count) as an appropriate distribution of health care resources, it is not obvious why such an interest exists. Donabedian (1971) distinguished two broad but rival ethical foundations for such a concern (which correspond to two prototype health care systems—called systems X and Y by Culyer, Maynard and Williams, 1981⁴⁷). The first values consumer sovereignty and market forces and treats health care like 'other good things' of life such as food, clothing, shelter and recreation. Those championing this view argue that what people are generally concerned about is the distribution of well-being, determined to a large extent by the way people choose to use their resources on basic goods, including health care (e.g. Fried, 1978, p. 127). People's investment strategies in basic goods determine their health outcomes—they may not be as healthy as others because of the investment strategy they adopt. Because good health is only one of the determinants of well-being, basic differences in individual health are therefore not matters of social or moral concerns—since people are entitled to spend their share of income in whichever ways they choose.

The second view supposes there is a more fundamental reason why we should be interested specifically in the distribution of health care resources. According to this view, there is something unique about health and about interventions to maintain and restore it. Health care (one of the principal mechanisms of intervening to restore or prevent deterioration in health) is fundamentally different from other good things of life such as beer, pens or potatoes—it is not a matter of

⁴⁷ These two respectively correspond to libertarian and egalitarian health care system. System X: 'has as its guiding principle consumer sovereignty in a decentralised market, in which access to health care is selective according to willingness and ability to pay. It seeks to achieve this sovereignty by private insurance; it allows insured services to be available partially free at the point of consumption.; it allows private ownership of the means of production and has minimal state control over budgets and resource distribution; and allows the reward of suppliers to be determined by the market'.

On the other hand, system Y:

'has as its guiding principle the improvement of health for the population at large; it allows selective access according to the effectiveness of health care in improving health ('need'). It seeks to improve the health of the population at large through a tax-financed system free at the point of service. It allows public ownership of the means of production subject to central control of budgets; it allows some physical direction of resources; and it allows the use of countervailing monopsony power to influence the rewards of suppliers'. (Culyer, Maynard and Williams, 1981, p. 134).

satisfaction of individual desires but something to which individuals have a fundamental claim by right, like access to the ballot box or to courts of justice. Its distribution therefore “should not depend in any way on income and wealth, though it will necessarily have to depend on the income and wealth of society in general (since resources are limited and health is not the only good thing, so health care is only ever going to receive a finite share of the total) and entitlement will also clearly depend on the conditions of ‘membership’” (Culyer, 1992). Concerning ‘rights’, a distinction has to be made between a legal right to health care, and the moral right to it (Veatch, 1982). Whether there is such a right as a matter of law depends on jurisdiction. The Sessional Paper Number 10 of 1965 (Kenya 1965) stated that the government was committed to each citizen’s ‘right’ to ‘equality’ before the law and certain other services, regardless of economic condition, political affiliation, religion, race, tribe, and so on. Amongst others, the Paper lists the areas to which such ‘rights’ apply as ‘equal opportunities’ to education and medical care services, also paraphrased as ‘freedom from want, disease and exploitation’ (pp. 2, 4⁴⁸). The Paper also states ‘The declared aim of the Government is to provide medical and hospital services, old-age and disability benefits. . .’ (p. 47). Because of resource constraints, the first step to the realisation of this objective in health care was the introduction of free services for outpatients and all children in 1965—groups considered previously under-served by the colonial health administration. Even during those days, the idea of cost-effectiveness seems to have existed, political dogmas notwithstanding:

One of our problems is to decide how much priority we should give to investing in less developed provinces. To make the economy as a whole grow as fast as possible, [resources] should be invested where [they] would yield the largest increase in net output. This approach will clearly favour . . . areas having abundant . . . resources. . . . A million pounds invested in one area may raise net output by £20,000 while its use in another may yield an increase of £100,000. This is a clear case in which investment in the second area is the wise decision because the country is £80,000 per annum better off by so doing and is therefore in a position to aid the first area by making grants or subsidised loans. The purpose of development is not to develop an *area*, but to develop and make better off the *people* of the area. If an area is deficient in resources, this can best be done by [among others], investing in the *health* of the people . . . (Kenya, 1965, pp. 46—47, *italics* in original).

A major omission is the failure to recognise that a true marginal analysis as advocated here requires *some* marginal valuation in each case. That aside, it is noteworthy that the Kenyan constitution—which guarantees the protection and preservation of ‘fundamental rights’ and other basic freedoms of the individual—does not mention health or health care (e.g. See Kenya, 1969a, Chapter V, sections 70—86). The rights of individuals to health care services, viewed in light of this pas-

⁴⁸ A similar statement, but which may have different connotations occurs on page 56: ‘Every effort will be made to ensure that equal opportunities are provided for people in less developed parts of the country’.

sage and the constitution may be interpreted as implying that individuals should have 'equal opportunities' to health care services in order to be 'free from disease' (see the discussion of access—the notion implied by the phrase 'equal opportunities'—and equity below).

However, in spite of the existence of such knowledge, these principles were not applied to health (as the last chapter showed concerning the distribution of resources among functions), and we shall later show that substantial inequalities in health and access to health care still persist. We also will proffer possible explanations for this persistence which need to be the focus of health policy if inroads are to be made.

Beyond the statutory right to health, there also appears to exist a moral claim to a right to health care in Kenya, although at first it was not clear to what types of services—for to say there is a right to health or health care does not necessarily imply a right to equal access for all to whatever services are available. Nor does it imply that the services have to be free. In most societies, such a right may be/and usually is, expressed in terms of a right to specified set of services, dynamically defined, but nevertheless not dogmatically equated with the best available (e.g. see Fried, 1983). That is to say, there is need to specify what services, if any, may be free, who shall have access to which (types of) services and on what terms—how equal should that access be at various levels given that time and money access costs are dissimilar between individuals? In the first national development plan (Kenya, 1966), the chapter on health also stated ' . . . it is incumbent on any government devoted to the social welfare of its citizens to provide adequate health facilities'—a political rhetorical razzmatazz that has proved hard to implement as a policy. More recently, it was stated that

Effective medical care, particularly when preventive in nature and directed to the rural areas, contributes significantly to national development . . . Public spending to maintain and extend costly urban-based hospitals will be curtailed, and the bulk of savings from the slowdown of capital projects in urban areas will be redirected towards small-scale projects at district and sub-district levels. . . . Preventive and promotive health programmes, if adequately supported, can be cost-effective (Kenya, 1984, p. 152—3).

The current development plan (Kenya, 1994a) also states 'Primary Health Care implies provision of essential health care *universally accessible to individuals and families in the community* through appropriate means, through their participation, and at a cost that the community and country can afford' (*added italics*, p. 231). Before giving an interpretation of this statement, it is important to point out that provision of free health care is not synonymous with equality of access for all. Second, universally accessible care is not the same as comprehensive care. Therefore, these statements might be interpreted to mean that the (moral) right to health care services applies to 'effec-

tive preventive care'. Where does this leave individuals in need of curative services? Policy documents are mostly silent on this, and therefore interpretations have to be inferred from inarticulate policy statements. The current development plan (Kenya, 1994a) puts it this way: '... resources will continue being redistributed from curative to preventive care and from urban to rural areas. Communities, as a result of being involved comprehensively in formulating and implementing health programmes and activities will ensure programme sustenance, and as a result, become more responsive to their [curative?] health needs' (p. 231). Thus, it seems clear there is no unlimited right to all conceivable types of health care that individuals may desire. In short, it therefore appears that in Kenya, the second view of health services—that health is a right of the individual and access to it should therefore not depend on individual income or wealth—is the prevailing philosophy (although there is a private sector where access to health care seems based on the first viewpoint), but that right is limited to preventive care services.

There is another reason why the Kenyan government might still acknowledge a right to health care in a way that is different from rights to other basic goods. For many basic goods the need of each individual is approximately equal within a certain narrow range, e.g. for clothing, shelter or even beer. Variations in individuals' consumption of these latter goods is largely a dependent on personal preferences. However, for health care (and education too), need is not as evenly distributed. An even distribution of health care services across individuals might lead to healthy individuals having more care than they really need whereas the less healthy may have considerably less than they need (irrespective of whichever definition of 'needs' is adopted—see section 5.3.2 below). Therefore, while it is true that the ultimate focus is on individual well-being, a strategy that allows incomes (and wealth) to determine quantities of goods and services available to individuals may not work for health care services. Thus, as a matter of social policy there must be some judgement about how health care services ought to be distributed. Taking this and other reasons discussed above as the reasons for interest in the distribution of health care in Kenya, the next sub-section examines the nature of equity advocated (again, using statements from policy documents). We shall later (in section 5.2) give content to this term as various interpretations can be attached to it depending on which theory of justice one starts with.

5.1.2 The approach adopted by government to realise the equity objective

Having established the basis of the concern for equity, let us now look briefly at the policies pursued to this end. To put the policies in context, we need to understand some historical perspective of health services in Kenya.

Development of health services before independence

The pattern of health care resource allocation that the first full-fledged government of independent Kenya inherited in 1963 owed its origin to about 70 years of British colonial administration in the country which at the turn of the century penetrated the territory for economic exploitation under the auspices of the Imperial British East African Company (IBEAC) in 1888.⁴⁹ That penetration brought with it western medicine into the territory but, in those days, the main motive of introducing western medicine was for the treatment of the personnel of the company. Later, in 1895 when the British Foreign Office assumed responsibility for Kenya, the IBEAC medical staff was taken over by the British government and this was the first step towards medical services supported and controlled by the state. But even then, the services were mainly for the administrators although the missionaries had introduced medical services to the indigenous people on a very limited scale—and these were extended to the indigenous people because of the threat of epidemics—a public health hazard, as well as the control of mosquitoes which presented a real threat to the health of the newcomers. Hence the first (western) medical services provided to the local people were mainly public health controls. But the coming of the first world war and the massive recruitment of indigenous males as porters opened a Pandora's box on the health situation of the indigenous people—it led the colonial authorities to reconsider past attitudes and policies and to improve the quality of medical care in Kenya. After the war, the colonial government constructed government dispensaries in the native reserves. Beck quotes the case of a medical officer who justified the increased medical expense by recourse to the humanitarian value of bringing medicine to the reserves (*ibid.*, p. 95). By 1922, the medical department was firmly established as a government department for the maintenance of health for the entire population and for the improvement of the living conditions in rural areas. The indigenous people's health was no longer the responsibility of the missionaries. Indeed, that responsibility then was accorded to the medical department alone, but the essential collaboration of other departments and the missionaries was not entirely ignored. As a result of that change in policy, medical expenditure increased but, by 1935, 'The department was unable to meet the African demand for the improvement in general welfare'

⁴⁹ The account of the history of health services in Kenya for the period before 1950 presented here is based on Beck, 1974.

(*ibid.*, p. 97). Medical problems took a turn for the worse during the world war two period, but the post-war period saw a renewed commitment which spanned not only the provision of services, but the construction of health centres and new and larger medical training centres and a hospital in Nairobi. As a result, (nominal) expenditures increased from an estimated £360,00 in 1945 to £580,000 in 1948. The onset of the struggle for independence and the state of emergency declared throughout the territory was a major setback to the development of health services.

The missionaries, as pointed above, also played an important role in the development of medical services in Kenya. Paralleling the colonial administration efforts was the development of missionary and church supported medical services. But, unlike the colonial administration government, the missionaries used the health services also as a way of luring the indigenous people to their faith, besides needing them for their own health—Beck writes that one Dr Arthur said in 1912 that the missionary should gain the confidence of the patient through his medical work and in this way prepare him for acceptance of the Christian message (*ibid.*, p. 98). They established hospitals and dispensaries among the indigenous peoples well before the colonial government and also pioneered in the training of Africans as medical aides—mainly as dressers. Another difference is that medical missions were quite successful at spreading medical care to the interior. Consequently, when the colonial administrators changed their attitude towards the health of the indigenous people, the missionaries benefited through increased subsidies in recognition of their important work, although later resource problems and doubts about the quality of services offered by missions strained the relation. This was, however, rectified through increased controls over mission directed services by the government medical department, even use of some of the more qualified missions as part of the state system. However, the development of the two systems remained separate and the situation remained so up to the time Kenya became independent.

Besides western medicine there was also a heterogeneous traditional medicine sector—which the colonial administration decided not to interfere with except an attempt to purge witchcraft practitioners, sorcerers and others who psychologically ‘intimidated’ society. As a result, many traditional practices were not interfered with.

The situation remained that way in spite of government commitment after independence to reduce and eventually eliminate inequalities both between regions and groups of people.

Prevailing equity conceptions and changes implemented since independence

Let us now look at the equity conceptions, then and now. In the immediate pre- independence period, as indicated above, there existed a largely fragmented health care system consisting of traditional medicine as well as science-based medicine. In the former, African traditions governed access to health care. In that tradition, there was mutual social responsibility over matters of disease and illness—members co-operated to do their very best for each other with the full knowledge and understanding that if society prospers its members will share in that prosperity. The social tradition ensured that those in need of health care services got them, even when unable to pay. The traditional medicine-man (the village herbalist) had a duty to treat sick members and to seek compensation later. For those unable to pay, members of their (extended) family had an obligation to contribute to the cost of treatment—by giving whatever material wealth they had for it. If these alone were unable, members of the clan to which the family belonged had such an obligation (by extension). Generally, it never went beyond the clan, although in times of calamities, all society was obliged to help. There was thus a general form of ‘social insurance’ against illness, where access was not on account of material or social standing, but according to need. Even today, this tradition persists, in the spirit of ‘Harambee’ fund raising⁵⁰.

On the other hand, access to scientific medicine was generally based on economic considerations. The scientific medicine system itself was fragmented. There was a network of government health services where access was at a charge, and missionary and church medical services, largely based on charity. In the former, there existed two compulsory schemes: the European Hospital Fund Authority catering only for Europeans, and the Kenya Hospital Fund Authority which catered for Indians and Arabs and those Africans who wished to join voluntarily. The missionary and church medical services sector provided services to indigent (mainly) rural populations, often on charity basis. Even these too were fragmented—along denominational lines, e.g., Protestant or Catholic.

Thus, for scientific medicine health care was treated like other commodities and access to it was largely dependent on ability to pay, although there appears to have existed some form of a caring externality on the part of the missionaries. This led to the development of a rudimentary health care insurance system catering for the needs of those who could afford the premiums—mainly Europeans, some relatively wealthy members of the Asian community and a few Africans. On the other hand, access to health care in the traditional medicine sector was socially guaranteed, provided a need for the same was evident, and payment was not a paramount issue.

⁵⁰ See footnote number 24.

(i) Equity and the public health services after Independence

After independence, the government considered the existing disparities in health services unacceptable and sought to consolidate the government services under a Ministry of Health, with a long term objective of providing free medical services to all the people, irrespective of their economic or other considerations (as discussed above). Economic constraints led government to introduce two important reforms in health care—one affecting the consumption of services, the other their financing. These were (i) the introduction of free outpatient medical services for out-patients and all children in 1965, and, (ii) the establishment of a compulsory national hospital insurance scheme—the National Hospital Insurance Fund (NHIF), for persons earning over K£600 per annum (or Ks 1,000 per month), later amended in 1972 to allow those willing low income earners to join on voluntary basis. The scheme operated on the insurance principle of spreading the risks and consequent costs of hospitalisation among all participants—i.e., community rating. Monthly subscriptions were set at K£1 (i.e., Ks 20) irrespective of income. Consequently the contributory mechanism adopted was highly regressive.

Beyond these initial changes, between 1966 and 1989 the equity issue seems to have largely been relegated to the background. Various development plans during this period indicate that the equity policy pursued was largely about increasing the quantity of health services, particularly facilities, in order to bring services to as many people as possible. For example, the building of more health centres, dispensaries and even hospitals was vigorously pursued. Also the training of more health personnel in order to lower the population/health personnel ratios. Besides these, other factors external to health care but that have a direct impact on health have also been incorporated in government policy including income redistribution through taxation, improved housing, environmental sanitation, education and so on.

However, during the 1980s, Kenya had a poor economic performance (as discussed in chapter one) that forced the government to change its outlook on financing of social services such as health. In what remains the most comprehensive policy document ever produced, the new policy was stated in the following terms:

The Government of Kenya has emphasised the provision of basic needs ever since independence. . . . In 1985/86, over one third of the total Government outlays will go for water, health and education, while of the recurrent ministry outlays, over 42 per cent is to be spent on these services. . . . These public expenditures on social services, along with the growth of incomes, have made an important difference in the quality of life for Kenyans. Since 1960, the recorded overall death rate has been cut in half . . . life expectancy has risen by more than ten years. . . . These and other improvements give more Kenyans the potential for productive involvement in the economy. . . . Of course, outlays on basic

needs cannot be judged entirely as investments in future growth and employment. They have some very immediate benefits . . . (Kenya, 1986, p. 11).

and elsewhere, it continues to state that

Government will retain substantial responsibility for basic needs such as education, health and water, although private and 'Harambee' efforts will remain important, and for traditional public services. In exercising this responsibility, *a balance must be struck between the desirability of asking beneficiaries to contribute increasingly to the cost of social services and the Government's concern that those least able to pay still have access to such services*

. . . Government's resources will remain limited Yet Government remains committed to doing as much as possible to ensure that both social and economic services are widely available to the public . . . *the only way out of this dilemma is to involve those who participate in Government services increasingly in financing the recurrent costs of these activities.* Government has long charged fees for certain services. The coverage and level of these charges will be reviewed where participation has fallen behind the level of costs and ability to pay, and *certain new charges will be considered*, . . . proposals will be made to increase the extend of participant's support . . . for services . . . (Kenya, 1986, p. 24, *italics ours*).

Guided by this new philosophy, in addition to pressures from increased demand in the face of reduced revenues, government sought to mobilise additional resources and utilise existing ones more efficiently. In response to the first challenge, in December 1989, the government introduced outpatient and inpatient user charges at all public sector facilities except dispensaries and in July 1990 introduced reforms to the National Hospital Insurance Fund.

In introducing user charges in government health services, it was also realised there was need for a mechanism to ensure the needy did not miss services due to inability to pay, and two instruments were incorporated—exemptions and waivers. There was a provision for complete exemption for those services that must be encouraged for social reasons, such as child health, family planning and AIDS treatment⁵¹. In addition, no extra charges were to be levied for treatment in specialised (public) hospitals. The exemptions were supposed to be automatic—patients satisfying the stated criteria were automatically exempted. The poor also were to be protected through the use of an explicit waiver based on ability to pay. The granting of the waiver on grounds of inability to pay is at the discretion of designated officers in each health institution. For example, the clinical officers or community nurses in health centres and the designated nursing, clinical or medical officer in hospitals. After the first consultation, the patient is referred to the area 'chief'⁵² with an exemption form for endorsement to certify the person's hardship. Once a person's inability to pay has been determined, the authorised officer issues an exempt certificate valid for one year. Persons dis-

⁵¹ The government pays for these services.

⁵² This term is used to designate the administrative officer in charge of a location in Kenya. A location is a (regional) administrative unit.

satisfied with the decision taken have the right to appeal. In spite of these measures, attendance at charging facilities initially declined following the introduction of fees. But public concern over lack of improvement in services and perceptions that the waiver system was not functioning properly prompted the government to lower maternity charges in August 1990 and to temporarily suspend all charges till improvements in service delivery and exemption procedures could be ensured. In 1991, it was observed that

Any system of patient fees for government-provided health services requires a process for ensuring that universal access to health service is maintained. There must be a system for exempting hardship cases and patients with limited financial resources. When Facility Improvement Fees [user fees were so baptised!] were introduced in December 1989, a waiver system was carefully developed. The system did not give fixed rules on how to determine whether or not a person is able to pay when he/she seeks care for the first time in a health care facility. The final decision [was] left to the judgement of the designated officer(s) of each facility . . . (MOH, 1991, p. 8).

This led to a refinement of the system including a specification of which services or people were exempted from fees—children under five years; patients from charitable and destitute homes and from homes for the mentally handicapped; prisoners and persons in police custody; ‘downward referrals’ from higher health institutions (patients in this category being exempt only for the first 30 days); civil servants and their dependants and inpatients readmitted for the same diagnosis within two weeks of discharge. Patients with chronic diseases (diabetes, epilepsy, asthma, tuberculosis (TB), and so on) and patients infected during an epidemic were not to be automatically exempted. The Facility Improvement Fund Operations Manual (MOH, 1992) explains further modifications of the exemption and waiver systems. Fees charging resumed in early 1993.

The National Hospital Insurance Fund (NHIF) was also reformed. The NHIF contributions were revised in 1990 from the constant Ks 20 per month per person to a graduated scale of contributions based on income ranging from Ks 30 (for an income of Ks 1,000 per month) to a maximum of Ks 320 per month (for those earning Ks 15,000 or more), effective from 1 July 1990. The rates are based on basic monthly income. The Ks 1,000 per month remains the entry point for contributors to qualify. When the Fund became operational in 1967, with monthly contributions of Ks 20 per month, only about 40,000 persons qualified for membership. In 1992, the Fund had a membership of 1.3 million people (mostly government and private sector employees, and a few farmers, fishermen and ‘jua kali’⁵³ artisans) and a dependent population of 8 million [Press Release to *The Standard*, April 28, 1992]. The benefits have similarly been revised. Initially when the fund was set, the rates ranged from Ks 35 to Ks 75 per day of hospitalisation, depending on the institution. The

⁵³ A colloquial (Kiswahili) term literally translated meaning working under the blazing sun, otherwise known as the informal sector.

latter was revised to a maximum of Ks 150 for some approved hospitals. Now they range from Ks 80 and Ks 600 a day, for a maximum of 180 days (of hospitalisation) a year. There is no deductible, and the sole 'copayment' consists of payment by contributors of the difference between the maximum amount set for a facility that the patient chooses to be hospitalised in, and the institution's per diem charges, if it is higher than the Fund's limit. This happens only when patients choose the more expensive hospitals. Only inpatient care is covered. But there are plans to widen the benefit package to include some ambulatory services. Participation by health institutions is by application which is followed by an appraisal for a decision. All government hospitals and health centres with (paying wings) beds, mission and accredited private hospitals, clinics and nursing homes are approved for the purpose of claiming benefits from the NHIF. There is a facility rate applicable to each category.

Why the waiver system is ineffective

There are reasons to suspect the waiver system may not be effective in realising equity objectives. It is possible that the user fee system as whole, though still not based on full-cost recovery, might be having adverse effects on equity as a whole, and might not be cost effective. The aims of the system are to raise revenue as well as to increase efficiency and improve equity. We have already shown the revenue raising capacity is far from meeting estimated targets (see chapter three, Section 3.4, p. 79). There is need to evaluate the over-all cost-effectiveness of this policy—its cost versus the revenues collected. Probably, if the costs of staff specifically recruited for its implementation, together with cost of other facility staff (time) taken in its administration are evaluated realistically, particularly in view of its impact on access to services, *it might not turn out to be cost-effective after all*. The government facilities are allowed to retain 75 per cent of the collections at the collecting facility. It is conceivable that this retention clause might discourage facilities from granting waivers to deserving cases as that lowers the amount available to the facility. An alternative system of sharing the revenue, for example, a capitated system, might be a better alternative. Second, the waiver form is an accountable document, but due to other problems within the system, some facilities occasionally have shortages of the waiver forms. Without them, facilities get excuses to charge all patients. Third, the system had assumed 'chiefs' and 'sub-chiefs'³⁴ would play an important role in identifying those in need of such waivers. But these were never informed about their expected roles. It was assumed they would know the people who cannot pay. In a country like

³⁴ A 'chief' is a civil servant (public administrator) in charge of an administrative unit called a 'location', while a 'sub-chief' is in charge of a cluster of villages grouped together to form a 'ward' or 'section'.

Kenya, where wealth is sometimes held in forms not easily assessed for income determination, this is too simplistic an assumption—there is no bases for assessing personal income for most rural subsistence farmers. Moreover, it was assumed the ‘chiefs’ and ‘sub-chiefs’ would be impartial. This might not be the case. Cases of political discrimination along political party lines are not rare in Kenya and there is no reason to believe such is not the case for health services.

The existence of the National Hospital Insurance Fund in its present form seems to be another source of concern for equity. People who contribute to this fund have an extra advantage over the rest in the sense they use the pooled resources to secure better quality care, mostly in the private sector. The implications are that a two-tier system effectively exists, a fact reinforced by the existence of a not insignificant private sector.

(ii) Equity in the private sector

Alongside the public health services, the private health services (including the Non-Governmental-Organisations’ services) sector was left to develop largely without interference from the government except for ‘normal’ controls to ensure services provided by the sector were of acceptable standards. Access to private services is mainly by out-of-pocket contributions as well as by private insurance, in addition to whatever the NHIF supports. Consequently, access to these services is determined by personal economic standing.

5.1.3 Conclusions

The question of equity in health and health care in Kenya has generally been treated ambivalently by Kenyan legislators and policy makers. As in most other developing (and some developed) countries, Kenyan health policy is rather vague on equity. Though there are policies that suggest Kenya safeguards and promotes equity in services such as health, we cannot say that each citizen is ‘guaranteed access on equal terms’, even if that was the intention. Government’s long term intentions are to increase availability so that most people have as unimpeded access to services as is practically possible. After independence, government was offering universal, general and free health services to target groups. Recent policy emphasises the systematic improvement of economic, social, cultural and working conditions which guarantee and/or promote good health. Government still ‘guarantees’ access for all citizens to preventive, curative and rehabilitative medicine, independent of their economic status or other similar considerations. There are now more pronouncements on equitable access to commodities that are health producing (and/or promoting)

—such as housing, education, clean water and sanitation, better environment and working conditions, and so on. These statements, taken together with the assurance of access to health services for all individuals, can be satisfactorily interpreted as adequate grounds for interpreting the value climate of the Kenyan health policy framework as safeguarding and/or enhancing access to commodities that promote health, including health services. But viewed in light of recent developments, that access need not necessarily be equal—it appears perfectly feasible that some citizens will (and do) enjoy a greater access to particular health promoting goods such as income and education. This conclusion is reinforced by the observation that the government is concerned that there should not exist situations where some people are deprived of a minimum acceptable access to these goods, especially health care services.

Moreover, looking at the history of Kenyan health policy since independence, it is evident that the basic equity objectives has changed. The initial objective was one of unimpeded access at the point of consumption—albeit for targeted groups. Zero price did not lead to the realisation of this objective (or so it will be shown below). The new policy still guarantees access at the point of use subject to diverse fees, taking into account the economic and social status of the citizens. Is there an apparent inconsistency inherent in the policy objectives or is there a change? Can equal access be guaranteed in situations where people have ‘unequal access’ to health promoting goods, despite the existence of policies geared towards eliminating or (at least) minimising the differentials in access to goods? It appears the answer must be affirmative. Policy has over time changed from one of simple equality to one that simply guarantees access to an ‘acceptable level of access’ to these goods. The important difference is that although the health policy indeed has equity implications about access to health services, the operational concept now is health itself, rather than health care. The actual change in the equity objective can be discerned from existing as well as proposed characteristics: in the past health care was universally available and free (although with token charges—again, albeit at least to targeted groups—for in-patient services). It was universal in that it was destined for all Kenyans without discrimination. General in that it covered preventive, curative as well as rehabilitative services. Free because it was financed by the state and users were not making any direct payments for use (or, even if they were, those payments were insignificant). Today, the situation is different. Access for those unable to pay is still guaranteed, but only on proof of inability to pay. So there is an element of discrimination.

In all, we may say that policy statements since independence have been rather vague with regard to equity. There have been vague commitments to the equitable distribution of resources and non-discrimination in treatment, but only few complete pronouncements usable as guiding objec-

tives. The actual policies indicate a concern for equity—mostly about regional resource allocation. But it is not clear what the aims of the present policy are with respect to equity—for it will seek to guarantee equality of opportunity of access for all citizens to health care services, yet, at the same time the government will seek to ‘make those able to meet part of or all the costs of their health care’ do so. If both statements are correct, then the correct objective would appear to be not just equality of access to health care, but equality to both private and public health care. This is of course an ambitious if not amorphous goal. Such equality of access cannot be guaranteed, short of a massive redistribution of income. The ambiguity in the statements arises because policy makers have opted for egalitarian formulation when one couched in terms of guaranteed basic level of services would seem more appropriate in light of other policy choices. The main problem is that policy is silent on access to private health care, which, by the very act of the new policy directions, is being encouraged. Changes in the National Hospital Insurance Fund (NHIF) could enable this apparent contradiction in government policy to be overcome or eliminated. It seems the original policy intention sought equity through the strive for a ‘threshold’ standard of care. Despite the lack of clarity of successive plans, this *continues* to be the underlying policy aim.

Two main conclusions may be drawn from the analysis presented in this section: First, that in Kenya, the *main* ethical basis of the public health services is that individuals have a ‘right’ to ‘equal’ access to health care, although not to all conceivable types of the services. Second, through inferential analysis of documents, four possible implicit interpretations of equity in the Kenyan health sector may be stated as:

- ① equal access to public health care services for equal need;
- ② equal access to health promoting commodities;
- ③ guaranteed minimum access for all; and,
- ④ unequal opportunity of access to private health care.

These interpretations suggest different objectives with greater or lesser difficulty of attainment and monitoring, depending on how the term ‘access’ is interpreted and how the objective function is specified. Accessibility is the absence of barriers (both monetary or non-monetary) that stand in the way of an individual desiring to use a service and the medical facilities that provide the care—i.e. the absence of barriers preventing ‘need’ from being converted into demand (Parkin 1980). Need and demand are not the same—need is the amount of care believed necessary by medical authorities while demand is the actual use, i.e. demand is met need (Feldstein, 1966). (Further interpretations of ‘need’ are discussed in section 5.3 below). It would therefore be useful if the equity

objective is stated explicitly, taking into account the different interpretations that may be attached to 'access'⁵⁵, to cover both public and private health care services. Unless this is done, any of the four possible interpretations might be taken as a correct yardstick for the measurement of collective decisions in achievement of equity—which, given the fourth interpretation, is questionable and inconsistent. The first and second objectives suggest there exists a specific concern for questions of distribution in the health domain that is distinct from other explicit or implicit objectives such as maximising survival, efficiency of services or guaranteeing consumer choice.

Thus there is need to evaluate whether health equity objectives are being met as well as the effects of policy choices in light of such goals. Second, the attainment of equity in health care appears to require equalisation of access between regions although there is an implicit concern for equity between socio-economic groups. Simply guaranteeing that the least well-off (whether areas or individuals or groups of same) have available a basic level of health care is insufficient. It will not, for example, guarantee equality of health, which is a more preferred objective (see below). Third, equity in the delivery of health care should be achieved through the equal access to the general

-
- ⁵⁵ Parkin (1980) distinguishes several notions of economic accessibility to health care as defined by
- (a) characteristics of the health care delivery system, which subsumes geographical (spatial) attributes and socio-organisation (non-spatial) attributes—defined in terms of the volume and distribution of medical resources. The organisation determines the terms of entry to the health care system and the structure use of the system once entered.
 - (b) characteristics of the population at risk, defined in terms of (i) predisposing factors—age, sex, religion, attitudes to health, and so on; (ii) enabling factors—income and wealth, insurance, and so on; (iii) the illness level, as perceived by the individual or by the delivery system.
 - (c) outcomes—which depends on utilisation rates and consumer satisfaction. The utilisation rate depends on the type of illness, site of treatment, purpose and time (continued next page).

Several notions of accessibility arise from consideration of these factors. Those barriers relating to the spatial distance between the individual and the provider facilities constitute physical accessibility barriers. These translate into the costs of travelling to make use of health care services using a particular mode of transport to reach the provider and also the cost of travelling to consume health care services—in terms of the opportunity cost of the treatment itself, i.e. the extent to which earnings (or other opportunity costs) foregone in seeking care act as a barrier to the consumption of health care. There are also time costs (foregone earnings). These factors are relevant to the specification of the equality of access to public and private providers alike. Price constraints are another barrier to accessibility. The price paid (either at the point of consumption of health care or via other pre-payment schemes) acts as a barrier to the consumption of health care. Given the differentials in incomes and other enabling factors that exist between individuals as well as regions, it is doubtful equality of access to either private or public services can be achieved. There is also another determinant of accessibility—eligibility, which defines the terms of access and to some extent what is being accessed—whether free of charge. Like any other commodity, consumption will tend to be inversely correlated with the price charged. This raises the issue of horizontal and or vertical equity. Besides, informational costs also play a vital role in accessibility. These refer to the extent to which the cost of obtaining information on availability, quality and so on of health care facilities is a barrier to consumption. The cost of obtaining such knowledge is arises from the opportunity to be treated at both types of providers—those unable to afford the charges of private providers will generally not have such information. A final accessibility determinant that is similar (in effects) to information constraints is the disutility of treatment—even if perfect information were available, patients might perceive barriers in terms of social stigma (perhaps of pain, embarrassment, and so on). These factors, individually and in combination suggest different policies for ensuring equity to services. In general, the examination of public documents yields ambiguity with regard all the key interpretations of the term 'access'.

range of services. Applying the rule to what are sometimes termed as ‘basic services,’ which excludes other available modes of care is not sufficient. Finally, the process of health production and health care delivery, i.e., availability, and final outcomes (mortality, morbidity, service utilisation) are relevant for evaluating the degree of equity in the system. In the rest of this chapter, we develop the conceptual framework that will be used as the basis for equity assessments, both in the current system, as well as for the proposals presented later on in the thesis.

The issues we would address in the rest of the chapter can now be briefly summarised:

- ① what is the basis for equity in the Kenyan health care system?
- ② Is it adequate to lead to a consistent policy on equity?
- ③ should this basis differ as between the private and public services?
- ④ Is it possible to incorporate efficiency?

We begin the investigation in the next section with a discussion of the bases of equity, both from the philosophical and health economics literature.

5.2 Philosophical and Health Economics Bases of Equity and Their Implications for Kenyan Health Policy

There is widespread agreement amongst writers that equity has something to do with ‘equality’ (e.g. Culyer, 1991a; Le Grand, 1984; Veatch, 1982). But although most notions of equity use the concept of equality, equality connotes cases where the shares (of the entity whose distribution is the point of interest) resulting from a distribution rule are *equal* whereas equity connotes a situation where they are *fair* or *just* (Veatch, 1982; Rawls, 1958). This distinction can be explained by a simple example. If, in a society composed of n individuals producing between them a certain (fixed) quantity of goods and services (call it Q), a distribution rule was devised that gave each individual a quantity Q/n , i.e. an equal division, then the allocation leads to equality (in shares). There are two reasons why such a distribution rule may not lead to desirable results. First, no two individuals in the system would want to consume exactly the same bundle of goods. In such cases, although individuals may have equal shares, they may not necessarily be happy—which reduces their welfare. Second, to achieve that equal allocation may (and often does) involve transferring resources (wealth) between individuals (from the more productive to the less productive), a mechanism that often interferes with (actually destroys) incentives to produce. If in such a society, a distribution rule (not necessarily the one we have described above) were devised under which

(for the moment ignoring the production side), the resulting shares were *acceptable to all, given the society's set of (moral) values (that define the maximand or objective of the distribution exercise)*, then the distribution would be considered *fair or just*—in other words, the concept of fairness is a value judgement based on social values. There are several (competing) theories purporting to explain what is a fair or equitable distribution of resources. The policy implications for health care resource distribution arising from each differ because of differences in focus (some focus on individual effort, others on personal ability, need, the individual's usefulness to society, yet others on proportionality—more on these anon—see section 5.3). We now review these theories, particularly their implications in the health domain, showing (as the analysis unveils), what form of justice is implied by each, and thus their implications for health policy in Kenya. Our purpose of reviewing these is not to critique them as such, but to see which one, if any, bears close semblance to the equity notion in Kenya.

5.2.1 Utilitarianism

In utilitarian welfare economics (e.g. Harsanyi, 1955), justice is inseparable from the goal of maximising aggregate utility, although in recent times, partly as result of developments in moral philosophy dealing with the notion of justice as of independent importance, the situation has been changing (especially due to accounts such as that of Rawls, 1972). Under utilitarianism *resources are allocated so as to maximise aggregate utility*—the sum total of individual utilities or to maximise social welfare. Actions are therefore judged on their consequences on aggregate utility⁵⁶, which requires the equality of the marginal utility of everyone. Utilitarianism is 'impartial' in its treatment of individuals, an attribute of 'fairness' or justice. If identical preferences exist, utilitarian allocations *may* be equitable in outcome (Culyer, 1980). But this is rare in the health sector and the application of the utilitarianism principle in the sector may lead to interpersonal inequities, where health resources are allocated to individuals with different abilities to transform them into outcomes, say health⁵⁷. In general, *resource allocations under utilitarianism rules result in efficient*

⁵⁶ Utilitarianism is actually an amalgam of three distinct principles: (1) welfarism—where consequences of actions are judged entirely on by the utility information related to a state (of distribution); (2) sum-ranking—where the goodness of a set of utilities of different individuals, taken together, is the sum of these utilities, thereby eliminating the possibility of being concerned with inequalities in the distribution of utilities (between individuals), and the overall goodness or 'social welfare' is simply the aggregate of the individual utilities; (3) consequentialism—where actions, rules and institutions are judged in terms of the goodness of their consequences (See Sen, 1987b).

⁵⁷ If policies promoting equity are costly in terms of output, the overall social utility frontier may have sections that are convex to the origin. For example, suppose there are two individuals, one healthy, the other chronically sick. The chronically sick person does not generate as much utility from a given amount of wealth as the healthy person. A utilitarian distribution of wealth starting from a position of equal income would require taking income away from the chronically sick person and giving it to the

distributions but not necessarily equitable distributions. This is because the basis of equality which realises maximisation of aggregate utility is the marginal utility—if aggregate utility is to be maximised, resources have to be allocated to those with higher marginal utility for their use. It is conceivable that those with such higher marginal utilities are not necessarily the people the health policy would rather they get the resources. Utilitarianism is therefore not a theory of justice, but about efficiency—it is actually devoid of any ethical basis of justice. Given the special consideration focused on disadvantaged members of society by the Kenyan health policy (as indicated by the *raison d'être* of the public health services), utilitarianism cannot form the basis for equity in Kenya (or indeed, elsewhere).

5.2.2 The difference principle

The maximin theory or 'the difference principle' (Rawls, 1972) rejects welfarism (utilitarianism) and judges distributions in terms of the advantage of the least well-off person or group—i.e. takes into account the inequalities in the distribution of utilities (implied by the distribution of primary goods). Rawls sets two principles of justice that should govern the assignment of rights and duties to regulate the distribution of social and economic advantages: (i) 'each person (or group) is to have an equal right to the most extensive basic liberty compatible with similar liberty for others', and, (ii) 'social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage, and (b) attached to positions and offices open to all'. He argues for the priority of the 'principle of liberty', which is about 'aspects of the social system that define and secure equal liberties of citizenship', i.e. the basis for social judgements. The 'second principle' regulates the distribution of social and economic advantages and, *inter alia* incorporates the 'difference principle', in which priority is given to advancing the position of the worst-off persons (pp. 67—75) with reference to indices of 'primary social goods'⁵⁸ (pp. 62—65)—'things that every rational man is presumed to want', including 'rights and liberties, powers and opportunities, income and wealth, and social bases of self-respect'. These are goods for which society should be

healthy person. Hence the distribution of income would become unequal. Here, utilitarianism results in inegalitarian distribution where none existed. This is relevant to health policy where resources are allocated to individuals who possess different abilities to transform them into outcomes, say health. In general, resource allocation under utilitarianism rules results in efficient distributions but will not necessarily be equitable.

⁵⁸ Rawls's index of primary goods incorporates five types of social goods:

- a set of basic liberties,
- freedom of movement and choice of occupations against a background of diverse opportunities,
- powers and prerogatives of office,
- income and wealth, and,
- the social bases of self-respect.

These are the things that every rational man is presumed to want.

held responsible for guaranteeing the individual a fair share (see also Rawls, 1982). In the 'initial position', or 'under the veil of ignorance' inequality is justifiable only to the extent that it improves the shares of primary goods of the least-well off persons, rather than their utility. This (second) principle therefore supports equity and efficiency claims. The first part of the principle supports efficiency; the second equity.

Interpreted within the context of health (Rawls classified health as natural social primary good—not directly under the control of the social structure (p. 62)), the Rawlsian maximin theory suggests that inequalities in (access to) health care may be justified only to the extent that they operate so as to benefit the least-well-off persons (Le Grand, 1987). People or areas with ill health should receive more medically effective—and thus health generating—resources to enable them to raise their health status to that of the average person (or areas), unless so doing would actually make them worse off. For Kenya this suggests health policy should be geared towards raising the health status of those (individuals or areas) with the worst health states, provided they are susceptible to health gains. The Kenyan health policy appears to have elements consistent with this type of thinking—the Sessional Paper No 10 of 1965 stated that 'The purpose of development is not to develop an area, but to develop and make better-off the people of the area' [among others] through 'investing in the health of the people' (Kenya, 1965, p. 47). There appears to have been a concern for the level and distribution of that health, initially the emphasis being on regions. In particular it was then noted that

If these ends are to be achieved, it is necessary to develop a formula for grant-in-aid and education and health allocations that take into account the needs and incomes of each province and district. Thus the Government must ensure that all people . . . have minimum provision of essential welfare services. A policy of making education, training and health facilities available to all provinces on the same financial terms means that people of the less developed provinces are penalised simply because they are already poor. (Kenya, 1965, p. 47).

