



HEFP

HEALTH ECONOMICS &
FINANCING PROGRAMME

**Health Economics for Developing Countries:
A Survival Kit**

Anne Mills and Lucy Gilson

HEFP working paper 01/88, LSHTM, 1988

Also published as EPC Publication No 17, LSHTM, 1988.

ACKNOWLEDGEMENTS

The material in this publication was first developed for the 1986 Seminar on Health Economics and Health Financing in Developing Countries held at the London School of Hygiene and Tropical Medicine in association with the London School of Economics and Political Science and the World Health Organization. The support of these organizations for this and subsequent seminars is gratefully acknowledged. In addition, support from the Overseas Development Administration has enabled the original material to be further developed and published in this form for wider dissemination: for this we are most grateful.

The original material was written not only by one of the current authors (Anne Mills) but also by several colleagues whose important contributions we would like to acknowledge. Their names and main contributions are as follows:

Geoff Hoare

- Health care: the state versus the market
- Concepts of economic efficiency
- Inputs, resources and costs
- Outputs, health and health indicators
- Health sector finance and expenditure
- Sources of finance for the health sector

George Cumper

- Economic development and health
- Financing economic and health development
- National Accounts and the health sector

Jenny Roberts

- Demand, supply and the price system

However, we accept final responsibility for the revised versions of the chapters presented here.

We are also grateful to Dianne Fishman for her editorial assistance and to our EPC colleagues for their support.

CONTENTS

	Page
Acknowledgments	ii
Preface	iv
Chapter 1: Health Economics and its Contribution to Health Planning	1
Chapter 2: Economic Development and Health	7
Chapter 3: Financing Economic and Health Development	17
Chapter 4: Health Care: the State versus the Market	23
Chapter 5: Demand, Supply and the Price System	35
Chapter 6: Concepts of Economic Efficiency	46
Chapter 7: Inputs, Resources and Costs	50
Chapter 8: Outputs, Health and Health Indicators	60
Chapter 9: Techniques of Economic Evaluation	64
Chapter 10: National Accounts and the Health Sector	77
Chapter 11: Health Sector Finance and Expenditure	83
Chapter 12: Sources of Finance for the Health Sector	89
Chapter 13: Budgetary Procedures: Budgetary Reform and Programme Budgeting	101
Chapter 14: Approaches to Financial Planning: Resource Allocation Planning and the Financial Master Plan	108
Bibliography	117
Glossary	126

PREFACE

Governments have accepted the goal of Health For All by the year 2000, but its achievement requires that resources are made available to national health systems and are used efficiently. Economic recession has exacerbated the problems of financing the health sector in many countries, yet the funds that are available within the health sector are not always used in ways that will have the maximum impact on the population's health.

Health economics is increasingly recognized as a discipline that has much to offer developing countries in addressing these problems, but how can it help? What economic concepts and tools can be applied to the health sector? A wider understanding of the discipline is required if it is to support health sectors, rather than remaining the preserve of a few specialists.

This publication provides an introduction to health economics for health professionals and students with no previous economic background. It aims to present basic economic concepts in a clear manner and to demonstrate their potential application to the health sector, particularly within developing countries.

Following an introduction to health economics and its contribution to health planning, six main areas are addressed:

- Economic development and health
- The role of the state in health care provision
- Economic evaluation concepts and techniques
- Economic information relevant to the health sector and its sources
- Health financing issues
- Financial planning and budgeting.

This publication can provide only an outline of the corpus of health economics, to whet the reader's appetite. It is, therefore, supplemented by an extensive bibliography that will enable the interested reader to pursue any of the topics discussed. A glossary of economic terms completes this health economics survival kit.

Interested readers should, in addition, refer to a basic economic textbook for the principal concepts that are discussed (e.g. Culyer, AJ Economics Basil Blackwell, Oxford 1985), and to a health economics textbook for their application to the health sector (e.g. Cullis JG and West PA The

Economics of Health: An introduction Martin Robertson 1979 or McGuire A, Henderson J and Mooney G The Economics of Health Care Routledge and Kegan Paul 1988).

The material in this publication was first developed as background reading for a short course on health economics and health financing in developing countries. A companion volume is being produced, in the form of a loose-leaf file, containing the course objectives, programme and exercises. It is intended to assist those wishing either to develop their own courses or to improve their health economics understanding by working through the exercises.

Chapter 1

HEALTH ECONOMICS AND ITS CONTRIBUTION TO HEALTH PLANNING

1. Definition of Economics

The best starting point for consideration of the contribution of health economics to health planning is a definition of economics. Samuelson, the author of one of the most widely read textbooks of economics, defines economics as:

"the study of how people and society end up choosing, with or without the use of money, to employ scarce productive resources that could have alternative uses, to produce various commodities and distribute them for consumption, now or in the future, among various persons and groups in society. It analyses the costs and benefits of improving patterns of resource allocation."

This definition does not restrict economics to any one kind of human activity: it applies to all activities where scarcity exists and there is thus a need for making choices. Indeed, economics is often described as the study of scarcity and choice.

The emphasis of the above quotation is on describing and analysing decisions to do with scarcity and choice. This area of economics is called positive economics and it is concerned with 'what is', or 'was', or 'will be'. In addition, normative economics attempts to determine what 'should be', not merely 'what is'. Normative economics thus has to make judgements about the norms, or standards to be applied and disagreement over normative statements cannot easily be settled by empirical observation. For instance debate over the desirability of a private market for health care is often as much concerned with issues of normative economics (such as the value to be placed on freedom of consumer choice) as it is with issues of positive economics (such as how a private market behaves in practice). Although positive economics may not state what 'should be', it is still relevant to policy-making. For instance, positive economics cannot decide what health objectives ought to be achieved but it can explore the implications of adopting different objectives and different policy options.

2. Definition of Health Economics

Health economics can be defined broadly as the application of the theories, concepts and techniques of economics to the health sector. It is thus concerned with such matters as:

- the allocation of resources between various health-promoting activities
- the quantity of resources used in health delivery
- the organization and funding of health institutions
- the efficiency with which resources are allocated and used for health purposes
- the effects of preventive, curative and rehabilitative health services on individuals and society.

Although health economics has only recently developed as a sub-discipline of economics, it has established an interest in many of the main theoretical areas of economics. Figure 1 attempts to indicate the intellectual span of health economics and the main fields studied within health economics. The concerns of the different fields are as follows:

Box A : what determines health? What is the relative contribution of health services, income levels, education, environmental factors etc?

Box B : what value is placed on health and how can it be quantified?

Box C : what influences the demand for health services (demand derived from the demand for health)? What is the influence of price, income, travel time, behaviour of health care providers etc?

Box D : what are the characteristics of the supply of health services? What are the costs of production, mix of inputs, nature of the markets supplying health care inputs such as labour, drugs, equipment? What are the payment systems for health service suppliers and how do these influence their behaviour?

Box E : what are the costs and consequences of alternative ways of improving health/delivering a health programme?