The last sentence in the quotation may be interpreted as indicating a concern for what we may here term as the least-advantaged areas (and, by extrapolation, the individuals living in those areas). It has been shown above that even today, there exists a concern for the worst-off members of society in Kenyan health policy.

In order to translate such a concern into equity policy that is operationally feasible, we need empirical measures that can be used to quantify either individual's health states or regional health status. There are a number of problems however. One is how to define the least advantaged person—is it in terms of their overall consumption of primary goods or in terms of health or health care?

It may also not be possible to distinguish between inequalities that benefit the least-well-off from those that do not. The principle implicitly suggests that inequality is justified only if it is to the advantage of the least well-off. Applied to health, there are problems, e.g. what if the worst-off (in terms of health) cannot be made better at all, or where some individuals have poorer health or inadequate health care consumption as a result of their own decisions?

5.2.3 Equality of opportunity

The two principles of justice advanced by Rawls apply to individuals who are 'normal, active and fully co-operating members of a society over the course of a complete life' (Rawls, 1982, p. 168). If the social institutions in this society guarantee equal basic rights and liberties, and a fair equality of opportunity, then the long term expectations of the least advantaged persons can be maximised. Without the assumption of 'normal people', Rawls's index of social primary goods is no longer suitable for assessing how well-off (representative) individuals are, since 'People with equal indices will not be equally well-off once we allow them to differ in health care needs.' (Daniels, 1985, p. 43). Therefore Rawls's list of *primary social goods* is too narrow once it is recognised that ill health has impact on equality of opportunity. Daniels proposes inclusion of health care in the list, by incorporating health care institutions and practices among the basic institutions involved in providing for fair equality of opportunity. The alleviation of impairment of normal functioning (used as a crude measure of the relative need for health care services) would thus be the focus of public (health) policy. The social obligation to meet health care needs derives from the more general obligation to guarantee 'fair equality of opportunity'—'Health care institutions will help provide the framework of liberties and opportunities within which individuals can use their fair income shares to pursue their own conceptions of good' (*ibid.*, p. 45). The health care institutions; and practices that are the relevant variables for this purpose are of four levels: preventive health care institutions; institutions which deliver personal medical and rehabilitative services that 'restore' the normal functioning (of individuals); institutions involved with more extended medical and social support services for the (moderately) chronically ill and the frail elderly; and health care and social services for those who can in no way be brought closer to the idealisation.

The fair equality of opportunity appears appealing since it provides a principled method of characterising the 'importance' of different strata of the health system, but it too has some problems, some of which have been appreciated by Daniels himself. First, for most commodities, the people's preferences for them can be accommodated by allowing the market for them to respond

to those preferences. Equity of access (to such goods) is assured if three conditions obtain: (i) the commodity is available at close to *true social cost*, (ii) the consumers are capable of making rational decisions; and, (iii) income distribution is approximately equal. Now, in the medical market, as Arrow (1963) has shown, the central problems of access are brought about by departures from the ideals of a competitive market, especially on the supply side, which amounts to the markets not delivering the services at their *true social costs*. The market is unresponsive to consumer preferences on the supply side and interventions may be needed to correct the problem, generally through interventions focused on capital expenditure policy, and on the structure of insurance arrangements or other means when the market is unresponsive. It may be assumed that sufficient income redistribution occurs to ensure that no one falls below the officially defined poverty line. These aspects are relevant in Kenya where there is a poor distribution of income, and there is no established 'official poverty' line—because the country does not have a welfare system. It might therefore be expected that even if such an account was to be used as the basis for health policy, some important mechanisms have to be established in lieu of income as the basis for determining access to the basic tier. Second, access to health care is more equitable the less the informational and financial barriers, or the supply anomalies that prevent access to health care services. Such factors are formidable in Kenya.

5.2.4 Sen's capabilities approach

The idea that the distribution of primary goods may be a better reflection of justice or fairness inherent in a social set-up has been questioned by Sen (1982, 1985), who argues the utilitarian and Rawlsian approaches focus on goods rather than what those goods can do for people. The utilitarian approach for example, in explaining what goods do to people (i.e. enable them generate utility, the ultimate end) uses a subjective measure. Non-subjective factors are however equally important. Hence, to explain what commodities can do to people, and how people use them to produce human activities, a much wider range of variables (beyond generating utility) should feature in the analysis. People relish commodities because they can be transformed into final outcomes (the ultimate being welfare). For example, people demand health care because it helps them generate (good) health. The ability of people to transform commodities into *human functionings*, such as ability to enjoy good health, is *the vital link between goods and final outcomes*. The commodities are transformed into more fundamental intermediate products, called characteristics—such a person being able to do certain things, e.g. move about, meet their own nutritional requirements, get clothed, participate in social life, etc., (Sen, 1982, 1985, cf. Lancaster, 1966). In demanding com-

modities, people are in fact demanding these characteristics. A demand for health care, for example, is a demand for clinical efficiency, caring by GPs, ability to restore good health, and so on. Equity objectives should therefore, Sen suggests, be directed at these basic capabilities. The process by which people transform these characteristics into 'good' health can be described as *functionings* (Sen, 1982, 1985), and can be represented as in figure 5.1 below.

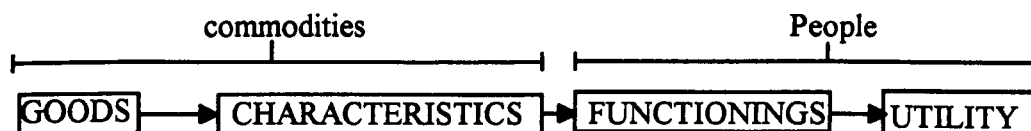


Figure 5.1: The chain from goods to utility

In traditional economics, although higher levels of utility are associated with better functionings, that link depends not only on the subjective aspects, but also on non-subjective elements. Some people have higher marginal utility per unit increase in income than others and to maximise utility, utilitarianism would recommend that such people be favoured. In reality, however, and following Sen's approach, people may want to incorporate other aspects, such as the intensity of need and other non-utility considerations. In terms of equity, Sen argues, the interest is less in whether a person is functioning in a certain way and more in whether the individual has the capability to do so. The guiding equity principle thus should be equality of basic capabilities. In health equity concerns, policy objectives may be stated in terms of people's access to, rather than utilisation of health care. The functionings approach requires that people should be able to use medical services when they choose to do so, rather than able to continuously do so. Health status depends on the capabilities to function which people have available to them. These are in turn determined by goods (health care, education, food, etc.), or more directly by the characteristics of the goods (such as clinical efficacy, knowledge of healthy living lifestyles, etc.). The characteristics of goods are related to the environmental factors (availability of medical care, good education, nutritious food, etc.) while the functionings are related to personal characteristics (such as age, socio-economic class, family size, etc.). Thus, people's ability to achieve good health depends on their access to health producing goods and their endowment of the health producing personal characteristics. With regard to equity, equality of capabilities (or functionings) implies equal access to health benefits.

The implications of Sen's approach for economic analysis can be visualised in line with a related model—the household production framework (Becker, 1965; Lancaster, 1966). This model emphasises the importance of inputs such as time and environmental constraints in the production of fundamental utility yielding commodities. The same is true of Grossman's (1972) model of demand for health by households. The fundamental commodities produced by individuals/households coincide with Sen's concept of *functionings*. The fundamental commodities are produced from market goods, environmental inputs and personal characteristics, which are the source of the capability set. But whereas the human capital tradition considers the link between functionings and utility as non-problematic, Sen emphasises that any two individuals or the same person at different times, may make identical choices when faced with the same capability set, and yet may experience quite different utility levels. Therefore the focus should be on the capability set. The empirical analysis involves determining the relationships of the capability set that are relatively universal and choosing determining variables and functionings that are relatively observable (Muellbauer, 1987).

5.2.5 Theory of entitlements

The libertarian theory of justice takes exception to the focus on outcomes. Notable among them is Nozick's (1974) theory of 'entitlements', which gives 'rights' complete priority and argues against any manipulation of outcomes to achieve social ends. He stresses that persons are equally 'entitled' to what they have provided it was acquired justly—through hard work, inheritance or through redistribution by government of wealth of people acquired illegitimately. That two individuals have different incomes is not sufficient to judge the distribution inequitable: whether it is equitable is entirely procedural. Any outcome arrived at through people's exercise of their rights must be acceptable since these have the moral force of rights. These rights include personal liberty, ownership rights over property (including the freedom to use its fruits, or to use it freely for exchange, donation or bequeathal to others). Attempts to redistribute resources, other than through voluntary charity constitute injustices. If this theory were to form the basis for equity in the health domain, it can be criticised for the 'extremism' (of mostly undesirable) results that it may lead to—the constraints imposed by rights can override other important considerations. In the health sector, it is often necessary to compromise rights so as to improve the welfare of the deprived members of society. There are people with some health conditions that they never *choose* to have, e.g., genetically acquired health anomalies. It appears Nozick would just consider this 'bad luck' which may raise sympathy, charity, and so on. In general, the theory attaches little weight to the unfortunate. Moreover, in Kenya, most of the health problems are due to lack (and neglect) of primary health

care interventions, such as immunisation, personal and environmental hygiene, and so on. People are unlikely to consider benefits other than those accruing directly to them if buying this type of care in the market if the libertarian philosophy was in force. Given the current epidemiological and economic conditions in Kenya, the libertarian approach to equity would therefore lead to a distribution that is biased against the poor and the not so healthy people. Due to these and other similar considerations, others have suggested some sort of safety net—a standard below which individuals should not be allowed to fall. This approach is often called the *decent minimum*.

5.2.6 The decent minimum

The decent minimum is a standard below which individuals should not be allowed to fall. This minimum standard may be defined in terms of final outcomes—health itself or in terms of a minimum standard of health care. The concept of a ‘decent minimum’ level of health care is relevant for multi-tier health systems that have a mix of public and private sectors, as in Kenya, where the public sector delivers health services at costs lower—usually subsidised—than the average price of the same in the private sector, and serves mainly the poor. But this does not suggest that public health services are somewhat lower quality compared to private services, only they are usually highly subsidised. So a problem of what a decent minimum is or how it should be determined arises. Whether outcomes or standards are used as the focus of a decent minimum specification, a number of problems would have to be addressed. In the former these pertain to health outcomes comparisons; in the latter, those to do with whether it is costs or quality that is to form the basis—if the latter, there would be problems of comparing health outcomes.

There have been suggestions (which do not address these problems) to explain ‘decent minimum’ in terms of a set of services considered to provide *the minimum quality of care* (e.g., Fried, 1983; Enthoven, 1980; Daniels, 1985) in terms of (i) specification of a general criterion by reference to which we can tell if the services are within the minimum or above it; (ii) description of a fair procedure for determining the minimum; or, (iii) a simple listing of the types of services to be included. Fried (1983) suggests that the ‘decent minimum’ should reflect some conception of what constitutes ‘tolerable life prospects’ in general, emphasise services like maternal and child health ‘which set the terms under which individuals will compete and develop’; he continues ‘. . . the notion of a decent minimum should include humane and . . . worthy surroundings of care for those whom we know we are not going to be able to treat them’. (p. 494). But he does not explain how/who should determine what constitutes tolerable or humane prospects⁹⁹. Enthoven (1980)

⁹⁹ Although it may be assumed this is a task for the political process.

suggested the list of basic health services that Health Maintenance Organisations (HMOs) offer should consist of—‘basic health services’—but left open the possibility of expanding or contracting the list which makes the list indeterminate and therefore of limited use in issues of health equity. Moreover, his approach is of limited relevance for Kenya where (private) health insurance is not a common (or widespread) mode of health care financing, mainly because of economic conditions. He suggests that such ‘structural’ barriers may be reduced by severing the link between plan and employment—through use of vouchers to enhance the purchasing power of rural areas. There are also doubts about the feasibility of this in Kenya where most established insurance companies operate mainly in urban areas. Moreover, few have diversified their activities beyond the traditional areas of motor vehicle and business related policies. Plus medical services are also biased to the urban areas. So even if the link between employment and insurance is broken through the use of vouchers, this alone would not guarantee enhanced health services to the rural areas. Facilities would need to be established and adequately staffed to offer relevant services. Finally, the bureaucratic and administrative structures, together with the attendant financial commitments required to effect this approach might be, to put it mildly, out of the scope of Kenya.

5.2.7 The non-envy criterion

Another interpretation of equity is the non-envy criterion (Varian, 1974; Pazner and Schmeidler, 1978; Baumol, 1986). Following Varian (1974), an allocation is equitable if no individual i envies (would rather have) another’s, j , share, which may arise if $u_i(Q_i) \geq u_i(Q_j)$, or $u_j(Q_j) \geq u_j(Q_i)$, for all i and j , [where $u(\cdot)$ is utility]. For example, if there were available a pound each of fish and beef, and I love beef but cannot stand fish, while you love fish but loathe beef, it is clearly foolish to insist on equal shares of both fish and beef for each of us, but the ‘equitable’ allocation which gives me the beef and you the fish leaves each of us contented (with our respective shares). This latter allocation is equitable. In more realistic life situations, there are many attributes in commodities that are hard to share out this way. I may not like so much fish, but still I could do with some. Or, consider health for example. If I was born with a certain genetic deformity while you were born a ‘perfect’ person, should you be disfigured so that I do not envy your looks (which otherwise I certainly would)? Or you have an angelic musical voice while mine is as coarse as that of a scratched record disk. In other words, equity is not a simple construct that can be established on the basis of individual preferences or characteristics of commodities alone. Another illustration with similar conclusions is the act of sharing a cake fairly—whereby for two people, one cuts it and

the other picks his/her share first (Le Grand, 1984). The resulting allocation should be acceptable to both, envy-free and fair.

This theory divorces the subjective external moral viewpoints of judging a distribution from the fairness or otherwise of it. Preferences of the individuals are the central players in shaping the fairness of distributions. But this may be its weakness. People's judgements about what is fair to them may be inaccurate—especially when such decisions concern their own health and use of health care. The conception is also at variance with everyday understanding of the term equity. There are situations considered equitable where envy persists. For example, in the cake cutting setting, suppose there was asymmetric information, say, the cutter knew that a part of the cake was more nutritious, and adopts a strategy which encourages the other person to choose the less nutritious part, the resulting distribution can hardly be termed equitable. In other situations there may be no envy but are nonetheless inequitable (e.g. the hard-done-by person who is a true saint who does not envy his extremely rich neighbour). It is also doubtful whether the approach can suitably be applied as a guide to equitable health policy. Its application in health care resource distribution can often lead to undesirable results—the method consistently disfavours those with chronic illnesses (Feldman, 1987). By concentrating exclusively on peoples' preferences, the method disregards the role of individual tastes, needs, etc. It has limited applicability (if any) to shaping health care policy in Kenya and it can hardly satisfy the envy-free allocation criterion.

5.2.8 The health maximisation principle (extra-welfarism)

Health maximisation is a recent conceptualisation of equity in the health economics literature that has been the focus of equitable health policy in the UK [e.g. see Bell and Mendus (eds.), 1988; Baldwin, Godfrey and Propper (eds.), 1990]. It is based on the premise that health services exist to promote health and the health of the community should be maximised (see for example Culyer, 1988a). It states that 'a distribution is equitable if and only if it serves to maximise [the community's] health' (Pereira, 1989).

Two strands of thought have emerged from the health maximisation principle. The first is the original conception developed to explain equity. It sought to explain why it is desirable for the public to fund the health care of those unable to do so for themselves. This, termed the 'caring externality' model of health care, is associated with the pioneering works of Lindsay (1969) and Culyer (1971, 1980). It can be explained briefly as follows. Peoples' utility functions are interdependent (in the sense developed by Hochman and Rogers, 1969).⁶⁰ Such interdependence in

utility functions arises from peoples' caring attitude about others in suffering. This is not a concern for equality (Culyer, 1971). It does not matter whether it is the rich or the poor who are in pain or suffering. Under-consumption by the 'needy' imposes an external disutility on others.

This approach emphasises the individual's health status (irrespective of their other status—economic, social, etc.) be the focus of concern; that this is absolute (no comparisons of relative health status); and, that there is need for external support/compensation to enable individuals restore their normal health status. In this regard the approach touches on equity and justice in health and health care, by an appeal to compassion or the caring attitude of people, which itself suggests a criteria for redistribution. But, like other approaches to equity, the formulation fails to be precise on terminology used. One cannot tell what kind of equity rule is being advocated by the approach—it is not clear whether the focus should actually be on health itself or in the equity of utilisation for equal need. If the focus is on health itself, the implications for health equity policy are complex. Since health itself is not a metric that is easily quantifiable, how is one to develop a consistent definition of equity? Since there are difficulties with the concept of health itself, then the optimisation of the consumption of the commodities that affect health or the allocation of such resources in a manner that maximises the health of the community are feasible alternatives. Either way, the focus is on maximisation rather than distribution.

The second strand of thought in the health maximisation principle, and one that generated a heated debate in UK health policy, is the cost-effectiveness analysis (CEA) or cost-utility analysis (CUA) commonly referred to as the Quality of Life (QALY) approach. It affords primacy to efficiency. Distributional questions are important only when they further the cause of efficient use of resources in health production. This approach especially found favour with health economists at the University of York in the UK, although the idea was originally developed by Torrence (1986) in Canada. It has been alleged that CEA/QALY approach is the most suitable means of achieving equity whilst maximising the health of the community (Williams, 1985; Culyer, 1988a, 1990). The proponents argue welfarism and utilitarianism as normative accounts for discussing the quality of life, though profound, paradoxically emphasise the importance of utility theory (Culyer, 1990), but leave unresolved some ethical issues, similar to those raised above for the 'caring externality'

⁶⁰ This contrasts with the non-envy approach in which peoples' utility functions are independent of each others'. But See Culyer (1989a) who has pointed out a possible problem with the caring externality framework:

While it is commonplace to distinguish between utility as 'welfare' and utility 'as an index of choice' in normative and positive analyses respectively, some have difficulty in distinguishing [in a welfarist context] between utility maximizing behaviour that is altruistic and that which is selfish: is not the one who maximizes utility by giving in some sense selfish?

model.⁶¹ The QALY approach goes beyond utility and incorporates the characteristics of the people (Culyer, 1990). This is why it is also known as the *extra-welfarist* approach.

The QALY approach significantly departs from the traditional Paretian welfare economics, and rejects the willingness and ability to pay as appropriate bases for rationing health care. It rejects *welfarism*—the view that social welfare depends on welfare (Culyer, 1989a, 1990; Williams, 1988). Culyer has argued that willingness to pay is an irrelevance and may be directly counter to this objective if it is positively associated with ability to pay, but ability to pay is inversely associated with ‘potential for health improvement’, as is often the case. The extra-welfarist approach

... provides a conceptual framework for handling extremely complex issues in a systematic fashion that exposes each aspect of the argument clearly. It is less important what the cost per QALY is, than that individuals with responsibility for resource allocation in health care have a means of working through the issues so that they can come to their own informed view about ... different resource allocations. It is in this sense, the method is intended ... as an aid rather than a substitute for thought. (Culyer, 1989a).

The (efficiency) policy objective underlying the QALY approach is the maximisation of the community’s health. A weighted measure of the individuals’ remaining part of life is calculated (the weights reflecting the expected quality of life in each year when the characteristics of the individuals are taken into account)⁶². The individual QALY scores are then summed up. It is this sum the health maximisation principle seeks maximised by allocating resources efficiently. Williams (1985) writes

Procedures should be ranked so that activities that generate more gains to health for every £ of resources take priority over those that generate less; thus the general standard of health in the community would be correspondingly higher.

Now, the contribution of QALYs in the equity debate seems unclear. Some writers have argued the QALY approach does not address the equity issue. Lockwood (1988) argued the QALY approach ‘is in principle liable to result in forms of allocation that are *unjust* or *unfair*’ (Lockwood, 1988, p. 45, emphasis in original). Other writers have voiced similar concerns [(e.g. Smith, 1987; Broome, 1985). Smith (1987) is concerned that the QALY metric is obtained through the use of arbitrary assessment of the values of people’s lives. Williams (1987) responds to this by pointing that ‘far from obscuring the need for value judgements, the procedure highlights the value judgements and offers techniques by which they could be more explicit, that the process of quantification

⁶¹ Such as whether the quality of life is an absolute or relative idea; whether distributional aspects of the quality of life and the standard of living are best looked at in terms of outcome distribution or by weighting the characteristics of people.

⁶² Sen’s notion of basic capabilities is similar to the ‘characteristics’ approach adopted in most works on QALYs—in that both content that utility focuses too much on the subjective responses to commodities and their characteristics, but not enough on what those commodities enable people to do—say move around, learn, enjoy good health, etc.

is, by virtue of its explicitness, open to criticism at every stage.' Broome (1985) argued the measure (QALY) ignores population dynamics—that there will be extra children born as a result of people having 'extra' time. Since these children would not have come into being had these people not received extra QALYs (say, due to health care), we need to incorporate those effects into QALYs—which is impossible, since we have no basis for valuing life.⁶³ But Culyer (1989a) pointed out that

... there is nothing to be gained in the context of resource allocation decisions—making from taking an ontological view of QALYs, or life or lives. One is not concerned with the inherent cherishable worth of people but rather with the value of resources that we might spend in order to gain better health or prevent (or postpone) death or change the prospects of either for the better. If we spend £2,000 per person to protect them from consequences of some risk that is fatal for say one in five hundred, it is merely arithmetic that we shall spend, on average, £1 million per life saved: only in this sense is a life 'worth' £1 million.

But Wagstaff (1991) and Culyer (1990) point that the critics of the QALY approach have tended to misinterpret the concept—often equating it with utilitarianism. Wagstaff (1991) himself contends that the QALY approach is not based on individuals' own valuations of their health, it regards a QALY as of equal value to everybody and that allows the approach to avoid the many pitfalls of utilitarianism, especially that which implies resources would tend to be allocated away from people who place a relatively low value on their health. Apart from this (crucial) difference, utilitarianism and health maximisation (through QALYs) are similar—'like utilitarianism, health maximisation leads to the conclusion that resources ought to be redeployed away from people . . . who have a low capacity to benefit from treatment' (Wagstaff, 1991, p. 20). Moreover, the QALY approach favours the younger. Also, the principle leads to the conclusion that in a fully employed economy, resources be redeployed towards people whose output is highly valued and who, as a result of treatment, are able to return to work soon.

Wagstaff (1991) believes the QALY approach embodies to a certain extent, 'a kind of equality' by the fact that a QALY is considered worth the same to everybody.⁶⁴ Therefore, 'the outcome of resource allocation via QALY is *automatically* equitable irrespective of the degree of inequality involved and the type of person who fares badly.' (Wagstaff, 1991, p. 27). This reinforces Culyer's (1989a) earlier assertion that it may actually be equitable to discriminate against

⁶³ This argument fits with Broome's (1978) earlier argument that the value of a statistical life is infinite since in the end that statistical life turns out to be someone's (or actual) life. But now we know people are willing to trade small risks for improvement in incomes etc. (Jones-Lee, 1982).

⁶⁴ This was explicitly stated by Williams (1974) long before the QALY approach became a focal issue:

'The acceptance of the fact that an additional year of healthy life is intrinsically worth, say £1,000, no matter to whom it accrues, would lead to a much finer, humanitarian and egalitarian health service than we have at present, . . .'

those whose capacity to benefit from medical care is limited—those with worse health (their poor health cannot allow them to earn sufficient income, etc.). In rationalising the QALY then, Culyer (1989a) has urged the reader

. . . to allow that the sickest in the society are by and large those whom the marginal product of health care in terms of QALYs is highest, that these are also the poorest, and that when (*ceteris paribus*) health service *per capita* rises, the marginal product in terms of health falls . . . [then] it evidently follows that efforts to equalize geographical distribution of resources, to channel more of them to the sick and more of them to the poor, might be seen, not as distributional policies to be justified by equity arguments but efficient policies justified by health maximization.

Thus in this case, the extra resources are going to the person whose health is worse and also happens to be poor. Thus it is possible to incorporate equity and efficiency objectives within a single policy procedure. But the validity or otherwise of the argument depends on whether or not it is true that the marginal product of health care in terms of QALYs is higher for the poor. Not all would agree though. Pereira (1989) has pointed out that the opposite could as well be true, where the rich, due to their environmental and personal characteristics, are better able to respond to treatment and thus gain more QALYs.

The most notable response to these criticisms has been the revision of the notion that a QALY is worth the same to all. The use of weights (to take account of QALYs going to different people) has been considered the most appropriate avenue (Culyer, 1990). Preliminary results from work in this direction suggests the use of weights to scale down or up some people's QALYs is agreeable, leading some to claim that whatever the prevailing notion of equity, distributional aspects can easily be integrated into the QALY approach simply by maximising a weighted sum of QALYs rather than the unweighted sum (Culyer, 1989a). But there have been objections to the approach adopted to determine the weights (Pereira, 1989; Wagstaff, 1991). Pereira (1989) was concerned that the survey actively encouraged the respondents to opt out for some form of discrimination as to which groups of people should receive treatments. That resources are scarce is no justification that discrimination will effectively operate (p 37). Wagstaff (1991) argues

It is, . . . , difficult to distinguish between what people regard as just (or equitable) and what they regard as desirable. The latter will depend not only on what people think is just, but also on their degree of compassion—or caring . . .—and their own selfish interests. (p. 29).

Wagstaff (1991) compares various definitions of equity to determine which among them gives rise to the notion that society might want to attach different values to QALYs going to different people.⁶⁵ Only the 'equality of health' definition appears to provide such a basis. But he argues

⁶⁵ The definitions considered are: equal treatment for equal need; equality of access; equality of

the objective is not sensible since a blind pursuit (of the objective) can lead to absurd conclusions⁶⁶. He proposes an alternative formulation of the QALY approach to overcome these difficulties when equality of health is used as the guiding equity objective. The formulation proposed uses a Social Welfare Function (SWF) defined over the health of the population, constructed so as to permit some trade-off between inequality and efficiency. Such a SWF captures both efficiency and equity considerations, and

... provides a way of examining the extent to which society wants to accept a lower per capita health status in order to achieve greater equity. The SWF approach also provides the non-technical information required in order to determine needs, since it indicates which health improvements are socially desirable. (Wagstaff, 1991).

Whereas some conflict is possible between welfare maximisation and the notion of equal treatment for equal need (both perspectives of equity), the same is not true if the SWF were defined over health levels.⁶⁷ Health care maximisation leads to the rule that resources ought to be deployed so that the marginal cost per QALY is the same across all types of health care. If we are to take account of inequality at all, resource allocation decisions cannot be made entirely on the basis of the marginal cost of additional QALYs. 'Consideration must also be given to (i) the expected health of A and B in the absence of treatment and (ii) the degree to which the society is averse to inequality.' The QALY approach, '... even in its weighted form—fails to reflect the aversion society apparently feels towards inequalities in health outcomes.' Thus Wagstaff concludes the QALY approach does not merit any claim to equity concerns.

Now, Wagstaff's attack on the equity implications of the QALY approach is appealing, but its application, as he himself concedes, will be hindered by the data requirements. Although he suggests possibilities of generating the data, there are doubts as to whether it will be possible to generate meaningful data values. For example, how does one go about estimating the values of the distribution weights? There is no universally agreed method that can be used to solicit responses from people, that does not, as he puts it himself, depend on what people think is just, or on the degree of compassion and does not reflect people's own selfish interests

On a more general context, the whole idea of health maximisation has data requirements which are beyond the capability of Kenya. It requires for example detailed financial and cost infor-

health; and equality and choice (Wagstaff, 1991, pp. 29—34).

⁶⁶ He gives an illustration where by a redistribution of resources (from an initially perfectly equitable but low health position) improves the health of two individuals, whilst introducing some (small) degree of inequality. The equality of health objective would favour the original position—in spite of the high health for each individual in the new allocation!

⁶⁷ Welfare maximisation identifies persons in equal need, but assigning those in equal need the same access costs will not necessarily guarantee that they receive the same treatment.

mation of particular programmes as well as specific aspects/operations, if one is to be able to generate meaningful QALY or other health concept information. Unless specific task forces are commissioned for this purpose, which means time and resources, it is not possible to get even data which will approximate anything but the required concept. It is even hard now to know the proportion of health resources going to primary preventive or secondary (mainly curative) health care, mainly because the paucity of resources forces sharing of the them over these facets of health care delivery. Another problem is the interpretation of the probability numbers associated with most instruments used to generate QALY information. As Loomes (1988) conceded, this is a problem with the relatively aged and the semi-literate, and this may be more acute in a developing country.

5.2.9 Conclusions

No single principle of justice appears to be the sole basis for health care policy in Kenya, largely due to unclear statement of policy aims. But it generally appears there is a strand of thought in all policy pronouncements favouring some 'decent minimum'. Moreover, it appears this decent minimum is largely couched in terms of access rather utilisation, which might be more appropriate. Since a comprehensive package of services is obviously not possible, on economic grounds, it is therefore necessary to ensure the utilisation of a 'minimum package of services' is guaranteed for diagnostic services that are essential to ascertain the need for medical care services. Access to these services has to be equal, though not necessarily free. Access to other services—such as curative and rehabilitative services might not be equal but should be proportional to need with charges varying according to ability to pay. The case for providing free or heavily subsidised preventive services has similarly not been well articulated and there is an unclear and uncommitted voice on the part of the government. These services clearly need to be treated in a fashion similar to that recommended for diagnostic services. Clearly, there is need for a more systematic approach to equity in Kenya that if possible leads to unique or general (and consistent) policy prescriptions. The next section is devoted to the development of such a framework for Kenya, one which also incorporates the principle of efficiency developed in the previous chapter.

5.3 Equity and Access to Health Care: Towards Operational Measures

This section is concerned with what is/ought to an equitable allocation of health care resources in Kenya, where there is a large governmental responsibility both for global resources made available for health services and for their distribution. We attempt to provide a framework

which may be used to answer several questions at the root of equity concerns: ‘what is an equitable distribution of the available health care resources in Kenya among Kenyans?’, ‘Is the present distribution of health care resources in Kenya equitable?’, and ‘How might the distribution be made—and made to remain—more equitable? As it might be now clear from the analysis above, in health care, although the need for equity is universally acknowledged, there is less agreement about what that equality is about, i.e., *equality of what?* Is it equality of outcomes, such as health, or is it equality of inputs (i.e. health resources)? The philosophical theories reviewed above were of little help in shedding light on this problem. If a coherent policy for equity in Kenyan health sector is to be possible, it is necessary to have some broad principles that guide policy formulation.

The objective of the Kenyan health care system (discussed in the last chapter) was shown to be the improvement of the health status of the population. Hence *health* will be taken throughout the rest of this chapter as the entity whose distribution we are concerned about. This means the distribution of health resources is only instrumental and not an end in itself. But since the health literature has other possible contenders, these too will be discussed, and the merits/demerits of each pointed out.

5.3.1 What constitutes equity in health and health care?

Above we have defined equity in terms of fairness (section 5.2). Like health itself, equity may be defined ‘negatively’ or ‘positively’: Most approaches to equity adopt the view that *inequality* in the way that individuals are treated is judged *inequitable* if it is arbitrary or it relates to irrelevant characteristics such as age, race, sex, religion, and so on. In a more positive way, equity in health care requires that patients (actual or potential) who are alike in ‘relevant’ respects ought to be ‘treated’ in a like fashion, and that patients who are unlike in ‘relevant’ respects ought to be treated in an ‘appropriately’ unlike fashion. This is reminiscent of the notions of horizontal and vertical equity in public finance economics. Horizontal equity requires similar treatment of like individuals while vertical equity requires the unlike treatment of unlike individuals (Culyer, 1991a, b, c). If for the moment, we take ‘need’ (we shall shortly define ‘need’ and other possible candidates) as the criterion used for allocating health care resources, the principles imply that like needs should receive like attention and resources (horizontal equity). Horizontal equity can be expressed in terms of health care production, inputs, process or output (Cullis and West, 1979; West, 1981; Culyer, 1991a, c). On the other hand, those with greater needs should receive greater attention and resources too (vertical equity). These principles can also be appropriately applied in terms of con-

tributions to the finance for health care. The horizontal equity principle implies that individuals with 'equal ability' to pay make 'equal contributions', whereas the vertical equity principle demands higher contributions from individuals with higher ability to pay. The principle does not necessarily relate payments for or receipt of medical care to willingness (and ability) to pay for it. Even if need was otherwise defined, the implications would not change—so long as whatever (new) distributional principle in use does not correlate with willingness/ability to pay.

But to apply these concepts in health, the meaning of 'need' (or other focus of equity) has to be clear. This is taken next before we review some of the frequently canvassed candidates for the respects in which people are to be considered alike or unlike.

5.3.2 'Need' and equity in health care

'Need' (for health care) may be defined from several perspectives. One view considers 'need' as an inherently instrumental concept—an entity is said to be 'needed' only if it is a *necessary condition* for the accomplishment of something else (Culyer, 1976, p. 14; Williams, 1974). This lends normative content to the *instrumental* view of need—depending on what it is that is ultimately needed (it could be health, for example). This may have far reaching (moral implications) for health care policy—e.g., needs ought to be met in full if the aim is to save a life; some trade-off *may* be allowed when one case is deemed 'more deserving' than another, etc. Thomson (1987) objects to this view, as normatively it entails an 'element of practical necessity'—implying inability to function without whatever it is that is needed, or that the consequences of doing without would be perilous.

An alternative concept of need, the *absolute* or *categorical* view, has been advocated by Wiggins (1991). This view introduces the notion of something being needed if a final objective cannot be achieved without it. The end result of whatever is needed is a dichotomous choice:

... a person needs *x* [absolutely] if and only if, whatever morally and socially it is (economically, technologically, politically, etc.) possible to envisage occurring within the relevant life-span, he will be harmed if he goes without *x*. (pp. 14ff.)

But in health this view fails to allow that variations in the degree of illness, etc., command different levels of inputs (urgency). Granting that health care resources are necessary because they enhance health, prevent its deterioration, and or prevent death, health is the ultimate entity that is needed. But even in this case health itself is not a need, but only those elements that are necessary for its attainment—resources, personal characteristics, etc. This returns us to the instrumental perspective of need but one which now focuses on the penultimate—medical care—which enables

transitions from poor health states (ill) to some desired end-states (healthy). But at the same time it implies a notion of need which is binary. This is unduly restrictive for it requires that whatever is needed must always have a positive productivity in terms of its contribution to health. There is considerable evidence in the health field to the contrary. Though health care is usually expected to lead to an improvement in health, it does not (nor is it expected) always to do so. Sometimes care is given not because any improvement in health is anticipated or hoped for (though nobody would mind if any occurred) but just to show a caring attitude. In other cases, one may not expect a pay-off in terms of better health than before, but rather better health than would otherwise have been the case—health maintenance rather than full recovery, reduction rather than elimination of disability, slowing rather than stopping deterioration (Culyer, 1992, p. 13). In other words, most results in health are not binary, there is room for better, for worse, etc. But this also raises further questions concerning ‘need’ and health care. Is it an attribute that is inherent in health care or outside it? If outside, where? In the person who purportedly ‘needs’ the services? or community? Culyer, Williams and Lavers (1972) placed this controversy in perspective when they wrote

... it is difficult to tell when someone says that ‘society needs . . .’ whether he means that *he* needs it, society ought to get it in *his* opinion, whether a *majority* of the members of the society want it or *all* of them want it. Nor is it clear whether it is needed *regardless* of the cost to the society. (p. 14).

Williams (1974) suggested that medical care or treatment is ‘needed’ *if* it would lead to improvement of one’s health status. This means we must consider (i) the scope for improving a person’s health through treatment (a *technical* judgement), and (ii) the extent to which the health improvement is desired (a *social valuation* judgement) (Wagstaff, 1991). This rules out ‘capacity to benefit’ (see below) as a method of assessing need [because] society, in addition, has to decide how health gains to individual (say, A) compare with those of another, B. Wagstaff (1991) argues if the objective of equity is interpreted in terms of the amount of resources required (to treat a person in order) to reach some desired improvement in health, then, equal treatment for equal need does not ‘*in itself*’ provide a basis for determining an equitable allocation of resources—there are other extraneous factors to be taken into account, especially the *social valuation* of that care. Culyer’s (1976) example of the need for water by an individual dying of thirst in a desert illustrates the difficulties of incorporating these other considerations:

Does the individual dying in the desert need a glass of water? Technically yes, if he is to live. Normatively yes, if others agree that he ought to have it. If they do not, he may *want* it as much as it is possible to want anything, but he does not, in our definition, need it. (p 16, *emphasis* in original).

Following this line of argument, society has a duty to decide how the health improvements of one individual are weighted against those of another (or others). This requires a determination of the desired level of health for the two individuals. And it means if need is to be interpreted along these lines (and medical care is just a means to improved health), the criterion of 'equal treatment for equal health' will not provide an adequate basis for determining an equitable allocation of resources (Williams, 1988, p 117; Wagstaff, 1991, p. 31).

An alternative formulation of need retains the instrumental notion but links it to resources and the various end-states they serve, so that the 'need' for medical care is defined relative to some pre-specified end-state (Culyer and Wagstaff, 1991; Culyer, 1992). Thus 'need' may be defined:

. . . simply in terms of improving (or maintaining) health. A need for medical care is then said to exist so long as the individual's capacity to benefit from medical care is positive . . .
. . . We define an individual's level of need as the amount of medical care required to reduce the marginal product of care (or equivalently the individual's 'capacity to benefit') to zero (Culyer and Wagstaff, 1991, p. 16).

Thus need is defined as the minimum amount of health care resources required to exhaust a person's capacity to benefit. Resource scarcity implies that if 'distribution is according to need', all requirements for medical care cannot be fulfilled. This raises the question of which needs shall be met equally. If the objective is equality of final outcomes, say, equality of changes in health of the individuals, several possibilities may arise. The policy would be egalitarian if people initially have similar needs and capacities to benefit from health care resources. In the more realistic cases, where people 'need' (require) same or equal amounts of health care resources to reduce their marginal capacity to benefit to zero) but differ in terms of their capacities to benefit (how much their health improves, say in QALYs), 'distribution according to need' in the manner described by Culyer and Wagstaff (1991) is inequitable, since persons with greater capacity to benefit from the resources receive larger improvements in health. 'Distribution according to need' in such cases introduces inequalities in health *where none existed before*. Another situation is where people have unequal needs to begin with (meaning also unequal initial health states). In such cases, the concept of vertical equity has to be invoked, meaning resources have to be divided in proportion to the degree of inequality existing in the initial situation. Although it is possible in this case for 'distribution according to need' to result in there being less inequality in health after treatment than before, it need not always be so. Whether or not the inequality gap is reduced (for example in the two-person case) will depend on whether the person with the greater need also has the greater capacity to benefit. In other words, 'distribution according to need' in situations where needs differ may result in greater inequality in health after treatment than was before treatment (Culyer and Wagstaff, 1991, p. 10,

see also Culyer, 1991). Elimination of inequality in health may require a more unequal distribution of medical care than that implied ordinarily by vertical equity. In general, 'distribution according to need' principle will result in inequality of outcome. Therefore the application of the rule in health care would constitute a departure from the principles of treating persons in equal need equally and treating persons in unequal need differently than to treat them the same, and to discriminate more heavily in favour of those with relatively large needs than warranted by the Aristotelian version of the equity principle.

5.3.3 Need, equity and efficiency

Defining need as the amount of resources required to exhaust a person's capacity to benefit immediately links the notion of need to the notion of effectiveness (and hence efficiency) as discussed in chapter two, since

If health care is not effective, it cannot be said to be needed. If the technology that would improve someone's health for the better does not exist, current services cannot be said to be needed . . . In deciding what needs shall be met, it is essential to be able to form a judgement about the likely size of the benefit (in terms of, say, enhanced health). So, if needs are to be fairly met (for example, equal treatment for equal need) it becomes important to be able to measure and make informed judgements about the capacity to benefit. . . . an important element of the capacity to benefit is that it must be seen in terms of changes in health status. An absolutely or relatively high mortality or morbidity rate does not in itself indicate a high capacity to benefit: that depends on whether there is a capacity for the rate to be reduced sufficiently by the application of the relevant resources for it to command a priority relative to other needs (Culyer, 1992, p. 12).

A number of implications follow from adopting this definition of need. First, need and ill health are not synonymous. Hence a measure of ill-health, of whatever magnitude, cannot imply need for health care. The only health care that can be needed is that which has potential gain (promotes or reduces/postpones deterioration in health)—if health care would make no difference, then it is not needed. The second implication is that capacity to benefit is not the same as 'need' since the former is defined in terms of outputs whereas need is the for resources required to exhaust the capacity to benefit. It is possible for example, for two individuals to have their different capacities to benefit exhausted by equal amounts of resources. The third implication of the definition is it emphasises what can be done for people and done efficiently, rather than what has previously happened or what their current state is. The fourth implication is that it may be equitable and usually will be desirable for some need to go unmet. Because resources are insufficient to exhaust all capacities to benefit, it may be deemed equitable and desirable to direct resources from some areas (needs) to others. The fifth implication is that equal access is an incomplete equity principle because

access implies “. . . ‘gaining admission’ to the system in order *for needs to be assessed* . . .” (Culyer, 1991a, p. 35, *italics* in original). There may however be inequality because of differences in the impediments that different patients face in trying to avail themselves of that assessment. Access too is not an end in itself—it is instrumental, just as need is. Beyond the assessment stage, utilisation takes over and this will, among other factors, depend on peoples’ capacity to benefit as well as how equitably the needs are met. Distribution according to need may not necessarily lead to equality of health (or less inequality) and it would rarely do so, except in situations where patients are assessed to be equally sick (equal levels of need for medical care) and have identical capacities to benefit. Thus there are possible conflicts between equity in terms of meeting needs and equity in terms of outcome (distribution of health). Indeed, ‘distribution according to need’, ‘distribution according to capacity to benefit’, and ‘distribution to minimise or eliminate inequalities in health’ are potentially conflicting distributive principles. Granted the fundamental reason for caring about the distribution of health care is that it helps improve health

it is perverse to select need, capacity to benefit or initial health as the characteristic in proportion to which health care is allocated, for none of these will necessarily produce more equal distribution of health that is the true ethical imperative . . . none is as egalitarian as it may initially have seemed. (*ibid.* p 38).

Distribution by initial health is the most likely to lead to the worst inequalities since resources are diverted from those with least capacity to benefit from them (who most times happen to be the sickest, requiring substantial health care inputs for even modest improvement to be realised). Distribution according to capacity to benefit will produce a perverse distribution when those with better health have higher capacities to benefit (and it is usually so). An equitable health care policy will seek to continually reduce (with a view to eventually eliminating) the inequality in health. It would still meet needs, but in proportion to the deviation of all area’s health away from the population average. This will mean those further away from the average should receive more and vice-versa. It will, in short, apply the principles of horizontal- and vertical-equity. Whilst the overall objective of equitable health policy remains one of meeting the health needs in such a fashion as to reduce the dispersion of health in the community, subject to resource limitations, conflicts will often arise between the ethical desiderata of equity and efficiency.

The principle of meeting needs equitably will thus not be any simple proportionate mechanistic approach. It will incorporate many relevant parameters. Judgements about some of the process oriented conceptions of equity—opportunity of access, equity of access, which apparently do not translate into equality of the respective objects, have to be considered for the contribution, however minimal, they may make to the ultimate objective of an equitable distribution of health.

This means, for example, when, considering the equity implications of introducing user-charges or redistributing (relocating) health care facilities to reduce the patients' time and transport costs of reaching those facilities, decisions have to be made whether such policies will help reduce inequalities in health. Those actions pushing the system in the direction of increased equity will be encouraged. It is the responsibility of policy makers, on behalf of, and accountable to the public, to make such judgements—about needs and the best ways of meeting them.

5.3.4 Equality of what?—Possible 'candidates' for equitable distribution

Mooney (1982) suggested that the objective of equity in health may be stated in terms of equality of any of the following:

- ① expenditure per capita;
- ② inputs per capita;
- ③ inputs for equal need;
- ④ access for equal need;
- ⑤ utilisation for equal need—or equal cost for equal need;
- ⑥ marginal met need; and/or
- ⑦ health itself.

These notions of equity have different policy implications for individuals' or territorial equity. If in different parts of the country, the cost of manpower and other goods and services that are inputs into health care differ, then adopting the notion of equity implied by ① and ② will certainly lead to situations where individuals differ in terms of their final health status⁶⁸, i.e. inequality in final outcomes, because equality of expenditure (or resources) says little, if anything, about the services received or their effects on health. Equity defined this way will thus be rather limited. In such cases, adopting ③ may be a better option. In such cases, a policy of equalising health resources allocated to different areas, unless matched with outcomes and other considerations (as, to some extent, is the case for the Resource Allocation Working Party [RAWP] formula used in the British National Health Service [NHS]—see DHSS, 1976). In this respect, ③ and ④ may not differ (in terms of equity policy implications) if to achieve equal access requires differential levels of inputs (e.g., *ceteris paribus*, if more resources have to be allocated to sparsely populated areas, etc. so as to guarantee equality of access). However, equal access is a wholly supply side phenomenon and means that individuals face the *same* costs (to themselves) of using a health care

⁶⁸ That is, if we assume for the moment, more inputs necessarily lead to more 'health' output, which is not necessarily the case for some treatment protocols.

facility—e.g. because they live equi-distant from the facility. Unless ③ is similarly defined, the two yield different implications. ⑤ refers to actual consumption rather than to access but consumption is dependent on valuation *in use* (as reflected in the individual demands for health care). If individuals incur different access costs or have different demands, different utilisation rates may occur as Figure 5.2 below shows (for two individuals, A and B, where utilisation is measured in terms of visits to the facility, and benefits may be interpreted as perceived health improvement after the visit).

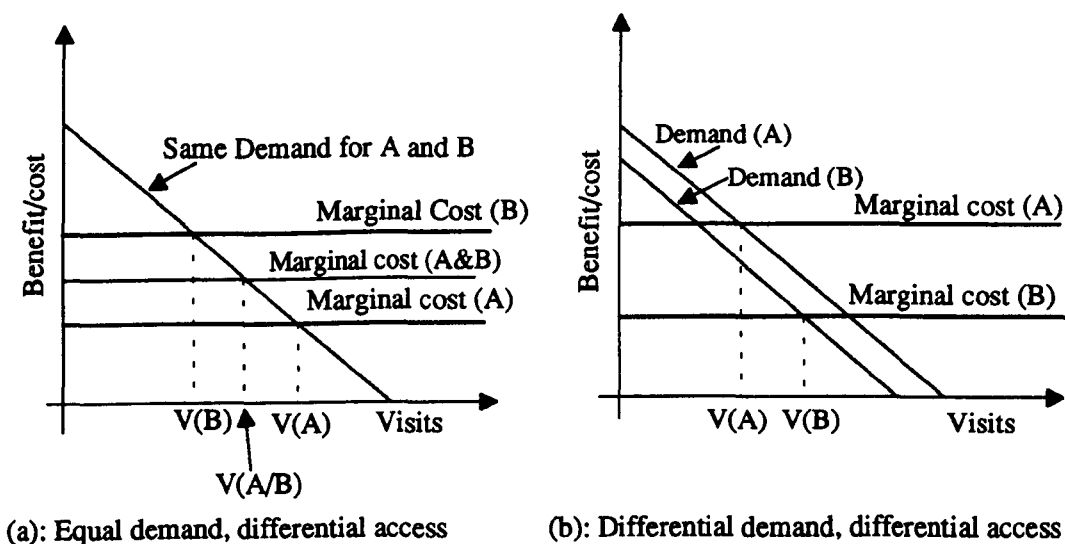


Figure 5.2: Differences in access costs/demands cause differences in utilisation

Equalising net need, ⑥, is possible if all regions are ranked in terms of their health care needs in the same order and each is rationally based on the principles of cost-benefit analysis (CBA—see chapter three). The objective of equal health, ⑦, is complicated by definitional problems, comparison problems and the expenses incurred in generating data for this purpose (Mooney and McGuire, 1987), but, a lot of work has been done focusing on this last as the objective of health policy. Concerning the distribution of health, the essential preliminary to any examination of the distribution of the health or ill-health is to decide on an operationally feasible definition of health.⁶⁹ Some economists have attempted to explain how these concepts can be applied to equity in health care. We eclectically review some of the suggestions.

⁶⁹ In defining health as an objective, one can focus on a number of definitions: for example, mortality or death rates among social classes—though not all illness are fatal, meaning this particular defini-

Mooney and McGuire (1987) argue for equality of access for equal need for individuals in terms of *fair shares*. Individuals should be able to demand and get health care services whenever they need them and in appropriate quantities (and quality). Their approach is about horizontal equity. Persons in equal health states should get approximately the same treatment, factors such as race, income, gender, place of residence, etc., notwithstanding. A fair allocation of resources under this principle suggests that each person should receive according to their need (Broome, 1988; Lockwood, 1988).

Equality of access as an equity objective has also received various interpretations. Le Grand (1982) suggested that access is best interpreted in terms of the time and money costs incurred by individuals in searching for and using health care facilities/services, measured in utility terms. Mooney (1983) likens equality of access to *equality of opportunity* (cf. Daniels, 1985, i.e., persons should be able to use health facilities when in need. The equality of access should apply only to those in equal need—equal utilisation for those in similar health states. But equality of access amongst those in equal need will not necessarily guarantee equality of treatment amongst those in equal need (Wagstaff, 1991). ‘Equal access for equal need’ should be interpreted to mean equal treatment for equal need—but this again has the problem of the meaning and method of establishing people’s (or regions) needs.

Mooney, Hall, Donaldson and Gerard, 1991, (hereafter MHDG, 1991) also interpret equity as ‘equality of access’, rather than the conventional notion of ‘equal treatment for equal need’—conventionally termed as *utilisation*, often used as a measure of equity in health care delivery. (They argue equity goals in policy documents mostly couched in terms of access rather than utilisation). The notion of equity based on equal treatment for equal need when people have different demand functions is an infringement of consumer sovereignty. In particular, it ignores the individual’s preferences, or, equivalently, applies uniform or ‘standard’ medical practices to all. But Culyer, Doorslaer and Wagstaff, 1991, (hereafter CDW, 1991), have pointed that even if people enjoyed the same access—according to the MHDG’s (1991) definition (say, facing similar costs when using the health services)—persons in equal need may end up consuming different amounts (or types) of health care because of differences in their demand functions. Health policy makers should be concerned why such differences in demand persist before making any judgement. This is neces-

tion will understate the true extent of mortality in the populations. Medical records are another source of information on health, but these are often incomplete or difficult to obtain (due to confidentiality), moreover, they are partial in the sense they cover only reported illness. The third source of health information is the sick-off days and illness as reported directly by people in survey interviews—but these are also not very reliable—they may have biases or untruths as well as under-reporting of certain (stigmatising) conditions such as mental illness and venereal diseases.

sary if they are to know whether such differences are due to differences in the marginal utility of income, or due to other more basic differences, or to simply because of differences in people's preferences for health care. CDW, (1991) point that whereas ' . . . policy makers *talk* about access to health care, policy *measures* are typically defined in terms of health care itself.' Thus, there is really no distinction between utilisation and access from this perspective, and this explains why many works on access have used measures of utilisation. This approach is similar to that adopted by altruistic models of externalities, which say nothing about equity and distributive justice (Culyer, 1980). They only explain what people regard as desirable (in this case, policy makers), which may be shaped to a certain extent by what they regard as just but, by and large, it also depends on their degree of compassion and their economic situation (cf. Mooney and McGuire, 1987; Margolis, 1982). If so, altruistic models are not the right instruments for investigating distributive issues, since

The whole point of making a judgement about justice is, after all, to frame it in a way that it is made independently of the interests of the person making it. That is precisely why Rawls (1972) and philosophers since him have been attached to the notion of the 'veil of ignorance' (CDW, 1991, p. 3).

However, 'equal treatment for equal need' is not necessarily inferior to the notion of 'equality of access', (as MHDG, 1991, put it)—especially in the eyes of policy makers)—just because it implies a departure from traditional welfare economics while the latter does not. MHDG, 1991, advocate for a concept that preferred by policy makers. But CDW, (1991) contend there may be a contradiction implied by this belief. While 'equal treatment for equal need' is inconsistent with Paretian value judgements, in reality, health policy makers *strive to sever* the link between ability to pay and receipt of health care. Since 'equal treatment for equal need' does just that, it should be favoured by policy makers too! Moreover, Gillon (1986) considered the applicability of the various types of social justice in health care and concluded that, of the various distributive principles, 'distribution according to need' is most favoured by physicians and other medical personnel. His interpretation of 'need' was 'ill-health'.

CDW, 1991, argue 'equal treatment for equal need' also appears to be consistent with the rejection of another value judgement of traditional welfare economics: that social welfare depends on, and only on, the utility of the various individuals who together make up society'—which is why policy makers believe that the business of health care is to improve health. This is the *extra-welfarist* approach according to which

health care is only a means to an end (viz. improving health) and the ethical justification of favouring one form of health care distribution rather than another (e.g., distribution according to need) has to be sought in the ethical justification of the associated distribution of health. The same is true of access, which in this view is also a means to an end,

Opting for one method of determining access . . . can . . . therefore be defended only in terms of ethical justification of the final distribution of health it gives rise to. (CDW, 1991).

The approach combines the fairness and efficiency of alternative distributional rules in terms of their implications for the distribution of (say, health) rather than in terms of how they fit with an inappropriate set of value judgements underlying modern welfare economics.⁷⁰

The definition of equity as equality of health is no less popular in the literature (see *inter alia* Black, 1980; DHSS, 1986; Whitehead, 1987). Any reduction in inequality (of health) is good and vice versa. Health care should therefore be distributed so that the end result is equal (shares of) health status between individuals or regions/groups. The definition provides a clear statement on what constitutes equitable distribution of health and health care—once the desired distribution of health is known, all that is necessary is to allocate the health care resources to achieve it. This will result in equitable allocation, given the people's (different) capacities to benefit from health care and the costs of the treatments.

5.4 Policy Implications for Kenya

We have seen that in Kenya, although policy statements are often stated in terminology of access, the basic objective of the health care system is health. Focusing on the latter leads to a situation where there will be less scope for conflict between efficiency (as discussed in chapter three) and equity. For the public health services, a number of practical policy implications follow from the combined analysis of the last two chapters. These fall into four broad categories: underlying (general) policy guidelines; policies to do with 'target groups'; information policy; and research policy.

⁷⁰ Another formulation of equity taking account of the non-health characteristics of individuals is the *equity as choice* definition (Le Grand, 1984, 1987). This suggests that inequalities are inequitable if they reflect inequalities in the constraints people face but are not if they reflect differences in tastes. This argument suggests we incorporate weights (say in health maximisation), these weights reflecting differences in peoples preferences. But Le Grand rejects such an approach arguing it would be inequitable for differences in health related behaviour (such as smoking) to have any bearing on the way people are treated by the system. He suggests instead people with such deviant health behaviours should be charged an annual premium to cover the expected costs of treatment arising out of their health damaging activities/behaviour, but should continue to receive the same treatment as other individuals. What Le Grand is saying in other words, is that although it might be legitimate and equitable for non-health characteristics to influence their rights to health care sector resources, any discriminations arising as a result should be confined only to the finance of the health care, but not the delivery of that care. In this perspective, there is no equity in the allocation of resources in the health care sector.

5.4.1 Underlying policy guidelines

Since there may be no single overriding equity principle to guide resource allocation, policy makers should be prepared to try a multi-pronged approach—for example, combining the caring externality approach with equality of health, and when necessary, seek to compensate for inequalities in non-health areas in terms of health care (and vice-versa). In distributing resources, need should be used as a necessary condition for equitable distribution of public health care resources. This means areas which are adequately serviced by private sector may not receive priority in public resource allocations, unless there are other overriding considerations—such as easing financial burdens of poor people in such areas, in which case other intervention mechanisms (to be discussed later) may be invoked. Equity requires equal and universal access for assessment purposes in the first instance (particularly for the primary and emergency services). There should therefore be as few disincentives (in terms of money costs or time) to these services. Access to ‘approved’ treatments—i.e., further consumption of health care resources (once the initial contact has established the existence of need) may be unequal, this dependent on the assessed needs of respective individuals. Finally, it should, as a rule, be ensured that the treatments that are ‘needed’—in the sense of there being evidence they are effective, are only those that contribute most to reducing inequalities in health.

5.4.2 Targeting resources to specific groups

In view of multiple deprivation, mainly in the form of hardships and adverse effects that some policies, such as user charges in public institutions might inflict on certain vulnerable groups, there is need for *functioning* policy and management structures whose province is the monitoring of the distribution of health and other relevant factors amongst vulnerable groups (such as the poor or at risk groups such as pregnant mothers, children, single parents, and so on. Simply because of their vulnerability, there is merit in seeking to identify those members of the community who are most at risk from changes in health policy and those who are least likely to take advantage of the available range of effective services. Geographical distribution will also be an important dimension because regions, like individuals, can be multiply deprived and also because health risks vary regionally, just as they over individuals. If, as is currently the case, policy were to move in the direction of greater decentralisation within budgetary allocations to districts, an important determinant of such allocations ought to be the differential patterns of (improvable) ill-health. Also, the potential contribution of different medical specialities to health differs considerably and discovering more about the significance of each for contributing to greater equality in health should be a priority. Fi-

nally, the current policy of reallocating resources from curative care to preventive and promotive services, apparently based on the popular fad that 'prevention is better than cure' ought to assess the relative impact on the distribution of health that will result from the relative expansions or contractions of the respective services. There has to be demonstrable effectiveness of the policy. Such a shift might be effective in raising the overall health status of the population, but again there is need to ensure resources are moved to areas where they secure maximum returns.

5.4.3 Implications for informational requirements

There will be need for information on the current distribution of resources and health in relation to target groups. Appropriate macro-summary statistical measures of inequality have to be developed and appropriately employed. These could then be used to make judgements about the effects of changes in the pattern of resource distribution on the health of individuals and especially target groups. There will also be need for information on the financial and opportunity costs of changes in resource patterns, including the overall funding from all sources. Information on time and monetary costs faced by patients in accessing the system and subsequently using it should play a major role in assessing the level of co-payments (i.e. user fees) to imposed, and this must take into account the long travel and waiting costs, and so on, incurred by some patients. This means it might be necessary to allow for area variations in these co-payments, unlike the present system where they are set at uniform rates throughout the country (if only for the sake of equity).

5.4.4 The need for further research

Further research will be needed in a number of areas, particularly that relates to collection of data and information on the aspects enumerated above. In this thesis we have, on account of statements from various government policy documents, ascribed the 'decent minimum' as the policy as the policy pursued by planners in the Kenyan health policy. The deficiencies (and strengths) of this perspective as far as equity in health care policy is concerned have been pointed out. The Ministry of Health may need to conduct its own research as to what notion of efficiency is to be initially adopted, but should progressively strive to move towards the adoption of the framework for efficiency and equity as outlined in the last and present chapters. Further epidemiological and economic work on medical and cost effectiveness, including the development and refinement of outcome measures of health for use both at the micro and macro levels is needed. This will entail research into the quantification of the health gains potentially to be gained from alternative delivery

strategies, and changes in programmes defined by targeted groups. The last chapter has indicated that preventable diseases and deaths continue to dominate our health status statistics. This should provide the initial set of clues as where it might be potentially rewarding to start research into 'avoidable' deaths. We suggest that the Ministry of Health immediately sponsor *ad hoc* research to make a comprehensive and authoritative review of the existing epidemiological evidence on the relative effectiveness of medical (including diagnostic) procedures, supplemented, where information is available, also by cost-effectiveness. Research into the actual distributions of health and sickness is also indicated. Preferably, such research should be conducted on a 'before' and 'after' basis, i.e. ahead of major policy initiatives so that the results can serve to evaluate the policy. Moreover, continued routine monitoring also demand research at intervals to monitor distributional changes over time, with a view to identifying not only their effects but also to assess the sources of changes as well as their desirability or otherwise. Finally, the Ministry of Health may also have to develop appropriate statistical measures of inequality and conduct surveys with empirical measures based on agreed concepts (of health, equity, access, and so on).

6. THE AGENCY RELATIONSHIP AND ITS RELEVANCE TO MODELLING A REFORM STRATEGY FOR THE KENYAN HEALTH CARE SYSTEM

6.0 Introduction

From the analysis of the previous chapters it is clear there is need for the Ministry of Health (MOH) to provide an enabling environment in which providers—both public and private—*have the right incentives to be efficient whilst being equitable in their operations*, i.e. one which leads providers to behave in accordance with the national health objectives. This chapter shows how an efficient and equitable system could be developed for Kenya.

The existing regulatory arrangements are not effective for promoting efficiency in use and equitable distribution of resources due to several factors. First, the Ministry of Health is organised on a hierarchical or 'pyramidal' structure with a communication network that is not effective for monitoring the performance of its own provider units (see chapter three), leave alone that of the private providers. Second, health personnel, particularly doctors, both in the public as well private sector, have little, if any incentives to be efficient, partly because such guidelines as exist are not fully enforceable, and also because of the scope for discretionary decisions that they have on account of the technological knowledge available to them and which allows them to exercise 'professional judgements' with respect to the utilisation of inputs as 'scientifically' required. Third, the payment to public sector doctors (and other health personnel in this sector) of a fixed wage and the cost-based retrospective reimbursement of public providers (which constitutes the financial environment surrounding the providers' decision-making process) cannot provide the incentives necessary for closing the gap that exists between technical knowledge and the motivation for improving provider performance. On the other hand, the private providers, driven by the desire to maximise profits, or due to resource inadequacy, may adopt practices that conflict with the national objectives of the health sector (see chapter three, section 3.2.2). Finally, the public budget is not distributed according to any rational basis, while private sector providers have little incentive to locate in 'needy' areas, to train own personnel, conduct research or to provide some types of care (such as preventive) or to specifically cater for the poor. If 'efficiency' and 'equity' are to form the basis for assessing resource utilisation in the health system, there is need to tackle these sources of problems in the health system

We first give a generalised theoretic account which shows why the present environment cannot motivate efficient behaviour in the operations of provider units. We will then summarise some factors that have to be borne in mind when designing reform proposals to change provider behaviour in such a setting and then outline the theoretical framework for reforming the system. For the public sector, this translates into designing a goal oriented budget policy, i.e. one specifying, within the resulting relationship, a management delegation relationship—through the budget—a policy *able to raise incentive-compatibility* on both sides of the relationship. How this can then be modified to take account of equity is also discussed. For the private sector, it translates into designing a policy where the objectives of the Ministry of Health (MOH) coincide with those of the private sector, i.e. there is no conflict. These relationships (between the MOH on the one hand, and the public and private providers, on the other), can be framed in terms of the principal-agent relationship. In both cases, we will show that the setting of an optimal reward structure, capable (in principle) of producing incentive compatibility—is adequate. For the public sector, (it will be argued) this is best cast in terms of some form of output-sharing, which, along with incentives for efficiency motivation (more of this anon), can raise the work-effort over and above minimum levels—by creating incentive compatibility and raising work-effort above some minimum (i.e. by removing X-inefficiency). Particular attention is paid to issues of performance monitoring and measurement to ensure that agents' operations serve to enhance overall system goals. These issues and their implications for system are discussed at length in this chapter. For the private sector, the principal-agent relationship is about *self-regulation*—in pursuit of *self-goals*—the problem becomes one of providing an environment within which, whilst pursuing own goals, the private providers also enhance the system's overall objectives.

6.1 Theoretical Approaches to Explaining Variable Efficiency in the Health Sector

6.1.1 Non-allocative inefficiencies that may sustain X-inefficiency in the health sector

Economists assume that variations in productivity within an industry are due to improper choice of input and output levels by firms. This may however be corrected by the operation of unimpeded market forces. The competitive market pushes the production units or firms to the production possibility frontier—those that fail to operate on it are simply pushed out of business. In this model therefore, variations in inter-firm efficiency are short-term phenomena, correctable by the market forces. Therefore resources tend, over time to be optimally allocated/utilised. This means, if firms are operating in an environment of competitiveness in both product and factor markets, profit maximisation leads to optimal utilisation of resources. If markets are not tampered with

—say, by government controls, taxes or other devices—or firms do not have objectives other than profit maximisation, this result would necessarily hold. But in reality this is rare. Often there are factors—internal to firm decision making process—that are frequently incompatible with profit maximisation. Policies based on the free market framework will not necessarily guarantee uniform (and efficient) levels of performance under such environments, because of intra-firm inefficiencies in the form of organisational inefficiency or motivational inefficiency—which are non-allocative inefficiencies—arising out of the existence of these non-profit maximising forms of decision making behaviour. Such factors may persist, especially in environments such as in the health care market where, because of the desire to secure fiscal socialisation of risk, health care is generally not only not viewed as an object of economic exchange among consumers and sellers and there is a desire for equitable rationing of the resources apportioned to it, but there exists a variety of interest groups—inside and outside the sector. Their cumulative effect is deviation from official goals which is not only attributable to managerial issues such as poor motivation, faulty co-ordination, bad co-ordination or inadequate resources, but also to the subjectively rational policies of interest groups attempting to pursue their separate interests. Besides, there is in general a tendency towards dynamic conservatism in health services—particularly in the practice of mainstream medicine—such that only decisions compatible with the dominant interest groups inside existing organisations are likely to be effective. In such cases, it is often easier to implement innovative plans by setting up new organisational structures rather than attempting to restructure existing ones, because the conservatism or system frozenness is difficult to overcome. Changes in the external environment—particularly in the market environment—can react forcefully against such conservatism and either force firms exhibiting such behaviour to change or go out of business.

The X-efficiency theory of Leïbenstein (1966) recognises that firms may exhibit characteristics that are inconsistent with profit-maximisation and allows the possibility of never reaching the production possibility frontier in some economic activities—due to a number of reasons. First, non-profit maximising behaviour may mean decisions are not based on a careful calculation of their potential effects. In such cases, the Yerkes-Dodson psychological law—[which says that at low pressure levels individuals will not put much effort into carefully calculating their decisions, but as pressure builds, they move more towards maximising behaviour; but if the pressure builds beyond a certain level, disorientation occurs and results in a lower level of performance (Leïbenstein, 1987)]—may be operative. This is particularly relevant in the health sector where most ‘professional decisions’ are based not so much on resource considerations as on some albeit incomplete/uncertain expectations about outcomes. It is not uncommon in most health care sys-

tems to find friction between planners and health professionals due to what the latter consider to be inadequate levels of funding. A second reason is what Leibenstein calls *inertia*— where functional relations are surrounded by inert areas, within which changes in certain values of the independent variables (e.g. increased financial budget or manpower allocations to a health facility) do not result in changes of the dependent variable (say, the number of cases treated). But theoretically, we would expect firms experiencing such *inertia* to disappear unless the capital markets and the product markets are not competitive. This qualification is true for firms in the health sector. In this sector, particularly in the public sub-sector, most providers do not rely on any organised market for their capital needs. Also, for well known reasons that have been widely analysed (Arrow, 1963; Evans, 1984; Culyer, 1989a, b), the health care market is hardly competitive. The existence of incomplete contracts (say, where the payment side is well specified but the effort side is not clearly defined) is another reason. Finally, the discretion that individuals have within certain boundaries may be used to counteract the allocative efficiency of profit maximisation behaviour of firms. This is especially true of activities where agents can manipulate constraints in their interest or perform in environments characterised by conflict, uncertainty and/or informational asymmetries. This is particularly relevant to causing X-inefficiency where the individuals have discretionary powers to choose the pace at which an activity is carried out. In such circumstances, the individual's performance cannot be controlled by imposing external guidelines—due to the informational asymmetry or the uncertainty. The individuals can choose to invoke these discretionary powers in preference to externally assigned objectives. Hence there is a latent Prisoner's Dilemma—where individuals have an incentive to move towards the minimum-tolerated effort level, while the firm has an incentive to move towards the minimum-tolerated working- condition-wage level. Often the result is an outcome which is intermediate between the Prisoner's Dilemma outcome and the optimal solution, where there is mutual gain for all participants (Leibenstein, 1987).

6.1.2 Budget restraint as tool for dealing with non-allocative inefficiencies in the health sector and its deficiencies

In some public activities—such as those of the health sector—the absence of markets enhances the possibility of existence of non-allocational inefficiencies as discussed above. But a pseudo-market can simulate the competitive result if allowance is made for the special characteristics of the output of health sector and the expenditure funding process involved in its provision. Competitive behaviour amongst health care providers is largely determined by both ownership (and the reasons for ownership—e.g., whether as shareholders interested in wealth maximisation

or with philanthropic inclinations) and how the revenues to fund their activities are generated. If the latter is performance related (e.g., related to quality and quantity of output), then the basics for efficient behaviour are already in place⁷¹. But where health providers are funded entirely from a general fund—through an external budget-based arrangement where revenues come out of a general tax fund or other sources not linked to the providers' performance—it is difficult to inculcate aspects of efficient behaviour. A budget can be considered as a means to delegate certain tasks, together with the authority to use a defined set of resources in order to perform those tasks. This opens the possibility of imposing tight financial restrictions while conferring freedom of resource allocation within this restriction. This is possible in a system where the health sector is largely controlled by the government.

Budget restraint has been advocated in public sector studies as the *price* element towards efficiency motivation in the sector (see Leibenstein, 1966). In the health care sector some studies have advocated the use of some form of a central budget mechanism so as to rationalise use of health sector resource (e.g., Rutten and van der Werff, 1982; Beske, 1982). A budget restraint transforms the provider's problem into a constrained optimisation problem and so serves to reflect resource limitation in the decision making process. The budget constraint gives the feasible sets of points for constrained optimising behaviour. Instituting a budget restraint policy aimed at reducing the relative inefficiency of facilities means that the amount of revenue allocated to a facility j is tied to some efficiency criterion (such as those discussed in chapter three), defined for all the facilities — $i = 1, \dots, j, \dots, N$. When health care providers are funded from a central pool through a process that is not performance related, no direct link appears between the institutional effects of the revenue and the expenditure sides of their activity. Besides, this type of funding arrangement has what Niskanen (1973) terms *illusionary effects* on the consumers, also the producers. It encourages consumers and producers to want to exhaust to zero the marginal benefits derived from demand and/or supply of the production units. The consumers believe that the tax revenue raised on their incomes 'entitle' them to the consumption of as much of the public health care as possible—up to the point at which the marginal benefits (to them) are exhausted. Predicting the pattern of individual rational behaviour under these circumstances is difficult since there is a 'composite effect trap' for increasing public expenditure. An increase in output demanded 'necessitates' increased revenue allocation to producers, which further encourages demand (due to availability), triggering further demands for increased revenue allocations, and so on, *ad infinitum*. Regulatory mechanisms for restraining the expenditure of producing units become necessary. Where no direct link

⁷¹ That is so in spite of the problems of the health care industry—the industry is traditionally non-profit making, where output and quality are hard to measure.

appears between the revenue and expenditure sides of the providers, and supply influences demand—as is the case in the health care sector (Evans, 1974a, 1974b; Fuchs, 1978; Cromwell and Mitchell, 1986⁷²) the suppliers may manipulate the demand side to attain *their* (suppliers') own specific objectives/targets. This can be a source of inefficiency. The funding body lacks the basis for refusing to grant requests for increased funding, except by resorting to arbitrary criteria, such as budget capping. Such arbitrary and indiscriminate budget capping penalises efficient producers and 'rewards' inefficient providers in the system.

The budget restraint introduces a *price* component on institutions' activities, and can be an implicit watchdog for efficiency that cause the desired changes in the public institution. But where this is not so, and the agents' supply factors influence demand conditions and therefore the performance of institutions, the possibility to link budget allocations to predetermined levels of performance could be a control policy that would provide efficiency motivation rationale against pressures for sequential expenditure increases. A competitive market ensures efficiency because all firms produce similar products and face the same factor and output prices, and the market forces of demand and supply cull out inefficient producers. When output is not sold under competitive conditions and revenue is not linked to performance, X-inefficiency may arise and persist.

In the literature on hospital costs, the considerable variation in the performance of comparable hospitals indicates that such variations cannot be explained solely by the notion that each hospital is unique. There has been interest in standardising these indicators so that comparisons can be made. But there are certain problems that make performance comparisons difficult or invalid. One of these is the true definition of hospital output and how this can be related to some notion of patients' health states. Much of the health economics literature points to the minimal contribution to health of health care services (Auster, Leveson and Sarachek, 1969; Grossman, 1972; Silver, 1972; Hadley, 1982; Sharif, Huq and Saleheen, 1993; Kalipen, 1993). Also, the output of hospitals—health—is a multidimensional concept that can be defined in physical, social and functional dimensions. This makes it extra ordinarily difficult to measure, even in principle, let alone the practical considerations of lack of statistics. The production of health care services is just but one method of attaining the objective of maximising health gains from available resources. Second, the apportionment of inputs to those multidimensional aspects that are the output of the health care industry is not easy due to the joint-production nature in the hospital sector. Third, some of the indicators may be misleading because they fail to reflect real improvements in patients health states,—e.g.,

⁷² Although what exactly constitutes demand inducement is far from settled, e.g. see Phelps, 1986, among others.

medical innovations that permit postponement of death rather than cure may be reflected in conventional performance indicators as reductions in hospital productivity.

6.1.3. Budget restraint and the 'proximate' behaviour of providers

Whether a budget restraint policy will move the system closer to pareto efficiency is dependent on the behaviour of the producers in hospital services industry. To determine and predict the extent of efficiency or inefficiency in the sector, a knowledge of hospital behaviour (and that of other facility levels) is necessary. In modelling hospital behaviour two relevant factors are considered regarding the unit of analysis:

- ◆ whether by focusing on the behaviour of the individual agents working in the hospital—whose actions ultimately determine the performance of the hospital—it is possible to extrapolate the behaviour of the whole organisation. If that is the case, one also has to consider the interrelations among the agents within the hospital unit to determine whether any specific group or sub-group more explains the aggregate behaviour of the whole organisation; and,
- ◆ whether the hospital as an organisation behaves as if it were a single consistent unit with a clearly defined set of behavioural rules that allow for joint analysis for purposes of modelling.

There are roughly two types of models of hospital behaviour corresponding to these two factors (Jacobs, 1974):

- ◆ (a) models that focus on agents' behaviour—either as physicians, administrators or as groups of trustees. These models assume the agent has a function to be maximised. The maximand could be income, profits, output or utility. The maximisation problem is usually constrained by some function defining the feasible set. The constraint may be income (Feldstein, 1970), or a function incorporating the marginal valuation of the individual's own time input that also may depend on income—leading to labour-leisure trade-off (Sloan, 1974), or some less well defined notion such as 'professional ethics' in terms of 'best practice' and patient welfare, etc., (Murray, 1974).
- ◆ (b) models that treat the hospital as an *organism* in itself. These focus on the behavioural and organisational aspects of hospitals—either from the view of hospitals as *bureaux* (Migué and Bélanger, 1974), non-profit making institutions (Newhouse, 1970), and

monopolies (Pauly, 1974; Goldfarb *et al.*, 1980), or from a bargaining type approach based on property rights (Clarkson, 1972; Harris, 1977).

Regarding the two factors and the corresponding models of the hospital behaviour available, none exactly represents the situation correctly for Kenyan public sector hospitals. Models in (a) are not an adequate representation of the Kenyan hospitals since doctors in public sector hospitals in Kenya are professional salaried employees—employed on terms determined by the government, including working time, as explained in chapter two. Besides, doctors can practice in own/private clinics outside the government working hours. In such cases, *shirking* may be a common feature. Also, doctors have a significant influence over the allocation of health care resources, arising from the professional training which imparts technological and professional knowledge to them, and empowers them to make decisions that have major resource implications. Hence they stand out as an influential group as far as hospital resource use is concerned (see chapter three, section 3.1). Yet, in Kenya, the amount of resources that they can control in any particular hospital is determined outside that organisation, most times without even consulting their opinion. The second group of models does not reflect adequately the situation in Kenya since physicians can flex muscles where resource allocation *within* the hospital is concerned—indeed, sometimes each physician would want to have as much resources as can be available to go to their particular departments/specialities within the hospital. Therefore we should expect them to shape the hospitals' decision process to a certain extent. Given this, what is the appropriate characterisation of the behaviour of the Kenyan public hospitals?

6.1.4 A model of hospital behaviour suited to Kenyan public hospitals

The Kenyan public sector hospitals can be represented as modified *bureaux*. Niskanen (1973) has defined a bureau as a non-profit making organisation (which may be government or privately owned), financed wholly or in part by a periodic appropriation or grant. The behaviour of bureaux can be portrayed in terms of (i) their characteristics; (ii) the type or nature of relations between the bureau and their environment; and, (iii) the maximand. The characteristics include (a) the inability of their owners (and, to a limited extent, the employees) to appropriate any part of the difference between revenues and costs as part of *their* income; (b) part or all the recurring revenue comes from source(s) other than the sale of output at cost price; (c) they specialise in the provision of goods and services that people prefer in larger amounts than would be supplied/bought if sold at their per unit cost price; (d) the goods they provide usually have high fixed costs of production, or cannot be adequately priced due to lack of property rights or due to inadequate marketing technol-

ogy. Bureaux are manned by appointed bureaucrats—(Niskanen likens them to civil servants)—full time employees. Niskanen, commenting about the behaviour of the bureaucrat asserts that:

The very problem that leads to the supply by bureaus (the difficulty of defining output) creates one of the more important problems of controlling bureaus in any condition for which the objectives of the bureaucrats are not completely consistent with those of the collective organization. The difficulty of defining the desired characteristics also makes it difficult to give appropriate instructions to the bureaucrat. When the objectives of the collective organization and the bureaucrat are consistent, the difficulty of instructing the bureaucrat can lead to a substantial variance of the achieved output around the desired output. When the objectives are not consistent, the difficulty of defining output and the consequent difficulty of instructing the bureaucrat can lead to an output that is systematically different from that desired (Niskanen, 1973, pp. 10-11).

The bureau's environment is defined by three parameter relationships: (a) its relationship with the funding organisation; (b) its relationship with suppliers of labour and other factors of production; and, (c) its relations with the customers of the services it produces. Bureaux have a dominant financier—usually a government department (such as the Ministry of Health in Kenya), which is in turn financed by general taxes. The funding organisation reviews the bureau's proposed activities and budget, approves it, monitors the methods and performance (rarely done in Kenya), and usually appoints the head of the bureau. The relationship between the activities of the bureau and those of the funding organisations are important—with the funding organisation relying on the bureau to supply the service whereas the bureau lacks other reliable source of funding. The whole relationship is in a way an awkward and personal one—characterised by incessant haggling between the two—sometimes with threats, other times with deference, as well as by gaming and appeals to a common objective by all concerned parties.

Unlike market type production relations, a bureau offers a total output in exchange for a budget, rather than units of output at a price. Usually, there is no explicit relation between the bureau's budget and the quantity of output. Instead, what is roughly expected of them is the provision of a certain level of service for the stated budget. *The demand for the bureau's product is revealed through the sponsor and will be related to the general population that the sponsoring organisation serves.* As such, the demand for the bureau's output is never directly revealed to it—only through the revealed preferences of the sponsor's demand for services can the bureau determine what level of activity to undertake. In the health sector this relationship is further complicated by the fact that profit maximisation is not even the appropriate objective, because of difficulties of quality and outcome measurement that make the efficiency and equity principles depend on managerial commitment of the provider units (usually committed to 'professional excellence' rather than overall social gains from scarce resources) and 'understandings' that are not quite legally enforce-

able contracts, even with public purchasers pursuing non profit maximising objectives. Also, on moral and ethical grounds, there is often a general feeling that health care services *ought not* be governed by market rules.

The sponsoring organisation is usually not a profit seeker, and the officers of the sponsoring organisation cannot appropriate as personal income part of the difference between the budget they would be willing to grant and the budget they actually grant the bureau. But there is a possibility that the bureau's officers can appropriate part of the bureau's expenditures as personal incomes.

Also, there is asymmetric information between the bureau and the sponsor. There is available to the bureau's officers more information—on the relationship between costs and production processes within the bureau, than to the officers of the sponsoring body. The bureau has a dominant position with respect to information for analysing the efficiency of its operations. Therefore, the sponsor may be passive—it only knows the budget it is prepared to grant for a given quantity of output but lacks either the capacity, incentive or the opportunity to obtain information on the minimum budget necessary to supply it.

The bureau hires most of its inputs from the competitive market and usually pays a single price for all similar labour and materials. But most often, bureaux, being monopoly (or so near) suppliers of the service, also are monopsony buyers of those specialised labour skills and material inputs used in the production of the monopolised services. This is partly true of the hospital sector in Kenya⁷³. The government hospitals are the main employers of medical school graduates and the largest buyer of pharmaceutical and medical equipment.

Some bureaux, besides the revenue they get from their sponsor, may raise revenues from the sale of output to individuals—usually at less than full cost. Such bureaux effectively face two distinct types of customers—with different demands for the *same* output—one type being represented by the sponsor, the other the customers who buy the bureau's output directly at a price⁷⁴. When the bureau is mainly dependent on revenues from per unit sale of services, its relations with its customers will be similar to those of the profit seeking organisation. But how far this relation approximates that of the profit seeker is also dependent on whether the bureau's sponsor is willing to compensate for any loss of revenues from such sales. A bureau whose sponsor is willing to compensate such losses will usually be indifferent to the interests of its customers, even if a large

⁷³ In fact, part of the requirements for registration with most medical professional bodies is proof of service for a specified period with an accredited institution after graduation.

⁷⁴ An example of the latter case in Kenya is where the hospitals have amenity wards, as in Kenyatta National Hospital or in some provincial hospitals, where the rates in the amenity wards may slightly differ from those charged in 'general' wards.

proportion of its budget comes from sales revenues. Effectively, the bureaux can practice price discrimination in such environments.

It is difficult to assign an objective maximand to a firm organised like described so far. However, Niskanen (1973) conjectured that the maximising behaviour of the organisation will be shaped by the behaviour of the agents that compose it. On this basis, he opined that agents will engage in 'purposive behaviour' in which some elements of their individual utility—rather than the general welfare and interests of the sponsoring body—enter the overall organisation's maximand. As such, the bureaucrat will tend to maximise the bureau's budget, since this will in turn affect those variables affecting their utility functions, the argument being that the higher the budget, the higher their associated utility—defined in terms of output, salary⁷⁵, reputation, power, patronage, and so on. Most of these factors are a positive function of the budget. Niskanen concludes that the bureaucrat will maximise the bureau's budget subject to the constraint that the budget must be equal to or be larger than the minimum total costs of supplying the output expected by the sponsor. There is thus X-inefficiency.