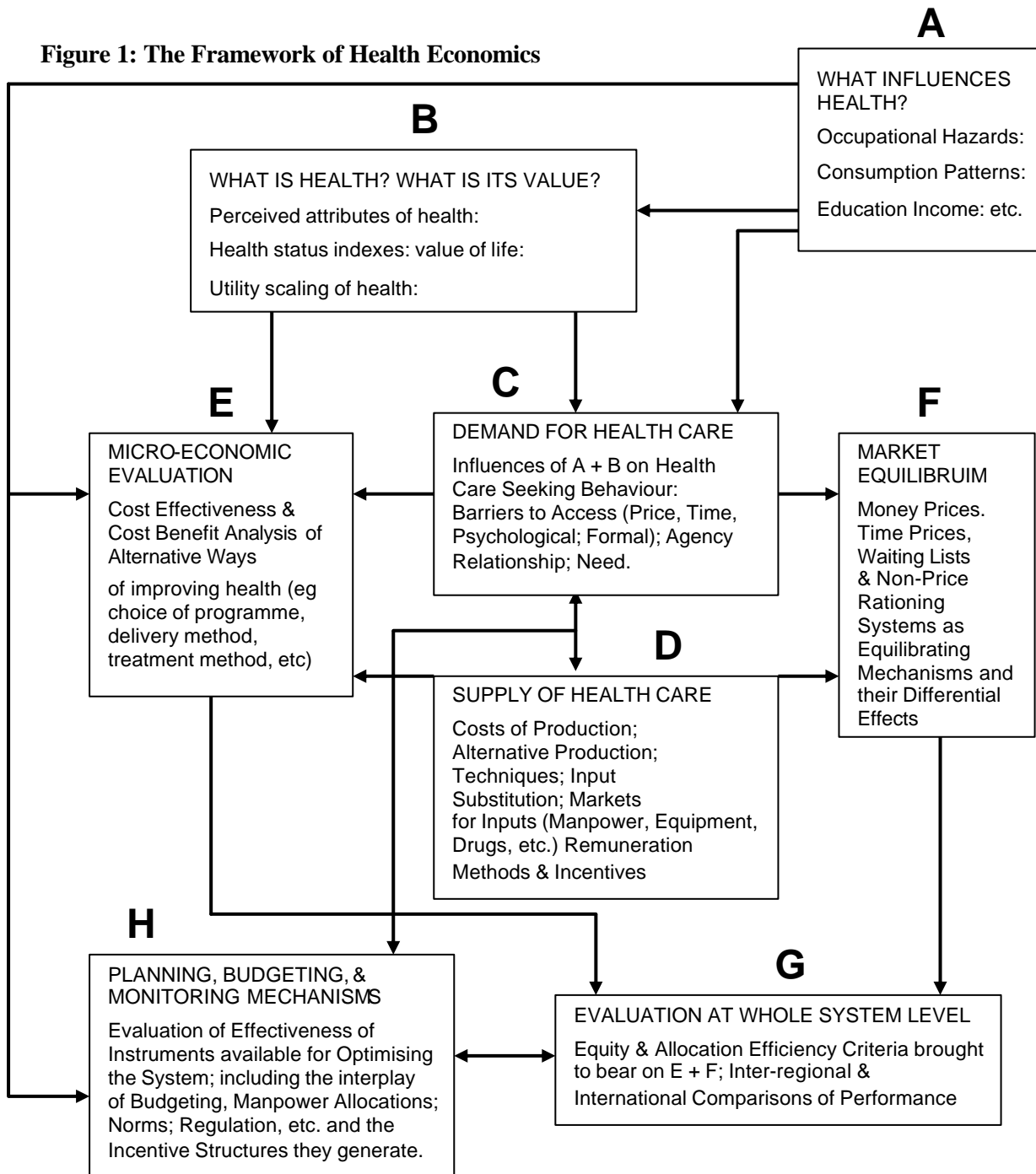
Box F : what are the results of the interplay of supply and demand for health services in terms of money or time price paid, rationing systems, who does/does not get health care?

Box G : what are the effects of different ways of financing and organizing the health sector in terms of efficiency and equity criteria?

Box H : what means are available to maximize the achievement of the objectives of the health sector (e.g. budgeting systems, planning methods) and how effective are they?

Theoretical and applied work has been done in all these areas, though in many cases the body of knowledge is still small relative to other sub- disciplines of economics.

Figure 1: The Framework of Health Economics



Source: Adapted from Centre for Health Economics, University of York.

3. The Contribution of Health Economics to Health Planning

Health planning is basically about choice: choice between one future or another; choice between various ways of achieving that future. Health economics is also interested in choice, so there is an obvious affinity between health economics and health planning.

Economic considerations play a key role in all aspects of life: in agriculture, housing, industry, trade and in health. In addition, the nature and level of a country's economic development is a major determinant of the health status of its inhabitants and is associated with the level of health service and health-related activities a country can support. Health policy and its implementation is thus strongly influenced by macro-economic considerations.

At the same time, the health of a population can itself influence economic progress. Health programmes have therefore come to be seen as part of a comprehensive strategy aimed at improving the social and economic welfare of populations. Such a strategy demands the selection of those programmes which improve health most efficiently: health services, the provision of other infrastructure such as water and sanitation, or actions aimed at improving nutrition, for example. Health economics can help to evaluate such choices.

The recent reappraisal of health policies in a number of countries has involved questioning the merits of many existing forms of care and of past strategies and priorities. Choices on how best to improve health exist everywhere, but such choices in poor countries are both crucial and difficult. Efforts to widen the choices to be considered for delivering health services and for encouraging health-promoting activities are therefore highly relevant.

They are particularly relevant in the economic context of lower income countries. Health services absorb a significant proportion of both government expenditure and family budgets. They also demand scarce foreign exchange for drugs, equipment and transport. Governments are actively seeking ways of containing costs, increasing efficiency and tapping additional resources. Health economics is attractive to them since it promises to help improve the allocation of health resources, increase their efficiency, identify more cost-effective technologies and evaluate alternative sources of health finance.

Table 1 attempts to elaborate the connection between economics and health planning in the following manner. The first column identifies a number of issues that are of direct relevance to planners (items A-G inclusive). Not infrequently the economist, in looking at such issues, needs to generate further questions which require answering before the issue can be tackled. These further questions appear in

the second column under the heading 'Prior Questions'. The final column is intended to show what economists can contribute in that area and offers pointers to those parts of economic theory that can best help to elucidate the planning issues.

It is important to emphasize, however, that health economics does not have all the answers. There are particular difficulties in applying some of the traditional conceptual and technical tools of economics to health. These difficulties may make health economics of great intellectual fascination to economists, but they hamper the application of health economics to health planning. While health economics certainly does have quantitative techniques to offer health planning, an equally important contribution is its distinctive mode of thought. The kind of approach characteristically adopted by the economist has been described by Culyer (1981) as:

"the desire to specify an unambiguous objective or set of objectives against which to judge and monitor policy; the desire to identify the production function; the recognition of the importance of human behaviour, as well as technology and the natural environment, in the causes, prevention, cure and care of disease".

The economist's views, of course, will represent only one input to planning and planners will weigh up views from a variety of sources in making their decisions.

VIII. Organizational behaviour

e.g. How can managers and health workers be encouraged to increase their efficiency?

1. Who makes the resource allocation decisions to and within the health sector, and what are their objectives?
2. What is the feasibility of reconciling the conflicting goals, values and interests of the various groups and individuals involved in the health sector?
3. What types of controls or incentives (monetary or otherwise) can be introduced to encourage efficient behaviour?

Managerial and behaviour theories of government, not-for-profit, profit and voluntary organizations

Notions of efficiency and the role of inducements (rewards and penalties)

XI. Project evaluation

e.g. Which health programmes or services should receive highest priority when allocating new funds?

1. Does the service do any good or have any discernible effect on health?
For whom?
2. What are the relative efficiencies (merits and demerits) of alternative health activities?
3. What are the distributional consequences of health activities (who incurs the cost, who receives the benefits?)

Micro-economic evaluation: cost benefit and cost-effective analyses. Notions of 'effectiveness' and the 'margin'; size and incidence of costs and benefits

X. Health Policy, equity and social justice

e.g. Does the operation of the health sector reflect the government's objectives
e.g. for equity?

1. How best can resources be matched to the population's needs, mortality and morbidity patterns, demands and utilization?
2. What impact do different health care systems have upon eligibility, access, take-up, and benefits received by target groups in the population?
3. What are the barriers, if any, to the provision of an equitable (fair) health service?

Optimum welfare criteria and the concept of the social welfare function

Inequalities and inequities in health care: definition and measurement issues

Effect of socioeconomic variables and physical access on utilization patterns

Chapter 2

ECONOMIC DEVELOPMENT AND HEALTH

1. Economic Development

Development means the regular unfolding of latent possibilities (as with the physical and mental development of a child). At the level of whole populations and nations, it refers to an historical process that has many aspects. Three important aspects considered here are: economic (which also has social and political implications), demographic and health.

1.1 The process of economic development

The idea of economic development is currently used in two ways:

- to arrange nations and communities on a scale from poor (less developed) to rich (developed)
- to refer to the process by which poor nations become richer.

Income per head ranges greatly between the nations of the world, from US\$150 in parts of Africa and south Asia to over US\$12000 in the United States, Switzerland and Scandinavia. There tends to be a systematic relation between income and other differences which include:

- pattern of production (eg. agriculture is more important in poor countries)
- pattern of trade (eg. poor countries tend to import manufactured goods, export primary products)
- energy use (eg. poor countries use less non-human energy per head)
- consumption patterns (eg. poor countries spend relatively more on food, little on consumer durables)
- degree of urbanisation (in poor countries a smaller proportion of the population lives in towns)
- demography
- health levels and health inputs.

Many of these differences between countries correspond to changes over time in the history of present 'developed' countries. Hence economists have built up a model of development as a process accounting both for the historical experience of developed countries and for present comparative data (the present rich countries are assumed to have started the process early, the poor ones late). This process is sometimes seen in terms of 'stages of growth'.

In the 1950s the driving force for this economic process was seen as investment (growth of physical capital such as roads, dams and factories). This leads to increased output, which in turn makes resources available for further investment - provided they are not swallowed up by population growth or increased consumption. The model was common to 'capitalist' and 'socialist' theories of development. The essence of development policy, following this view, is to invest more while holding down the rate of increase of population and consumption.

However, events in the 1960s and 1970s made this view of development seem oversimplified because 'investment' involves not only physical capital but also finance, technology and social organisation, each of which has special problems; and external trade and finance impose important constraints on development (taken up in chapter 3).

Economic development is, therefore, no longer seen simply as a process of injecting capital into the economy with automatic benefits in terms of production and human welfare. It is recognized that it may even have negative side-effects. Because capital is not a 'magic bullet', planning for development becomes much more complicated; it must take into account all aspects of the economy including health.

Although economic development has its own momentum, there has been agreement since the 1950s that it can at least be stimulated and its effects controlled, and that this involves some form of planning. There is a theoretical distinction between three approaches:

- 'laissez faire' where intervention is limited to providing favourable conditions for private industry, the benefits of which 'trickle down' to the whole population
- 'democratic socialist' where state intervention is more active and much attention is given to distributing the benefits of development through social services
- 'Marxist' where the crucial step is the transfer of power from a ruling class (capitalist or feudal) to the party representing the people; distributing benefits then becomes merely a technical planning problem.

In practice, whatever the approach adopted, most poor countries succeeded between 1950 and the mid-1970s in expanding national production, slowing population growth (so that average income rose in almost all countries), and improving the conditions of life sufficiently to produce a general increase in the expectation of life.

On the other hand, these successes were not inconsistent with an increasing absolute income gap between the richest and poorest countries, an increasing absolute number of poor and illiterate in the

world, and a recognition that the groups and classes involved in development continually used their political power to protect and further their interests. For example: developed countries try to preserve existing patterns of trade and finance, so that gains flow from the 'periphery' to the 'centre'; elites in developing countries try to corner the gains on development and multinational corporations use their bargaining power for the same purpose.

Further, development has been a highly uneven process. Some 'newly industrialized countries' (NICs) seem to be well on the way to developed status (e.g. Singapore, South Korea). Other countries appear to be economically static (e.g. Ethiopia) and international 'poverty belts' exist in sub-Saharan Africa and, to a lesser extent, in southern Asia.

Since 1975 economic growth everywhere has been lower or even negative following the large oil price increases of 1973-74, a rise in other product prices, and substantial borrowing by developing countries to address the resulting balance of payment problems (see Chapter 3). However, the cost of debt servicing has only exacerbated these problems and the 'debt crisis' of the 1980s has been the ultimate result - with developing countries unable to re-pay loans and international banks required to re-negotiate them in order to avoid a collapse in the international banking system. Not only have these problems inevitably meant that investment in social services could not be maintained at previous levels by many (non-oil importing) developing countries, but they have also come under direct pressure, for example from the International Monetary Fund, to cut their public expenditure as part of a package of 'structural re-adjustment' measures designed to address the crisis.

1.2 Demographic development

Economic development is paralleled by a process of change sometimes called the 'demographic transition'. For much of human history populations have grown very slowly because high birth rates (40-50 per thousand) have been offset by almost equally high death rates. At present the richest countries are again in a situation of slow population growth (less than 1% per year) but with much lower birth and death rates (10-20 per thousand).

In the transition from one situation to the other, birth and death rates do not generally move in step. Death rates fall first, and in combination with high birth rates produce a period of very rapid population growth (currently 3-4% per year in sub-Saharan Africa). Only when birth rates fall does the population begin to stabilize. This may not happen for a period of decades, giving time for the population to grow to several times its original level.

The demographic transition, like economic development, is believed to have its own logic and momentum, but is also partly controllable. The crucial factor, the speed at which birth rates follow the fall in death rates, is believed to depend partly on the fact that greater child survival causes parents to

desire less children, but also on deliberate policies of population control. This is the justification for the huge annual expenditure on national and international population control programmes.

1.3 Health development

'Health development' is a convenient name for the process by which populations move from a low level to a high level of health. The nature of these changes shows most clearly in the case of infant deaths. At low levels of health (Infant Mortality Rate of, say, 150 per thousand births), most infant deaths are associated with communicable diseases, particularly diarrhoeal and respiratory conditions. At intermediate levels the common infectious diseases (many of which have a nutritional and environmental element) begin to give way to a range of (mostly perinatal) conditions which require institutional care. At high levels of health (IMR 10 or less) infant deaths are reduced to a core of congenital conditions and expensively institutionalizable diseases.

For children and young adults the same basic pattern is found, although at lower absolute levels of mortality. Among the infectious conditions, diarrhoeal diseases fall in importance relative to the classic diseases of poverty (such as Tuberculosis) and of the tropics (such as malaria). Maternal mortality becomes important as the counterpart to perinatal conditions, and falls very sharply with the improvement of general health. Accidents are also significant causes of death. Although they have much the same incidence at all levels, in poor countries they tend to be agricultural and in rich ones, industrial or mechanical.