The model outlined so far does not 'adequately' depict the behaviour of managers in most public hospitals in Kenya. A more appealing statement of what will shape the manager's behaviour is that advocated by Migué and Bélanger (1973) who argue since profit maximisation and competition are not the driving motives in production, the hospital and the decision makers—represented by physicians and management—are free to give preference to other goals. Following these 'other' objectives has consequences on production levels and unit costs of hospital services. The hospital management derives personal satisfaction out of allocating hospital resources to various expenses. In this respect, the management finds itself in a situation similar to that of consumers having to decide how to allocate their budget among various goods and services. In allocating the available budget, the margin of discretion enjoyed by the management is equal to the excess of revenue over the minimum cost of producing the minimum level of output that can be tolerated (by the sponsor). This is the actual budget of the management, as opposed to the hospital's budget, which by definition is the cost of producing the actual output (according to the sponsor). The management's budget is what becomes available to be apportioned to various utility generating expenses—*plus* the cost of producing the output expected by the sponsor.

⁷⁵ This may not be very relevant in Kenyan (public sector) hospitals since the salary is usually invariant, irrespective of the budget size, but we may think of personal allowances and so on as having a similar effect.

The level of output produced and the unit cost of production will depend on the management's choice between several possible uses in allocating the discretionary budget. The choice becomes one of either producing a higher level of output and reduce the utility generating expenses by operating at close to minimum costs, or to maintain production at a lower level and use the resulting budget surplus to cover *[strictly] non-essential expenses* that bring personal satisfaction. Producing the minimum allowable level of output—or in terms of the terminology of section 6.1.1 above, the minimum tolerated work effort—maximises the amount allocated to other desired expenses. This has the effect of shifting the unit costs of production upwards, and total reported costs would then equal total budget at all levels of output. On the other hand, producing the maximum output at minimum cost avoids all unnecessary expenses. The hospital's final choice will obviously be influenced by the preferences of management and physicians. If the hospitals' primary goal is to distribute the services it produces as widely as possible (output maximisation) —it will make every effort to attain productive efficiency by minimising production costs in order to maximise the consumption of its services. Now, if this was the case, inefficiency would not be the problem that it is in hospitals and we must therefore discount the hypothesis that quantity maximisation is the driving force in Kenyan public hospitals⁷⁶. It does not seem to accord with casual/observed facts. In Kenya, this is further complicated by the presence of two lines of authority within the hospitals: some employees—mainly the non-medical—are under the authority of the administrative secretary (the permanent secretary in the MOH), while the medical personnel are under the authority of the Director of Medical Services (as explained in chapter two). Unless the two are co-ordinated, conflicts in interest which work against the over-all efficiency of hospitals can be expected to abound.

If the management is tempted to incur unproductive expenses to satisfy personal preferences, the objective of maximum output will not be reached. Here, the hospitals can be perceived as producing two kinds of goods: goods consumed by the patients and goods enjoyed by the management as utility generating expenses. The two objectives conflict—hence hospitals cannot attain minimum costs of production—and we can only conclude that the hospitals will operate at maximum cost—this means they will spend all the allocated budget—this conclusion will be shown to hold (in general) below—section 6.2 below.

An empirically verifiable hypothesis that follows from this analysis is that there is an absence of a relationship between technical or economic efficiency requirements, the goals pursued by the hospital and its decision makers, and the environment in which they operate (in which there is no

⁷⁶ This is despite the fact that the objective of the health policy, as already shown, might be so interpreted.

drive to maximise profits—there is no competition). This is the main source of inefficiency in public hospitals in Kenya. The financing arrangements and the semi-monopoly nature of the hospitals frees them from competitive pressures to be efficient, and this denies the MOH an alternative source of information by which to gauge the efficiency of hospitals—thus compounding the monitoring problem inherent in the nature of hospital output. There is weak external control.

6.2 Towards an Alternative Policy Model for the Kenyan Health Sector

6.2.1 The appropriate setting—Some factors to consider

The environment in which actual budgetary interrelationships occur, particularly as regards the scope of decisions and powers that providers have, is important *as it can be moulded in a manner that internalises the problem of incentive compatibility*. For the public sectors providers, this could be grounded in the budgeting process (as a form of an efficient reward structure). If there is to be efficiency (and improvements in the equity situation), there is need effectively to regulate the relationship between the centre (the Ministry of Health) and the periphery (the various providers that make budget requests from the centre). As previously described, the current system is regulated statutorily within a hierarchical system, with (nearly) unity command and spanned control of management—implemented through central directives at various levels of the ministry hierarchy. In this type of structure, ‘a focus on markets alone would ignore most, if not all, of the processes whereby health care resources are actually assembled and allocated’ (Evans, 1983). In this context, if the MOH is to have some impact on public provider behaviour, it needs to have the power to enforce its directives on policy targets that it sets for providers to achieve. This has proved virtually impossible because of the actual relationship between providers and the MOH, which is more of a ‘principal-agent’ nature than a hierarchical type of relationship. The following factors can be proposed as contributing to the former type of relationship:

- ◆ Following the hospital model discussed above, the MOH will, in general, always have less than full information about the performance of providers—there is asymmetric information concerning the nature of output and technological information which may often be used by providers as justifications for the actual input utilisation. It may prove too costly for the centre to generate sufficient information to verify these relationships.
- ◆ In addition, problems of behavioural dependence accompanied by informational dependence further weakens the communication network required for the efficient operation of hierarchical structures under conditions of no clear observability of outcomes in relation to inputs employed by providers.

- ♦ As a result providers, operating under less than completely specified contracts, can dominate the relationship (due to the asymmetric information which the MOH can do little about) and this makes the MOH virtually ineffectual and merely passive in the relationship. In such a context, only the identity of interests between the MOH and providers' interests can guarantee the achievement of the system's targets—which is unlikely whenever rational agents take advantage of the behavioural and informational asymmetries to support the power with which they are vested.

In this type of environment, the co-ordination of transactors in the health market towards common objectives, in the presence of conflicting interests, requires a much richer exchange of information and incentives than is provided through market exchanges (Hurwicz, 1973).

In general hierarchical structures by nature do not specifically recognise the lack of identity between individuals and their organisations since economic institutions in these contexts perform in a sort of neo-classical 'black-box'—in which rewards are expected to stimulate the individuals' productive response, preventing any form of economic 'shirking'. This view of economic reality has obvious deficiencies—it fails to recognise the significance of the factors listed above, and in particular that, within the hierarchy, fixed wages are not related to marginal productivities and therefore cannot *per-se* provide incentive motivation to agents (Lazear, 1979). Consequently, an allocation mechanism, designed to achieve efficiency (and equity) cannot accomplish these goals unless the mechanism is informationally feasible and compatible with 'natural' incentives of the participants. Incentive compatibility is the concept introduced by Hurwicz (1972, p. 320) to characterise those mechanisms for which participants in the process would not find it advantageous to violate the rules of the process. The idea may be traced back to the 'invisible hand' of Adam Smith who claimed that in following individual self-interest, the interests of society might be served (Culyer, 1985; ten Have, 1988).

Changes in resource allocation procedures, system organisational structures, introduction of competition, and regulation are among the leading mechanisms favoured by health care sector reformers as possible vehicles for enhancing incentive compatibility (amongst participants) in order to achieve efficiency and equity in the health sector. These principles suggest mechanisms that would accomplish the goals if all participants follow the stipulated rules, and, like other problems of incentive compatibility in economics, informational feasibility and incentive compatibility are often major concerns (Ledyard, 1987). The informational concerns arise because asymmetric information prevents the attainment of the (efficient) competitive outcome (e.g. see Akerloff, 1970; Rosen, 1985;

Kreps, Milgrom, Roberts and Wilson, 1982; Atkinson and Stiglitz, 1980). Incentive compatibility issues are about ensuring that participants in the process would not find it advantageous to violate the rules of the process, otherwise the mechanisms would perform other than as intended, resulting in inefficient allocation of resources.

In a planned system, such as Kenya's public health services, the problem of lack of incentive compatibility can be considered as a sequential one. When management has only imperfect information about the abilities and willingness of individuals to work, and the outcome of the agent (but not the agent's work-effort) can be observed, the problem of incentive promotion and the efficient reward structure to achieve it needs to address the question of efficient funding. The setting provides a constant perverse incentive for reducing the agent's work effort instead of moving it towards some targeted outcomes. Alchian and Demsetz (1972) have pointed out that in these circumstances individuals may rather freely adjust their work effort with respect to realised reward, so as to equalise the marginal rate of substitution in consumption. In other words, under less than complete contracts, each agent will feel compelled to take more leisure, because the effects of relaxing their rates of substitution between output and leisure will be less than the effect on the true rate of substitution. Thus any contract that pays a pure wage in these circumstances would not provide incentives to honour the agreement, since the worker could take their wage as given and select an effort level which tends to zero.

Given the environment presently prevailing in the Kenyan health care sector, the agency relationship appears to be the most appropriate setting for analysing the relationship between the MOH and the public providers. Some changes in the current statutory and financial arrangements to accompany the regulation for the allocation of overall responsibility to the decentralised agencies will obviously be necessary as it will provide the basis for promotion of efficiency (and equity) improvements. Two issues need to be dealt with:

- ◆ delineating the postulates in this relationship that support a model of incentive compatibility for the achievement of some common (system-wide) goals, and,
- ◆ identifying the appropriate (optimal) reward structure to be effected in the funding arrangements.

These issues are analysed in the next sub-section, the first within the framework of principal-agent relationship, where it is shown it is possible to have incentive compatibility through a modification of the central funding procedures that grants overall responsibility to the decentralised providers. This is considered in a rather generalised context, although with specific emphasis on

hospital units within the health care system. The second issue will be resolved in the context of development of an intra-sectoral financial allocation agreement, verifiable and agreeable to participants in the relationship, after the potential effects of a change in the funding procedures aimed at ameliorating system efficiency and equity are outlined and analysed. This will be used later as the basis for proposals to reform the resource allocation procedures (next chapter).

6.2.2 Funding and performance incentives under the agency relationship

We have stated elsewhere the state ought to ensure health services are available to all the population. In light of the discussion above, the relationship between the Ministry of Health and the providers of health care can be framed in terms of the agency relationship under which the latter (described as the agents), act on behalf of the former (the principal) in health care rationing decision-making process. In this set-up, the principal is the one statutorily granted the responsibility for ensuring health care services are provided to the population. Viewed in this light, the principal may take diverse designations—it would be the MOH if we have in mind the whole country, but it could also be a provincial or a district health authority that is responsible for health care services within a particular region in the country, or, indeed, any other decentralised level of management (say, in charge of a particular type of service, group of the population, and so on). The identity of the particular level of decentralised management does not change the nature of the principal, but it has consequences for the type of monitoring which can be exercised, given the amount of information available/required—at each of those organisational levels—for effective control.

The agent, on the other, can be the lower level organisations responsible for health services in a certain region (*vis-à-vis* the MOH), or providers—hospitals, health centres and dispensaries—that act on behalf of the principal. The physicians or other health personnel responsible for resource allocation to individuals may be considered as the ultimate units of analysis, because it is them who supply the work-effort in terms of medical service provision (in collaboration with other [passive] inputs) to produce ‘output’.

Viewed from the perspective sketched above, the agency relationship becomes the appropriate basis for analysing resource allocation and use since it immediately brings to the forefront the need to understand the environment which surrounds the relationship. In general, the principal enjoys the outcome of the agent’s activity and needs to appropriately ‘remunerate’ the agent in order to motivate and compensate for the disutility suffered by the latter. The reward structure must cope with both elements of compensation and motivation, taking into account the issue of informational

asymmetry raised above, which leads to 'adverse selection' and 'moral hazard' problems, in order to orient their efforts towards some planned targets of performance. These problems are relevant for both the public and private sector providers. Before we discuss the basis of system of incentives that can be adopted to do exactly that in the Kenyan case, let us first outline the essential ingredients of the reward structure within the principal-agent relationship.

Optimal reward structure in a single agency relationship

Following the discussion above, the health care providers (both public and private) can be viewed as 'agents' acting on behalf of the state (the principal) which has the responsibility to ensure adequate health care services are available to the population. The state has several areas of interest where incentive-raising is required. First, it wants to induce consumers to consume health care services without necessarily resorting to frivolous consumption, which is wasteful. Second, on the supply side, it has interest in ensuring that providers provide the right amounts of health care services. The interaction between the 'principal' and 'agents' in this set-up can be likened to that between the employer and an employee.

Following Lazear (1987), the principal-agent relationship may be defined in simple terms as follows. The employer seeks, through the clever use of an incentive contract, to get the worker to operate efficiently, which enables the firm to achieve lower costs and hence become more competitive. Competitive firms can offer higher wages to workers and in so doing draw efficient workers from the less competitive firms that cannot afford higher wages. The objective function faced by such a firm may be stated as:

$$Max_F. F(Q, E) - C(E) \dots\dots\dots (6.1)$$

Where Q is output and is defined as the *numeraire* and E is worker effort. The firm announces a compensation schedule $F(Q, E)$. $C(E)$ can be thought of as the monetary cost associated with supplying effort level E . Perfect competition in both product and factor markets implies that firms must maximise the worker net wealth as in (6.1) subject to the zero-profit constraint

$$Q = F(Q, E) \dots\dots\dots (6.2)$$

An incentive-compatibility problem arises because workers take the compensation scheme $F(Q, E)$ as given and choose effort levels to maximise expected utility. Once a worker has accepted the job, his/her problem is

$$\text{Max}_E. F(Q, E) - C(E) \dots\dots\dots (6.3)$$

The worker's effort supply function can then be obtained by solving the first-order condition associated with (6.3) to be:

$$C'(E) = \frac{\partial F}{\partial Q} \cdot \frac{\partial Q}{\partial E} + \frac{\partial F}{\partial E} \dots\dots\dots (6.4)$$

i.e. the worker sets the marginal cost of effort equal to its marginal return to him/her. The transformation of effort into output, $\frac{\partial Q}{\partial E}$, depends on the production function, e.g.

$$Q = E + v \dots\dots\dots (6.5)$$

so that output is the sum of effort, E , and luck (v). An incentive contract selects $F(Q, E)$ subject to the zero constraint (6.2), taking into account that the worker behaves according to (6.4). Various types of incentive contracts are subsumed by $F(Q, E)$ —the polar cases being salary on the one extreme, and piece-rate on the other (see Lazear, 1986). A salary is a compensation that depends only on input so that $F(Q, E)$ takes the form $S(E)$, such as an hourly rate. Irrespective of the amount produced, the worker receives only a fixed rate that depends only on the fact he/she supplies E of effort. The main deficiency of these types of incentive contracts is the difficulty of measuring E , or the input). On the other hand, piece-rate compensation depends only on output so that $F(Q, E)$ takes a form such as $R(Q)$. No matter how much or little effort the worker puts, remuneration depends only on the number of units produced.

A number of other intermediate forms of incentive contracts exist (see Lazear, 1987 for more detailed treatments) and one form which is particularly relevant to us is the one where payment is by *relative output*. Two variations exist in the literature. The first characterises labour markets as 'tournaments' where workers are pitted against each other. The one with the highest output gets the winning prize (e.g. in the form of high wages) and the other the losing prize (low wage). By increasing the spread between winning and losing prices, incentives are provided to work hard. The optimum spread induces workers to move to the point where the marginal cost of effort just equals the marginal (social) return to it. This approach has two advantages. First, it only requires that relative comparisons be made. It might be easier to observe that one worker produces more than another than to determine the actual amount that each produces. Second, it 'evens out' common noise—attributable to factors that have nothing to do with worker effort—so that the 'best' worker is still identified, i.e., relative comparisons are unaffected. This ensures that both risk averse and risk neutral workers are not penalised for factors beyond their control. The main disadvantage of this form of incentive contract is that it can only encourage workers to behave effi-

ciently if they are risk neutral. But workers can increase their probability of winning by subverting other workers' efforts—in other words, it discourages co-operation and team spirit.

The second approach suggests that if output levels can be observed, then payments can be based, at least in part, on team average (see e.g., Hölmstrom, 1982). Using team average allows firms to better address risk aversion. A peer average picks up disturbances that are common to all the firms and allows each to cater for the tastes of risk-averse workers. The problem with this approach, particularly in the health care market, is that output is not only difficult to measure, but in addition, there are problems of quality, which may be even harder to measure. Incentive contracts based on output quantity induce workers to go for speed and in the process, ignore quality. Of course, if quality can be observed, then the worker can be compensated appropriately for quality and quantity. The appropriate compensation schedule would essentially be the consumer's demand for the product as it varies with quantity and quality. These aspects are particularly relevant in the health sector, as the next section will show.

The next sub-section discusses how a similar approach can be used to raise incentive compatibility in the health care system, with special modifications which will be pointed out as the analysis unfolds, and explores the problems that may arise (specific to the health care) under each approach.

6.2.3 An application to centrally-funded health care facilities

Following Ledyard, 1987, we may describe an economic environment (e) as those features of the allocation process that might be taken as given, including a description of the agents, the feasible allocations they have available and their preferences for those allocations. The economic environment describes the agents, their feasible allocations, and their preferences for those allocations. Let $i=1, 2, \dots, n$ denote each of (in this case, public sector) health care providers, and A be the set of feasible allocations $\alpha = (\alpha^1, \alpha^2, \alpha^3, \dots, \alpha^n, \alpha^i \geq 0$ is a typical element of A , so that $\sum_{i=1}^n \alpha^i = A$, where A is the available (budget) resources to be allocated among the agents. Each agent (here taken as a health care provider) is assumed to have a selfish utility function $u^i(\alpha^i)$. The environment is described by $e = (I, A, u^1, u^2, \dots, u^n)$, where I defines the available production technology. Information is initially dispersed since only i , and only i , knows u^i . The specific knowledge that i has can be identified as i 's characteristic e^i . The economic environment in this case can be described in terms of u^i , since that information resides only with i , hence we can set $u^i = e^i$. Although it is possible to argue where providers submit periodic reports about their activities (as all

public facilities in Kenya are required to do by the Ministry of Health), the ‘omniscient’ (central) planner, since s/he has all the information pertaining to the i ’s is in a position to evaluate u^i , this is hardly the case (for reasons to be discussed anon).

Following Hurwicz (1960), one possible allocation mechanism requests information from each of the i agents and proceeds to compute a feasible allocation. It requests information from each agent according to a response function $f(r^1, r^2, \dots, r^n)$, where each agent reports $r^i = f(r, e^i)$, if others have reported r , (here, the r ’s can be interpreted as budget requests, where r is standard—common-knowledge or reported—information). Let $\mu(e, f)$ be the set of equilibrium response functions in the environment e . The feasible allocations can then be computed using an outcome function $\psi(r)$ and the net result in an environment e is the allocation $\psi[\mu(e, f)] = \alpha$ if all i follow the rules, f (Ledyard, 1987). This can be interpreted as follows: all (public) health care providers send their budget requests that are expressed as functions of prices, computed on the basis of aggregate demands of health care consumption in their respective regions. In equilibrium situations, each provider would then be allocated a budget according to the reported price and demand intensity. Allocations computed on this basis would be pareto-optimal.

The problem in the above budget setting mechanism is to identify mechanisms that will be *self-reinforcing*—i.e. for which no provider is encouraged to cheat (Coles and Malcomson, 1979). If a provider gains nothing (and possibly loses) by cheating (e.g. through use of the undetectable characteristic $[u^i]$ that is unobserved externally), then the allocation mechanism is said to be *incentive compatible*. Formally, an allocation mechanism is said to be incentive compatible in all environments (e) whenever *there is no agent i and no environment ε in e , and no characteristic ε^{*i} such that $\varepsilon|_{\varepsilon^{*i}}$ is in e (where $\varepsilon|_{\varepsilon^{*i}}$ is the environment derived from ε by replacing ε^i with ε^{*i}), and such that*

$$u^i\{\psi[\mu(e, f)], e^i\} > u^i\{\psi[\mu(e^i|_{\varepsilon^{*i}}, f)], e^i\}$$

where $u^i(\alpha^*, \varepsilon^i)$ is in i ’s utility function in the environment ε . That is, no provider can pretend to have a characteristic different from the true one in order to get an allocation in excess of the correct one—all providers have an incentive to report correct information. In such a case, the rules are compatible with their (providers) motivations. In other words, incentive compatibility involves identifying conditions under which, in our case, performance standards can be recreated by an allocation mechanism under the hypothesis that individuals will follow their self-interest when they participate in the implementation process. We propose that in the health sector such incentive

compatibility mechanism may be achieved through enhancing 'goal compatibility'—viz, shared objectives by both providers and purchasers of health care services.

By transforming the problem of health care resource allocation into one of output sharing, we can view it as a bargaining situation which is amenable to the tools of Game Theory since the existence of transaction costs, informational asymmetries and the characteristics of health care fit into the framework outlined above. In particular, a solution to the problem requires the search for some contract agreeable to both parties in the agency-relationship. In the context of the health care sector, this can be realised through the separation between demand division—that generated by the medical staff in facilities, and the supply division of the central administration. Under these circumstances, budget based contracts can be an efficient response to the dichotomy between actual management and the control/ownership of the provider units. Fama, 1980, notes that this type of solution requires only that

- (a) the activity should be capable of being carried out under minimum degree of discipline among agents (for instance, in competing for a fixed amount of resources), and
- (b) some 'market' for the recognition of agents' capacities with regard to their performance exists (e.g., the principal is willing to compensate according to performance)

The second requirement means that the agents' compensation consists of two parts—one defined over the subset of favourable outcomes, and another defined over unfavourable ones. Following the framework of reward structure outlined above, this scheme ensures agents are not penalised when performance is affected (particularly adversely) by factors outside their control. Consequently, the contract is specified with reference to a standard benchmark. As a result, the outcome is made dependent upon the relationship between actual and pre-stated performance standards, where 'performance' is defined over some observable attributes of the outcome, derived from the agents' activities.

This general setting for incentive enhancement could be optimal for the Kenyan public health care system since it leads to a prior commitment to share a pre-stated amount of resources (the total budgetary funds availed to the sector) with a possibility of monitoring 'minimum' levels of effort resulting, for instance, from average benchmarks. It would have the effect of providing incentives to the participants for behavioural changes by allowing some form of competition between providers—arising from the differential system of reimbursement that is to be based on some observed

levels of performance, which is able to motivate a change in the agent's work effort over and above the minimum.

But how can this framework be implemented to induce desired effects in the previously described relationship between the Ministry of Health and the public providers? The public health care sector in Kenya, as those in other health care systems, serves not only to relieve (mostly poor) people of the financial risks associated with illness and the cost of health care, but also has two other equally important functions (cf. Evans, 1987). By detaching contributions from expected risks—through their inclusion in the general tax system—it is able to redistribute wealth from those with low *ex ante* expectation of loss to those at high risk. In this way it integrates redistributive policies in a way that the private sector cannot. Second, it serves as a *collective* purchasing agency for consumers/patients. It can therefore be used to influence directly both the mix and quantities of different services provided, as well as their prices. In the process, the system can affect the distribution of wealth between providers and users/reimburers of health care. In addition, it can also serve to shape the technology, the decisions as to how health care will be provided, and the overall output of the delivery system. But, and more relevant to our present work, since in addition, the Ministry of Health—under whose portfolio the public health services fall—has a duty to ensure citizens have access to health care services at affordable terms, it should have powerful incentives not only to control its own providers, but also to respond to (as well as control) private providers' interests in the system. The political and administrative structure of the reimbursement process is critical in determining whether incentives to control providers interests would be effective.

6.2.4 Implications for improving incentive compatibility in the health sector

Our analysis so far indicates that the incentive structure in the Kenyan health care system—both at macro and micro levels—is not geared towards desired outcomes in terms of efficiency and equity. This lack of synchronisation of the incentive structure and general policy aims in the system can be divided into four major areas of concern: (i) general incentives in the health care system as a whole (as provided through the structure put in place by the government; (ii) incentives to providers—both physicians and hospital managers; (iii) incentives to consumers; and, (iv) incentives to insurers. These roughly correspond to the broad groups of transactors in the health care sector. In general, the performance of a social system is a function of both its structure and the rewards it offers to participants (Dunham, Morone and White, 1982). It is desirable to understand the complexity of the responses of the transactors to the pattern of rewards, but equally important

is the need to identify the participants themselves with the wide variety of institutions (Evans, 1987). But problems of the industrial structure of the health industry raise a fundamental social policy problem—about the design of institutional frameworks so that resource allocation decisions are assigned to people or groups of people who possess the necessary information to make optimal decisions and who have appropriate incentives to ensure appropriate correspondence of private and social objectives (Evans, 1983). In contemplating alternative structures that can be implemented in order to improve the allocative and technical efficiency in the health care system, we need to realise that different methods assign resource allocation authority differently, reflecting the inherent informational asymmetries in the health care market and the perverse incentive created both by this characteristic and by the social responses to it.

The framework discussed above suggests there is need to introduce a spirit of competitiveness in the allocation of resources to the various end-users in health care. Under ideal conditions, the competition should be driven by the consumer who, when in need, decides where to seek what services from, and so providers would have to respond according to ‘revealed’ consumer preferences. But for reasons widely acknowledged the patient-consumer in Kenya, as in most other systems, is not in a position to significantly influence the providers. Experience elsewhere suggests that providers and insurers are the most appropriate targets for economic incentives rather than fairly ignorant consumers since in most incentive schemes—case management, clinical budgets with or without diagnostic related groups (DRGs), competitive medical plans, or prospective rate setting—the agency role of both providers and financial intermediaries (insurers or health authorities) is reinforced by putting them financially at risk and making them accountable for their decisions (Nonneman and Doorslaer, 1987). Putting the general principle of reinforcing the agency role into practice in the Kenyan public health care setting would require a redefinition of the role of the Ministry of Health. At present, the ministry merely acts as a financial intermediary between the tax payers and the public providers. The budgets of the providers are set centrally and there is little room for competition—either by price or otherwise. Deficits are passed onto the government because most (public) providers are not accountable themselves.

There is therefore need for a different type of incentive structure to ensure there is increased efficiency. This is where the role of financial competition becomes important. Following the theory discussed in the previous section, there are two major ways in which the public providers’ role could be strengthened by putting them financially at risk. In the regulated approach provincial and district health authorities would get fixed budgets from the MOH budget from which they would have to cover their beneficiaries for a prescribed set of risks. They should obtain more autonomy in

negotiating payment contracts with individual providers. They also should be relatively free to negotiate remuneration packages for individual employees. Groot (1987) has demonstrated that tensions are created in situations where salaries of specialists and other health personnel are omitted from the budget formula and that this may be an obstacle to the emergence of more cost-effective practice patterns. Some strategic structural innovations are necessary if the framework is to be applied successfully. Since the budget carries with it a sense of authority, it should also reflect a sense of responsibility among the budget administrators. This demands there be a devolved administrative structure in which it is relatively easy to assign responsibility and blame in the event of non-function. Besides assigning administrative responsibility, there is also need to assign policy responsibility. By this we mean determining the appropriate levels of policy formation. While the central government, through the Ministry of Health will be able to intervene at any level of delivery, there is reason to believe the service units at different levels should be allowed some latitude in determining the service mix they would offer the local population. This is to be dependent (to a large extent) on local conditions. However, it will be important to ensure the balance between primary and curative care services does not get distorted as providers operate in a 'business-like' fashion. The Ministry can introduce incentives in the budget process to ensure those aspects it deems critical are not over-looked by the providers who bid for resources.

The other approach, which would require more profound changes in the system would be to make public (or even including private) providers compete for clients for a given budget. In this case, in order to prevent preferred risk-selection, certain regulation would be necessary to ensure providers are equitable in dealing with patients. This would require the setting of minimum benefit packages and improved supply of information.

In the rest of this chapter we review the theoretical and empirical evidence of the potential effects of competition in general, and selective contracting in particular, on some performance measures in the health sector. In reviewing the empirical evidence, we draw mainly on evidence from the developed countries—mainly the US system, the British National Health Service as well as evidence from the Western European health care systems—because it is in these countries where competitive health care reforms have been implemented with varying degrees of success. However, such limited evidence as may exist on the implementation of competitive strategies in the developing countries will also be discussed. The purpose of this review is to make an assessment of the kind of regulatory environment that is likely to lead the Kenyan health care system towards a suitable environment for pursuing the twin objectives of increased efficiency and improved equity.

6.3 Costs, Outputs and Competition

This section examines the effects of competition and regulation on some selected aspects of performance such as cost, quality, length of stay and capacity utilisation. The section focuses mainly on hospitals, the biggest user of any health system's resources, including in Kenya.

A profit-maximising hospital in a competitive environment, selling an undifferentiated single service, will always select an output level that is higher and a lower selling price compared to a profit-maximising monopolist hospital, their relative sizes (in terms of capital infrastructure) notwithstanding. In both cases, the unit cost, at whatever output level selected, would be minimised. The monopolist hospital may however earn surplus economic profits compared to the competitive hospital because its monopoly status enables it to pursue objectives such as those discussed in section 6.1.1 above, aspects that may translate into X-inefficiency. For example, if monopoly confers a degree of 'managerial slack', the surplus economic profits realised may be converted into types of costs that enhance the utility of managers—forms of expenditure that represent opportunity cost but that do not represent the minimum cost per unit of output (minimum marginal cost)—or are passed onto capital owners or employees in the form of remuneration that is higher than is necessary to keep resources in that particular use (Culyer and Posnett, 1990). Such expenditures no longer correspond to the true opportunity cost of resources.

Thus, in the competitive situation, each hospital acts as a profit maximising 'firm', responding to the preferences of the fully informed and knowledgeable consumers. Each hospital is regarded as being so small, and one among so many, that it cannot exercise control over any aspect of the market, except in its own (internal) cost structure. Without collusion, hospitals are forced to compete *only* on the basis of price, since consumers seek only hospitals with the lowest prices. Hospitals therefore have an incentive to operate at minimum cost in order to attract custom. Those not operating at least cost have this reflected in higher prices, to which consumers respond by switching their custom to elsewhere within the health care industry. The monopolist hospital is free from such pressures and can afford to operate at higher than minimum costs.

Similar comparative conclusions obtain for output-maximisation subject to a no loss constraint with respect to costs. Under both competition and monopoly, the output maximising hospital will set price equal to the average cost and standard results obtain. But predictions of standard economic theory may be inadequate because hospitals do not adequately fit within this framework.

Hospitals are often non-profit making organisations that are different from other organisations in a number of respects⁷⁷. First, there are often two lines of authority—medical and managerial. It is normal for both doctors and management to formulate hospital policies. The policies so formulated by each group need not coincide. The doctors may not even be employed by the hospital although they make decisions regarding the admission and treatment of patients. Given the ownership of the hospital may also be different, there is a management triangle (consisting of physicians, administrators, and trustees). Such a management triangle is bound to have some effect on hospital resource allocation. For example, doctors, ‘taking on the role of advocacy, and administrators, with a more global view of hospital activity and costs, may even respond differently to the same method of hospital reimbursement’ (Donaldson and Gerard, 1993, p. 118). This raises the question of who in the hospital reacts to incentives, how they react, and whether different groups of people within the same hospital react independently or in conjunction—an aspect that depends much on the physical structure, power and relationships within the hospital and the environment in which health care is provided (see below).

Quality of services is another aspect distinguishing hospitals from other organisations (e.g., see Lee, 1971). The predictions of standard theory will differ depending on how much emphasis is put on ‘quality’ as a determinant of resource allocation. If it is assumed the quality of care is given and unaffected by output levels, and the hospital receives either wholly or in part, revenues to finance its activity from patient care, expansion in hospital activity does not change the outcomes.

The following Figure (6.1) illustrates the point. When the hospital must finance all output expansion from internal revenues (part *a* of Figure 6.1), output will be increased up to the point A. On the other hand, if allowance for loss or a subsidy is granted, output will be expanded up to the point where the new constraint is reached (point A' in part *b* of Figure 6.1). In both cases, output is expanded up to the point where marginal cost equals marginal revenue.

⁷⁷ See Culyer, 1993 for a discussion of other factors differentiating hospitals from ordinary organisations.

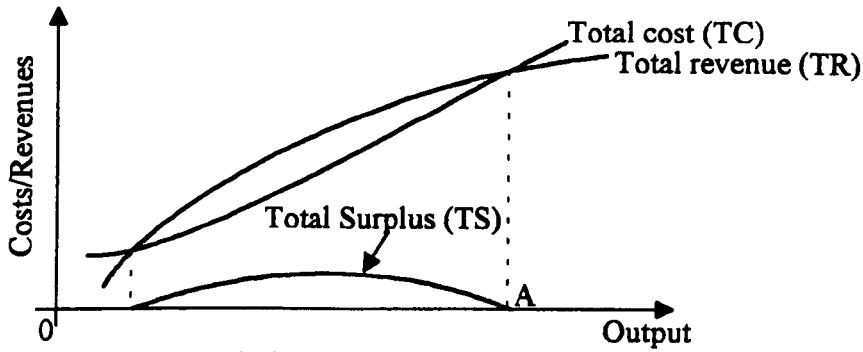


Figure 6.1a: Output-Maximising Hospital with no Subsidy

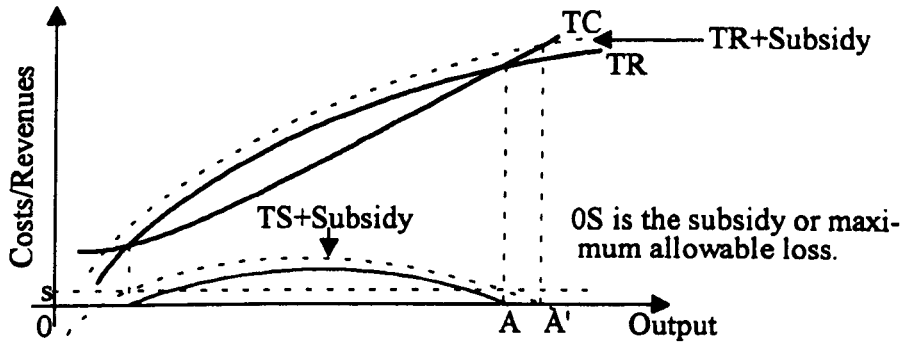


Figure 6.1b: Output-maximising Hospital with a Subsidy or Permitted Loss

Figure 6.1: The output-maximising hospital under different conditions

However, if variations in the quality of care as a determinant of expenditure is allowed, some problems may arise. Health care may be divided into two categories: necessities and supplementary products (Rice, 1966). Supplementary products are amenities, not strictly necessary for health care to have desired effects (say, in terms of patient health outcomes). If hospitals seek to maximise a function that contained both quantity and quality as its arguments (e.g., see Newhouse, 1970), it may become impossible to distinguish the increases in cost (of inputs) that are genuinely quality increasing from those that may be associated with the supplementary products—and a possibility of X-inefficiency arises (Culyer and Posnett, 1990). In such cases, the cost per unit of output may be higher than technical efficiency requires, for a given level of quality—particularly if quality was measured in terms of patient health outcomes *only*.

Another aspect that changes the results of standard theory when applied to hospital behaviour is the way health care expenditures are financed, particularly the use of health care expenditure insurance (or health insurance in short). If health insurance arrangements are such that there is full retrospective cost reimbursement for whatever health care services are provided to patients, we would expect hospitals that are profit-maximisers to seek to attract patients not through price com-

petition, but through *non-price* competition, since whatever costs are incurred would be reimbursed in full (Culyer, 1993; Culyer and Posnett, 1990). This frees hospitals from competitive pressures and allows them to pursue objectives that may strictly not be in accordance with profit-maximising behaviour—such as elevating the ‘prestige’ of the institution, increased security and job satisfaction for decision makers, including for the physicians. The pursuit of goals such as creation of a ‘prestigious’ institution can bring in particularly undesirable aspects of competition between institutions. Prestige is generally related to inputs. The wider the range of services available and the more expensive and highly specialised the equipment and personnel in the hospital’s service, the greater the status of the hospital (Lee, 1971, p. 49). But prestige is a relative thing. One hospital’s perception of prestige is a function of the (prestigious inputs) of *other* hospitals. If one hospital moves ahead by acquiring new machinery, such as a computerised tomography (CT) scanner, others may believe they have fallen behind. A new environment is created in which their prestige is no longer at maximum. New inputs may then have to be acquired to help them to *the* top perch in this new environment, mostly without even evaluating the contribution that the new equipment will make (e.g., in terms of actual health outcomes of their clients) to their operations—it is more of a psychological than economic game. Innovation by hospital ‘A’ ‘creates’ a gap for a competitor hospital ‘B’, between desired and actual prestige. Besides, other non-price modes may be employed by hospitals in such an environment including claims of higher quality of care (advertised to patients in terms of the additional inputs and procedures—more diagnoses and treatments per patient—which may [in most cases] be [wrongly] perceived by patient-consumers as welfare enhancing but which may not actually affect ultimate outcomes), convenient locations, better levels of clinical and non-clinical care (e.g., five star hotel type of services), shorter waiting times, (unconvincingly) courteous staff, and longer in-patient stays (‘just to be sure everything is okay’). All these forms of non-price competition are cost-raising, and, provided revenues come from priced services and insurance schemes, a rise in the desired status on the part of the hospital leads to its raising the price in order to raise the extra revenue required to fund the competition for prestige, and so on. So long as the costs can be passed to the insurer, there are no incentives to limit them and there is no mechanism for ensuring that these dimensions of hospital activity are optimal (Culyer and Posnett, 1990). Matters might even be more muddled if the physicians are paid on a fee-for-service basis, as this will tend to reinforce the adverse effects of the full cost retrospective payment system. The overall implications are that more inputs are used than is strictly necessary for the required task; there is idle capacity (meaning higher costs per unit of output); and factor use that may be insensitive to factor prices. All these factors lead to X-inefficiency. In short, the effect of competition on costs is uncertain, but it can easily lead to product differentiation and,

consequently higher costs. When (non-profit maximising) hospitals maximise quantity and quality of a service, unit costs are minimised at each activity level but activity levels will generally exceed the optimum because consumers and producers alike lack incentives to evaluate the opportunity costs of the services (sought or provided). Thus there is moral hazard (see Culyer, 1993; Donaldson and Gerard, 1993, chapter 3).

In the United States, retrospective reimbursement was the most prevalent form of hospital reimbursement before 1982, and available evidence suggests that hospitals behaved in the manner described above. For example, Robinson and Luft (1985), using 1972 hospital data found that hospitals operating in more competitive environments (as judged by hospital density) reflected higher costs per day and per case than those operating in less competitive situations after adjustments for case-mix, scale, ownership and so on and there was a tendency for costs to increase monotonically with the number of neighbouring hospitals. Duplication of clinical services, variations in efficiency of provision of specific services, higher reserve margins, higher prices and higher quality have also been found to be phenomena common to competitive environments (Luft, *et al.*, 1986; Wilson and Jadow, 1982; Joskow, 1980, 1983; Woolley, 1989; McPherson, 1989; Wennberg and Gittlesohn, 1982)—factors which have been attributed (by Culyer, 1993) to the inadequacy of the underlying technology that determines the production and cost functions⁷⁸.

While these phenomena obviously complicate assessments the efficiency of competition in the health system, theoretically, it appears that most apparent inefficiencies of provider competition owe their origins by and large to the payment mechanisms on the demand side (*ibid.*), particularly the use of health insurance—i.e., to the form of competition. Comprehensive insurance coverage removes the incentives for consumers (whether patients or physicians) to select suppliers on the basis of cost, and generates pressures for suppliers to compete on a non-price basis. The greater the degree of competition for patients, the greater the extent to which economic rents will be eroded in cost increasing expenditures. Under a retrospective based reimbursement scheme, hospitals have no incentives to minimise costs in order to protect net revenues. The combination of third party payment and retrospective reimbursement creates a situation in which the objective of consumers to obtain the highest quality care available is accommodated by suppliers facing an essentially open-ended budget constraint. There has therefore been interest to investigate the ways in which competition between health care providers can be made to work in such a fashion as to ensure (i) that the product mix, output rates and quality approximate their policy targets and (ii) that what-

⁷⁸ See Culyer and Posnett, 1990 and Culyer, 1993, for a review of evidence along this line, from these and other studies.

ever is produced is produced at least cost. Most policies that have been devised or implemented to 'tame' competition in light of the above two desirable aspects fall under the rubric 'managed care'.

6.3.1 Competitive reforms, regulation and managed care

Competitive reforms and managed care

Most competitive reforms that have appeared in the health sector reform literature are closely linked to managed care. But competitive reforms are generally not synonymous with managed care (Arnould, *et al.*, 1993a, 1993b). Competitive reforms are generally used to describe those policy reforms that strive to improve the performance of the health sector markets by harnessing competitive forces—either through direct imposition of competition or by bringing market forces to bear indirectly through a process of 'yardstick competition' (see Schleifer, 1985). The purpose of competitive reforms is not to eliminate health care regulation such as there may be, but to restructure the role of the government (or more accurately, payers), to harness competitive incentives—through the restructuring of the regulatory framework towards increased reliance on market mechanisms rather than command-and-control devices. Managed care reforms are mechanisms—some competitive and some regulatory—used to enhance competitive policy reforms or purely regulatory reforms or socialised reforms. They can be instituted in both centralised systems (e.g. the British National Health Service) as well as in the more market oriented (such as the US and other similar) systems.