At older ages the gap between countries with low and high general levels of health tends to narrow. The difference in mortality rates for infectious diseases is much the same as at younger ages, but infectious diseases are much less important relative to chronic and degenerative diseases, which show few consistent differences between countries. Hence in all countries the typical diseases of old age are cancer, heart disease, stroke, diabetes, arthritis and mental conditions.

1.4 Interactions

What is the relation between economic, demographic and health development? The simplest answer is a straight causal sequence: economic development provides the extra resources for better nutrition, better housing and sanitation, health services and technology. These lead to lower mortality which triggers off demographic development.

But this answer is too simple. First, some of the relationships involved are circular. For example, unless the demographic transition is completed many of the gains from economic growth will be absorbed by a high rate of population increase. Economic development promotes better health, but better health, by reducing the burden of sickness and uncertainty, facilitates economic development

(although this has proved very difficult to demonstrate). The relationship between economic, demographic and health development is a complicated system of interacting variables.

Second, in this system some of the arrows between variables point both ways. For example, some nutritional differences between poor and rich countries make for better health; others, it has been argued, lead to specific 'Western' diseases (such as cancer, heart disease).

2. Production and Health

Economic development is usually accompanied by changes in patterns of production which interact with the health of the population. These production changes include:

- a shift in the balance between manufacturing industry and agriculture
- agricultural production for cash sale rather than subsistence
- rising levels of energy use, associated with new, more capital-intensive technologies.

The most obvious examples of interaction are unfavourable to health, possibly because:

- capital-intensive production shifts the balance of power in favour of employers, who exploit their position
- social controls break down and are not replaced by effective political controls (planning, safety)
- workers are not trained to handle new technologies safely.

Sometimes the damage to health is direct and concrete e.g. Bhopal in 1984 when a release of fumes from a chemical factory killed 2000 people, Chernobyl in 1986 when a partial meltdown of a nuclear reactor spread detectable radiation over most of Europe.

In other cases the effect is more indirect (and disputable). For example, irrigation leads to increased agricultural output but creates a suitable environment for the spread of waterborne diseases such as schistosomiasis (as occurred with the Aswan High Dam); the use of insecticides in agriculture encourages the emergence of resistant strains of malarial mosquito, making malaria more difficult to control.

Possible countermeasures at the national level include more careful impact assessment of new investments, and greater social responsibility among, and control of, individual producers. The cost of these countermeasures also has to be borne in mind.

However, from the economist's point of view development is always potentially favourable to health. It makes resources available that can be used for all the intersectoral actions which can contribute to better health, and also to more and better health services. Yet, for development in a particular case actually to be favourable to health it clearly must not generate any of the dangerous side-effects discussed above. Two more general questions must also be asked:

- how is the income from production distributed among persons and groups?
- how is that income spent (by individuals or society)?

Finally, it should be remembered that the health care industry is itself a form of production. It changes in the same direction as the rest of the economy. That is, it becomes more productive but also more capital-intensive with development, and can generate dangerous side-effects through the ignorance of clients, indifference and exploitation by producers and the failure of effective social control over its activities.

3. Distribution and Health

The proceeds of economic development (resources for health services, other kinds of goods and supplies favourable to health, income in general) are not necessarily equally distributed between persons and groups. Strict equality (equal shares for all) is probably not possible for health services or in general; what is sought is equity, fair shares for all obtained through the avoidance of inequalities which are not necessary or socially acceptable.

The goal of equity is supported on two grounds: as a matter of general social policy, or because it is believed that equitable distribution produces better health results for a given input of resources. This belief is justified because health services, like many forms of production, are subject to diminishing returns. That is, the more resources that are applied to a given population, the less the increase in output (health) obtained from adding one more unit of input.

It is conceivable that development may increase the available resources but may increase inequality to such an extent that large groups of the population are worse off than before, so that there is no improvement in overall health levels or other aspects of welfare. Hence there is a debate between those who believe in the 'trickle down effect' (benefits from development spread themselves naturally throughout the population) and those who believe that special measures are needed to preserve equity during development. For instance, it has been argued that the 'Green Revolution' is an example of unequal development. Based on new varieties of wheat it increased agricultural production and incomes in India (the Punjab) but the gains went to the operators of large and medium-sized farms, leaving the small farmers and landless worse off.

Through an examination of the main aspects of inequality in developing countries some general relationships between development and equity can be seen. (It must be remembered that the degree of equity always varies in pre-industrial countries, depending in particular on the amount of free land available.)

Urbanization - in developing countries towns are usually richer and have better social services and health indicators than rural areas. Development may increase the differential because new industries, high- technology services and administration are located in the towns. On the other hand, larger urban markets often mean high food prices and thus increased farm incomes - an example of the 'trickle down effect'. Development is also accompanied by a growing urban proportion of the population. On balance, the effect of development may be to reduce urban-rural differences, but only after a long period and with great variations between countries.

Sex differentials - in many developing countries, women have less access to formal employment, cash income and public services than men; these differentials are much narrower in developed countries. Health indicators often show the results of women's disadvantages, and in extreme cases women have a lower expectation of life than men (although the opposite is true for most countries). These disadvantages may also impede general health development - for example, the health of children is linked with the level of education of women. As fertility is also linked with female education, sexual inequality may impede demographic development as well. Finally, sexual inequality tends to limit the supply of nurses and some other types of health workers. The broad conclusion is that if development reduces sexual inequalities it is favourable to health.

Factor incomes - one way of looking at people's incomes is in terms of the kind of factor of production that they control (i.e. land, labour or capital), the amount of the factor that they have and its price (rents, wages, interest rates). In market economies, and indirectly even in planned economies, the scarcer the factor, the higher its relative price. Development involves changes in the balance between factors and so far as it results from investment (i.e. an increasing stock of capital), the price of capital should fall and income inequality be diminished. On the other hand, if a technological advance is made which needs capital to put it into effect, it is the existing capitalists who are best placed to take advantage of it, and so push up the rate of profit. Which tendency prevails in the long run? The evidence is hard to evaluate, but two broad indicators suggest that on the whole development evens out factor incomes: the share of the national income going to labour is higher in developed countries, and real interest rates (net of inflation) are lower.

Overall, therefore, economic and social inequalities and hence health differentials (measured, for example, by mortality rates) are probably less in developed than in developing countries. In the UK

class differentials in mortality still exist but they may be diminishing. They are certainly less than those demonstrated, for instance, between landowners and landless in Bangladesh.

4. Consumption and Health

4.1 Private consumption patterns

Patterns of consumption tend to vary in a predictable way with income, whether differences are considered between national averages or between income groups in the same country.

A useful way of summarizing these income/consumption relationships is through the economic concept of elasticity of demand. The income elasticity for a particular category of consumption, say food, can be defined roughly as the percentage change in expenditure on that item associated with a 1% change in the consumer's income. (See Chapter 5.) An income elasticity of more than 1 means an item of consumption takes an increasing share of total expenditure as incomes increase (termed a 'luxury' good). Medical care typically has an income elasticity of more than 1 (say, 1.3), although measurement is complicated by the role of the public sector in the provision of health care.

A typical poor country will spend 23% of national income on health care (public and private), whereas the richest countries spend up to 10%. Such statistics raise the question of what share of national income a poor country should devote to health care. WHO has suggested a minimum of 5%; but applying this in practice raises further questions:

- if countries can hold down the price of health service inputs, they can get more health care (in real terms) from a given expenditure. Attention recently has been given to reducing the price of drugs (e.g. through essential drug policies), but the most important input price to health care is the wages and salaries of professional and other health workers. These vary widely between poor countries, being influenced not only by the level of economic development but also by local demand and supply and by competition in the international market
- if expenditure on health care is increased, what other kinds of consumption will be reduced to make this possible, bearing in mind that many of these other categories contribute to health?

From a health perspective, levels and patterns of food consumption are obviously important. Food as a whole tends to have a rather low income elasticity (about 0.7), accounting for perhaps 75% of income in the poorest countries and 15% in the richest. With increasing incomes, there are shifts in the balance between categories of food; basic cereals tend to have the lowest elasticity, while oils and

fats and meat have high ones. Further, at higher incomes, food consumption includes an increasing element of refined and pre-cooked foods.

What is the significance of this for health? From one point of view, these trends are positive: starting from a diet which is largely carbohydrate, it seems that with increasing incomes consumers naturally move to one with a better balance in terms of proteins and vitamins. It has been argued that nutritional improvement (in quantity and quality) explains much of the improvement in health in industrial countries in the 18th and 19th centuries.

On the other hand, it can be argued that the changes in diet that are favourable to health at low levels of income may become harmful at higher levels. Among the factors blamed for various health problems in rich countries are excessive fats, salt and additives, and too little fibre. Further, it can be argued that considerations of prestige and imitation of foreign patterns encourage the spread of 'rich country' dietary patterns to countries at quite low levels of income. Hence the concept of 'Western' diseases.

4.2 Public consumption patterns

All countries have a public sector - i.e. some services are provided by agencies of government, central or local, rather than by private suppliers - and only the most extreme market economists argue that public provision of some functions is unnecessary. Nevertheless the share of the public sector in the economy varies widely. Has this any connection with levels of development on the one hand and with health on the other?

It might appear that regardless of level of development, countries are free to make a political choice between a 'socialist' pattern (in which all production, or at least the commanding heights, is under public control), a 'market' pattern (in which public production is minimal), or any intermediate stage of 'mixed economy'. But the actual situation reflects two conflicting trends. On the one hand, poor countries which aim at economic development see great advantages in a socialist approach - the possibility of coherent planning, equitable sharing of the results of development and continuity of policy. On the other hand, it is more difficult in the poorest countries for government actually to control the economy. Much of the economy may consist of small-scale agriculture and industry, which is always hard to control; tax systems have to be kept simple; communication difficulties limit government effectiveness and so on. Many of the poorest countries are nominally socialist, but typically only 15-20% of the national income comes under government control, against 40% in even non-socialist developed countries. Hence in many nominally socialist countries government does not in practice control all production, but concentrates on industries crucial to the development plan, plus provision of basic social services (e.g. India). Alternatively, or in addition, functions are pushed down to the local level for administration and financing (e.g. China).

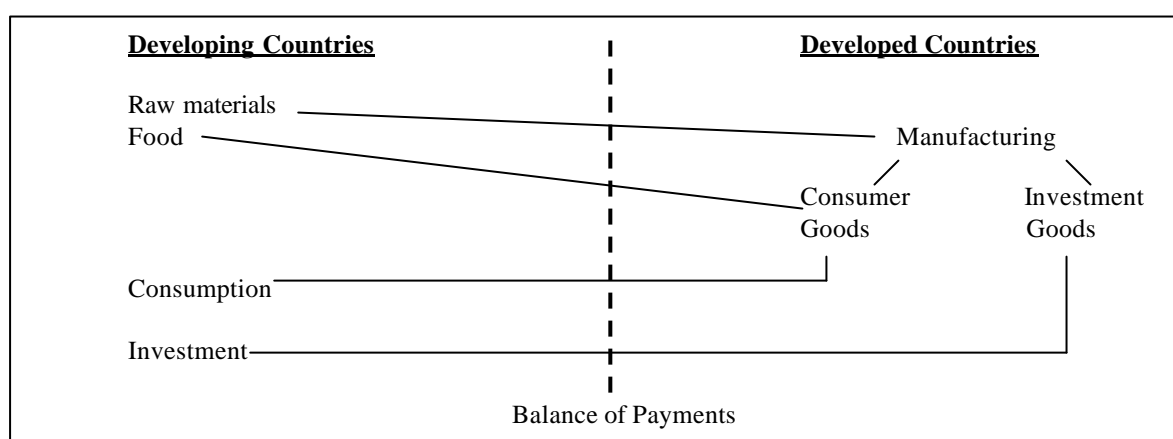
This pattern of ownership applies to health services as much as other forms of consumption: in most developing countries, a substantial share of health care is provided privately, or at a local level. Is this bad? It is very hard to show statistically any consistent difference in health indicators between countries with 'socialist' and 'private' systems of health care provision. Perhaps we need to distinguish between the many kinds of preventive and promotive work, which are unlikely to be provided effectively on a private basis, and other services where the difference in efficiency between public and private may not be great.

Chapter 3

FINANCING ECONOMIC AND HEALTH DEVELOPMENT

1. Trade and the Balance of Payments

At the end of the colonial period developing countries were mainly exporters of food and raw materials and importers of manufactured consumer and capital goods. Even after the postwar period of development (1945-75) this pattern still applies to many countries.



It has certain implications for economic growth in developing countries:

- if there is no structural change in their economies, output can grow only as fast as the demand from developed countries
but
- if there is structural change, the manufacturing function can be (partly) shifted from developed to developing countries (import substitution, light and then heavy manufacturing exports)
but
- this needs investment (factories, power, etc.) before manufacturing can grow - hence increased imports of capital goods, usually paid for by borrowing from developed countries.

The balance of payments of a developing country is typically made up as follows:

<u>Receipts of foreign currency</u>	<u>Spending of foreign currency</u>
Exports of primary products Other exports (including tourism)	Imports of: consumer goods material & fuel capital goods
New borrowing from abroad	Interest and repayment of previous borrowing.

Receipts must be balanced with spending. This imposes a limit on the rate of growth of developing countries: measures to promote growth tend, in the short run, to increase the excess of imports over exports, either through greater imports of consumer or capital goods, or the diversion of resources from export production to the domestic market. The aim of a developing country must be to achieve the maximum rate of growth consistent with long-run balance of payments equilibrium.

The problems of doing this have been complicated for developing countries in the postwar period in several ways:

- increased production of food and raw materials in the developed countries (e.g. European beet sugar competing with cane; synthetic fibres competing with cotton)
- fluctuating price and demand for primary products (an old problem)
- the oil price increases 1974 - 1981
- depressed demand in developed countries (especially since 1979)
- uncertainty about prices and exchange rates leading to high interest rates.