Managed care systems strive to improve the performance of the health care system by actively 'managing' the patient's choice of provider care (Arnould *et al.*, 1993b). Managed care is designed to intervene in decisions made by health care providers to ensure that only appropriate and necessary services are provided and that those services are provided efficiently (Langwell and Menke, 1993). The novelty of managed care systems is that contractual arrangements for provision are made with selected providers to provide a comprehensive set of health care services to a designated population or group of same, usually at negotiated prices. Financial or other incentives may be used to 'steer' patients towards these providers and the providers have ongoing accountability for their clinical and financial performance through formal quality assurance and utilisation reviews (Hoy, Curtis and Rice, 1991).

The essence of managed care is that providers become accountable to their patients or their representatives (usually payers—the state or other insurers/payers) for quality, effectiveness and the cost of care (Lynn, 1991). The interventions used are diverse—they could be in the form of in-

centive contracts or again they could be in the form of regulatory or administrative directives such as utilisation reviews and protocols. The organisational innovations that have facilitated the implementation of managed care strategies have mostly originated in the American health care system. Some have evolved in response to the needs of the private sector such as the preferred provider organisations (PPOs) and health maintenance organisations (HMOs). Others, such as selective contracting, are public sector initiatives—that have occurred where elements of managed care have been directly carried out by the government—as mechanisms for competitive reforms.

The implementation of managed care and competitive reforms has taken diverse forms in different countries, the differences being dictated by underlying social system structures. In the USA, the main ones being the diagnosis and related groups (DRGs) based Medicare Prospective Payment System (PPS) and the introduction of more competition in the health care market through selective contracting and other innovations such as Health Maintenance Organisations (HMOs) (e.g., see De Lew, Greenberg and Kinchen, 1992; Enthoven, 1987). Other countries are also experimenting, or considering experimenting with DRGs (Wiley, 1992). The UK and some European countries have proposed systems of reimbursement that encourage competition (Saltman and von Otter, 1992; Culyer, Maynard and Posnett, 1990). Others, such as Canada and some European countries use global budgeting (Evans, *et al.*, 1989; Wolfe and Moran, 1993). Some developing countries have also implemented reforms such as devolved budgeting and competitively contracting out of some aspects of health services (McPake and Banda, 1994). The theoretical aspects of these reforms are briefly reviewed below, before evidence on their effects is presented.

Prospective payment systems

In the United States, the drive for reform was as a result of a desire to constrain hospital cost escalation particularly after the introduction of Medicare and Medicaid. Besides attempts to sharpen consumer incentives, efforts have concentrated on two broad strategies: changes in the way hospitals are reimbursed accompanied by price controls on hospital charges, and the introduction of a greater degree of explicit price competition into the market through selective contracting. In most schemes designed to constrain hospital inflation, hospitals contract with financial intermediaries to operate within predefined budgets. Historic hospital costs and utilisation patterns, adjusted for inflation and changes in patient workloads are used by the funding agency to form the basis for estimating hospital workloads. The estimated level of workload is then costed in terms of a fixed price per type of case to determine the hospital's budget in the forthcoming year. These

budgets can be administered globally for the whole hospital or at departmental levels with teams of clinicians (this being termed as clinical budgeting). In order to encourage hospitals or departments to remain within budgets, those whose costs exceed the fixed rates receive no supplementary funding while those with lower costs are able to retain the surplus, provided such surpluses are spent on patient care. This approach is an attempt to break the direct link between payment and costs. A payment system which pays hospitals according to self-reported costs provides weak incentives to be efficient. Setting a fixed price based on the average cost of all suppliers in the industry breaks that link and generates incentives for producers to minimise costs.

The Prospective Payment System (PPS) is one example of policies that use this approach. It uses diagnostic and related groups (DRGs) to categorise hospital inpatient activity and to set price per case. DRGs group patients according to diagnosis and resource use and the reimbursement rate per case is set prospectively for each DRG (diagnostic related group) category according to the average cost for that DRG. Thus price per case is constrained by the funding agency and the hospital is free to decide on the quality and quantity, in terms of length of stay, procedures administered or the number of cases admitted.

It may be expected that under this system hospitals will be encouraged to minimise costs in order to maximise hospital net income. However, how hospitals behave will be determined to a large extent by market structure (Culyer and Posnett, 1990, p. 23). Monopoly hospitals may be expected to continue to minimise costs under a fixed price payment system, provide the minimum level of amenity consistent with demand and to convert any earned surpluses to profits or other utility raising expenditures. Competitive hospitals also would strive to minimise costs but in this case it is possible that non-price competition will continue to exert a significant influence on final outcomes. For hospitals where the PPS rates are higher than the average cost of treatment, competition for patients may lead to increased expenditures on amenity and quality enhancements until the difference between the fixed rate and the hospital costs is eroded. Theoretically, such hospitals could expand their market shares over the long term, driving out those competitors unable to break-even (as their market shares decline)—and so an industry that was once competitive achieves monopoly status.

On the other hand, were the PPS rates to be lower than average costs, competing hospitals would be expected to try to reduce their expenditures per case and this process can be expected to continue until costs per case are equal to the PPS rate. In this case, it might be expected that amenity costs and hospital resource use will fall, which will be reflected in falling cost per admission and

per case. This decline could be achieved through any one of several possibilities: reduced hospital stays, substitution of less expensive inputs for costlier ones, reductions in quality of care or any combination of these (Donaldson and Gerard, 1993, p. 122). Also, hospitals may select-out those patients within each DRG who will be less costly to treat (Culyer and Posnett, *op cit.*). Hospitals may also have an incentive to select-out the less severe DRGs and treat them on an outpatient basis, rather than admit them as inpatients. The effect of this behaviour would be reflected in an increase in the average cost per inpatient case since the inpatient mix contains a relatively higher proportion of the more expensive (and more severe) cases. It is also possible that hospitals will—either deliberately or inadvertently—misclassify cases by ‘shifting’ some into DRGs which makes them appear more complicated than they actually are in order to attract a higher prospective payment than would be otherwise.

In summary, in a PPS, prices are not the result of negotiated contracts, nor does the system attempt to steer patients. It is competitive reform that applies ‘yardstick competition’ where prices are not determined by a competitive process but set prospectively by an oligopolistic buyer. The notion underlying a PPS is to place providers in direct competition in the absence of direct competition. It seeks to create incentives for providers to behave as though the market were competitive and ‘forces’ them to minimise costs. If a fixed-price PPS system is to generate cost savings, it will do so only for those hospitals for which the fixed rate is lower than current costs. But it is important to know exactly what factors have contributed to reduced costs before claims of hospital efficiency improvements under PPS can be substantiated, since efficiency may be sacrificed if the cost savings are achieved at the expense of quality of care or some patients remain untreated or if treated, are not treated appropriately. Reducing costs through reduced lengths of stay, earlier patient discharge, or reduced in-hospital resource use *enhance efficiency only to the extent that patients are not adversely affected and that additional costs are not imposed on other suppliers* such as primary care physicians or other segments of the community (Culyer and Posnett, 1990, p. 27) In section 6.3.2 below we shall review some of the empirical evidence that has accumulated assessing the effects of PPSs along these lines.

Competitive contracting

Besides changing the basis of hospital reimbursement and price controls on hospital charges, *explicit price* competition has been introduced into the health care market through competitive contracting (Arnould and De Brock, 1986). The PPS system reviewed above is of course one

form of competitive contracting. Other forms include *selective contracting* and *competitive bidding*, which are often linked. These two are innovations that seek to identify efficient providers (either public or private or both) or those willing to bear some *or all* the responsibility of the financial risk that goes with service provision. Through the process of bidding, sometimes followed with direct negotiations between providers and funders, contracts to provide services are awarded to providers who appear promising enough to provide substantial quantities of services at a lower cost. The bids may specify price only or may include a variable quality component if this has not been specified initially. These innovations clearly fit within the description of managed care because the choice of provider is actively managed, but also represent competitive reforms since they also entail restructuring regulations to facilitate market competition.

Compared with fixed-price PPS, selective contracting is more flexible since it allows providers with excess capacity (e.g., empty beds) to make use of it through selective price discounting, whilst offering payers the opportunity to control the range and quality of contracted services. Besides, in a truly competitive environment, the fixed-payment contract schedule becomes the effective upper-bound to prices rather than the norm. But contracting has the disadvantage that its effectiveness is dependent on the existence of genuine competition—‘without competition, contracting becomes impotent’ (Culyer and Posnett, 1990), and also the fact that there always is the fear of a potential disruption of services in case of bankruptcy or insolvency by the winning contractor(s)⁷⁹. The emphasis on competition may be a critical factor in Kenya as we have already shown that besides the uneven distribution the existing service network, it is not adequate. Nevertheless, this does not rule out the possibility of selective contracting in Kenya as will become apparent when we review some of the evidence on contracting in section 6.3.2 below.

The incentives generated by selective contracting are to a large extent dependent on the form of risk sharing (between providers and payers) arrangements. The uncertainty of health care cost coverage means the payer is uncertain of both the number of people from the covered population that will require treatment and hence the total cost of treatment during the period. An ideal contract from the point of view of the payer is the one that specifies in advance the total expenditure, such as capitation contract under which much of the uncertainty is shifted from the payer to the provider. Under this arrangement, the provider agrees to provide medical care for a specified population group for a specified period of time, irrespective of the actual number of cases that get treated or their cost of treatment. In this case, the provider has the incentive to minimise costs of

⁷⁹ For this reason, financial stability and quality assurance are often more crucial than prices in determining the winning bids (see e.g., McCall *et al.*, 1987).

treatment, but may also have incentives to minimise admissions since additional workloads merely add to costs without adding to revenues. There is need therefore on the side of the payer to establish quality review procedures in order to discourage excessive cost-cutting and ensure that access is not denied to patients and the access is to acceptable levels of care. This calls for regular financial reviews and medical audits.

On the other hand providers would prefer open-ended contracts, stipulating a core of services or base component where prices are related to specific costs since this protects their total revenues and rewards extra workloads. Under competitive situations, open-ended contracts encourage providers to negotiate (with payers) discounts on normal rates and also have the effect of encouraging them to diversify the range of services they provide and so costs per day or per case may be reduced. This might be good where, as in Kenya, it is desired to ensure services are available to as many 'needy' people as possible, since providers will have no incentive to serve any particular type of clients. But in the absence of strict utilisation review procedures (which may be costly), the impact on total costs may be uncertain. This is where the application of the concepts discussed in chapter three becomes necessary. Indeed, one of the controversies of the benefits of contracting out of health services (or other publicly provided social services) is how easily could the cost savings and better value for money (the two aspects much stressed by those in favour of contracting out) to be gained from the exercise be offset by reductions in the quality of service, given their short term nature (Hartley and Huby, 1985). The empirical evidence on the benefits and shortfalls of competitive contracting is reviewed in section 6.3.2 below.

Global budgeting

Another reform that has been used in health care systems successfully (mainly to contain all or portions of health care costs) is global budgets, the best example of such a success being Canada (see e.g., Evans, 1987). There appears to be a strong link between the role of government in provider payment and the use of global budgets, which is intuitively appealing since the direct link between funding sources and the providers makes global budgets relatively easy to negotiate and administer (Wolfe and Moran, 1993). Where there are many funders, setting global budgets and determining what proportion will be paid by which funding agency would be far more difficult. Where used, global budgets tend to specify information on the following characteristics (*ibid.*):

- ◆ Type of provider covered by the budget—whether physicians or hospitals, or subsets of these.
- ◆ Type of expenditure covered by the budget—whether it applies only to operating costs or to capital expenditures as well, or the latter are reviewed and approved separately.

- ◆ Type of service covered by the budget—whether all or only particular types of services are covered.
- ◆ The process by which the budget is set—whether by negotiation between the funding agency (agencies) and the providers or are set by the payer (e.g. insurance company or the government) alone.
- ◆ Action taken if budget is exceeded—in most cases end of year overruns are not covered as matter of fact in order to control costs.
- ◆ Budget financing source(s).
- ◆ Geographic boundaries to which the budget applies, if any. The level of geographic specificity is a function of the degree to which negotiation is a part of the budget-setting process. Regional (rather than national) negotiations often result in better triage between payment levels and local community health needs.

Besides the notable success of Canada (the second highest per capita spender on health care in the world after the US) in controlling health care costs through global budgets, the effects of global budgets have hardly been quantified. But to the extent that global budgets are a form of retrospective reimbursement, all problems associated with this form of reimbursement (as discussed previously can be expected to abound) whereas if they are set prospectively, they approximate the PPS system discussed above. However, in most countries that previously used global budgets, e.g., the United Kingdom and some European systems, dissatisfaction with the health systems' performance has led to evolution of different forms of regulatory frameworks because there existed no pricing mechanism through which the supply of health care resources could be matched to demands. For example, in the UK, quasi-markets' or 'internal markets'—artificial markets—have been created through rules and regulations with a view to separating the roles of the responsibility for the finance and provision of health services (Secretaries of State, 1989), and use is made of aspects of competitive contracting.

The next section looks at the empirical evidence about the performance of the system innovations reviewed above with a view to drawing conclusions about the lessons learned elsewhere that may be instructive in designing a system reform proposal for Kenya.

6.3.2 Incentives and system performance under competition: What empirical evidence reveals

Substantial empirical evidence which has emerged since the early 1980s comparing retrospective and prospective reimbursement systems—including selective contracting between payers and providers—(largely in the USA) tends to lend credence to some of the concerns raised about the PPS, particularly on how reduced costs (per day, per case) are achieved and the impact on the quality of care (for reviews of this evidence see for example, Donaldson and Gerard, 1993, chapter 8; Culyer and Posnett, 1990; Culyer, Donaldson and Gerard, 1988; Arnould, Rich and White,

1993). Here we offer only a selected review of this literature to highlight effects of competition and regulation on hospital performance and patient health outcomes.

Rosko and Broyles (1987) examine the effects on cost per admission, cost per day, length of stay, and cases treated using data from a group of 84 hospitals in New Jersey experimenting with two PPS reimbursement mechanisms relative to a comparison group of 76 hospitals in Eastern Pennsylvania that were still reimbursed on a cost-based retrospective system. In New Jersey, the retrospective payment system was initially replaced in 1974 with a PPS that used a formula to calculate a prospective patient *per diem* payment for Blue Cross and Medicare patients. This system was known as SHARE (standard hospital accounting and rate evaluation). After 1978, the *per diem* rate was replaced by charges based on DRG categories and the system extended to cover all patients. Their analysis shows the SHARE programme reduced costs per day by 9.1 per cent, cost per admission by 9.8 per cent, while the DRG system reduced costs per day by 9.8 per cent and costs per admission by 14.1 per cent, relative to the cost-based retrospective system. However, both the SHARE and DRG systems increased admissions—the former by 8.8 per cent, the latter by 11.7 per cent. Hospital average lengths of stay were reduced by 6.5 per cent under the DRG system but were unaffected under the SHARE programme. Neither payment system was monitored for the impact of the shorter length of stay or lower costs per admission and per day on the quality of care or the health status of patients.

Melnick, Zwanziger and Bradley (1989) and Zwanziger and Melnick (1993) reported similar results on cost per case after the introduction of selective contracting in competitive markets in California. However, in less competitive markets, costs per case tend to rise, albeit at a lower rate than before (Robinson and Luft, 1988). Most hospitals participating in prospective payment schemes experienced reductions in the annual rate of hospital cost inflation. But there is also some conflicting evidence. For example, Guterman and Dobson (1986) found that admissions had fallen (by 3.5 per cent) but their study did not control for possible extraneous countervailing factors (and they are aware of this) such as a simultaneous increase in hospital admissions in response to empty hospital beds. Sloan *et al.*, 1988, in a controlled before-and-after study assessing the effect of Medicare PPS on the use of medical technologies in hospitals found that the length of stay of Medicare patients in intensive care units did not change in response to the introduction of PPS but the use of routine testing declined, while the use of many non-surgical procedures only increased at a lower rate or actually decreased, relative to the non-Medicare patients control group.

Whether the changes associated with PPS, particularly the reductions in costs (as measured by various performance indicators) represent real gains in efficiency has been questioned by some economists (e.g., see Culyer, 1993; Donaldson and Gerard, *op. cit.*). Under the DRG PPS, costs may fall because hospitals exercise greater selectivity by admitting relatively lower cost patients in each category—a form of cream skimming. For example, Ginsburgh and Cater (1986) found that the case-mix (an index reflecting the proportion of high cost weighted DRG-, relative to low cost weighted DRG-patients) was 9.2 per cent higher in 1984 than in 1981, although much of this change could be attributed to changes in documentation and coding under the Medicare PPS programme. But the evidence on cream skimming under DRG PPS, commonly termed as ‘DRG creep’ is not conclusive and it is not yet known whether ‘DRG creep’ is a one-time-off effect. Other alternatives that may lead to reduced costs under the DRG—though not necessarily total costs—are the earlier of discharge patients, ‘perhaps, with less regard to the adequacy of convalescent facilities at home or other community support services’ (Culyer, 1993); and shifting of patients to modes of care not included in the PPS. Whether these practices have any effect on quality of care or patient health outcomes has also been of interest to economists. In a before- and- after analysis of the effect of Medicare PPS on patterns of hip fracture, Fitzgerald *et al.*, (1987) found that while the mean length of hospitalisation fell (by about 38 per cent) and the number of therapy sessions received declined (by about 50 per cent), the number of patients being released to nursing homes more than doubled. Moreover, of those released to nursing homes, 39 per cent were still in nursing homes six months after discharge from the hospital, compared to 13 per cent before the introduction of PPS. Fitzgerald *et al.*, 1987, have conjectured the results suggest deteriorating care, which leads to overall increased costs to society. Carrol and Erwin, 1987, in a similar study also found that patients are on average more sick on discharge (under PPS) than previously, although there is little evidence to suggest their changed overall health care package resulted in worse health outcomes. Morrisey *et al.*, 1988 collaborate the view that the introduction of Medicare PPS increased the likelihood of an earlier discharge from a hospital into sub-acute care, with a decline in discharges to patient’s own home, and reduced length of stay. Sagar *et al.*, (1989) examined the location of death of elderly people after the introduction of Medicare PPS. Their main finding was that the increase in deaths in nursing homes was greater than expected and was accompanied by a decline in the percentage of deaths in hospitals.

Therefore, if a PPS is to be implemented in Kenya, it is important that it should have in-built incentive structures and/or regulatory mechanisms to ensure that these undesirable effects are obvi-

ated. In the next chapter we outline some possible ways of alleviating these (undesirable) effects in one of the proposals suggested for reforming the resource allocation process.

Some empirical evidence assessing the effects of competitive contracting reforms has also accumulated. In one of the earlier assessments of the process and benefits (in terms of reduced costs) of contracting out of health services and other local government authorities' services in the United Kingdom, Hartley and Huby (1985) surveyed firms that were contracted *to provide the same level of service as previously supplied by in-house units* by local government authorities and district health authorities. Their evidence suggested that the yearly savings from competitive tendering averaged 26 per cent on a scale ranging from potential savings of 68 per cent to extra costs of 28 per cent. 'Since the average value of contracts in the survey just exceeded £m1, that range represented considerable sums of money', they concluded. The survey also sought opinions on 'savings expectations' from health authorities and firms that were not involved in contracting. 'Out of 59 replies, 23 expected savings of 11-20 per cent, and 27 expected savings of 10 per cent or less. Health authorities were thus consistently underestimating likely cost savings. Firms on the other hand estimated an average of 30 per cent saving on actual contracts, a figure which is somewhat higher than the average actually experienced or estimated by local authorities and the NHS' (*ibid.*). Their main conclusion was that the findings 'provide ample justification for the introduction of competition', but also cautioned that if competition is to achieve its full potential, the right conditions must be provided, in particular that:

- ◆ contracts should be awarded to the lowest bidder, on the basis of more than one specification.
- ◆ Genuine efficiency improvements require competitive tendering for different levels of service, so that the public sector buyer can obtain accurate information on the cost of a little more or a little less. This enables the buyer to determine whether the existing level and quality of service is worthwhile or whether a higher or lower level would be more beneficial.
- ◆ This should be a fixed-price contract, with penalty clauses for private firms and their equivalent for in-house units. Without fixed price contracts and budget constraints, private firms and in-house units have every incentive to 'buy into' a contract by offering optimistic estimates on price, quality and delivery.
- ◆ Poor performance must be penalised. That requires a willingness on the part of health and local authorities to enforce penalty clauses for poor quality and late delivery. In principle, in-house units which are successful in competition need to be subject to the same penalties and incentives offered to private contractors.
- ◆ There must be genuine rivalry between those seeking to be awarded contracts. That requires opportunities for new firms to enter the market. Without the possibility of new entry, there is a danger that cartels and collusive arrangements will emerge resulting in private monopolies replacing public monopolies.
- ◆ The results of competition should be publicly available. Details of the winning and rival bids should be published: rate-payers have a right to know how their money is being

spent. Moreover, publication reduces the danger of political patronage in the award of contracts, and the information also contributes to the competitive process: winners and losers need to know how they compare (*ibid.*)

Since this study, clinical services increasingly came under scrutiny for their potential to be contracted out and a model of 'managed care' in which all clinical services are subjected to this process has been published by the United Kingdom government (Secretaries of State, 1989). There, district health authorities (DHAs) have become 'purchasers' of health services and have decide whom to contract with to provide services necessary to meet all the service needs of their district populations. They no longer play any direct role in the provision of services. Public hospitals (which are responsible for higher tiers of the National Health Service (NHS)), some major acute newly independent 'trust' hospitals (i.e., those hospitals with over 250 beds which have volunteered to opt out of the direct control of the NHS to be governed by boards of directors responsible to management boards that report to the Secretary of State (for health affairs)), and private hospitals may compete for 'block funding contracts' for providing services to a specified number of cases (on a three-year rolling basis), with additional cases funded on cost-per-case basis.

Other European countries have also implemented (albeit in slightly different forms) models of managed care systems based on competitive contracting. For example, the 'Dekker' reforms in the Dutch health care system includes proposals for an 'internal-market' whereby insurers will purchase services from suppliers of health and social services on the bases of cost and quality (van de Ven, 1989). Similar proposal have been implemented in the Swedish health care system (von Otter and Saltman, 1991; Saltman, 1992). The (expected) effect of these (internal-market) systems in Europe is largely *a priori* and evidence concerning their impact has yet to emerge, but it should be expected to affect prices, output, and quality, as similar experiments elsewhere have shown.

In an oft-quoted experiment on selective contracting in California, the state financed Medicaid programme was empowered in 1982 to solicit bids from suppliers to provide hospital care to Medicaid beneficiaries. Johns (1985) estimated that in 1983-84 savings in excess of \$700 million had been realised. Paringer and McCall (1991) have also reported that under a bidding system in the state of Arizona, Medicaid equivalent costs increased by 34.2 per cent from 1983 to 1989, compared to 60.7 per cent increase for traditional Medicaid programmes. McComb and Christianson (1987) have also reported on the successful implementation of a bidding process in certain Wisconsin counties to set capitation rates for Medicaid eligibles.

Competitive contracting out aspects of health services is also emerging as a possible policy alternative to public provision and has been applied mainly for non-clinical components of health

care services in the developing countries, but the evidence is very scanty⁸⁰. In Uganda, one hospital is said to be contracting out meals for staff, elevator services and management and maintenance of the steam and boiler houses. In Zimbabwe, laundry services are contracted out, while pilot studies on contracts for non-clinical services have been carried out in Mexico. The extreme contracting model which has occurred in the developed countries has yet to be explored in developing countries, although it is reported by McPake and Banda (1994) that the British dependent territories in the Caribbean are basing their reform programmes on the British model, while In Zimbabwe and Pakistan, contracting out clinical services is already occurring. In Zimbabwe, the government is contracting out clinical services for non-mine employees to mine hospitals in the mine areas where a (government) district is not available, while in Pakistan, the contracting out of secondary and tertiary services is under consideration. Similar measures for contracting out for clinical services are being considered in Mexico and South Africa. Some of the contractors are formally contracted to provide services on a fee per patient day basis, others receive a subsidy for treating non-private patients, while some receive global (or block) budgets. But unlike in the above quoted studies, no indication of the extent of savings or other benefits, if any, have been given concerning the process of contracting-out in developing countries, without which it is hard to evaluate the potential effects of contracting in these countries. India is also said to grant subsidies to non-governmental organisation (NGO) health facilities provided they operate an exemption mechanism for patients identified as poor. In Kenya too, some NGO health facilities receive subsidies from the government although there are no formal agreements that they should serve specific groups of people, but their services are generally available to most people, particularly the poor in rural areas. This might be interpreted as an informal contract, where the NGO facilities 'act in good faith', in return for the subsidies received from the state. Clearly there is potential for contracts with non-governmental, particularly with the not-for-profit health agencies in developing countries, and this is a potential area for exploitation in Kenya.

6.3.3 Conclusions

The analysis in this section (6.3) has been concerned with reviewing the theoretical and empirical aspects of the behaviour of hospitals under competition. The theoretical analysis shows that there are evident hazards in transferring the results of the economics of perfect competition to the health care sector because of the peculiar characteristics of health care, particularly informational

⁸⁰ The evidence on aspects of contracting out aspects of health services in developing countries presented here (except the reference to Kenya) is derived from McPake and Banda, 1994, who have summarised this information from various sources.

asymmetries and the lack of market forces (due to their non-profit maximising nature) that would compel hospitals to minimise costs. Because of the peculiar characteristics of health care, competition provides incentives that often produce results that may have the opposite effect to what is usually expected—the inefficient use of resources. The main factor making competition impotent is the form of reimbursement mechanism used to finance hospital activity. Retrospective cost-based reimbursement in particular has been shown to be the single most important factor contributing to inappropriate resource use in health care. But by modifying the nature of competition, mainly through reforming of the reimbursement process, it is possible to harness competitive forces in order to promote greater hospital efficiency. Three main reforms of the reimbursement process (the prospective payment system, global budgets and competitive contracting) implemented particularly in some developed countries have been reviewed for their theoretical innovations, as well as the empirical evidence about their performance. The empirical evidence indicates that whereas these do not completely eliminate the undesirable effects of competition (and some like PPS introduce other undesirable effects), competitive contracting appears to have the least (undesirable) effects. The empirical evidence also shows these approaches have yet to receive widespread acceptance in developing countries. In contemplating the relevance of these approaches to reforming the health care sector, our conclusion is that it is possible to introduce competition in the Kenyan health care system, provided government creates an environment that exploits competitive forces wherever possible and builds into that environment features that overcome the undesired effects of competition. In the next chapter we will have much to say about the character of this environment as well as outline some necessary safeguards.

6.4 Summary and Conclusions: Implications for Health Policy Formulation in Kenya

This chapter has shown that the present command-and-control structure of the Kenyan health care system that is based on a 'pyramidal' hierarchical organisational set-up is inadequate for promoting the objectives of efficiency and equity. It was argued that due to the peculiar characteristics of health care, direct control and other similar regulatory mechanisms are ineffective for synchronising the incentives of the participants, particularly on the supply side due to informational asymmetries that can be used by the providers to undermine and sometimes to counteract the imposed regulations/controls. An alternative policy model, based on the 'agency relationship' was then proposed and it was shown that within that framework, it is possible to synchronise as well as enhance incentive compatibility through 'goal compatibility', viz., shared objectives by both the

providers and purchasers of health care. The analysis shows that introducing some form of competition between agents (providers) can lead to 'goal compatibility'. The rest of the chapter then reviewed the theoretical and empirical evidence of the potential effects of competition in general, and selective contracting in particular, on some performance measures in the health sector, in order to make an assessment of the kind of competitive environment that will provide a suitable environment for pursuing the twin objectives of increased efficiency and improved equity.

The general conclusion that may be drawn from the analysis of the chapter is that there is need to ensure there is incentive compatibility, structured so as to ensure that goal compatibility between the providers and purchasers of health care services is achieved. While this may not be important in ordinary markets, it is important in health care, mainly because of the difficulties of measuring and/or monitoring the quality of health care by individuals and other purchasers (due to informational asymmetries). For this reason, there is need for publicly expressed demand for meeting local population health care needs. The analysis in this chapter shows that one way of achieving this is via the budget setting mechanism. How this achieves efficiency and equity will be dependent on the incentives for competition and for equitable distribution of resources that are incorporated in the budget. The next chapter outlines some proposals incorporating various types of incentives for encouraging providers to cost their programmes efficiently, to train own personnel, and so on.

7. PROPOSALS FOR RESTRUCTURING THE KENYAN PUBLIC HEALTH CARE SYSTEM

7.0 Introduction

The exposition given so far of the current Kenyan health care system has exposed many weaknesses, particularly in the public services, which are based on public assistance, including:

- ◆ inadequate funding for particular services (which has adversely affected the efficiency and effectiveness of those services), as well as overall under-funding for the system (see chapter two, section 2.2.2 and chapter 3, section 3.3.4);
- ◆ inequity in the distribution of health care resources, including health personnel (chapter 2, 2.2.3, and chapter 3, 3.3.4);
- ◆ a general lack of incentives for providers and payers to operate efficiently. In the public sector, the financing process financially ‘punishes’ economy and efficiency. In the private sector there is insufficient functioning of market forces (see chapter 3, 3.3.2);
- ◆ a general lack of co-ordination between the public and private sub-sectors, which has in the past led to duplication of services. Besides impeding efficiency and desired care substitutions, this induces undesirable ‘revolving door effects’ where similar forms of care are financed differently, e.g., hospital services being financed by general tax revenues in the public sector whereas in the private sector, these services are financed either by insurance or by direct out-of-pocket payments (see chapter 3, 3.1.3 and 3.3, chapter 4, 4.4);
- ◆ financing that is strongly oriented towards (i) institutions (and even more narrowly, towards hospitals) rather than services, and (ii) inputs rather than outputs (see chapter 3, 3.4 and chapter 4, 4.3);
- ◆ ineffective control over providers, particularly those in private sector (see chapter 3, 3.2.2).

In addition, other factors such as advances in medical technology, a rapidly rising population, and the emergence of new health problems (such as AIDS) continue to make increased demands on available resources. If the system is to be able to cope with these challenges as well as others as they emerge, it has to be flexible and responsive to changes. This chapter presents four

(alternative) structural changes recommended for the public health care system to achieve that flexibility and also proffers some suggestions about the private sector.

The rest of the chapter is organised as follows. Section 7.1 summarises the criteria used to assess the performance of each proposal. In section 7.2, four proposals for restructuring the public health services system are outlined and analysed in light of the set criteria. Section 7.3 outlines measures that can be implemented to improve self-regulation in the private sector following the agency framework discussed in the previous chapter. Section 7.4 summarises and concludes the chapter.

7.1 The Criteria for Option Evaluation

The following are the simple but explicit criteria by which the proposals are to be judged and trade-offs between them made more explicit. Our point of reference is the current health care system, taking into account the pressures and problems faced by it, the level of resources that are available to obviate those problems and the realism of the suggested means of obviating them.

7.1.1 Equity

The first criterion, which we designate as *equity*, takes into account Kenya's long term objective of health for all which aims to decrease the inequalities in health status between regions of the country and between social strata, through the reduction of the incidence of diseases, the rate of disability, premature mortality, and extension of life expectancy at birth. This will be realised under any option that:

- ◆ improves the distribution as well as the quantity of facilities available to the population, so helping to reduce the distances that patients cover to access health care facilities,
- ◆ has in-built mechanisms sufficient to ensure that the poor and certain categories of at risk population groups have access to free or subsidised comprehensive care of acceptable quality, and,
- ◆ leads providers to discriminate only on the basis of 'consumption according to need'.

7.1.2 Efficiency

The second criterion derives from the recognition that pressures experienced by the economy in the recent past had adverse effects on the government's ability to maintain sufficient funding for health care services and forced the re-appraisal of the efficiency with which resources are to be/being used in the health sector. Effectiveness, quality of services and consumers' choice have also become important considerations for improved system performance. Therefore our second criteria, *efficiency*, will encompass all these aspects—any option that assures efficiently provided and medically effective services (both in a technical and economic sense, as discussed in chapter four) that are of high quality with increased choice is to be the preferred option, other things equal. Efficiency will be interpreted in the sense discussed in chapter four, i.e. securing maximum output from the available (current) resources. The effectiveness issues are about whether the policies are helping the system to achieve the intended objective—improved health status. The issue of consumer choice, which may be linked to availability, may not be central to reforms of the public services (for reasons already discussed), but if in addition to efficiency, a proposal leads to an increase in the choice set that is made available to patients (and on terms that do not violate the first criterion compared to others), that option will be preferred over others.

7.1.3 Ease/difficulty of implementation

The third criterion concerns the pragmatic realities of implementing change in the system. There are several dimensions to be taken into account here. First, the option has to be economically viable—that means affordable and therefore implementable. Second, it has to be 'politically digestible', although it need not be 'politically palatable'—that is, it takes into account the policy aspirations of government policy. Third, the changes implied by the option have to be socially acceptable. The public has to be shown to have reasons to want to use the proposed system of delivery. Such would be the case if the option shows clearly how the needs of specific groups with special health needs are catered for—e.g. pregnant mothers, disabled people, and children who may have health needs that differ from those of the average person in the health system. Fourth, the negative effects of system change should involve a minimum of disruption. Preferably, the disruptive aspects of the change should be more than compensated for by the anticipated benefits. Fifth, the change should be accepted by the health professionals—that is, it should encourage the professionals and other staff to work for its success rather than against it. Finally, the option has to be practical in the sense that there exists, or one can be devised with little cost, the managerial and technological capacity necessary to realise it. These six factors shall constitute our third set of criteria for pitting options against each other, and we shall refer to them simply as *pragmatic realism*.

7.1.4 Adequacy

Finally, an all embracing criterion—*adequacy*—is used in a straightforward sense to assess whether a particular option is sufficient for the achievement of the desired objects as described above. The option that requires the least ‘other necessary changes’ to realise those objects shall be the preferred option.⁸¹ Therefore, the option that is adopted/preferred should address the current problems of the health care system, achieve the most cost-effective balance in light of the stated criteria, and should not produce a new range of problems that are insurmountable, or that have more disadvantages than benefits.

Besides these, the proposals are also analysed in terms of their acceptability to the workforce, morale of all health workers, effects on employment, and informational requirements.

7.2 Proposals for Restructuring the Delivery of Public Health Care Services in Kenya

The implementation of efficiency and equity in health care is a demanding task whose outcomes are rarely perfect. It is for this reason we opt for an evolutionary reform where changes are implemented gradually as the health care system develops institutions and mechanisms necessary to support new structural changes. We start from the recognition that the (present) public system has incentive structures that are inappropriate or deficient for the promotion of efficiency and equity. There is need to address the question of what type of incentives (and other structural changes that) would be needed, and to whom (or where) and when they should be directed in order to realise increased effectiveness, decreasing costs, operational efficiency, allocative efficiency, and equity.

In sub-section 7.2.1 we present four proposals each of which can be adopted by the public health delivery system to realise (with varying degrees of comprehensiveness vis-à-vis) the factors listed above. The four proposals are (i). the current system with ‘modest’ internal changes, (ii). adoption of ‘global budgets’, (iii). implementation of prospective reimbursement, and finally, (iv). the separation of the purchasing and provision roles through the implementation of provider markets. These proposals have been arranged both in terms of their (i) increasing impact on efficiency, and, (ii) increasing difficulty of implementation. The last two are the only reforms likely to increase

⁸¹ In the health economics literature, the adequacy criterion is usually addressed at the macro level from the stand-point of whether the level of health care spending is adequate. This often is approached in terms of international comparisons. But as the IHSM, 1988, has argued, this is a moot point as ‘it is impossible to assess whether a given country’s level of health care is adequate because there will always be unmet need.’ (p. 19). At the micro-level, adequacy is defined in terms of whether consumers (subjectively) judge the services adequate (*ibid.*).

competition in the system. The other two are 'rearrangements' (to varying degrees) of existing system. These proposals are discussed next.

7.2.1 Option 1: The current system with internal changes

Under this first proposal, the Ministry of Health information system, following the discussion of section 6.2.4 in the previous chapter, will be used to develop local information data bases that will be useful in generating some indicators of performance to be used to gauge the performance of each facility with respect to movement towards predetermined objectives. This requires the development of performance indicators that can help increase the efficiency of resource use. For example, indicators of the average cost of providing a certain type of care such as maternity services may be developed for all hospitals with a maternity ward, and hospitals would then be required to emulate the lowest cost hospital, provided it is established that the low cost has not been achieved by trading off quality for cost. This might be difficult without information about the technical relationship between output and inputs. Such a relationship may be developed by use of 'Delphi panels' or through other methods such as regression analysis, system analysis, and so on.

The integration of clinicians in the budget development process is important because they are among the major allocators of health care resources. Their participation might involve the development of information on outcomes expected of various medical interventions, quality of care, and appropriate case-mixes. This would then allow cost-effectiveness analyses to be applied in determining the best ways of using the limited resources. The usefulness of peer reviews has also been recognised by the MOH. The application of the results of such reviews may lead to better use of resources than is the case presently.

The payment of health personnel could be altered (at present they are paid a fixed salary) to incorporate a 'performance related component' by which staff receive a combination of capitation fees, fees for some items of service (this is possible under the present arrangements where it is possible to admit private patients in amenity wards in some public facilities), and some allowances that are based on 'good practice'. This would encourage standard setting and make performance review more acceptable to health professionals (rather than view it as an hidden agenda designed to intrude into their professional autonomy).

Another internal reform that is worth attention is contracting out those in-house activities currently undertaken using ministry of health personnel to the private sector through the use of the

tender system, following the guidelines on competitive contracting given in sections 6.3.1 and 6.3.2 in the previous chapter. In that chapter we have already shown these are being used elsewhere and provided the expected benefits exceed the opportunity costs, it should be encouraged. But as this would be a new ground, this should at first only be restricted to those activities that have no bearing on health outcomes of patient care such as catering and laundry services. The move towards this direction could start on experimental basis in selected institutions, provided enough legal clauses are put in the contracts to ensure there are no loopholes that might lead to some hospital managers misusing the tendering process for individual gains. When enough experience related information has been accumulated, then contracting out can be introduced into clinical services.

Assessment in terms of the evaluation criteria

There is little in this first proposal to suggest there will be improved equity and therefore it falls short of the requirements of the first assessment criterion—equity. The financial incentives (to health personnel) may lead to some improvements only in terms of throughput, but we cannot tell *a priori* whether any disadvantaged groups would gain out of this change or not. However, the performance related pay may be set in such a manner as to encourage such groups to receive increased attention of the providers (see option two below on how such aspects may be incorporated in setting pay levels for personnel). Increased efficiency may be expected although this will not be phenomenal. But, although this proposal appears simple and straightforward, it has massive informational requirements—to develop performance indicators, assessment of alternative ways of providing care in order to determine the lowest cost option will need to take into account factors such as variations in case-mix complexities across providers. The creation of the required data base necessary for development of performance indicators will divert resources from current provision, the extent of which is hard to judge at this point. This will require personnel trained in statistical and epidemiological analysis, among others, which makes extra demands on current resources. It therefore appears the proposal makes more demands on the system than it ‘promises’ to deliver, particularly on the counts of efficiency and equity. However, given the complementary nature of it vis-à-vis other proposals, aspects of it may become ‘suitable inputs’ into other proposals, as shown below.

7.2.2 Option 2: Introduction of reimbursement using 'global budgets'

A. Introduction of reimbursement using a resource allocation formula

(a) The setting of institutional recurrent budgetary needs

Under this proposal, in addition to some of the internal changes suggested in proposal one above, resources would start to be allocated to regions—provinces or districts—on the basis of area 'needs'. There would be need to develop regional health (need) indices. Mortality and morbidity are among some of the leading contenders. A combined index of mortality and morbidity would be the ideal index but due to scantiness of data, simple proxies of need might have to be adopted—for example, the use of mortality data that more reflects the health status of each region. Geographic discrepancies in still births, infant, neonatal and post-natal mortality rates would be suitable candidates here, because the highest proportion of deaths and morbidity in the population occurs in these early childhood ages. In the absence of such refined data, childhood mortality data might be used. If overall area mortality rates were to be used, the sensitivity of the index and comparability between geographic areas might be improved by 'standardising' the mortality rates, as well as developing a composite (hybrid) index that reflects the effects of different causes of mortality (including sex composition), which are bound to reflect ecological factors that influence death rates in different regions. The result of such an exercise would be a ranking of regions according to 'need'. Once this has been done, the next stage would be to try to link the health resources distribution to regional health status. To do this, it would be necessary to relate productivity and economic efficiency in health production in different areas. The assessment of economic efficiency could be facilitated by an understanding of the technological and economic (as well as non-economic) components of 'health production' in the regions. A cross-sectional production function, relating inputs (resources) to outputs (health status—proxied by measures described above) in the different regions could then be used to generate 'resource allocation coefficients' that could be used to evaluate the extent to which differences in regional health status may be related to the distribution of resources. Such a model could provide the basis on which future budgets would be based as it gives the basis on which resources might be 'justifiably' transferred from one region to another—it implicitly provides a measure of the relative productivity of resources vis-à-vis the objectives of efficiency and equity. But it might be necessary to make some adjustments to the outcomes of the above exercise to cater for net-cross boundary patient flows, speciality costs and the costs of some teaching facilities. The general factors that the formula will take into account are as follows (most of these might be incorporated directly in the estimation process if regression analysis were used):

- ◆ factors that help account for variations in need for health care across geographical areas: these include, for example, the population size, its demographic characteristics, morbidity, and the socio-economic characteristics of the population. These factors, used in conjunction, may indicate the 'relative deprivation' (of health care services) faced by the population in a certain area. Demographic characteristics would be used to adjust (weight) regional populations in relation to need for health care. Mortality data may better reflect need for health services as opposed to data on hospital outpatient (OPD) attendance, admissions, etc., which vary according to the regional distribution (supply) of facilities. Then the regional standardised mortality rate used as a proxy of morbidity may be a comprehensive summary of relative need for health care in the area.
- ◆ any extra workload imposed on an institution due to its status (*viz.*, teaching hospital, specialised, etc.).
- ◆ cross-boundary patient flows (this would have to be done separately),
- ◆ market forces—to incorporate the relative variations in the costs of locally purchased inputs, e.g., food-stuffs, non-qualified/specialised labour, etc. This adjustment might turn out to be insignificant, in which case it may be dropped.
- ◆ other factors: e.g., population distribution, availability of private care facilities, etc. Sparsely distributed populations raise service costs, as do populations that can only be reached through special efforts, e.g., using mobile clinics to serve the areas inhabited by the nomadic peoples.