Even before the present depression, the importance of the world trade and monetary system as a constraint on development was a matter of international discussion. It was recognized that no one country alone could do much to improve the prospects for economic growth. There are many possible approaches to solving the basic problems:

- the 'imperial' solution: one developed country trades with a number of developing countries which it controls politically. This is clearly politically obsolete
- the 'free trade' solution: remove all restrictions on trade, currency movements and immigration, and let supply and demand operate. This has a long history and much theory behind it. But success would depend on all governments following the rules in spite of short run economic and political costs: this seems highly unlikely
- the 'multinational' solution: control of trade across national boundaries in one or a group of commodities by large businesses. This can produce stability of price and supply (e.g. the oil markets before 1973) but at the cost of exploitation and loss of sovereignty
- 'self-sufficiency': minimize foreign trade, build up the home economy from indigenous resources. The short-run costs are high and even in the long run only feasible for large developing countries (e.g. China, India - both apparently now retreating from this policy)

- 'common market' arrangements: countries of the same economic level trade together under agreed long-term conditions. This has worked best for developed countries (EEC, Comecon)
- a 'New International Economic Order': a combination of international agreements and the setting up of functional agencies under UN auspices to deal with specific problems. This solution was endorsed by the UN in the 1970s, although there has been little progress on it since the onset of the depression. Some agencies are already established: for the foreign exchange system (IMF), for investment capital (IBRD). Additional agencies would deal with stabilization of primary product markets, sharing of technology, etc. The problems include conflicts with national sovereignty, the weakness of the UN as an executive agency, and conflicts of interest between and within developed and developing countries.

Failing some international solution, the balance of payments deficit remains a key problem of economic growth for developing countries.

2. Financing the Balance of Payments Deficit

If a country's export earnings are not enough to cover the cost of its imports and other foreign currency commitments, the foreign exchange needed to fill the gap has to be provided in some form by foreigners. It matters a great deal who does this, for what purpose, and what price and other conditions they impose in return.

The greater part of the money borrowed abroad by developing countries is commercial - that is, it is lent by private individuals and institutions for profit. In many cases the borrower is also a private person or institution. The circumstances of private-to-private lending may vary greatly, from direct investment by foreigners in a local business under conditions where they share the risk of success or failure, to short-term loans to importers who must repay by a fixed date. Private-to-private lending can be justified provided the production it generates covers the foreign exchange commitment to interest and repayment, through higher exports or lower imports. But if this condition is not met it represents a possible burden on the foreign exchange resources of the borrowing country, for which the government is ultimately responsible. The shorter the term of the loan the more difficult it is for the government to deal with the consequences if anything goes wrong.

Private-to-private lending normally brings profit to both parties, although the interest rate required may be very high because of the strength of the lender's market position and the uncertainty attaching to investment in many developing countries. Private-to-government lending may also be justified if the borrowing can be applied to produce gains for the whole country - for example, infrastructure or

social development. But in this case the borrowing government's responsibility for the foreign exchange consequences is much more direct.

Outside the field of commercial investment, the governments of developed countries (and more recently of the capital-surplus oil countries) may also contribute to the foreign exchange needs of developing countries. Here the motivation is nominally non-commercial, and loans are only one of the forms of official development assistance, which also includes grants and technical co-operation. Such assistance may be either multilateral or bilateral.

Multilateral assistance involves funds or technical advice provided by international agencies who are financed by national governments but operate nominally in independence from them. The most important source is the UN system, with agencies concerned with world economic organization (World Bank, IMF), the preparation of development projects (UNDP), and specialized fields of assistance (e.g. WHO). Other international groupings also provide multilateral assistance (Comecon, EEC, regional development banks).

Bilateral assistance is negotiated directly between a donor and a recipient government. Unlike multilateral assistance it may include provision for special terms of trade between the countries concerned, and in fact much of the development assistance of Eastern bloc countries is believed to take this form (e.g. Russian purchases of Cuban sugar).

Although the central purpose of foreign aid is to provide developing countries with foreign exchange, many of its problems spring from the way in which both donor and recipient governments use it for other purposes. The former use it to promote their cultural influence and prestige and the commercial advantage of their nationals (for example, through tied loans which must be spent on the products of the donor country). The latter also use it for political and prestige purposes, and to raise money for government use beyond the limits imposed by their taxation system.

Finally, aid can be obtained from non-governmental organizations (NGOs). This aid is non-commercial and nominally free of government control (although few NGOs would ignore the sensitivities of the governments concerned). NGOs have usually operated at the local project level rather than the level of programmes or policies but this situation may be changing. They have a good reputation for innovative and emergency work.

3. Foreign Aid in the Health Field

Although it is difficult to find consistent and recent data on foreign aid, some estimates may help to show its limited role in relation both to the general foreign exchange problem of developing countries

and to health in particular. In 1977 the developing countries' annual foreign exchange requirement (excess of imports over exports) was about US \$45 billions; development assistance amounted to about \$16 billions, of which only \$4 billions was multilateral. In the health field, a number of estimates in the late 1970s suggested that the additional annual expenditure needed in the developing countries to achieve 'Health for All' (modestly defined) was of the order of US \$20 billions for health services, and a similar amount for other health-related services. The fraction of development assistance going for health is, however, quite small - perhaps 5% for health services and another 10% for other health-related activities. In absolute financial terms the contribution of foreign aid to health development is quite limited. It is therefore important to examine how it is distributed between purposes, types of expenditure and countries.

Even though the ultimate justification of foreign assistance lies largely in the recipient country's foreign exchange situation, its immediate justification may be in terms of specific inputs from overseas - not only imports, but also technical knowledge (in the form of 'experts', or of the training of nationals in the donor country or of appropriate forms of organization). In the health field the emphasis varies widely from one form of aid to another. In emergency food aid, for example, the commodity element is dominant. In other fields, such as malaria control or immunization campaigns, the organizational element may be the most important.

For various reasons, only a part of the nominal total of assistance funds becomes available for direct expenditure on health problems in developing countries. After deduction of the sometimes considerable administrative overheads, a substantial part of bilateral assistance funds remains with nationals of the donor country, as export guarantees, payment for technical services, etc. Multilateral agencies reserve considerable sums, often quite justifiably, for headquarters services of various kinds.

Bilateral assistance, in particular, tends to be biased toward capital projects, which in the health services field means the building and equipping of hospitals and health centres. From the point of view of the donor these are easier to administer, more conspicuous and more profitable in terms of exports, whereas from the receiver's point of view they involve less external control than a dependence on foreigners for current expenditure (only acceptable to quasi-colonial territories such as Puerto Rico and the French Antilles). For similar reasons water supply and education (school-building) are favoured among health-related activities. This capital bias is hard to reconcile with policies emphasizing Primary Health Care.

The distribution of health aid between regions and countries does not correspond to need. Bilateral aid is often frankly political. Multilateral aid has fewer political limitations but has no generally accepted rationale for the distribution of funds: WHO, for example, has experimented with distribution schemes based on health and income indicators, but applied to a base which is historically determined. WHO also accepts the general UN designation of certain countries as 'least developed'

and needing special treatment, but has limited means to provide this treatment without harming the interests of existing aid recipients, particularly in a period of economic stringency. Further, there is a bilateral element even in the operation of WHO since the regions may administer regional funds outside the regular budget - for example, the considerable funds flowing from the United States to Latin America.

The external aid available for health is therefore both small in relation to need and subject to distributional biases that limit its efficiency. It is very important for a developing country contemplating the use of external assistance to consider carefully the issues involved in order to make the best use of the aid available.

Chapter 4

HEALTH CARE: THE STATE VERSUS THE MARKET

1. Introduction

"Is health care in some way fundamentally different from food, clothes or any other commodity which is distributed through the market? Are there special characteristics distinguishing medical care from these other goods? Why should society's objectives be better achieved by non-market provision?" (Le Grand and Robinson 1976 pp. 32-33).

These questions are typical of those that are frequently raised in the debate on the nature of the commodity 'health' and the most appropriate organizational context for its optimal production, allocation and delivery. They reflect the two aspects of the debate: first, technical arguments about the functioning of the health care market; and second, more fundamental arguments about the nature of society's objectives. The second concern more basic issues because only when society's objectives are established can 'optimal production, allocation and delivery' be defined and understood.

However, the current debate focuses around the technical issues. For example, there are those who argue that health care is not different and can be allocated via market mechanisms in the same way as other goods and services. On the other hand, there are those who argue that health care is in some ways different, and requires the modification or replacement of market mechanisms. This focus reflects the establishment of efficiency as the central societal objective and the understanding of 'optimal' as 'most efficient'. However, a second important objective is that of equity, which is still a central policy objective within many government health care systems. It remains important for practical policy-making, therefore, to consider whether, and to what extent, the operation of the health care market achieves equity. The changing balance between equity and efficiency as policy goals helps to explain why the extent to which governments have intervened in the free operation of the health care market varies from country to country and has changed over time within countries.

In considering these issues three groups of questions will be addressed in this chapter.

- first, technical issues: what is the rationale for the market and the price mechanism? What are the requirements for the market and the price mechanism to function? What are the economic characteristics of health care? Is health care different? Are the conditions of perfect competition satisfied in the case of health care?

- second, what are society's objectives? Are they considered or realised through the functioning of the market?
- third, what do the answers to these questions imply for the organisation of health care services?

2. Markets and the Price Mechanism

What is a market? What do markets do? How do they work? The features of markets are usually described in relation to consumer choice and producer competition:

"A market economy may be defined as an economic system in which the mechanism of variable prices functions freely to signal consumer preferences and, through its effect on profitability, to encourage the allocation of resources - manpower, capital and raw materials - so as to satisfy those preferences. The consumer exercises choice by voting with his purse. It is an impersonal system which permits decentralised initiative in the use of resources; this in turn promotes competition and efficiency while maximising the range of consumer choice." (Centre for Policy Studies, Why Britain needs a social market economy, Barry Rose, Chichester 1975.)

A central feature of the market is the price mechanism (see Chapter 5). On the demand side, price is a measure of how much income must be sacrificed in order to obtain a commodity. The more something is valued, the more consumers will be willing to give up to obtain it. On the supply side, prices indicate to producers the strength of these consumer values. The price at which a producer can offer a commodity for sale reflects both his efficiency, and (to the extent that the prices of the resources used in production measure their scarcity value) the opportunity cost of the commodity as well.

Within the context of the market it is argued that the price mechanism, if allowed to operate free of constraints, will result in an optimal pattern of allocation, distribution and exchange. This is because:

- it transmits information about consumer preferences and the strength of those preferences; and also information about resource costs, about scarcity, about the efficiency and opportunity costs of production
- it provides incentives both to produce what is most highly valued in society, and to produce it in the most efficient manner

- given an initial distribution of resources or of income, prices determine who in society gets what goods and services.

In other words, prices signal what society wants, how much it wants it, what this costs in terms of resources or alternative options sacrificed and how efficiently producers are able to satisfy these wants. 'Optimal' allocation in this sense is efficient: i.e. a situation where no reallocation can make one individual better off without making at least one worse off (allocative efficiency). The price mechanism, in this argument, also has the advantage of securing efficiency in resource use: i.e. achieving maximum output at given cost or a given output at minimum cost (operational efficiency).

This type of structure stands in contrast to non-market systems which require conscious decision-making about what will be produced, how, when, where, in which ways, and for whom. Some writers have argued that the very complexity of a modern industrial economy makes it more difficult for a central planner to comprehend all the knowledge and information necessary for efficient management than in smaller and simpler economic systems.

The claimed advantage of the price mechanism as a rationing system is that it allows the decentralized coordination of a large number of separate activities. All of this happens automatically without the need for conscious planning or centralized decision-making. Or does it? What are the requirements for markets to function in this manner? Do markets operate this way in practice? The theoretical analysis and rationale for markets are usually explored using an idealized model of the real world known as perfect competition. For the price mechanism to operate in the way described above, a number of conditions must be simultaneously satisfied. These conditions are known as the conditions of perfect competition.

2.1 The model of perfect competition

The conditions of perfect competition are as follows:

- a large number of buyers and sellers, each small in relation to the total number so that they are unable to control price or output. With only a few sellers, competition may not operate
- no barriers to entry: producers are free to enter or leave the market, while factors of production are mobile
- no significant economies of scale which would give a price advantage to large-scale producers and imply a tendency towards monopoly. A natural monopoly describes the situation where it is more efficient in terms of production at least cost to have just one large producer, but this eliminates the competitive controls on price, quantity and quality

- no product differentiation or brand names; products are homogeneous, without quality differences etc.
- assumption of self interest: producers aim to maximize profits and consumers aim to maximize utility (i.e. benefits)
- no externalities or spillover effects in production or consumption. An activity creates an externality if people who are not involved in the decision to produce/consume are affected by it
- no risk or uncertainty: there exists perfect knowledge of prices, of products, of the implications of consuming or not consuming a product etc.

Market failure refers to the situation in which these conditions (necessary to achieve the market-efficient solution) are absent, or are contravened in one way or another.

3. Is Health Care Different?

Health care is not different from other goods in the sense that like other commodities it is scarce, and therefore requires allocation and institutions to organize its allocation. A number of characteristics are, however, often mentioned as distinguishing health care from many other goods and services:

- health care has both a consumption and an investment element (i.e. good health yields direct consumption benefits from being healthy, and contributes to greater production and investment)
- the costs of health care can be high in relation to income, and ill-health may also affect earning capacity
- patients have a direct involvement in the production as well as in the consumption of health care
- consumption or non-consumption decisions may well be irreversible (i.e. decisions may lead to death or permanent disability).

Although these characteristics are unusual they do not in themselves mean that health care cannot be treated in the same ways as other commodities, nor do they necessarily imply a particular form of organization. For example, insurance schemes have developed in order to minimize the adverse effects of high health care costs and the effects of illness upon earning capacity.

However, the health care market does not, in practice, function according to the theory of perfect competition. The main market failures are discussed below.

3.1 Consumer rationality and consumer sovereignty

Rationality implies consistency and in particular that decisions are consistent with the principle of utility maximization (i.e. that consumers use their economic resources in order to maximize their utility or benefit). In a few instances in health care, rationality would appear to be absent or impossible. Those who are mentally ill, and who reject or who do not recognize their need for treatment, are incapable of pursuing rational ends. Others, such as those who are unconscious, are temporarily unable to exercise rational choice. Health care may, therefore, in some circumstances be a merit good which must be distributed by the government because it will be under-consumed if left to the willingness-to-pay of individuals (who are not always rational in their demand for health care).

3.2 Risk and uncertainty

The need for health care is difficult to predict. If the cost of health care was small in relation to income this would not be important but it is not. Costs associated with illness are uncertain and often large. Insurance is of course the mechanism that has developed in order to cope with the problems of risk and uncertainty, and in many countries insurance systems exist to cover the expenditures arising from ill- health. Insurance systems do not necessarily imply state involvement, although many existing schemes operate with varying degrees of government support or control.