For the foregoing to perform as expected, a good information system is necessary. This provides not only the flow of information between the different levels of health care delivery but also for accounting purposes, as well as the basis for decision-making at different levels. This information system should *collect*, *analyse* and *publish*, both health and social data, by regions, economic/social classes, and so on.

The following example illustrates how a simple financing system (with an equity objective of equal per capita inputs for equal needs⁸²) which results from the consideration of the factors discussed above might be implemented. For each district, information on (i) demand/utilisation, (ii) health situation in each region, and (iii) existing health coverage would be collected. This information could then be used to determine the distribution of health care expenditures as follows (see Figure 7.1 below).

⁸² See chapter 5, section 5.3.

case be designed to favour the more disadvantaged districts. For example, the population aged less than one year might be weighted with the percentage of coverage by health centres and the percentage that has received recommended infant vaccination according to the Kenya Expanded Immunisation Programme (KEPI). If the coverage by health services in different age groups is relatively high in a certain district, that district should end up receiving less resources, the health situation notwithstanding. (Recall the principles of efficiency discussed in chapter four).

The application of these criteria to the formulation of expenditure distribution policy will vary according to the type of expenditure under consideration, i.e. the weights would vary according to whether it is preventive services, curative services or services by the private sector. The weights used will have to be compiled from information that is specific to the type of expenditure under consideration. Similarly, appropriate weights would be required for application to capital development expenditure allocations (see section (b) below), which could be then structured so as to hasten the pace at which increased services availability proceeds (when private sector services are also taken into account) in the poorly served areas, so that regional health status converge to the national average.

(b) Setting capital budgets

The capital allocation process can follow similar principles to those used in the recurrent expenditure approach. It will be necessary to set targets for capital facilities and compare these with the existing facilities in order to determine relative regional deprivation. To do this, it will be necessary first to value each region's (province or district) capital stock. Such a valuation should be as comprehensive as possible—encompassing the value of existing land, buildings and equipment—an exercise that can make use of stock schedules that already exist (which should be updated where necessary). Plant and equipment that is relatively old—appropriate accounting procedures should be used to determine the age at which plant and equipment should be considered obsolete—should be valued at its replacement cost and depreciated accordingly in order not to over-value the existing stocks where they are likely to become a bottleneck in the near future. On the other hand, plant and equipment which, though still in active use, is considered obsolete in terms the depreciation rules should be valued at its (current) replacement cost.

The second stage of the capital allocation process would be to set notional stock targets for each region by 'distributing' the value of the total capital stock pro rata to appropriately weighted regional populations. Preferably projected future populations should be used in the weighting

stage. The difference between the target (notional) stock and the existing (actual) stocks reveals the over- or under-target position and shows the degree to a region is above or below its capital requirements, given its population. Capital allocations can then be set in such a way as raise the under-target regions to their target levels.

There is one major problem to be confronted though. It should be borne in mind that most capital in place is relatively rigid. While the distribution of new capital allocations can be varied at will, existing infrastructure cannot be similarly changed. Consequently, it would be necessary to determine the minimum levels of capital which should be allocated to each region without regard to any relative shortfall in order to ensure services continuity. The relative increases/decreases in capital allocations to change the regional distribution should then be determined independent of this minimum requirement level.

(c) Some difficulties in the use of a formula to allocate resources

(i) The difficulty of incorporating efficiency in the resource allocation formula

The use of a formula to allocate resources has the primary objective of redressing inequity in the system, by eventually leading to equal opportunity of access to health care for people at equal risk⁸⁴, besides implicitly promoting self-sufficiency. But, as we have already shown in this thesis, the efficiency of resource commitments in alternative uses has to take account of the efficiency with which those resources are used. If it turns out that the neediest areas (according to the formula) are the least efficient by the criteria of the efficiency of providing services, a real possibility of conflict arises. If this inefficiency is penalised by denying resources to these (inefficient) regions, this would have grievous effect on patients as their access would be further curtailed. A possible strategy prof-fered in the UK to resolve a similar problem is to institute a system of inter-district payment for work performed by districts other than the home one (Paton, 1985). The less efficient districts contract with the efficient ones to perform for them the tasks they are less efficient at. This allows money to change hands, *yet it is still spent on services for people in the district that received it on account of its relative deprivation, and is spend on services that are efficiently provided.* The sys-

⁸⁴ Although this might not be realised in actuality. Priorities vary from one area to another and services may similarly vary. Even if we assume service variation is not due to differences in efficiency or the effectiveness of the same services in different areas, individuals in different regions might still experience differential opportunities of access to particular types of services despite the fact that they are at equal risk, owing to the orders of ranking diseases in the different areas. The formula has to provide incentives through the manipulation of the resource allocation process or directives to areas concerning services. Granted that all health problems cannot be tackled simultaneously owing to resource constraints, resources can be allocated to regions on the basis of a set of priorities based on a ranking of 'general disease groups' in a way that broadly ensures that there is equal opportunity of access across regions for those at equal risk to treatment in each category rather than for every disease.

tem thus allows reallocations to 'needy' areas to continue, despite the shortage there of facilities, adequate expertise and/or lack of enough time to provide certain services quickly. The contracting-out need not be with other public providers only, it could include private providers within or outside the district. The policy thus leads to improved flexibility between capital and revenue allocations and allows sound economic decisions to be made—e.g. as between provision at home and contracting with other providers. Thus the aim of using a formula for resource allocation (equity) is met without necessarily the eschewing of the efficiency of resource use.

But the system has certain implicit problems. The districts that contract-out for services might get locked into a vicious cycle unless some way is found allowing such districts to build the capacity to provide those services in the long run. There is also the issue of whether the inter-district payments would have to be for marginal or average cost. Besides, the problems associated with a fee-for-service mode of payment might also arise, irrespective of the basis chosen. Finally, the issue of who pays for transport costs has to be resolved. From the point of view of efficiency, the cost of transporting patients for treatment counts. But from the point of view of assessing the efficiency of resource use by various districts, it makes economic sense to have the districts whose patients are being treated outside their boundaries meet these costs (although from a social point of view it does not matter who meets these costs).

(ii) The problem of cross-boundary flows

The other major difficulty to be confronted relates to cross-boundary flows. Although as stated above, the determination of a district's allocation (of capital or revenues) treats the district as the natural unit or catchment area and therefore uses its population as the base-line for assessing resource and service needs, service planning and resource allocation must realise that ordinarily people move across district boundaries to receive treatment. In some cases such movements are so extensive they cannot simply be ignored. Even if districts were to strive to be self-sufficient, it might not be desirable to iron out such flows since on efficiency grounds, some services can only be meaningfully provided if the substantial economies of scale (and sometimes economies of scope too) are realised (as opposed to quick access). Where such economies of scale and scope exist, it means the districts that are better provided or have surplus capacity have lower marginal costs. The use of any such district's facilities by others could provide an efficient short-term respite (although the continued dependence on such capacity by the latter diminishes the incentives or speed at which capacity equalisation, or more balance provision between districts proceeds in the long-run).

With these caveats in mind, it is recommended that the resource allocation formulae for both capital and revenue should have 'in-built' mechanisms to adjust allocations to take into account these cross-boundary flows—by adjusting district populations for these flows. This adjustment cannot be taken care of under the method discussed in (i) above since these flows are not activated by those in charge of health affairs in a particular district—rather they are dictated naturally by gravity or decay models relating to particular institutions' catchment areas.

(d) Assessment in terms of the evaluation criteria

The allocation of resources using a formula largely caters for equity, as already noted. If emphasis is put on equity, the option promises large gains. Besides, in terms of some of other evaluation criteria, the option fares well. For example, in terms of pragmatic realism, the option would find favour with government as it promises to tackle the equity issue sufficiently—it enables resources to be redistributed to areas with greatest need—one of the stated government's objective. Second, it is unlikely to lead to significant service disruption (if any) as there is nothing in it to suggest it would face any opposition from health professionals. Although it would definitely lead to a reallocation of health personnel, it is unlikely to have grievous effects on the level of employment. Finally, if, as suggested above, the approach incorporates most of the internal changes contained in proposal one, there would be some improvement in the efficiency with which resources are used and therefore a cost-effective balance in service provision would be realised. *But the main problem would be need for even more information*, for besides the informational needs implied by the 'modest' changes suggested under option one itself, *other* information would be required:

- ◆ to assess the economic efficiency through use of (regional) health production functions,
- ◆ on cross-boundary flows,
- ◆ valuation of capital and its distribution—methodological issues that this valuation raises notwithstanding,
- ◆ on population demographic and socio-economic characteristics,
- ◆ the distribution of mortality and morbidity by age groups, by sex and by regions,
- ◆ facility utilisation showing utilisation in relation to population characteristics—e.g., hospital attendance and admissions by age groups, by sex and by regions, and so on.

These increased information demands diminish its feasibility since even more resources will be diverted from current service provision to administration of the fixed budgets. Also the option has immense demands for trained managerial resources. But it has to be pointed out that at some

point in time, these sacrifices have to be made if the system is ever going to have the data base suitable for assessing performance in terms of economic efficiency and social dimensions (equity).

B. An alternative methodology for estimating institutional budgetary requirements
(a) Setting recurrent budgets

As an alternative to the method of allocating revenues outlined above, the Ministry of Health could consider budgeting in a format which shows clearly the allocations for personnel, drugs and supplies, non-medical supplies, transport operating and maintenance costs, building and equipment maintenance costs, utilities and so on as separate categories of expenditure for each facility rather than the current method. Chapter four showed that it is possible to reorganise the budgeting process along these lines (see section 4.3 of that chapter). It would then be relatively easy to track trends in these categories, and following the rough guidelines outlined in chapter four, it would be able to see where deficiencies exist and take corrective action⁸⁵. The main advantage would be to ensure that items that are crucial to overall efficiency are not under-funded, whilst it will be possible to incorporate equity considerations as the following shows.

(i) Setting health personnel requirements for facilities

According to the analysis of chapter three, health personnel is the most expensive resource, taking an average of about 60 per cent of the recurrent budget in salary and related allowances. It makes sense to use such an expensive resource efficiently. It might be necessary to re-evaluate the budgeting process to take into account differences of health personnel requirements for different types of facilities in different locations—rather than adopting some arbitrarily set '(nation-wide) staffing norms'. This would ensure that the appropriate mix of doctors, nurses, pharmacists and pharmaceutical technologists, radiologists and radiology assistants, laboratory technicians, and so on, is deployed *according to regional needs*. This would especially be necessary for lower level facilities—especially dispensaries and those health centres without beds where, taking into account different disease zones, it might be necessary to staff them taking factors such as nurse time per patient, the annual number of patients, and the main causes of facility visits for each ecological zone. Probably, the time taken may vary with types of diseases.

⁸⁵ The guidelines, based on 'international experience' suggested that:

- salaries account for not less than 40% or more than 80 % of total recurrent costs,
- drugs and medical supplies share to be no less than 10% or more than 35% of the budget,
- transportation operating and maintenance expenses to be between 5% and 15%, and,
- maintenance and repair of buildings and equipment to be budgeted to for at between 1.5% and 15 %, respectively, of replacement value.

For health centres with beds, laboratories and pharmaceutical technologists, estimates of average time taken per patient also have to be taken into account. For instance, to determine the requirements for laboratory technologists it would be necessary to identify the main types of tests carried, which could vary with the type of disease, and probably region. However, since most types of analyses have standard procedures that do not vary with region—for example, urine analysis, blood tests, sputum analysis, and so on—only the number carried of such tests may vary between areas and this variation should be used to determine the required staff levels accordingly. Such a procedure could be used to establish the health and other personnel requirements for different types of facilities in different locations.

The following methodology, if used to develop staffing needs for different facilities should not only lead to better use of staff resources (i.e. efficiency), but also will apparently redress the maldistribution of personnel. The method takes into account standard workloads—takes into account utilisation, rather than the staff-population ratios inherent in phase 1 above, that are likely to ignore the patterns and levels of morbidity (unless corrected for as described).

Each facility would have to identify the most important task or tasks of each category of health worker in terms of hours per 'normal day' taken in various tasks. Particular attention will be paid to the distribution of staff time between inpatient and outpatient activities. This might be in terms of time taken processing admission, time taken in inpatient wards, time taken with outpatients, and so on. This would give the distribution of time in a 'normal' day of work, between those activities, and would have to be adjusted for time-offs, annual leave entitlements, and non-scheduled events such as 'sick-offs' and time taken by staff away from the facility on administrative duties, attending seminars, training courses, and so on. The workload standard so developed should be based on average time requirements in these functions that will maintain good professional standards for the job. Such estimates will obviously need professional inputs of the staff concerned with external checks where possible to ensure some staff are not too liberal in allocating their time use patterns.

Alternatively, the staff time taken in each alternative task might be used as the basis for calculating staff requirements, assuming no task sharing, using a similar approach. This should not be hard to accomplish as most of the data requirements can be easily determined from records or simple (say one to three months *ad hoc* exercises, although annual data which shows seasonal variations in workloads might be more informative)⁸⁶. The figure so derived would then be divided into

⁸⁶Mwabu (1989) found that utilisation of facilities in rural areas in Kenya varied according to seasons.

the estimated annual workload (e.g. estimated number of inpatient in the coming year) to calculate the number of staff in that category required to handle the estimated workload. The figures so derived can then be derived will give the required number of staff in each category. This figure, multiplied by the average pay of the health worker category (for the particular level of health facility) and summed up for all health worker categories for each facility yields the total budgetary requirements for the facility. This approach gives the 'required' staff for each facility, for each category, which can then be compared with actual staff deployed to the facility and appropriate redeployment (between facilities) can then be implemented. The ratio 'actual'/'required' would then give an indication of which facilities are under greatest work pressure—those with smallest values for the category of personnel would be the ones with the greatest shortage of staff while staff would need to be reduced for facilities where the ratio exceeds unity.

The same procedure could also be used to calculate requirements for non-health staff personnel, e.g. clerical staff, orderlies, and so on. The aim of the exercise should be structured to ensure that the quality of care being delivered by staff corresponds to the minimum standard quality required for each type of facility. This means, for example, the workloads calculated for nursing staff in health centres will presume that all health centres will operate on the same standards of nursing care, so all health centres will be required to use the same standard workloads to calculate their nursing staff requirements. If necessary, the practice in the private sector could be examined to see whether there are variations, provided quality of service is comparable. We suspect for some categories, the exercise will reveal the MOH needs to employ more staff to meet the national requirements, but the results of the exercise can be used to distribute more equitably the available staff before more can be recruited (assuming funding will be forthcoming from The Treasury to meet recruitment needs for categories of staff where acute shortage is evident). Possibilities of substitution should also be examined.

There will also be need to revise the estimates from time to time as professional standards change, as well as taking into account differences in technology in various facilities. For example, facilities with microcomputers can be expected to be more efficient in processing patient records, producing reports, and so on, compared to those using manual methods. It should however be appreciated that this method of developing staffing needs will not resolve issues such as where local demand patterns do not correctly reflect needs due to extraneous factors such as where facilities have low intakes of patients due to poor quality associated with them or inability to pay for services. This latter factor need not be due to the level of user charges, it might be due to physical costs of access to the facilities that result in low utilisation, and so on.

(ii) Setting pharmaceutical and non-medical supplies budgets

Shortage of pharmaceutical and non-medical supplies has been one factor that has adversely affected the efficiency of the public health services and there is need to rationalise their budget allocations. Kenya is a vast country with diverse disease ecological patterns. This could serve as the point of departure. Using ecological disease zone patterns, the major causes of ambulatory visits and OPD health facility visits and hospitalisation in each ecological zone (these might be refined and narrowed further if necessary so as to be more precise) would be identified. Then the *main treatment options for each disease at different levels of facilities would be identified*⁸⁷. This would require constituting panels of clinicians who devise and deliberate on the options in order to identify a 'standard procedure' or several options for treating each condition. This would take into account factors such as differential inputs necessary to treat the same condition in different locations, and it might also involve comparing various alternatives. For example, the treatment of malaria might take into account the fact that some malaria strains in Western, Nyanza and Coast provinces are resistant to routine therapies, and so on. Having decided on the 'effective' procedures, it would then be possible to calculate the actual cost of treating a single case of an episode, e.g. of malaria. It would then be possible to calculate the cost of each treatment option for each disease. Then using data on current facility utilisation, the annual estimated number of cases per disease, adjusted for population growth, disease trend patterns, sex-composition, and so on, can be derived. The process might be simplified by concentrating on the top twenty leading causes of visits and hospitalisation, but there is no reason why it should not be as comprehensive as possible. Moreover, in order not to severely constrain the hands of clinicians, in cases where the cost of treatments do not show large variations (may be on account of inefficient clinical practices), a compromise might be reached by taking a weighted average cost of treatment costs. This would require information showing the number of patients treated using any particular treatment regimen. The setting of the medical and non-medical supplies budgets along these lines will ensure that facilities do not run out of these items too early in the financial year.

⁸⁷ These options, which would be identified on the basis of their effectiveness and cost, would then constitute the *only* treatment options clinicians would be expected to use unless there are compelling reasons for other alternatives.

(iii) Transport operating, buildings and equipment maintenance budgets

These can be set on the same principles as the budgets for personnel requirements, but would take into account the nature of the institution and the volumes of patients needing transport to other facilities in case of transport operating expenses. For maintenance purposes, the Ministry can contract out some of the services such as building maintenance. The Ministry of Public Works, which is currently responsible for maintenance of public buildings can be among those tendering, but if it wins, its performance should be subjected to the scrutiny that other winners would have faced. The same applies to equipment maintenance, where the Ministry's own maintenance unit can be allowed to bid on terms that apply equally to all.

(iv) Utilities and other non-medical items budgets

For most other items, the setting of the budget would be relatively easy if the main categories above have been budgeted for. For example, in those areas with electricity and telephones, it should be possible to set realistic budgets for these items since the prices exist and any changes in them can be easily incorporated, by getting information about anticipated changes in pricing from the relevant corporations. The exercise is easier in that the providers of these services operate on the same financial accounting periods as the Ministry of Health.

(b) Capital budgets

The setting of capital budgets under this alternative might adopt principles similar to those described under sub-section A (b) above.

(c) Assessment according to the evaluation criteria

Compared to the 'formula approach', this alternative is less demanding on the information that is required for implementation—except for capital budgets, most of the information required is easily available or can be generated by use of relatively short period *ad-hoc* studies. But while the efficiency of revenue resources is likely to improve in the short run, it might take a longer time to improve the equity situation, particularly inter-regional variations in capital stocks (given area needs) since the revenue budgets would continue for a while to be set on the basis of the existing capital infrastructure.

7.2.3 Option 3: Introducing publicly financed competition in the health services through prospective reimbursement budgets

Under this proposal, providers will be reimbursed on the basis of pre-defined units of workloads. Units of workloads could be defined in two ways: (i) in terms of the cost of a particular type of patient day in hospitals, so that for example, patients in the intensive care unit command a different rate as compared to those in the maternity ward, and for maternity ward patients, finer distinctions can be made between those who have a normal delivery and those with complications or those with a caesarian-section operation, i.e. the use diagnosis related groups (DRG) type of information; or, (ii) in terms of a population- and utilisation-based index. In both cases there will be need to develop sufficient data base to enable all case mixes to be costed accurately.

(a) Use of DRG information to set prospective budgets

The DRG type of information will not strictly be used for reimbursement purposes as such, as has been the case in the US, but for general planning—in the setting of budgets that go to various providers. The purchasing agency will review the available information on disease etiology, come up with probable distributions that may be used as the basis for setting budgets. It may be necessary to deliberate with the providers on the levels of (fixed) fees for each item and providers are then paid according to that schedule for the ‘anticipated’ work.

Diagnosis-Related Grouping (ICD Codes)		Average cost per day	Relative cost per day	Estimated workloads (cases)	Budget provided to cater for workload
No.	Detailed description		a	b	(axb)
1	Infective diseases (001-136)	c_1	c_1/x	w_1	B_1
2	Nutritional metabolic (260-279)	c_2	c_2/x	w_2	B_2
3	Acute URI (460-465)	c_3	c_3/x	w_3	B_3
4	Pneumonia (480-486)	c_4	c_4/x	w_4	B_4
5	Complications of pregnancy (631-639)	c_5	c_5/x	w_5	B_5
6	Skin diseases (680-709)	c_6	c_6/x	w_6	B_6
—	—	—	—	—	—
—	—	—	—	—	—
	Total			W	B
W is the facility's total workload while B is the total budget provided for that workload for this particular facility.					

Table 7.1: Setting hospital budgets using DRG information: an illustration

The payment for each facility will be based on the national cost averages of treating different patient types, generated using morbidity-based cost estimates⁸⁸. In the case of hospitals, these national cost averages would be based on the use and cost of acute inpatient services for individual DRGs relative to the average cost per episode of hospital stay. For example, if the overall (national) average cost per patient stay in hospitals is x , and the (national) average cost per hospital stay episode of a patient in DRG group i is c_i , the ratio c_i/x will be applied to the estimated workload that a facility j anticipates to handle during the period under consideration to determine the total payment to that facility for that category of patients (see Table 7.1 above).

(b) Use of population and utilisation based standards

Prospective budgets can also be set to reflect the performance of facilities gauged by general 'population-based' and 'utilisation-based' standards. The population-based standard will be defined (at the national level) as the simple product of the rate of facility utilisation per capita (e.g. the rate of hospital admissions per capita) and the corresponding utilisation cost per capita (such as hospital cost per admission). This calculation should then be applied to 'anticipated' ambulatory or inpatient services workloads for the various facilities. The utilisation cost per capita rates will of course be adjusted by age distribution of the population as well as by facility utilisation rates by age-groups (figure 7.2)

The use of this procedure for calculating hospital activity costs can be illustrated as follows. Hospitals, depending on their status, perform roughly four major tasks:

- ◆ diagnostic and therapeutic
- ◆ administrative and supportive services
- ◆ nursing, and
- ◆ educational (research and teaching).

These services can be used as the basis for calculating actual and expected costs by breaking them into their respective sub-components as shown in the figure below (the list is for illustrative purposes only and may be incomplete), for both inpatients and outpatients.

The method would lead to a per patient cost that depends on the cost of labour, material inputs (such as medical and surgical supplies), prices, and the intensity of service use under each of the major components of hospital services. Reimbursement rates to various providers would then

⁸⁸ Where necessary adjusted in order not to incorporate existing inefficiencies in the system.

be a weighted average number, the exact value of which will be dependent on the type of diagnostic category.

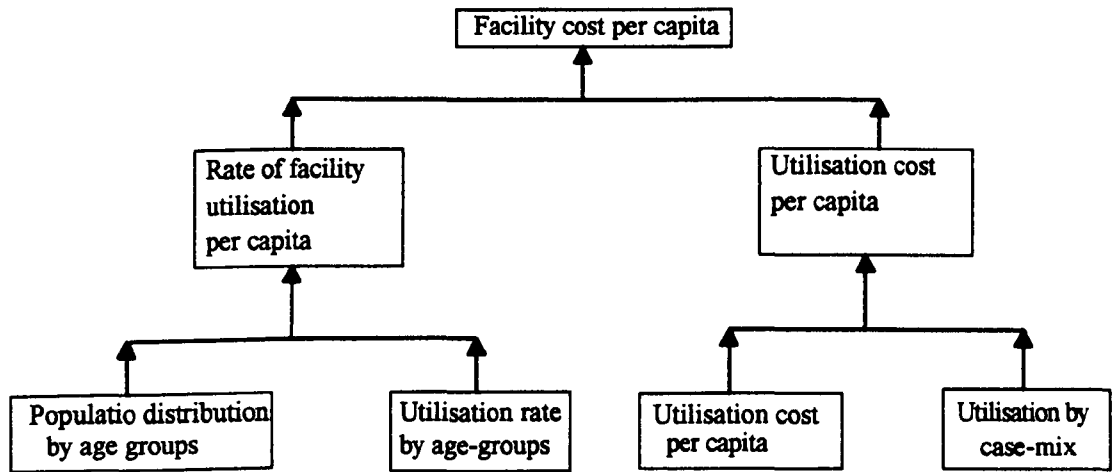


Figure 7.2: The determination of facility cost per capita

The application of this method for prospectively reimbursing providers may introduce competition if the purchaser allows both public and private providers to vie for the purchaser's limited budget. If the public hospitals are allowed to keep any surplus revenues they generate due to the adoption of cost-saving measures, they will have incentives to improve efficiency, where they substitute less expensive inputs for the expensive ones, reducing any unnecessary lengths of stay, and so on. Moreover, if they do not provide the accepted quality of care, they will be unable to attract patients.

To minimise the possibility of deterioration in quality (usually in the form 'regression to the mean') it will be necessary to audit quality in the system so that providers do not adopt cost-minimising methods which lower the quality of services. Another problem associated with prospective reimbursement is whereby providers shift the burden to other agencies or between budgets—e.g. the shift of burden from the providers to the family through early discharges, with possible consequence on full recovery, or by providers concentrating on relatively cheap to treat cases.

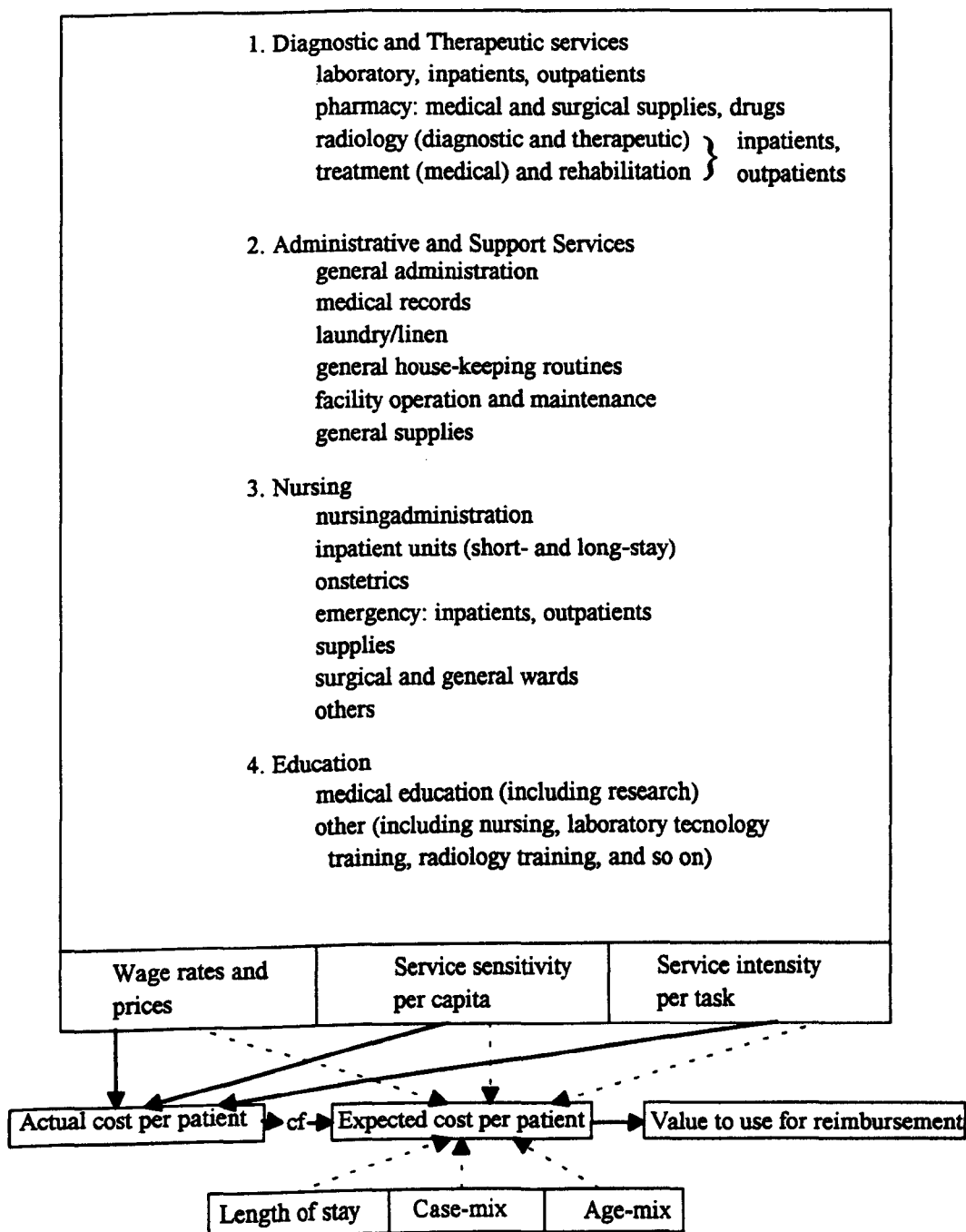


Figure 7.3: Calculation of hospital cost—components to consider

Since the method is by its nature open-ended, it will be necessary to plan for target levels of activity in which the numbers and types of admissions are fixed in advance, but enforceable clauses should also be incorporated in the contracts to ensure that patients are not turned away on flimsy excuses if an institution has over-spend its budget. This will place an element of risk on providers and therefore encourage them to be more efficient. But such a requirement raises issues of managerial control as the resource use patterns and intensities are the ambit of physicians—admissions, treatment regimens, and so on, whose interests may not necessarily coincide with those of the hos-

pital administrators, for example. If physicians are aware of the financial consequences for their institutions, they might feel obliged to adopt methods which maximise surplus income, if they stand to gain from such surpluses. This method may therefore be hard to implement as it may lead to conflicts that are hard to resolve.

(c) Assessment in light of the set criteria

In terms of the set criteria, the introduction of prospective payment by item of service has certain advantages that have far reaching implications, particularly on efficiency. Besides, the method provides a good opportunity for the government to start to divest itself of the provider role, if it so wishes. Since providers will be reimbursed on the basis of pre-defined units of workloads, it will not be necessary for the government to provide services. As the funding agency, all it needs to do is to decide on an appropriate price for a particular service (such as an operation or a course of treatment), and then reimburse the provider of that service at the *pre-set* amount of money, irrespective of whether that service is provided for less cost. This will encourage providers to search for and use cost-saving techniques. This (and other problems that may have adverse effects on equity as already noted [see chapter 5, section 5.3.1], e.g., adverse selection and cost shifting) may however raise problems, necessitating quality of care audits and other regulatory mechanisms. However if the government decides to divest itself of the 'provider role' changes in employment patterns may be expected, but this should not necessarily mean service disruptions since all that will happen is a 'movement' of labour between the 'newly constituted' providers.

The only disadvantage is that implementation of the option would require detailed costing information in order to set 'appropriate' payment levels for specific services—in the form of the development of diagnostic related groups (DRGs) or other form of resource management techniques.

7.2.4 Option 4: Introducing publicly financed competition in the health services through the separation of the financing and provider roles

Since the 1980s Kenya has implemented liberalisation policies targeted at various sectors of the economy. Recently, there has been some evidence that these policies are having some positive results including a fall in the rate of inflation, increased foreign exchange availability and a lower growth in the money supply (Kenya, 1994b). However, besides the introduction of the few 'reforms' discussed briefly in chapter one (section 1.2) and in chapter five (section 5.1.2), there has

been no attempt to extend these policies to the health sector—largely because of the concern and uncertainty of their impacts on the vulnerable groups in the population. However, with careful planning, it is possible to liberalise the health care sector and still ensure all the undesirable effects of competition (as discussed in the last chapter) are largely avoided. This fourth proposal seeks to move the health sector in this direction by recommending the separation of the purchaser and provider roles. Although the government has a principal responsibility in ensuring the *appropriateness* of quality of care (in terms of financial and geographical accessibility of care; sufficient supply of care, according to the needs of local populations; shifting of responsibility back to the consumers, providers and insurers; effectiveness and efficiency of care in health care provision), this does not mean that it has to necessarily be involved directly in the provision of health care. Under conditions set by the government and stimulated by appropriate incentives, market and self-regulation could help achieve the goals set out for health care. Through selective use of more market and self-regulation incentives, flexibility, efficiency and substitution of care can be increased by use of measures targeted at care providers, insurers, and to a limited extent, consumers, it is possible to make the parties involved in the health care sector play a prominent role in a restructured system.

Under this proposal, which is the only truly *perestroika*, 'modified' social rules and arrangements that govern the production and consumption of health care under competitive situations would be invoked. These can operate at one or more of five levels: the individual, provider, purchaser, payer or at the system level. Only those reforms that would introduce competition between the public health sector providers, purchasers and payers are outlined—in the last chapter we saw that these are the participants in the health sector more suited for targeting economic incentives at.

The purpose of introducing a competitiveness spirit in the public health sector is to make the public providers behave more like private sector providers by allowing more flexible (and dynamic) organisational forms that can adjust easily in response to changing conditions, both inside and outside the health sector. The (important) difference, however, is that the provider institutions will continue to be public-owned and operated by the state (albeit indirectly), so that they remain publicly accountable—in short, much of the red-tape bureaucratic organisational structure will go. Officers running these institutions will make decisions instinctively, without having to wait for 'authority to do something' from the centre. This will remove most of the obstacles that presently bog the system down. This means those running it would behave more like private providers—will be motivated by the drive to do well in the 'business of health care provision' in a competitive environment, but only to the extent public health objectives are not compromised. This ensures the system will not lose touch with its social basis. The system will therefore be able to *fine-tune*, at

appropriate levels of decision-making the behaviour of the main actors in the sector—particularly hospitals and professionals, the former through the creation of attractive incentives based in the environment within which they operate, the latter by providing appropriate alternatives—and also by providing flexibility (e.g. of cost variation) through bargaining at this level where information is (or could be) most easily obtained. This in turn will ensure a better matching of available resources and patients' individual needs, since a better rapport between the professional and the patient is created.

The object of this reform strategy is to change the organisational structure and institutional reimbursement procedures so that it is possible to influence the desire of the public providers to attract and satisfy patients' needs and operate effectively and efficiently. We first outline the structural changes envisaged, before explaining the process that will be used to increase competitiveness.

(a) Restructuring the system through giving prominence to primary health care

The population of Kenya is largely rural based and any reform—be it in health or other sector—must take this fact 'as given' and unlikely to change soon. Furthermore, the government can expect itself to continue to be a major financier (and possibly provider) of primary and community health services. Given this, primary health care should be the corner-stone of any proposed reform. Primary health care services in the current system are provided mainly through health centres and sometimes in the outpatient department (OPD) sections of hospitals and dispensaries. These service sources are the first points of contact with (most) patients in need of advice or treatment, also are the 'gate-valves' to specialist hospital services, and access to them should be as unimpeded as can be made possible. For these reasons, the objective of reform should be to make PHC services more responsive to the needs of the consumer, to raise the standards of care, to promote health and to prevent illness, to give patients the widest range of choice in obtaining quality primary care services, give value for money and to enable clearer priorities to be set for PHC service personnel in relation to the rest of the health sector. To realise these, the following structural reforms in this important area of the sector are recommended.

Chapter four showed there is a potential for saving if the referral system were effective (see section 4.3). To reduce the burden of unnecessary and inappropriate referrals to hospitals, and by that cut down on the expenditure on the latter, it is essential that health centres and dispensaries become the focal points of combating ill-health, also for health promotion. Services in these facilities should be improved so as to reduce the temptation of patients to go to higher level facilities. Cur-

rently, this group of facilities handles the bulk of the system's workload, both curative and primary care services, yet, they are 'residual-recipients' in the budgeting process, as already explained (see chapter three, section 3.2.1). Their budgets should be structured to reflect their importance in the system, as explained below.

(b) Other necessary changes in the organisational structure of public hospital services

(i) Changes in the status of public hospitals

In line with the proposed changes to primary health care, which redefines the role of health centres and dispensaries, the role of hospitals will also have to be redefined. Individual public hospitals will cease to be dependent administrative units, and adopt a new image—akin to that of public firms. They will no longer be funded through an automatically allocated budget. Instead, they will be expected to support themselves partially (and in the long run, entirely) with the revenues they will generate in providing services to clients (the government or the private sector— see below).

Most of the staff presently employed by hospitals are civil servants. It will thus be necessary to change the employment contracts so that hospitals can hire and fire personnel, set pay scales (probably with guidance from centre), and decide how to increase productivity and reduce other transaction costs. In this environment, if the hospital funding arrangements are changed to include a modified capitation based component, with institutional budgets and personnel salaries determined by actual workloads (and hence indirectly by demand), institutions would get budgets depending on their share of the public market. The role of the Ministry of Health under the new arrangement will be confined to regulating the operations of hospitals and to making new investment decisions regarding large capital allocations, running the primary care facilities— dispensaries and health centres, and funding of research and medical education.

(ii) Changes in the way the National Hospital Insurance Fund operates

The National Hospital Insurance Fund (NHIF) presently runs hospital reimbursement scheme for all its contributors. The contributions are in the form of compulsory 'earmarked taxes'. The state presently provides all the manpower and financial resources to administer the scheme although most contributors receive their care mainly in the private sector, as evidenced by the fact that over 70 per cent of total payments are for care received in the private sector. In its present form, the NHIF does not merit the label 'national health insurance'. The way the funds are reimbursed to contributors (who are in need) contracts the principles of a national health insurance pro-

gramme. In its present form, it just affords the better-off (essentially the wealthy) the opportunity to get good quality, specialised care that is subsidised at much lower prices than the actual cost of these services in private (mostly for-profit) hospitals. It is perplexing that one of the steps the government took in 'reorganising' the health sector in 1989 was to raise the levels of reimbursements to private (mostly for-profit hospitals)—thereby further reducing the direct costs of hospital care for those who can afford the co-payments usually associated with the use of private hospitals where the costs exceed the NHIF reimbursement rates. This move mostly benefits the rich since in actuality the costs of care by the wealthiest groups in the country are *partly* financed by the low middle income groups who cannot afford these co-payments and have to resort to the public services (like the rest of the non-inured people), yet they must contribute to the NHIF, because such contributions are mandatory. This contradicts the principles of national health insurance, besides working against the principles of equity as outlined in chapter four. The scheme should be run so as to pool resources (from each according to their ability to pay) and redistribute resources 'fairly' (i.e., to each according to need).

Taking the foregoing into account, it is proposed that appropriate legal amendments to the NHIF Act be tabled in parliament to make the NHIF an independent public institution (parastatal) that will support its operations entirely from contributions that it will solicit from 'willing' contributors in ways that other sickness insurance funds operate or be disbanded all together. In the same amendment, changes to allow *all* the funds generated from the national health insurance contributions to be directly available to the Ministry of Health (without any pro-rata reductions in the budgets allocated to it by The Treasury) to be used to help reduce the gap between the urban and rural areas, and between curative and preventive services. In what we might term as a 'second-best' option, if all the funds cannot be availed to the ministry, then the amendment should seek changes that at most allow only a (small) part of this revenue to be used as payments for medical services for the insured while the rest is given to the ministry for reallocation. In this case, the Fund should be required to institute a reimbursement scheme that is inversely related to personal incomes (at least that part earned from employment). Since the current system has all the information on income earned from employment, this change should not be hard to implement. Employees use NHIF cards whenever they or their relatives are hospitalised and hospitals use the information on the card to bill the Fund. The Fund should set the rates that an individual is to be reimbursed according to a scheme which ensures those in lower income brackets get a bigger relief when using the card, and put those rates on the cards. If such a scheme were implemented, the current situation where the lower middle income workers subsidise the upper middle and upper income groups would be

greatly improved, although it will not eliminate such subsidisation all together. It would be expected that employers would be honest with the details they send to the Fund about personal incomes.

If the first option is adopted (reallocating all the funds from NHIF contributions to the MOH) and the Fund transformed into something akin to a sickness fund, this change will not only expose the institution to some form of competition, but also bring in more money to the public health services—about 20 per cent of the ministry's recurrent budget (roughly the amount the ministry hoped to raise by instituting user-charges, but which we have seen is becoming increasingly elusive to get). Issues such as the employment status of the workforce currently working in this 'department' of the ministry will of course have to be tackled. Workers (most of who are civil servants) should be given the opportunity to decide whether to be absorbed in other departments of the ministry or even other ministries or to remain under the new institution.

Besides tackling the issue of employment, the government should also be concerned about the directions the private health insurance industry will now take without the checking influence the Fund has previously exerted on health insurance in the system. It is likely private insurers will seize this opportunity to fragment the health insurance further by developing a multitude of insurance plans that are unco-ordinated, except to the extent market forces are operative, which in itself cannot guarantee fairness in the long run, particularly for some people who might be dissatisfied with the public services but would be unable to get medical plans they can afford within the private sector. To prevent such an occurrence, the government, through appropriate legislation, can strive to encourage the establishment of health maintenance organisations that will contract with employers (and individuals as well) to provide services. In this way, the equity aspirations of the system would still be promoted, since those with a higher ability to pay (but may have less need for health services) will still be contributing to the improvement of the public health services (through the earmarked taxes), yet they would still have opportunities to purchase health insurance packages that suit their needs, if dissatisfied with what is on offer from the public services, without necessarily the eschewing of their social obligation to support the unfortunate.

There are several avenues open to implement such a strategy. One option is use the separate economic divisions existing in the labour market—along the lines presently used to determine membership/affiliation to trade unions to establish health insurance purchasing groups. Workers in a particular trade can be advised that if they so desire, they can contribute a certain proportion of their wages for health insurance coverage, probably depending on their family size and other 'rele-

vant factors', probably supplemented by the employer and the trade union. Then the relevant trade union will contract with a health maintenance organisation for providing health care to its members. Another alternative could use the workman's compensation act. Employers are required by law to insure their workers (not all employers though—only those with a workforce exceeding 20) under the workman's compensation regulations. Changes might be introduced in this law by amending it in such a way to enable it to replace the National Health Insurance Fund that exists today, by encouraging employers through carefully worked-out tax concessions on employers and employees. Another possibility is explored the use of the co-operative movement (which is relatively widespread, and again mainly organised along divisions in the labour market). The suitability of these as avenues for purchasing health insurance for workers are areas that need to be studied closely.

(c) How competition can be injected into the new system

Assuming the changes proposed above in place, the operation of the 'new' system can now be explained. The purpose of the proposed strategy is to break the link between health care financing and provision as currently exists within the Ministry of Health into separate entities—i.e. to create distinct purchaser and provider components. This will have two effects. First, it will introduce an aspect of transparency in the operations of the health sector. Second, it will largely free the labour market as some of the labour now engaged by the Ministry will be free to associate with any provider to form bidding consortia. The Ministry of Health will largely divest itself of the health production/provision aspect (except in PHC) and concentrate on the regulatory and purchasing functions.

*(i) Tackling equity and efficiency in the new system
How the overall budget would be split*

The Ministry would, using criteria such as those discussed under proposal two above, set the allocations that will be available to provinces. This means the total resources availed by the government to the public health services sector (through the MOH budget) would be allocated to the regions in a two-stage process. In the first stage, using the criteria outlined under option two above, the total MOH budget would be allocated to the provinces on the basis of their area needs. In the second stage the provincial budgets would, also using similar criteria, share their budgets between the districts within them, i.e. the districts would get shares of the provincial health budgets

according to their needs. In reality, therefore, districts become the actual budget holders. This two-stage allocation process will therefore cater for equity issues.

Districts as budget holders: How they'll allocate resources to ensure efficiency

The district total budget would be split into two distinct components: (i) a curative services component that will be allocated primarily to hospitals, and (ii) a primary health care services component that will be in most cases capitated. It would be up to each district to determine the size of each component, but as pointed out elsewhere in the thesis, since the health problems of most districts are of the primary health type, this fact should be reflected in the share allocated to PHC activities.

Allocating the curative services budget resources

The districts will in turn allocate the available curative services resources through a process that combines capitation and bidding. It is recommended that the central mechanism for both provider institution reimbursement and personnel salaries be changed to combine a predetermined 'base level' component and a 'variable' component that will be tied to various performance indicators. This 'base plus adjustments' formula will be constructed of different components for institutions and staff, and for the capital, primary care or community health service sub-sectors of the system. The objective in all instances is to mix a degree of financial certainty (the base) with a degree of financial risk (the variable component) in a manner that will stimulate operating efficiency and responsiveness to patients without endangering the broader effectiveness-linked social objectives of the overall system. It would be unwise to make all the finances contingent upon the success of bidding alone as there are areas where alternatives do not exist (even when allowance is made for private not-for-profit providers to bid), but by making it plain that any other provider can move into the area, set up activity and bid for service delivery, the government would be sounding a warning to inefficient bidders, including public sector bidders. This means even in areas where now there is only one feasible bidder from an area—may be because of economies of scale or scope—the facility presently enjoying natural monopoly would be aware of the threat of entry by new producers (including private sector competitors).

The use of flexible budgets would transform the public provider units from bureaux (as discussed previously in chapter five) to 'public firms'—'free' to design their internal structures and make judgements concerning scale advantages, employment, contracting services or buying inputs

from other sources (presumably the cheapest) in the health production chain. Rather than use bureaucratic control mechanisms, co-ordination of the production activities will now be guided by the market and providers will increasingly operate as delivery (rather than production) units. If public hospitals are allowed to act independently, some may decide to reduce their complexity and overloads by concentrating upon clinical services and purchasing any ancillary and support services (e.g., mortuary, catering, cleaning and laboratory services, etc.) from other public hospitals or independent dealers (private sector). Depending on the nature and or type of service, hospitals might go for 'hard-' or 'soft-contracting.' In the former, each hospital would operate autonomously and pursue its interests vigorously. Contracts then would have to be complete and exhaustive. In 'soft-contracting,' there will be closer identity of interests between the parties and so formal contracts need not be as complete. But here there is danger of apportioning costs and revenues. So, except for those services that scale considerations rule out hard-contracting it is recommended that hospitals will be expected to engage in prospective bids for public resources based on plans that they will provide the MOH, who, presumably will have established a 'plan evaluation and review' unit. That means there must be a way of assessing or evaluating the performance of each hospital or provider.

A 'flexible budget' achieves this because appropriate changes are made in the political and organisational patterns already in place within the system, so that fiscal linkages are forged across the service sub-sectors within the system—hospital, primary and community health services, to encourage treatment at the least intensive, but still appropriate level of care. The institution of competitive flexible budgeting does not mean there will be any larger budget over-all, or that it will vary with the number of facilities. 'All facilities' will compete for what will essentially remain a prospectively defined resource pool. It will not be a system of open-ended retrospective reimbursement as in pluralist health care systems. Competition merely translated into zero-sum competition among the existing facilities for a politically determined pool of aggregate capital and operating resources. The key to effective functioning of this system is the clarity and appropriateness with which each existing health sub-sector is divided into specific zero-sum markets. The competitive (and flexible) budgets that are finally allocated to various provider units within each service sub-sector should reflect the demographic as well as the geographic characteristics.

Allocating resources for primary care at the district level

We recommend that these facilities be given more emphasis in the financial and resource allocation process by giving them capitated budgets (which can be determined in the manner de-

scribed under proposal 2 above)⁸⁹. Since these providers (health centres and dispensaries) are more widespread than the hospitals, *giving them such capitated budgets and allowing them to purchase 'required hospital services' from the cheapest⁹⁰ sources* (where/when necessary) will serve the needs of a greater proportion of the population (and therefore the country) better.

A programme of expanding the range of services available in these lower level facilities is necessary. This programme must consider that it will be necessary to move doctors from hospitals to work in health centres (and sometimes dispensaries). The health centre might be the appropriate level to begin at. Upgrading the health centres' and dispensaries' premises and equipment will need to be undertaken along with introduction of new services such as minor surgery, diagnostic clinics, and so on. Over the long term, when sufficient expansion (of the physical facilities and personnel at this level) has taken place, people will have the facilities (and probably doctors) to choose from, and the present system whereby people's health history is not kept in one institution (if at all kept) will cease. At that stage, it will be possible to relate the doctors' salaries (incomes) to the number of people under their jurisdiction (e.g., on their lists), with appropriate incentives for doctors practising in sparsely populated areas. This system can weed out bad doctors—those inefficient, uncommunicative or simply disinterested in their patients and/or provide poor or non-existent quality of care—and in the long term will lead to the emergence of the 'family doctor' concept in Kenya, where each household will be 'registered' with a doctor who would take care of its primary health care needs/problems. Such doctors would effectively act as 'gatekeeper' to hospital services.

It will be necessary to devise a system of incentives that would encourage doctors to move to the health centres (and to the rural areas) and to remain there. The current employment practice, under which doctors and other health personnel are employed as civil servants—outside of the control of the employing institution—lacks appropriate incentives. We suggest that this be discontinued and instead a different scheme, with appropriate incentives (say, fixed-term contracts relating pay to levels of work, quality of care, and so on) be introduced. We propose these fixed contracts be structured to include in part a capitation component which will be dependent on (i) the experience of the physician, (ii) the number of patients that a particular physician manages to attract to their health centre (patient list size), and (iii) per cent of patients on the list that are seen

⁸⁹ Of course, giving these facilities 'capitated budgets' would not in itself reduce 'unnecessary referrals' unless appropriate incentives are also given to the health personnel—particularly clinicians—not to 'pass on' cases they can competently handle. Besides, 'self-referrals' will not be affected by the use of 'capitated budgets', unless a mechanism to deter people from 'self-referring' is instituted.

⁹⁰ It will be necessary to set up a procedure to ensure that the source of hospital services contracted with *is indeed the cheapest source* in order to minimise and/or eliminate possibilities of misuse of this 'privilege' of choosing a 'preferred' provider.

during the year. Other factors such degree of coverage of PHC activities such as immunisations, the level of onward referrals, and so on can also be considered. Appropriate weights for each of the components that are finally settled up on will need to be structured so as to encourage prevention and health promotion. Alternatively let market forces do the job where full privatisation occurs.

There will of course be need for co-ordinated effort between the health and other sectors (*viz.*, housing, education, environment, etc.) to ensure the gains achieved in health are not undermined by lack of progress elsewhere. There also will be need for strengthening training and continuing education for general practitioners, nurses and other health personnel. These are essential if the proposed reforms are to have the desired effect. The main advantage of this proposed structural change is that it will strengthen the referral process in a way that has not been possible before.

The process of splitting the district health budgets between curative and PHC services advocated above would change the resource allocation process so that PHC activities are no longer 'residual' recipients in the health budget allocation process. In addition, the process introduces a dimension of interface between the private and the public health services sub-sectors that has not previously existed in the system—since it would now be possible to contract private providers and pay them public resources to provide health care services, depending on the efficiency of their proposed contracts (see below).

(ii) The process of contracting

The process of contracting, which will mainly apply for curative services, can be for an initial specified period, say, one or two years, *subject to review and re-contracting* to ensure the performance of incumbent providers does not lag behind market performance. In so doing, the market will be used as regulator. In the trial period though, contracts need not be detailed (in terms of price, quality, or volume), and should only apply to public providers. The (public) hospitals can enter into agreements with the ministry's representatives (the 'purchasing commissions'—representing the ministry's interests in various districts) to provide care for a specific period (without necessarily specifying the contractual relationships), initially through use of 'block grants'. These contacts will generally define short-term relationships between the providers and the purchasers. This would leave open a possibility for purchasers to later revise the contracts, should it become necessary—on account of non-performance by public providers—to specify cost, and volume (and even place). It will leave open the option for 'purchasing commissions' to place contracts with private sector providers if dissatisfied with the performance of (public) providers, and even more im-

portant, to match health care demands with 'needs' and to purchase only effective care, although at first they won't in general yet know what this is, but the demand for such information will thereby be created.

Thus, although the current revenue flows will not be changed initially, anticipation (of change) will be created among the (public) providers, and by extension, the physicians working in them too, that they could be changed in the future. This anticipation can be expected to serve as a further incentive for physicians and other hospital personnel to operate co-operatively and efficiently in order ensure the success of their institution during the next contracting phase.

Besides the contracts with the ministry's 'purchasing commissions,' (public) hospitals will be expected to supplement their revenues by contracting with the primary health care providers, the NHIF (if it would still be there) and/or other private sector insurers, and other hospitals. Initially, public hospitals will be poor at attracting private patients, as they do not have private rooms or 'good' food—amenities that presently attract patients to the private sector providers—although even at present some private providers send their patients to the public hospitals for conditions they cannot tackle (e.g. maternity homes which send maternity cases with complications to public hospitals), it might be expected there will be possibilities for 'trade' along these lines. But ways must be found to identify patients that have been referred to the public hospitals by privately practising clinicians. This requires changes in hospital admission regulations. At least, for maternity cases, regulations may be set requiring patients to have 'booked' admissions at some prior time—this is not an unexpected 'illness' after all. Otherwise they would have to show where they expected to go for delivery (in which case negotiations between the original booked institution and the institution carrying out the delivery can be entered into) or the patient pays the full cost of the service. Ways of dealing with similar cases for other diseases have to be found.

(iii) Implementing the change—the time framework

The changes discussed above aim at strengthening primary health care, through the development of a network family physicians. Prevention should be the focus of health activities rather than treatment of disease, and the family (the smallest but probably the most important social unit) the focus of intervention. The proposed structural reform will:

- ♦ provide patients with a diversity of providers to choose from—essentially, by allowing them to choose either private (where these have contracted for provision of services to public sector patients) or public providers as their preferred provider.

- ◆ introduce incentive driven contracts for facilities and key personnel, based on capitation or other appropriate incentives and real target payments for meeting annually agreed health objectives established between the Ministry of Health, and by the districts and their hospitals and primary health care services,
- ◆ maintain the current nominal co-payments and strengthen the exemption system at the point of delivery for hospital services,
- ◆ enable to put in place a national process of quality and accreditation for health professionals and health care facilities that would operate at all levels, but would be managed locally.

This structural change can be tested by establishing on a trial basis some independent units, say one per province/district, separated from the national bureaucracy and able to operate much more flexible local working practices and incentive driven relationships with specialists and other health personnel. The success of this would then determine the pace at which other facilities are co-opted into the scheme. Moreover, even for the providers chosen as 'guinea-pigs', the transition from fixed budgets to flexible budgets should be phased gradually. We recommend the 'variable component' (of the budget) which represents 'uncertain revenue' be as small as possible at first and this be increased gradually, performance being evaluated against a pre-set target budget each year. The aim of the phase-in period is to minimise disruptions to services.

Such a transition would, whilst controlling charges from the centre, 'allow facilities the freedom within certain limits' (guidelines to stipulate these) to vary charges either way depending on local conditions. The extent to which very hard pressed and under-funded local facilities increase their contributions to community health services would be rewarded by further centrally provided financial incentives, using a formula that has an in-built inverse relationship. Such incentives would be related to established annual health objectives. For this to work, it is essential that the ministry devises a formula that takes into account the demographic and socio-economic factors into account when redistributing finances so that economically disadvantaged facilities do not continue sinking into oblivion while those in economically better-off areas continue to thrive. But care should be taken not to discourage those facilities doing well in terms of revenue collection.

The critical element is that during a transitional phase, of say, between five and ten years, different models of reform that change patients', and professionals' and policy makers' attitude can be tested and evaluated together. The change in the health care system should be transitional to build

into a system that runs on the strengths of the current service and Kenyan cultural norms, rather than great leaps into uncharted territories.

(d) Some issues to be addressed in the future

The liberalisation of the health sector, particularly along the lines suggested here, has to be approached cautiously since, although impressive in intentions, we have no benchmarks, particularly from other countries in similar stages of development, about the potential outcome of this reform process. Even in the developed countries where reforms along these lines are already underway, there is yet to emerge hard evidence about the extent to which heightened efficiency and continued preservation of the social values underlying most health care systems—such as equity and comprehensiveness of services—are compatible. Besides, the capacity of contracting within publicly operated health systems to achieve efficiencies greater than the added transaction costs has yet to be demonstrated (Saltman and von Otter, 1992). Moreover, the aspect of contracting itself is a relatively new phenomenon in Kenya, which has yet to take hold even in private sector business. Consequently, the Ministry of Health, which has no experience of writing contracts—even simple soft-contracts—has no examples to go by. But the recent strike by public hospital doctors, which forced the ministry to transfer some patients for treatment in the private sector might have served as the beginning of a learning process along these lines.

Besides, the ministry has little or no basis for evaluating quality and the performance of hospitals—because these have not been practised in the past and also because of poor information about the relationship between resource use and outcomes. The ministry should therefore undertake intensive efforts to develop methods to monitor and evaluate service quality and outcomes (both for public as well as private sector activities).

Another issue that will need to be considered is the linkage between the operating and capital budgets. The ministry should devise mechanisms to ensure that the market allocation of resources reinforces rather than undermines the broader social objectives of the publicly operated health care system by taking new investment to areas where it is needed (on equity grounds) but also ensuring that efficiency is not compromised.

7.2.5 An assessment of the proposals—a summary

An *ex-ante* assessment of the proposals presented in sub-sections 7.2.1 to 7.2.4 is necessary before the adoption of any particular option is endorsed. Section 7.1 presented some of the crucial consideration to govern the choice that is finally made. The table below presents a self-explanatory conceptual matrix of option assessment in terms of efficiency, equity, difficulty of implementation and adequacy.

Proposal → Evaluation criterion ^a ↓	Option 1 ^b	Option 2 ^c		Option 3 ^d	Option 4 ^e	
		Alternative 1	Alternative 2		Alternative 1	Alternative 2
Effect on equity	possible	high	high	questionable	high	medium
Effect on efficiency	low-medium	high	high	medium	high	high
Difficulty of implementation	low	medium	medium	medium	high	medium
Adequacy (in terms of meeting the health sector's objectives)	low	medium	medium	medium	high	medium
Other effects: — revenue generation [*] — effects on worker-morale — effect on employment ^{**} — increased service volumes — effect on quality — informational needs — other (unknown effects)	limited slight slight unchanged minimal high possible? ^{***}	none medium low high high high —	none medium low high high medium —	possible medium medium medium questionable high possible?	high high high high questionable medium-low possible?	medium high high medium questionable medium-low possible?
Notes:						
a: All options are being evaluated relative to the current system.						
b: The current system with internal changes such as the use of performance indicators, integration of clinicians in the budgeting process, changes in the way health personnel are paid (to include a 'performance related' and a 'variable' component—see main text), and contracting out some 'in-house' activities in the short-run, and clinical services in the longer-run.						
c: The introduction of reimbursement using a resource-allocation formula. Under 'Alternative 1', both revenue and capital resources are allocated taking into account population and area 'needs', and builds onto all the changes suggested under (a) above (see main text). Under 'Alternative 2', whilst building on changes suggested in (a) above, resources are allocated to facilities also taking into account population and area needs, but by major categories of expenditure—personnel, drugs and other medical supplies, non-medical supplies, transport operating and maintenance cost, buildings and equipment maintenance costs, utilities, and so on.						
d: Introduction of public financed competition through use of prospective reimbursement budgets.						
e: Separation of purchaser and provider roles through the creation of quasi-internal markets (to promote competition), and use of block budgets and bidding to allocate resources to public providers, with a possibility of incorporating private (not-for-profit) providers. 'Alternative 1' is where <i>all</i> the funds collected under the NHIF act are to be used entirely by the MOH to improve service provision in the rural areas and/or redistribution between preventive and curative services. 'Alternative 2' is where only <i>part</i> of the NHIF funds go to the MOH for redistribution as under 'Alternative 1', whilst the other part is used to reimburse the contributors who are insured under the scheme on an 'inverse-basis', given their income.						
* in terms of either revenue saving (which is assumed to be redistributed, e.g., as under 'e' above) or increased revenue availability (to the public sector) relative to the present system.						
** in terms of disruption (relative to existing practices).						
*** indicates possible existence of some as yet unknown /unanticipated effects .						

Table 7.2: Public health sector reform proposals assessment matrix

In light of the assessment criteria, all proposals will lead to an improvement over the current system. But proposals two and four come out as the most suitable, but the latter is more difficult to implement, involves a lot previously untried aspects, and may disorient the health labour market, although it is definitely the better of the two in the long run. Besides, it would require a lot of policing by the Ministry of Health in order to protect policy principles as well as to have some control over the direction in which the pattern of provision evolves. Given this, we would recommend the adoption of option two first (despite its high informational requirements) in order to develop the necessary infrastructure and the environment suitable environment for launching option four. This should be left in place until it is evident that the regional discrepancies in availability have converged towards the national average—say, for between five to ten years. If this route is taken, option 3 might be used as the intermediate ‘test-bed’ to see how the system would respond to exposure to competitive based incentives—a phase which might run for a further five years.

The separation of the purchasing and provider roles (as outlined in option 4) appears to be the major alternative to the present system in the longer term since that will address some of the current problems such as the poor linkage between funding and workload, efficiency (and variations of it within the system) and over- and under-utilisation of the capacity of facilities.

7.3 Some ‘Observations’ about the Private Sector

This thesis is mainly concerned with the reform of the public health care system, but as we have seen, the private sector also plays a not insignificant role in the health care system. It would therefore be erroneous (and unwise) to concentrate on the public sector alone as this would omit an important participant in the health system. Consequently, this section provides, although in general terms, suggestions concerning the private sector.

The agency-relationship framework sketched in section 6.2.2 in the previous chapter can also be used to show how the relationship between the state and private sector providers can be enhanced. We alluded to this aspect in section 6.3.2 of that chapter. Elsewhere (in chapter four) we showed that the private sector has characteristics and/or objectives that often conflict with those of the government in the health care sector, in particular the concern for profits, inadequate attention to public health, lack of interest in training personnel (besides causing a high attrition of highly trained personnel from the public services sector), poor quality of health care services by some of the private providers, and the poor integration of the private sector with the public services. A tri-

age of these problems can be alleviated by an improvement of the relationship between the private sector and the Ministry of Health (MOH). Generally the Ministry of Health can use the various legislation at its disposal to stipulate binding rules and regulations, and where necessary directly enforce them over the private sector, or it can enforce them indirectly through 'third parties' such as professional associations, and so on. However, given the inadequacy, and/or mostly lack of information about the operations of the private sector, the ministry can instead try to achieve similar objectives without coercion. In our considered opinion, this is a preferred approach because of three reasons. First, although the ministry might have basic regulation concerning various aspects such as training; registration; minimum standards for premises, their geographical location and quality of services to be provided; and incidence of dangerous or unethical practice, it lacks the capacity to police the private sector to ensure adherence to these regulations. The lack of funding resources—particularly the time, transport and skills—required to effectively police the private sector behaviour is a major constraint in this respect. Corruption (among the government regulators) may be another inhibitive factor—some officers supposed to enforce various aspects of health policy may instead seek illicit rents in order 'to keep a blind eye on some aspects of non-compliance'. Second, the use of third parties, such as professional organisations, may fail to realise the expected result since third party regulation enforcers may take a passive role, often due to professional self-interest—the medical profession is anyway known for its solidarity, particularly where one of their members is under criticism, and also due to the ethos of clinical freedom. Thus, although the professional organisations may have relatively easy access to information about provider behaviour, and have the professional capacity to evaluate them (as compared to the ministry), their objectives may be *too similar to allow impartial, or independent assessments, unless* it would be to their advantage to do so. Third, policing the private providers through the regulatory framework requires massive information (besides resources), not only types of information that can be easily acquired such as the number and identity of providers, their qualifications and locations, but also (internal) information about how activities are carried out in the private sector, including the number and types of patients seen, employment, the quality of care offered and so on—information that is not only difficult to get, but that is also difficult to assess (in terms of its validity).

Taking these factors into account, we suggest that it would be to the advantage of the ministry, and all other concerned parties, to design incentives structured in a way that it would be to the advantage of the private sector, including the professional organisations thereof, to conduct themselves *responsibly and in accordance* with the aspirations of the national health policy, rather than as atomistic elements motivated by private gain, within the system. More research about the sector

and the type of incentives that should be provided to achieve this is urgently needed in order to ensure that the private sector does not (continue to) undermine the gains achieved by the public sector in terms of the various health sector objectives. This research should focus on the nature incentives required to make private providers generate and provide *accurate* information about their activities, provide some of the services (mostly public health type of services) they are presently disinterested in, engage in training (rather than 'poach' expensively trained personnel from the public sector), encourage them to locate and provide services in areas with greatest need, and, generally, to ensure their priorities coincide with those of the state in their activities. This research should initially focus on those aspects of private sector operations thought to be out of tune with the aspirations of the national health policy, particularly what is needed to curb practices that cause more harm than good, use of unqualified personnel in the provision of sophisticated treatments, refusal to grant treatment to emergency and other desperate cases, curbing of unnecessary medical interventions, and generally improving quality of care provided in the sector.

7.4 Summary and Conclusions

The current health care system, particularly the public services sub-sector, suffers from in-built inefficiency because there is no direct link between funding and workload. The reimbursement system is in principle a retrospective full cost reimbursement system with an in-built bias towards a lack of concern for efficiency and effectiveness. This chapter, primarily concerned with ways of restructuring the public health care system, has proffered four alternative proposals.

The first, introducing moderate changes within the current system—such as the use of performance indicators, the integration of (principal) medical personnel in the budget setting process, changes in the way the medical personnel are remunerated and contracting out some non-medical activities such as most in-house activities (such as laundry and catering activities in the short run, and the extension of contracting-out to medical services in the long run—have been found to have only a marginal impact on the system, although they demand substantial information.

The second, the introduction of a reimbursement scheme that separates the allocation of revenue and capital resources and relies on various indicators of 'need' has been found to be one of the most promising short-run solutions to some of the current problems, particularly inefficiency and inequity. However, the option, like the previous one, has a nearly nil capacity to generate extra resources to the public health services sub-sector. It is also demanding on new information.

The third proposal discussed was the adoption of a prospective payment system. This method combines clinical budgets and DRGs in order to ensure improved economy in the use of resources. Two approaches have been proposed. The first uses inpatient case mix information to classify cases into DRGs and hospitals will be paid a uniform fixed prospective payment for each case, adjusted for relevant factors such as teaching status. The second approach suggested uses hospital activity (such as nursing, administration, educational, therapeutic and diagnosis) to set the *per-diem* rates that hospitals will receive prospectively for anticipated workloads associated with each activity type. First, each hospital is given a reference group, defined in terms of similarity—taking into account any special characteristics such as teaching status. The *per-diem* rate will then be set taking into account these special characteristics of the facility, and the costs of a particular reference group will be used to reimburse activity. This reference cost *per-diem* would therefore reflect what a particular facility in the reference group would incur if its per unit costs were equal to the average of its reference group for all cost items. It is expected that the adoption of such a procedure, combined with the removal of the existing restrictive norms on various medical and para-medical staff and the (usually arbitrarily set) ceilings on specific cost items that have in the past served as effective constraints on the efficient use of resources would give hospital managers more autonomy for more efficient use of resources. The method rewards producers for reducing costs per case. It forces hospitals and their medical staffs to co-operate to control expenditure, to relate clinical and financial information, to consider cost-benefit trade-offs and to control quality. The proposal allows the possibility of extending these prospective payments to private (not-for profit) providers who meet certain criteria. This (threat of potential entry of other providers) injects some aspect of anticipation for competition, and, it may be expected, would force physicians and institutions to voluntarily impose on themselves quality reviews and other controls they may not ordinarily succumb to if the state tried to impose them.

The fourth and last proposal advanced is the separation of the provider and purchasing roles of the Ministry of Health. Specifically, this proposal suggests that the ministry should divest itself of participation in hospital services provision and concentrate on provision of primary and preventive health care services. The ministry should give these services more attention than it does at present in budget allocations. Another suggested change affects the operations of the National Hospital Insurance Fund and the way the funds contributed to it are allocated. In particular, the proposal favours the idea of turning all the funds collected under the Fund to the ministry for reallocation to rural areas as well as to primary and preventive services, reallocations that will serve the greater needs of the population as compared to present practices, where it appears the Fund's operations

benefits more the middle upper and upper income groups than the rest of the population, an aspect that in itself is not only a travesty of the principles of equity, but is contrary to the notion of a 'national health insurance' programme.

The chapter ended with suggestions about how the problems of the private sector should be dealt with. In particular, it was suggested that the agency- relationship framework is a suitable approach to resolving problems posed by the sector, although it was pointed out that more research is needed to determine the nature of incentives that would cause the operational aspirations of the sector providers to coincide with those of the public sector.

8. CONCLUSIONS

8.0 Summary

This dissertation has been concerned with the design of a framework for reforming the Kenyan public health care system, which is currently based on a 'Stalinist-type' of model in which (supposedly) uniform services are provided via a centralised system that spans provinces, districts, and sub-districts. Available statistics indicate the system has performed relatively well in terms of broad crude indicators, given the founding health policy objectives. However, beyond these crude indicators, no one had any idea how well it was performing. But the poor economic performance and other factors (particularly the implementation of structural adjustment and economic stabilisation policies—demanded by the World Bank and the IMF—in the 1980s) that placed stringent controls on the government's participation in the economy, resulted in a gradual run-down of state subsidies to various social services, notably health and education. The public health care system was particularly adversely affected by these changes and by the end of the decade the proportion of government budget allocated to the health sector had declined from about 8 per cent just a decade earlier to 5 per cent, and by then it was obvious the health system needed reforming, as service provision failed to expand as anticipated. This stirred interest in examining how well the system was performing with respect to established objectives whilst at the same time imposing the urgency of seeking alternatives to the public funds required to sustain the public provision of health services in order to prevent the deterioration of services. On the basis of evidence from several studies, in 1989 the government introduced user-fees in all public facilities above the dispensary level and increased most inpatient charges, mostly to supplement the public revenues, but also to improve the quality of services. Contributions to the National Hospital Insurance Fund were also increased—from the previously regressive Ks. 20 for all qualifying contributors into a more progressive schedule in which contributions now vary with income.

It was argued these changes are inadequate to transform the health system into one that can serve the rising population efficiently and equitably, now and in the future. The thesis that the organisation of the public health services and the existing (more or less guaranteed) retrospective cost reimbursement institutional budgets and the present health personnel remuneration practices that go with it (which incidentally provide strong *disincentives* for higher productivity and creativ-

ity in the efficient use of resources) have combined to create a *poor* environment for efficient use of resources, and cannot lead to equity has been sustained by a review of the present system.

An evaluation of the present system leads to the conclusions that

- ◆ there is urgent need to explicitly focus on setting strategic direction and objectives that make the organisation of operating activities the responsibility of professionals;
- ◆ in the public sector, production should be decentralised and decisions delegated;
- ◆ evaluation and feedback should be a regular feature of all production;
- ◆ patients should have the opportunity to choose between alternative types and or providers of care;
- ◆ the responsibility for finance and production in the public sector should be separated, with more frequent use of external purchases where these promise better service than public provision;
- ◆ there is need to change the engagement contracts of the public health sector employees if any incentive related gains are to be extorted out of them.

Beyond these (implicit) structural shifts at the system level, there is need for substantial changes in the decision making processes inside individual provider organisations. The major impetus to effectiveness should come from improved social dynamics (incentives and communication). Command and control relationships should be replaced by interactive incentive- driven relationships. Issuing instructions (that are expected to be implemented) should be replaced by dialogue—by a market style exchange between producers and consumers by participatory discussion among employees inside organisations, and an emphatic and respectful discourse with patients. In short, there is need to change the existing structural arrangements within the public health service. Starting from this premise, the dissertation proceeded to design a basic institutional framework/system reform that can help change the current situation.

This framework focused on efficiency and equity. On efficiency, it was noted that although it is ethically desirable that every Kenyan should achieve the best level of health care the available resources can provide, resource scarcity imposes the need for choices between alternatives to which resources may be committed. This means that whatever resources are allocated for the purpose of promoting health have to be used in such a fashion as to *maximise their impact on the nation's health*. That is prescribed as the objective the system should strive to achieve—the maximisation of health gains from the resources allocated for the purposes of promoting health. The framework has four elements which are cumulatively comprehensive—the system should:

- ① only provide medically effective services—by which is meant providing only those services for which there is credible evidence about their ability to alter for the better the patients' course of illness—a notion subsuming both technical efficiency and medical effectiveness;
- ② provide services that are cost-effective—which requires that whatever services are being provided should be provided at least cost. This requirement provides the criterion for discriminating between alternative technically efficient methods (which incidentally, following ①, have also to be medically effective).
- ③ concentrate on health services that offer the highest pay-offs in terms of health gains. Since medical interventions rarely have equal effectiveness, the concept of marginal cost effectiveness should be the relevant criterion for choosing between alternative medically effective interventions—a concept which subsumes all the above.
- ④ provide such services as are implied by ①—③ above at appropriate scales. This stage guarantees overall efficiency since it is at this stage where it is decided which actions are worth allocating resources to in health care rather than to others outside the health sector.

The dissertation showed there is potential for efficiency improvement in the Kenyan health care sector. It also demonstrated that in general the allocation of resources is not consonant to the existing health problems.

On the issue of equity, it was shown that the current policy (on equity) fails to address the 'correct' issues and will therefore not lead to any significant improvements in the equity situation *unless* changes in the orientation of the health policy are implemented. It was shown that although health care—according to pronouncements in various government policy documents on health and health care—is viewed as an individual's right, an analysis of *actual* policies hardly underscores this. Four implicit interpretations of equity in the health sector are possible, given the current policy:

- ① equal access to public health care services for equal need;
- ② equal access to health promoting commodities;
- ③ guaranteed minimum access for all to public services; and,
- ④ opportunity of access to private health care determined by personal resources.

Taken together, these interpretations suggest *different notions of access to health services, and therefore different policies for ensuring equity in health care services*, some of which contradict the notion of a 'right' to health care. Following a review of various philosophical bases of equity it was concluded that the 'decent minimum' is the best interpretation of equity that the Kenyan health policy should adopt. This interpretation was given further substance by reviewing the meaning of equity in health care (including alternative interpretations of such concepts as 'need', 'want' utilisation and so on,) and their relation to efficiency and it was concluded that the decent minimum should be interpreted in terms of 'utilisation' rather than 'access'.

A framework to achieve the above was then proposed. In the first part a generalised theoretic account discussed some reasons for inefficiency within an industry, and some factors that—even in competitive situations—may cause the economically efficient (competitive) equilibrium not to be realised. Focusing on the hospital sector for illustrative purposes, two general models of hospital behaviour were examined and found inadequate for explaining the Kenyan (public) hospital sector. An alternative model of Kenyan hospital behaviour was outlined which shows that the present environment cannot motivate efficient behaviour in the operations of providers units. The economic incentives to operate efficiently (or more correctly, inefficiently) within the present system were then discussed at length. It was argued that due to the peculiar characteristics of health care, direct control and other similar regulatory mechanisms are inadequate for synchronising the incentives of the participants in environments such as exists in the Kenyan hospital sector—particularly on the supply side—due to informational asymmetries that can be used by the providers to undermine and sometimes to counteract the imposed regulations/controls. The general conclusion was that there is need to provide incentives structured so as to ensure that goal compatibility between the providers and purchasers of health care services is realised.

An alternative policy model based on the 'agency relationship' was then proposed and it was shown that within that framework, it is possible to synchronise as well as enhance incentive compatibility through 'goal compatibility', *viz.*, shared objectives by both the providers and purchasers of health care. The analysis showed that introducing some form of competition between agents (providers) can lead to 'goal compatibility'. The theoretical and empirical evidence of the potential effects of competition in general, and selective contracting in particular, on some performance measures in the health sector was then reviewed, in order to assess the kind of competitive environment that will provide a suitable environment for pursuing the twin objectives of increased efficiency and improved equity.

Following this, four alternative proposals have been advanced. The first—introducing moderate changes within the current system, such as the use of performance indicators, the integration of (principal) medical in the budget setting process, changes in the way the medical personnel are remunerated and contracting out some non-medical activities such as most in-house activities—(such as laundry and catering activities in the short run, and the extension of contracting-out to medical services in the long run)—has been found to have only a marginal impact on efficiency and equity in the system, though it has enormous informational demands for implementation. Moreover, its implementation involves substantial diversion of resources from current provision.

The second—the introduction of a reimbursement scheme that separates the allocation of revenue and capital resources and relies on various indicators of ‘need’—has been found to be one of the most promising short-run solutions to some of the current problems, particularly inefficiency and inequity. However, the option also has an immense informational requirement, and has a nearly nil capacity to generate extra resources to the public health services sub-sector.

The third option proposed was the adoption of a prospective payment system. This method combines clinical budgets and DRGs in order to ensure improved economy in the use of resources. Two approaches have been proposed. The first uses inpatient case mix information to classify cases into DRGs and hospitals will be paid a uniform fixed prospective payment for each case, adjusted for relevant factors such as teaching status. The second approach suggested uses hospital activity (such as nursing, administration, educational, therapeutic and diagnosis) to set the *per-diem* rates that hospitals will receive prospectively for anticipated workloads associated with each activity type. First, each hospital is given a reference group, defined in terms of similarity—taking into account any special characteristics such as teaching status. The *per-diem* rate will then be set taking into account these special characteristics of the facility, and the costs of a particular reference group will be used to reimburse activity. This reference cost *per-diem* would reflect what a particular facility in the reference group would incur if its per unit costs were equal to the average of its reference group for all cost items. It is expected that the adoption of such a procedure, combined with the removal of the existing restrictive norms on various medical and para-medical staff and the (usually arbitrarily set) ceilings on specific cost items that have in the past served as effective constraints on the efficient use of resources, would give hospital managers more autonomy for more efficient use of resources. The method rewards producers for reducing costs per case. It forces hospitals and their medical staffs to co-operate to control expenditure, to relate clinical and financial information, to consider cost-benefit trade-offs and to control quality. The proposal allows the possibility of extending these prospective payments to private (not-for profit) providers who meet

certain criteria. This (threat of potential entry of other providers) injects some aspect of competition, and, it may be expected, would force physicians and institutions to voluntarily impose on themselves quality reviews and other controls they may not ordinarily succumb to if the state tried to impose them. This option also has immense informational requirements, and if implemented may require continuous policing to ensure equity objectives are not trampled upon in pursuit of efficiency.

The fourth (and last) proposal advanced is the separation of the provider and purchasing roles of the Ministry of Health. Specifically, this proposal suggests that the ministry should divest itself of participation in hospital services provision and concentrate on provision of primary and preventive health care services. The ministry should give these services more attention than it does at present in budget allocations. Another suggested change affects the operations of the National Hospital Insurance Fund and the way the funds contributed to it are allocated. In particular, the proposal favours the idea of turning all the funds collected under the Fund to the ministry for reallocation to rural areas as well as to primary and preventive services, reallocations that will serve the needs of greater proportion of the population as compared to present practices, where it appears the Fund's operations benefits more the middle upper and upper income groups than the rest of the population, an aspect that in itself is not only a travesty of the principles of equity, but is contrary to the notion of a 'national health insurance' programme.

Some suggestions about how the problems posed by the existence of the private sector should be dealt with are also given. In particular, it was suggested that the agency-relationship framework is a suitable approach to resolving problems posed by the sector, although it was pointed out that more research is needed to determine the nature of incentives that would cause the operational aspirations of the sector providers to coincide with those of the public sector.

8.1 Conclusions/Recommendation

Following the assessment of all the proposals, it has been found that in the long-run the separation of purchaser/provider roles is the most promising proposal for restructuring the system. But since the immediate implementation of that option would introduce 'shocks' of a type previously unknown to the system, we recommend the adoption of option two at first (despite its high informational requirements) in order to develop the necessary infrastructure and the environment suitable environment for launching option four. Although it has been shown this option will involve

diversion of resources from current provision, it is recommended due to its capacity to redress the inequalities in the system. This is considered a necessary precondition for the introduction of competitiveness in the health system. Once implemented, it should be left in place until it is evident that the regional disparities in availability have converged towards the national average—say, for between five to ten years. If this route is taken, option 3 might be used as the intermediate ‘test-bed’ to see how the system would respond to exposure to competitive based incentives—a phase which might run for a further five years. The separation of the purchasing and provider roles (as outlined in option 4) will then be relatively easy to implement as the problems of the linkage between funding and workloads, efficiency (and variations of it within the system) and over- and under-utilisation of the capacity of facilities would have been smoothed. This proposal would separate organisationally the funding and provision of health care (as explained in the main document). The provision of health care would largely be privatised, but not the demand for it. The advantage of this separation is that competition and efficiency gains would be created combined with the intrinsic equity of a ‘central allocation’ system. Given the difficulty of correcting the capital infrastructure discrepancies, this time-framework is thought not overtly long, but it will need the commitment of successive governments if in the long run it would be possible to implement the purchaser/provider roles.

8.2 Weaknesses and Strengths of the Thesis: An Overall Assessment

The principal focus of this dissertation was stated as the design of a basic institutional framework/system reform for the delivery and financing of personal services as well as preventive medicine in Kenya encompassing hospital and clinic based services, training institutions for doctors, nurses and other paramedical professionals, the role of research in supporting health services, and the role of various public agencies concerned with preventive programs. A reflection on what has been achieved or not achieved indicates that two of these major objectives (i.e. training and the role of various public agencies concerned with preventive programmes) have not received as detailed an analysis as their importance warrants. The only excuse for not doing so is that these objectives span several ministries—for example, training is partly funded by the Ministry of Education and partly by the Ministry of Health. Substantial reforms have already been implemented by the Ministry of Education concerning various aspects of education including cost-sharing. The Ministry of health should co-ordinate with that Ministry to streamline its training programmes, particularly those that largely fall under its jurisdiction—such as the training of nurses. However, there are

some training implications arising from the proposals that are directly the responsibility of the Ministry of Health. This particularly concerns the need for training managers and clinicians and other personnel required to develop, interpret, and use the various performance indicators and other information on medical- and cost-effectiveness of services.

Similarly, most programmes with an element of preventive and promotive care fall under different ministries, especially the Ministries of Health, Education, Water, Labour and Man Power Development, Housing, and Environment and Natural Resources. Any sensible analysis of reforms should take the role played by each of these various participants into account—an aspect beyond the scope of the present work.

The other notable deficiency in the present work has been lack of quantification of most aspects of the suggested reforms. For example, under the resource allocation by formula proposal, it would have been more informative if it had been demonstrated what differences it would make to switch to such a system (compared to the present system). Lack of suitable data to quantify many aspects suggested for use was the main reason for this omission—for example, while it is relatively easy to get data on the distribution of population by region (provinces, districts, and so on), there is hardly any data on aspects such as utilisation by age group, sex groups, and so on, that can be meaningfully used to calculate appropriate weighting factors. Consequently, most such issues are relegated to the next section, as areas that require further research.

Other than the above, I have painstakingly developed a framework for structural health reform of the Kenyan health care system that provides the ingredients required to transform the public health care system into one that is both efficient and equitable. Even if one were to ignore the recommendations for reform outlined here, a search for any other alternatives must build on the discussion of efficiency and equity as provided here, if it is to be genuinely problem-solving oriented. I want to believe that (and indeed I do!) that—besides the original account of the system and the problems inherent in it outlined in this dissertation—this is the most significant contribution of this work to the Kenyan health care system.

8.3 Research Implications of the Suggested Structural Reform

A number of areas where more research is needed have been identified at various points in this dissertation. These are briefly summarised below.

- ◆ In-depth study of the factors influencing the demand for health care, to particularly identify the factors that have largely contributed to rising demand.
- ◆ Productivity of resources in various uses—their impact on health.
- ◆ Cost-effectiveness of alternative treatment regimens.
- ◆ Assessment of the relative effects of overall level of pay, grading structures and industrial relations under ‘facility level’ versus ‘national’ bargaining—effects of different forms of rewards to the public health care staff to identify those that induce the staff to respond to local health needs.
- ◆ The level of detail, in terms of price, quantity, quality and nature of services needed when specifying contracts.
- ◆ The productivity of different types of health personnel in relation to outputs/outcomes.
- ◆ The anticipated distributive effects of the structural change.
- ◆ The role of the private sector—how its contribution to the objectives of the health care system may be improved.

Of course more research issues will become evident as the implementation of the reform unveils other problem areas.

— REFERENCES —

- Abel-Smith, B., 1967, *An International Study of Health Expenditure*. Public Health Papers No. 2. Geneva: World Health Organisation.
- Akerloff, G., 1970, *The market for lemons: Qualitative uncertainty and the market mechanism*. *Quarterly Journal of Economics*, Vol. 84, pp. 488—500.
- Akin, J S., N. Birdsall and D. de Ferranti, 1987, *Financing Health Services in Developing Countries: An Agenda for Reform—A World Bank policy study*, The World Bank, Washington, DC.
- Alchian, A. A., and H. Demsetz, 1972, *Production, information costs and economic organisation*. *American Economic Review*, Vol. 61, 1972.
- Arnould, A. R., R. F. Rich and W. D. White, (eds.), 1993, *Competitive Approaches to Health Care Reform*. Washington DC: The Urban Institute Press.
- Arnould, A. R., *et al.*, 1993a, *Competitive reforms: Context and scope*, in Arnould, Rich and White, 1993, chapter one.
- Arnould, A. R., *et al.*, 1993b, *Competitive reform and the future*, in Arnould, Rich and White, 1993, chapter five.
- Arnould, A. R. and L. M. De Brock, 1986, *Competition and market failure in the hospital industry: A review of the evidence*. *Medical Care Review*, Vol. 43, No. 2, pp. 253—292.
- Arrow, K., 1963, *Uncertainty and the welfare economics of medical care*. *American Economic Review*, Vol. 53, pp. 941—973.
- Atkinson, A. B., and J. E. Stiglitz, 1980, *Lectures on Public Economics*. New York: McGraw-Hill.
- Auster, R., I. Leveson and D. Sarachek, 1969, *The production of health: An exploratory study*, *Journal of Human Resources*, IV, Fall, pp. 411—436.
- Baldwin, S., C. Godfrey and C. Propper (eds.), 1990, *The Quality of Life: Perspectives and Policies*, London: Routledge and Kegan Paul.
- Barr, N. and D. Whytes (eds.), 1993, *Current Issues in the Economics of Welfare*. Current Issues in Economics Series. London: Macmillan.
- Baumol, W. J., 1986, *Super-fairness: Applications and Theory*, Cambridge, Massachusetts: The MIT Press.
- Beck, A., 1974, *History of medicine and health services in Kenya (1900—1950)*, in Odingo and Onyango, 1974, pp. 91—106.
- Becker, G., 1965, *A theory of the allocation of time*, *Economic Journal*, Vol. 75, 493—517.
- Bell, J. M., and S. Mendus, 1988, *Philosophy and Medical Care*, Cambridge: Cambridge University Press.

- Beske, F., 1982, *Expenditures and attempts of cost containment in the statutory health insurance systems of the Federal Republic of Germany*. In McLachlan and Maynard, (eds.), 1982, pp.233—264.
- Black, D., 1980, *Inequalities in Health: Report of the working group chaired by Sir Douglas Black*, London: HMSO.
- Bonte, J., 1974, *Patterns of mortality and morbidity*, in Odingo and Onyango, (eds.), 1974.
- Broome, J., 1978, *Trying to value a life*, *The Journal of Public Economics*, Vol. 9, pp. 91—100.
- Broome, J., 1988, *Goodness, fairness and QALYs*. In Bell and Mendus, 1988, *op. cit.*, pp. 57—73.
- Carroll N. V., and W. F. Erwin, 1987, *Patient shifting as a response to medicare prospective payment*. *Medical Care*, Vol. 25, pp. 1161—1167.
- Clarkson, K. W., 1972, *Some implications of property rights in hospital management*. *Journal of Law and Economics*, Vol. 15, pp. 363—384.
- Cochrane, A. L., 1972, *Effectiveness and Efficiency: Random Reflections on Health Services*. London: The Nuffield Provincial Hospital Trusts.
- Coles, M., and J. M. Malcomson, 1979, *Contract theory and incentive compatibility*, in Hey, 1979.
- Cowell, F. A., 1986, *Microeconomic Principles*. Philip Allan Publishers.
- Cromwell, J., and J. B. Mitchell, 1986, *Physician-induced demand for surgery*. *Journal of Health Economics*, Vol. 5, pp.293—313.
- Cullis, J. G., and P. A West, 1979, *The Economics of Health: An Introduction*. London: Martin Robertson.
- Culyer, A. J., 1971, *The nature of the commodity health care and its efficient allocation*. *Oxford Economic Papers*, Vol. 23, 189—211.
- Culyer, A. J., 1976, *Need and The National Health Service: Economics and social choice*. London: Martin Robertson.
- Culyer, A. J., 1978, *Need, Values and Health Status Measurement*. in *Economic aspects of health services*, London: Martin Robertson.
- Culyer, A. J., 1980, *The Political Economy of Social Policy*, Oxford: Martin Robertson.
- Culyer, A. J., 1985, *Economics*. Oxford: Basil Blackwell.
- Culyer, A. J., 1987, *Cost containment in Europe*. In OECD, 1987, *Health Care Systems in Transition*. Paris: Social Policy Studies No. 7.
- Culyer, A. J., 1988a, *Inequality in health services is, in general, desirable*, in D. G. Green (ed.), 1988, *Acceptable Inequalities?: Essays in pursuit of equality in health care*. London: Institute of Economic Affairs, Health Unit Paper No 3.

- Culyer, A. J., 1988b, *Health Care Expenditures in Canada: Myth and Reality; Past and Future*. Canadian Tax Foundation.
- Culyer, A. J., 1989a, *The normative economics of health care financing and provision*. **Oxford Review of Economic Policy**, Vol. 5, pp. 34—58, also reprinted in McGuire, A., P. Fenn and K. Mayhew, 1991, *Providing Health Care: The Economics of Alternative Systems of Finance and Health Care Delivery*, Oxford: Oxford University Press, pp. 65—98.
- Culyer, A. J., 1989b, *Health, Health Expenditures and Equity*, University of York, Centre for Health Economics, Discussion Paper No. 83, April.
- Culyer, A. J., 1990, *Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life*, in Baldwin, Godfrey and Propper, (eds.), 1990.
- Culyer, A. J., 1991a, *Equity in Health Care Policy*, University of Toronto/University of York, (*mimeo*).
- Culyer, A. J., 1991b, *The promise of a reformed NHS: An economist's angle*. **British Medical Journal**, Vol. 302, pp. 1253—1256.
- Culyer, A. J., 1992, *The morality of efficiency in health care—Some uncomfortable implications*. **Journal of Health Economics**, Vol. 1, no. 7, pp. 7—18.
- Culyer, A. J., 1993, *Health care insurance and provision*, in Barr and Whynes, (1993, chapter 8.
- Culyer, A. J., A. Maynard and A. Williams, 1981, *Alternative systems of health care provision: An essay on motives and beams*, in Olson, M., (ed.), 1981, *A New Approach to the Economics of Health Care*. Washington and London: American Institute of Public Policy Research, pp. 131—150.
- Culyer, A. J., A. Williams and R. Lavers, 1972, *Health indicators*. In A. Schofields and S. Shaw, (eds.), *Social Indicators and Social Policy*. London: Heinemann, pp. 94—118.
- Culyer, A. J., and A. Wagstaff, 1991, *Need, Equality and Social Justice*, University of York, Centre for Health Economics, Discussion Paper No 90, October.
- Culyer, A. J., and A. Wagstaff, 1993, *Equity and equality in health and health care*. **Journal of Health Economics**, Vol. 12, pp. 431—457.
- Culyer, A. J., and J. Posnett, 1990, *Hospital behaviour and competition*, in Culyer, Maynard and Posnett, 1990.
- Culyer, A. J., C. Donaldson and K. Gerard, 1988, *Financial Aspects of Health Services: Drawing on Experience*. The Institute of Health Services Management: Working Party on Alternative Delivery and Funding of Health Services, Working paper No. 3.
- Culyer, A. J., A. Maynard and J. W. Posnett, (eds.), 1990, *Competition in Health Care: Reforming the National Health Services*. Macmillan Press.
- Curry, S., and J. Weiss, 1993, *Project Analysis in Developing Countries*. London: Macmillan.
- Dahlgren, G., 1991, *Strategies for health financing in Kenya—the difficult birth of a new policy*. **Scandinavian Journal of Social Medicine**, Vol. 546, Supplement No. 46, pp. 67—81.

- Daniels, N., 1985, **Just Health Care**, Studies in Philosophy and Health Policy. Cambridge: Cambridge University Press.
- Dasgupta, A. K., and D. W. Pearce, 1972, **Cost-Benefit Analysis: Theory and Practice**. Basingstoke: Macmillan.
- de Ferranti, D., 1983b, **Some Current Methodological Issues in Health Sector and Project analysis**. The World Bank: Population, Health and Nutrition Department, December.
- de Ferranti, D., 1983a, **Background Information for Analysis of Financing and Resource Allocation Issues in Health Sector and Project Work**. The World Bank: Population, Health and Nutrition Department, December.
- De Lew, N., G. Greenberg and K Kinchen, 1992, *A layman's guide to the US health care system*. **Health Care Financing Review**, Vol. 14, (1), pp. 151—169.
- DHSS, (Department of Health and Social Security), 1976, **Report of the Resource Allocation Working Party**, London: HMSO (Her Majesty's Statistical Office).
- DHSS, 1986, **Review of the Resource Working Party Formula**. Report by the NHS Management Board. London: HMSO.
- Donabedian, A., 1971, *Social Responsibility for personal health services: An examination of basic values*. **Inquiry**, Vol. 8, pp. 3—19.
- Donaldson, C., and K. Gerard, 1993, **Economics of Health Care Financing: The Visible Hand**. Economic Issues in Health Care Series. London: Macmillan, .
- Dunham, A. B., J. A. Morone and W. D. White, 1982, *Restoring medical markets: Implications for the poor*, **Journal of Health Politics and Law**, Vol. 7, pp. 488—501.
- Eatwell, J., M. Milgate and P. Newman (eds.), 1987, **The New Palgrave: A Dictionary of Economics**, London and Basingstoke: Macmillan.
- Ellis, R., J. M. Kirigia and G. Mwabu, 1990, *South Nyanza district study: An empirical exposition of health care users*. Nairobi: USAID, (Mimeo).
- Enthoven, A. C., 1980, **Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care**. Reading, Massachusetts: Addison Wesley.
- Enthoven, A. C., 1987, *The US health care economy: From guild to market in ten years*. **Health Policy**, Vol. 7, pp. 241—251.
- Evans, R. G., 1974a, *Supplier induced demand: Empirical evidence and implications*. In M. Perlman, (ed.), **The Economics of Health and Medical Care**. New York: John Wiley and Sons.
- Evans, R. G., 1974b, *Modelling the economic objectives of the physician*. In R. D. Fraser, (ed.), **Health Economics Symposium: Proceedings of the first Canadian conference**, September, 4—6.
- Evans, R. G., 1983, *The welfare economics of public health insurance: Theory and Canadian practice*. In L. Söderström, (ed.), **Social Economics**. Amsterdam: North-Holland.

- Evans, R. G., 1984, *Strained Mercy: The economics of Canadian medical care*. Toronto: Butterworths.
- Evans, R. G., 1987, Public Health Insurance: *The collective purchase of individual care*, *Health Policy*, Vol. 7, pp. 115—134.
- Evans, R. G., 1990, *Tension, compression and shear: Directions, stresses and outcomes of health care cost control*. *Journal of Health Politics, Policy and Law*, 15: 101—128.
- Evans, R. G., J. Lomas, M. L. Barer, R. Labelle, C. Fooks, et al., 1989, *Controlling health expenditures—the Canadian reality*. *New England Journal of Medicine*, 320, pp. 571—577.
- Fama, E. F., 1980, *Agency problems and the theory of the firm*. *Journal of Political Economy*, Vol. 88, No. 2.
- Feldman, A. M., 1987, "Equity" in Eatwell, Milgate and Newman (eds.), 1987.
- Feldstein, M., 1963, *Economic analysis as operational research and the National Health Services*, *Oxford Economic Papers*. March.
- Feldstein, M., 1965, *The effects of case-mix on hospital costs*. *Medical Care*, Vol. 3, pp. 95—103.
- Feldstein, M., 1970, *The rising price of physician services*. *Review of Economics and Statistics*, Vol. 52, May, 1970.
- Feldstein, P. J., 1966, *Research on the demand for health care services*. *Milbank Memorial Fund Quarterly*, Vol. 44, No. 3, Part 2, pp. 128—165.
- Fitzgerald, J. F., L. F. Fagan, W. M. Tierney and R. S. Dittus, 1987, *Changing patterns of hip fracture care before and after implementation of the prospective payment system*. *Journal of the American Medical Association*, Vol. 258, pp. 218—221.
- Fried, C., 1978, *Right and Wrong*, Cambridge, Massachusetts: Harvard University Press.
- Fried, C., 1983, *An analysis of 'equality' and 'rights' in medical care*. In J. Arras and R. Hunt, (eds.), 1983, *Ethical Issues in Modern Medicine*. Palo Alto, California: Mayfield Publishing Company, pp.490—496.
- Fuchs, V. R., (ed.), 1972, *Essays in the Economics of Health and Medical Care*. New York: NBER.
- Fuchs, V. R., 1978, *The supply of surgeons and the demand for operations*. *Journal of Human Resources*, Vol. 13 (Supplement), pp. 35—56.
- Gillon, R., 1986, *Philosophical Medical Ethics*. Chichester: Wiley.
- Ginsburgh, P. B., and G. M. Cater, 1986, *Medicare case-mix index increase*. *Health Care Financing Review*, Vol. 7, pp. 51—65.
- Goldfarb, M., M. H. Hornbrook and J. Rafferty, 1980, *Behaviour of the multi-product firm: A model of the non-profit hospital system*. *Medical Care*, Vol. 18, No. 2.

- Golladay, F., and B. Liese, 1980, **Health Problems and Policies in the Developing Countries**. World Bank Staff Working Paper No, 412.
- Groot, L. M. J., 1987, *Incentives for cost-effectiveness behaviour: A Dutch experience*. **Health Policy**, Vol. 7, pp. 175—188.
- Grossman, M., 1972, **The Demand for Health: A theoretical and empirical investigation**, New York: NBER.
- Guterman, S., and A. Dobson, 1986, *Impact of the medicare prospective payment system for hospitals*. **Health Care Financing Review**, Vol. 7, pp. 97—114.
- Hadley, J., 1982, **More Medical care, Better Health? An economic analysis of mortality rates**. Washington DC: Urban Institute Press.
- Harberger, A. C., 1978, *On the use of distributional weights in social cost-benefit analysis*. **Journal of Political Economy**, Vol. 86, Supplement, pp. 87—120, April.
- Harris, J. E., 1977, *The internal organisation of hospitals: Some economic implications*. **Bell Journal of Economics**, Autumn.
- Harsanyi, J. C., 1955, *Cardinal welfare, individualistic ethics, and inter-personal comparisons of utility*. **Journal of Political Economy**, Vol. 63, August, pp. 309—21.
- Hartley, K. and M. Hubby, 1985, *Contracting-out in health and local authorities: Prospects, progress and pitfalls*. **Public Money**, Vol. 5, September, pp. 23—26.
- Heller, P. S., 1978, *Issues in the allocation of resources in the medical sector of developing countries: The Tunisian case*. **Economic Development and Cultural Change**, Vol. 27, pp. 121—144, October.
- Hey, J. D., (ed.) 1989, **Current Issues in Microeconomics**. Basingstoke: Macmillan.
- Hölmstrom, B., 1982, *Moral hazard in teams*. **Bell Journal of Economics**, Vol. 13, Autumn, pp. 324—40.
- Hochman, H. H., and J. D. Rogers, 1969, *Pareto optimal distribution*. **American Economic Review**, Vol. 59, No. 3, pp. 542—557.
- Hoy, E. W., R. E. Curtis and T. Rice, 1991, *Change and growth in managed care*. **Health Affairs** (winter), pp. 18—36.
- Hurwicz, L., 1960, *Optimality and informational efficiency in resource allocation processes*. In K. Arrow, S. Karlin and P. Suppes (eds.) **Mathematical Methods in the Social Sciences**. Stanford: Stanford University Press, pp. 27—46.
- Hurwicz, L., 1972, *On informationally decentralised systems*. In R. Radner and C. B. McGuire, (eds.), **Decision and Organisation**. Amsterdam: North-Holland. pp.297—336.
- Hurwicz, L., 1973, *The design of mechanisms for resource allocation*. **American Economic Review**, Volume 83, Papers and Proceedings, May, pp. 1—30.

- IHSM (Institute of Health Services Management), 1988, **Working Party on Alternative Delivery and Funding of Health Services: Final Report**. Institute of Health Services Management
- Irvin, G., 1978, **Modern Cost-Benefit Methods. An Introduction to Financial, Economic and Social Appraisal of Development Projects**. London: Macmillan.
- Jacobs, P., 1973, **A Utility Maximising Approach to the Analysis of Non-Profit Maximising Hospital Behaviour**. Unpublished D. Phil. Thesis, Department of Economics and Related Studies, University of York, England.
- Jacobs, P., 1974, *A survey of economic models of hospitals*. **Inquiry**, Vol. XI, June.
- Johns, L., 1985, *Selective contracting in California*. **Health Affairs**, Vol. 4, (Fall), pp. 32—48.
- Jones-Lee, M. W., 1982, **The Value of Life and Safety**. Amsterdam: North-Holland.
- Joskow, P. L., 1980, *The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital*. **Bell Journal of Economics**, Vol. 11, pp. 421—447.
- Joskow, P. L., 1983, *Reimbursement policy, cost containment and non-price competition*. **Journal of Health Economics**, Vol. 2, pp. 167—174.
- Kalipen, E., 1993, *Determinants of infant mortality in Malawi: A spatial perspective*. In **Social Science and Medicine**, Vol. 37, No. 2, pp. 183—198.
- Kenya, Republic of, 1965, **African Socialism and its Application to Planning**. Sessional Paper No. 10.
- Kenya, Republic of, 1966, **National Development Plan, 1966—1970**. Nairobi: Government Printer.
- Kenya, Republic of, 1969a, **The Constitution of Kenya**. Nairobi: Government Printer.
- Kenya, Republic of, 1969b, **National Development Plan, 1969—1974**. Nairobi: Government Printer.
- Kenya, Republic of, 1984, **National Development Plan, 1984—1988**. Nairobi: Government Printer.
- Kenya, Republic of, 1986, **Economic Management for Renewed Growth: Sessional Paper Number 1 of 1986**. Nairobi: Government Printer.
- Kenya, Republic of, 1989, **Demographic and Health Survey, 1989**. Kenya Population Council.
- Kenya, Republic of, 1992, **Economic Survey, 1992**. Central Bureau of Statistics, (CBS), Nairobi.
- Kenya, Republic of, 1994a, **National Development Plan, 1994—1997**. Nairobi: Government Printer.
- Kenya, Republic of, 1994b, **Economic Survey**. Central Bureau of Statistics, Ministry of Planning and National Development.

- Kirigia, J. M., P. Fleuret and D. J. Byrne, 1989, **Effects of Health Care User Fees: Evidence from Meru District**. Nairobi, Kenya: UNICEF/USAID, (Mimeo).
- Kreps, D., P., Milgrom, P. Roberts and R. Wilson, 1982, *Rational co-operation in the finitely repeated Prisoners' Dilemma*. **Journal of Economic Theory**, Vol. 27, August, pp. 245—52.
- Lancaster, K. J., 1966, *A new approach to consumer theory*, **Journal of Political Economy**, Vol. 74, 132—157.
- Langwell, K. M., and T. Menke, 1993, *Controlling costs of the US. health care system: Trends and prospects*, in Arnould, Rich and White, 1993.
- Lazear, E. P., 1979, *Why is there mandatory retirement?* **Journal of Political Economy**, Vol. 87, December, pp. 1261—84.
- Lazear, E. P., 1987, *Incentive contracts*, in Eatwell, Milgate and Newman, 1987,
- Le Grand, J., 1982, **The Strategy of Equality: Redistribution and the Social Services**. London: George Allen & Unwin.
- Le Grand, J., 1984, *Equity as an economic objective*, **Journal of Applied Philosophy**, Vol 1, Part 1, 39—51.
- Le Grand, J., 1987, *Equity, health and health care*, in **Three Essays on Equity**, London School of Economics and Political Science: SunTory Toyota International Centre for Economics and Related Disciplines (STICERD), Welfare State Project, Discussion Paper No. 23.
- Ledyard, J. O., 1987, *Implicit contracts* in Eatwell, Milgate and Newman, 1987.
- Lee, M. L., 1971, *A conspicuous production theory of hospital behaviour*. **Southern Economic Journal**, Vol. 38, July.
- Leibenstein, H., 1966, *Allocative efficiency vs. 'X-efficiency'*. **American Economic Review**, Vol. 56, pp. 392—415, June.
- Leibenstein, H., 1987, *X-efficiency theory*, in Eatwell, Milgate and Newman, (eds.), 1987.
- Lindsay, C. M., 1969, *Medical care and the economics of sharing*. **Economica**, Vol. 36, pp. 351—362.
- Lindsay, C. M., 1976, *A theory of government enterprise*. **Journal of Political Economy**, Vol. 84, No. 5.
- Little, I., and J. Mirrlees, 1974, **Project Appraisal and Planning for Developing Countries**. London: Heinemann Educational Books.
- Lockwood, M., 1988, *Quality of life and resource allocation* in Bell and Mendus (eds.), 1988.
- Loomes, G (1988): *Disparities between health state measures: An explanation and Some Implications*, Paper Presented to the UK Health Economics Study Group meeting, Brunel.
- Luft, H., J. Robinson, and D. Garrick., *et al.*, 1986, *The role of specialised services in competition among hospitals*. **Inquiry**, Vol. 23, pp. 83—94.

- Lynn, T. J., 1991, *The performance of managed care: An insurer's perspective*. **Health Affairs**, Vol. 10, (winter), pp. 625—58.
- Margolis, H., 1982, **Selfishness, Altruism and Rationality**. Cambridge: Cambridge University Press.
- McCall *et al.*, 1987, *Evaluation of Arizona health care cost containment system, 1984—85*. **Health Care Financing Review**, Vol. 9, pp. 79—90.
- McComb, J., and J. Christianson, 1987, *Applying competitive bidding to health care*. **Journal of Health Politics, Policy and Law**, Winter, pp. 703—22.
- McLachlan, G., and A. Maynard, (eds.), 1982, **The Public/Private Mix for Health: The relevance and effects of change**. The Nuffield Provincial Hospitals Trust.
- McPake, B. and E. E. N. Banda, 1994, *Contracting out health services in developing countries*. **Health Policy and Planning**, Vol. 9, (1), pp. 25—30.
- McPherson, K., 1989, *International differences in medical care practices*. **Health Care Financing Review**, Annual Supplement, pp. 9—20.
- Melnick, G. A., J. Zwanziger, and T. Bradley 1989, *Hospital behaviour under competition and cost-containment policies in California, 1980—1987*. **Health Affairs**, Vol. 8, (summer), pp. 129—39.
- Migué, J. L., and G. Bélanger, 1974a, *Toward a general theory of managerial discretion*. **Public Choice**, Vol. 17, Spring, pp. 27—47.
- Migué, J. L., and G. Bélanger, 1974b, **The Price of Health**. Canada: Macmillan.
- Mills, A., 1990a, *The economics of hospitals in developing countries. Part I: Expenditure patterns*. **Health Policy and Planning**, Vol. 5, No. 2, pp. 107—117.
- Mills, A., 1990b, *The economics of hospitals in developing countries. Part II: Costs and sources of income*. **Health Policy and Planning**, Vol. 5, No. 3, pp. 203—218.
- Mishan, E. J., 1971, **Cost-Benefit Analysis**. London: Allen and Unwin.
- MOH, 1990a, **Preventive and Primary Health Care Resource Gap Study**. Nairobi: Study co-funded by Abt Associates Inc. and Thunder and Associates Inc.
- MOH, 1991, **Health Care Financing Programme: Position Paper**. Nairobi: Health Care Financing Programme, Ministry of Health, July.
- MOH, 1992, **Facility Improvement Fund: Operations Manual**. Nairobi: Health Care Financing Secretariat, July.
- MOH/GOK, 1993, **Strategic Action Plan for Financing Health Care in Kenya**. Nairobi: Kenya Health Care Financing Project, March.
- MOH/World Bank, 1986, Report on '*Expenditure and financing of the health sector in Kenya*,' Nairobi: Ministry of Health, February.

- Mooney, G., and A. McGuire, 1987, *Distributive justice with special reference to geographical inequality in health care*. In A. Williams, (ed.), **Health and Economics**, London: Macmillan, pp. 68—71.
- Mooney, G., J. Hall, C. Donaldson and K. Gerard, 1991, *Utilisation as a measure of equity: Weighing heat?* **Journal of Health Economics**, Vol. 10, 475—480.
- Mooney, G. H., 1982, **Economics, Medicine and Health Care**. Harvester Wheatsheaf.
- Mooney, G. H., 1983, *Equity in health care: Confronting the confusion*. **Effective Health Care**, Vol. 1, pp. 179—185.
- Morrisey, M. A., F. A. Sloan and J. Valvona, 1988, *Medicare prospective payment and post-hospital transfers to subacute care*. **Medical Care**, Vol. 26, pp. 685—698.
- Muellbauer, J., 1987, *Professor Sen on the standard of living*, in Sen, 1987a.
- Murray, J. A., 1974, *The behaviour of physicians: Key to health care sector*. **Journal of Behavioural Economics**, Vol. 3, No. 1, pp. 245—272.
- Mwabu, G. M., 1989, *Non-monetary factors in household choice of medical facilities*. **Economic Development and Cultural Change**. Vol. 37, No. 2, pp. 383—392.
- Mwabu, G. M., J. K. Wang'ombe and V. N. Kimani, 1989, *Demand for health services in rural Kenya*. Paper presented in the International Health Policy Conference, Manila, Philippines, 31 July—4 August.
- Mwabu, G. M., J. K. Wang'ombe and V. N. Kimani, 1991, *Health service pricing reforms and health care demand in Kenya*. Nairobi: University of Nairobi/Kenyatta University, (Mimeo).
- Newhouse, J. P., 1970, *Towards a theory of non-profit institutions: An economic model of a hospital*. **American Economic Review**, Vol. 60, pp. 64—74.
- Newhouse, J. P., 1977, *Medical care expenditure: A cross-national survey*. **Journal of Human Resources**, Vol. 12, pp. 115—125.
- Ng, Y., 1979, **Welfare Economics**. London: Macmillan.
- Niskanen, A. W., 1973, **Bureaucracy: Servant or master?** The Institute of Economic Affairs.
- Nonneman, W., and E. Van Doorslaer, 1987, *Economic incentives in the health care industry: implications for health policy making*. **Health Policy**, Vol. 7, pp. 109—114.
- Nozick, 1974, **Anarchy, State and Utopia**. New York: Basic Books.
- Odingo, S. and Z. Onyango, (eds.), 1974, **Health and Disease in Kenya**. Nairobi: Heinemann.
- OECD, 1987, **Financing and Delivering Health Care: A comparative analysis of OECD countries**. Paris: OECD.
- Paringer, L., and N. McCall, 1991, *How competitive is competitive bidding?* **Health Affairs**, Vol. 10, No. 4 (winter), pp. 220—230.

- Parkin, D., 1980, **Distance as a User-Cost Affecting Accessibility to Health Services**. Unpublished D. Phil. Thesis, Department of economics and Related Studies, University of York, England.
- Paton, C., **The Policy of Resource Allocation and its Ramifications**. Nuffield Provincial Hospitals Trust, Occasional Papers, 2.
- Pauly, M. V., 1971, **Medical Care at Public Expense: A Study in Applied Welfare Economics**. New York: Praeger Publishers.
- Pauly, M. V., 1974, *The Behaviour of non-profit hospital monopolies: Alternative models of the hospital*, in Havighurst, **Regulating Health Facilities Construction**. AEI.
- Pazner, E., and D. Schmeidler, 1978, *Egalitarian equivalent allocations: A new concept of economic equity*, **Quarterly Journal of Economics**, 92, No. 4, pp. 671—687.
- Pearce, D. W., and C. A. Nash, 1981, **The Social Appraisal of Projects: A Text in Cost-Benefit Analysis**. London: Macmillan.
- Pereira, J., 1989, **What Does Equity in Health Mean?** University of York, Centre for Health Economics, Discussion Paper No 61, August.
- Phelps, C., 1986, *Induced demand—Can we ever know its extent?* **Journal of Health Economics**, Vol. 5, pp. 355—365.
- Rawls, J., 1958, *Justice as fairness*, in **The Philosophical Review**, 67, pp. 164—94.
- Rawls, J., 1972, **A Theory of Justice**, Cambridge, Massachusetts: Harvard University Press.
- Rawls, J (1982): *Social unity and primary goods*, in Sen and Williams, (eds.), 1982, 159—185.
- REACH (Resources for Child Health Project, REACH)/KNH (Kenyatta National Hospital), 1988, **Kenyatta National Hospital Board: Plan of Action for Reform**. Comprehensive Report, Nairobi, May.
- REACH/MOH, 1988, **Nairobi Area Study**. Final Report: Nairobi: USAID/GOK.
- REACH/MOH, 1990, **Provincial and District Health Services Study: Nakuru District**. Main Report. Nairobi, September.
- Rees, R., 1989, *Uncertainty, information and insurance*, in Hey, 1989.
- Rice, R. G., 1966, *An analysis of hospital as an economic organism*. **The Modern Hospital**. Vol. 106, pp. 87—91.
- Robinson, J., and H. Luft, 1988, *Competition, regulation and hospital costs, 1982 to 1986*. **Journal of the American Medical Association**, Vol. 260, pp. 2676—2681.
- Robinson, J. and H. Luft, 1985, *The impact of hospital market structure on patient volume, average length of stay, and the cost of care*. **Journal of Health Economics**, Vol. 4, pp. 333—356.
- Rosen, S., 1985, *Implicit contracts: A survey*. **Journal of Economic Literature**, Vol. 23, September, pp. 1144—75.

- Rosko, M. D., and R. W. Broyles, 1987, *Short term response of hospitals to the DRG prospective pricing mechanisms in New Jersey*. *Medical Care*, Vol. 25, pp. 88—94.
- Rutten, F., and A. van der Werff, 1982, *Health policy in the Netherlands: At the cross-roads*. In Mclachlan and Maynard, (eds.), 1982.
- Sagar, M. A., D. V. Easterling, et al., 1989, *Changes in the location of death after passage of medicare's prospective payment system: A national study*. *New England Journal of Medicine*, Vol. 320, pp. 433—439.
- Saltman, R., 1992, *Recent health policy initiatives in the Nordic countries*. *Health Care Financing Review*, Vol. 13, No. 4, pp.157—167.
- Saltman, R. B., and C. von Otter, 1992, **Planned Markets and Public Competition: Strategic Reform in Northern European Health Systems**. London: Open University Press.
- Schleifer, A., 1985, *A theory of yardstick competition*. *RAND Journal of Economics*, Vol. 16, part 3, pp.219—27.
- Scitovsky, T., 1952, **Welfare and Competition**. London: Unwin University books.
- Secretaries of State, 1989, **Funding Contracts for Health Services**, London: HMSO, Working Paper No. 2.
- Sen, A K, 1982, *Equality of what?* in Sen, A K, 1982, **Choice, Welfare and Measurement**. Oxford: Blackwell; Cambridge, Massachusetts; MIT Press.
- Sen, A. K., 1985, **Commodities and Capabilities**. Amsterdam: North-Holland.
- Sen, A. K., 1987a, **The Standard of Living**. Cambridge: Cambridge University Press.
- Sen, A. K., 1987b, *Justice*. In Eatwell, Milgate and Newman, 1987.
- Sen, A. K., and B. Williams, (eds.), 1982, **Utilitarianism and Beyond**. Cambridge University Press.
- Sharif, A. M. H. R., S. M. M. Huq and M. us-Saleheen, 1993, *Spatial patterns of mortality in Bangladesh*. In **Social Science and Medicine**, Vol. 36, No. 10, pp. 1325—1330.
- Silver, M., 1972, *An econometric analysis of spatial variations in mortality rates by age and sex*. in Fuchs, (ed.), 1972.
- Sloan, F. A., 1974, *A microanalysis of physicians' hours of work-decisions*. In M. Perlman (ed.) **The Economics of Medical Care**, International Economic Association, London: Macmillan.
- Sloan, F. A., M. A. Morrissey and J Valvona, 1988, *Medicare prospective payment and the use of medical technologies in hospitals*. *Medical Care*. Vol. 26, pp. 837—853.
- Smith, C. J., 1987, *Equity in the distribution of health and welfare services—Can we rely on the state to reverse the inverse-care law?* **Social Science and Medicine**, Vol. 23, No 10.
- Squire, L., and H. van der Tak, (1975), **Economic Analysis of Projects**. Baltimore: The John Hopkins University Press.

- Sudgen, R., and A. Williams, 1978, **The Principles of Cost-Benefit Analysis**, Oxford University Press.
- Tait, A. A., and P. S. Heller, 1981, **International Comparisons of Government Finance**. International Monetary Fund, August.
- ten Have, H., 1988, *Ethics and economics in health care: A medical philosopher's view*. In Mooney, G., and A. McGuire (Eds.) **Medical Ethics and Economics in Health Care**. Oxford: Oxford Medical Publications.
- The Standard**, Nairobi: April 28, 1992.
- Thomson, G., 1987, **Needs** London: Routledge and Kegan Paul
- Torrance, 1986, G. W., (1986): *Measurement for health state utilities for economic appraisal*, **Journal of Health Economics**, Vol. 5, 1—30.
- Torrance, G. W. (1976): *Social preferences for health states, An empirical evaluation of three techniques*, **Socio-Economic Planning Science**, Vol. 10, 129—136.
- UNICEF (United Nations Children's Fund), 1993, **The State of the World's Children, 1993**. Oxford University Press.
- United Nations, 1986, **World Development Report**, Washington DC., 1986.
- van de Ven, W. P. M. M., 1989, *A future for competitive health care in the Netherlands*. NHS White Paper Series, Occasional Paper No. 9, Centre for Health Economics, University of York.
- Varian, H., 1974, *Equity, envy and efficiency*, **Journal of Economic Theory**, Vol. 9, 63—91.
- Veatch, R. M., 1982, *Ethical dimensions of the distribution of health care*. In J. van der Graag, W. B. Neenan and T. Tsukahara (eds.), 1982, **Economics of Health Care**. New York: Praeger Studies in Grants Economics, Praeger Publishers pp. 3—24.
- Vogel, R. J., 1990a, **Health Insurance in Sub-Saharan Africa**. The World Bank: Pre-Working Paper Series, Population, Health and Nutrition Department, August.
- Vogel, R. J., 1990b, *An analysis of three national insurance proposals in sub-Saharan Africa*, **International Journal of Health Planning and Management**, Vol. 5, part 4, 271—285.
- von Otter, C., and R. Saltman, 1991, *Towards a Swedish health policy for the 1990s: Planned markets and public firms*. **Social science and Medicine**, Vol. 32, No. 4, pp. 473—481.
- Wagstaff, A., 1991, *QALYs and the efficiency-equity trade-off*, **Journal of Health Economics**, Vol. 10, No 2, 21—41.
- Weekly Review**, June 3, 1994, *The medical profession: Doctors issue strike threat*. Stellagraphics, Nairobi, Kenya, pp. 19—20.
- Wennberg, J and A. Gittlesohn, 1982, *Variations in medical care amongst small areas*. **Scientific American**, Vol. 246, 100—111.

- West, P., 1981, *Theoretical and practical equity in the National Health Service in England*. **Social Science and Medicine**, Vol. 15, pp. 117—122.
- Whitehead, M., 1987, *The Health Divide: Inequalities in the 1980's, A Review*, London: Health Education Authority, March.
- Wiggins, D., 1991, *Need, Values, Truth*. Oxford: Basil Blackwell.
- Wiley, M. M., 1992, *Hospital financing reform and case-mix measurement: An international review*. **Health Care Financing Review**, Vol. 13, (4), pp. 119—133.
- Williams, A., 1973, *The cost-benefit approach: Bastard science? and/or insidious poison in the body politik?* In Wolfe, J. N., (ed.), 1973, *Cost-Benefit and Cost-Effectiveness: Studies and Analysis*. London: Allen and Unwin. pp. 30—60.
- Williams, A., 1974, *The cost-benefit approach*, **British Medical Bulletin**, 30, 252—256.
- Williams, A., 1985, *The economics of coronary artery by-pass grafting*, **The British Medical Journal**, 291, 326—29.
- Williams, A., 1987, *Quality adjusted life years*, **The Lancet**, (i), 1327.
- Williams, A., 1988, *Ethics and efficiency in the provision of health care* in Bell, J M and S Mendus (eds.), 1988.
- Wilson, G., and J. Jadow, 1982, *Competition, profit incentives and technical efficiency in the provision of nuclear medicine services*. **Bell Journal of Economics**, Vol. 13, pp. 472—482.
- Wolfe, P. R., and D. W., Moran, 1993, *Global budgeting in the OECD countries*. **Health Care Financing Review**, Vol. 14, (3), pp. 55—76.
- Woolley, J. M., 1989, *The competitive effect of horizontal mergers in the hospital industry*. **Journal of Health Economics**, Vol. 8, pp. 271—291.
- World Bank, 1993, *World Development Report: Investing in Health*, Washington D C.
- Zwanziger, J., and G. A., Melnick, 1993, *Effects of competition on the hospital industry: Evidence from California*. In Arnould, Rich and White, 1993.

6.3 Costs, Outputs and Competition

This section examines the effects of competition and regulation on some selected aspects of performance such as cost, quality, length of stay and capacity utilisation. The section focuses mainly on hospitals, the biggest user of any health system's resources, including in Kenya.

A profit-maximising hospital in a competitive environment, selling an undifferentiated single service, will always select an output level that is higher and a lower selling price compared to a profit-maximising monopolist hospital, their relative sizes (in terms of capital infrastructure) notwithstanding. In both cases, the unit cost, at whatever output level selected, would be minimised. The monopolist hospital may however earn surplus economic profits compared to the competitive hospital because its monopoly status enables it to pursue objectives such as those discussed in section 6.1.1 above, aspects that may translate into X-inefficiency. For example, if monopoly confers a degree of 'managerial slack', the surplus economic profits realised may be converted into types of costs that enhance the utility of managers—forms of expenditure that represent opportunity cost but that do not represent the minimum cost per unit of output (minimum marginal cost)—or are passed onto capital owners or employees in the form of remuneration that is higher than is necessary to keep resources in that particular use (Culyer and Posnett, 1990). Such expenditures no longer correspond to the true opportunity cost of resources.

Thus, in the competitive situation, each hospital acts as a profit maximising 'firm', responding to the preferences of the fully informed and knowledgeable consumers. Each hospital is regarded as being so small, and one among so many, that it cannot exercise control over any aspect of the market, except in its own (internal) cost structure. Without collusion, hospitals are forced to compete *only* on the basis of price, since consumers seek only hospitals with the lowest prices. Hospitals therefore have an incentive to operate at minimum cost in order to attract custom. Those not operating at least cost have this reflected in higher prices, to which consumers respond by switching their custom to elsewhere within the health care industry. The monopolist hospital is free from such pressures and can afford to operate at higher than minimum costs.

Similar comparative conclusions obtain for output-maximisation subject to a no loss constraint with respect to costs. Under both competition and monopoly, the output maximising hospital will set price equal to the average cost and standard results obtain. But predictions of standard economic theory may be inadequate because hospitals do not adequately fit within this framework.