The chief problem with the operation of insurance in a market system is in the way risks are treated. Private insurance systems must at least cover their costs. To do this they set a premium based on the observable characteristics of those applying for insurance, and the estimated costs of treating certain conditions. Some individuals have characteristics that make them bad risks (such as the very young or very old, those with existing chronic conditions, and smokers); some conditions imply high treatment costs as a result of expensive procedures or long periods of treatment (such as the problems arising from chronic illness, haemophilia, or old age). A private insurance system is unlikely to cover these individuals. Some form of state intervention is therefore essential to ensure that all individuals secure access to health care, irrespective of their age, initial state of health, or ability to pay.

3.3 Information

For many commodities the consumer has some understanding of the product, or can acquire such information by experience. With medical care, patients have little idea of the effectiveness, of the quality, or of the consequences of having or not having treatment. Individuals may not even realize that they are ill. Furthermore, consumer entry into the market is infrequent, knowledge acquired from

past experiences becomes rapidly outdated, while the urgency of some conditions precludes time-consuming and often costly information gathering. There is little incentive for producers to provide information, and although patients may try to obtain more information through 'second opinions', doctors are traditionally reluctant to provide conflicting information or to disagree with colleagues. The irreversibility of much medical care emphasizes the importance of making the right decisions based on adequate information. Many professional groups and agencies have developed to provide information but for ethical reasons there are often controls imposed on the advertising of services.

These problems have in part contributed to the unusual relationship between producers and consumers in health care. The doctor is consulted and acts as an agent on behalf of the patient. The consumer chooses to delegate decisions about consumption to the doctor, thus demand for health care is often initiated by the supplier (demand may even be supplier-induced). An obvious danger of this is that consumers may be exploited: depending on the price of health care, and the method used to pay for it, doctors may stand to benefit from generating demand for their own services.

3.4 Externalities and public goods

There are numerous instances in health care where the consumption or production behaviour of one party spills-over to affect another party. Inoculation against infectious diseases, and public health measures such as sewerage and drainage are the most frequently quoted examples of these physical external effects. There may also be a caring externality underlying the provision of health care - this is discussed later.

A particular type of externality is known as a public good. The two distinguishing characteristics of this type of commodity are non-rivalness in consumption, and non-excludability. In simple terms this means that once the good is provided, it is possible for additional individuals to consume it without reducing the consumption of others, and that at the same time, it is impossible, or at least prohibitively difficult, to exclude others from consumption. The classic example of a pure public good is the lighthouse. The practical implications of such goods are that individuals have an incentive to understate their valuation of such commodities, and/or once they are provided, not to pay for them.

Certain health care services are characterized by the problems associated with public goods. Malaria control through environmental management (e.g. cleaning ponds) is an often-used example where one person's consumption does not impede another's, and where it would not be practical (or would be prohibitively expensive) to exclude non-payers from the benefits. Such services, if they are to be provided at all, will require some form of collective action.

The practical significance resulting from the presence of externalities and public goods is that left to market mechanisms alone some commodities will not be produced at all, whereas others will be produced in inappropriate amounts.

3.5 Competition and barriers to market entry

The medical care market is not conducive to free competition. A number of professional groups have developed to control entry of suppliers to the health care market. The arguments in favour of this control are that it is important to maintain standards of practice, and to reduce the uncertainty regarding professional competence. The disadvantages are that such controls, by reducing supply, tend to increase costs.

A number of barriers to entry exist - professional licensure, licensing of drugs and pharmaceuticals, controls on the establishment of new facilities and/or services. Professionalization may have increased the numbers of highly trained personnel and reduced uncertainty about competence, but it is essentially anti-competitive.

3.6 Economies of scale and monopoly

There are some instances of economies of scale and tendency to natural monopoly in health care. Examples include pharmaceutical firms, and hospitals. More often, the market will be characterized by a limited degree of competition between a few large producers (i.e. oligopoly). Price competition may be reduced in these circumstances by collusion, and competition limited instead to, for instance, advertising of brand names.

4. What are Society's Objectives?

The current debate about the most appropriate organizational context for the optimal production, allocation and delivery of health care has focused on whether efficiency can be achieved through the market. Yet equity remains an important goal for most societies. Although difficult to define, equity is associated with social justice and distributional fairness. The balance between the goals of efficiency and equity within any society is essentially an ideological issue. Three main perspectives have influenced this balance, to different degrees and at different times. They suggest alternative social arrangements for the treatment of human beings, and concern the importance of the individual, the nature of individual liberty, the functioning of the free market and the role of the state. They also suggest different views about the importance and nature of social justice.

4.1 Libertarian views

Libertarians analyse society in terms of its individual members and give importance in policy-making to the protection of individual freedom, defined narrowly as the absence of restraint (but including political liberty, free speech and economic freedom). The free market is seen as the most beneficial means of the production and distribution of goods because it is both efficient and protects individual freedom.

State intervention in the market is not acceptable for two reasons: 'natural rights libertarians' argue that it is morally wrong, and 'empirical libertarians' argue that it is wrong because it will reduce the total welfare in society. To the natural rights group, the allocation of earnings and goods by the market is just because the market procedures are themselves just (efficient and protective of individual liberty). Private property (wealth) is, therefore, outside government control and taxation is seen as a form of theft.

Empirical libertarians assert that the pursuit of social justice through market interventions is fruitless because something is just or unjust only if it has been caused by the action or inaction of an individual or individuals. Allocation via the market cannot, therefore, be judged in terms of justice because the market is an impersonal force. Moreover, they argue that the pursuit of social justice by governments is actively harmful (reduces welfare) because it entails the subjugation of individuals to government control and so undermines individual liberty. They do concede, however, the necessity of some taxation for the provision of a limited range of public goods such as poverty relief and law and order, where no method of private supply can be found.

4.2 Liberal views

Liberal theory also analyses society in terms of its individual members and emphasizes individual liberty. The definition of liberty, however, embraces economic security or need and so goes beyond that of libertarians. As a result, although the market theoretically allows the efficient distribution of goods, state intervention is deemed appropriate when the market fails to safeguard the security of all society's members. The market can fail to meet individual need because those who already have property rights (wealth and power) are favoured in comparison with those with fewer property rights. There is, therefore, a conflict between the efficiency and liberty that the market theoretically secures and the need to protect the economic security of some individuals. Taxation and income redistribution can be appropriate policy tools but should be used with reference to their effect on the trade-off between efficiency and equity.

This trade-off is most clear in the utilitarian position, in which the aim is to distribute goods (including rights, freedoms and political power) so as to maximize the total utility (welfare) of the

members of society. Such maximization involves both the efficient production and allocation of goods and their distribution in accordance with equity. However, primary importance is given to efficiency and so utilitarianism can sanction what some would call injustice because it can justify harm (failure to protect economic security) to the least well-off if this maximizes total utility. Utility can be associated both with the process of distribution as well as its result, and the primary importance of efficiency may be lessened where utility results from the process of allocating goods equitably.

An alternative liberal view (Rawlsian liberalism) makes equity and justice the primary aims of policy, resting on the principle of 'the maximin' i.e. maximizing the position of the worst-off in society. In practice, the approach implies that there is a legitimate redistributive role for the state, and goods are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the least favoured.

4.3 Collectivist views

Collectivist theories are varied but have three central goals: equality, freedom and fraternity. Analysis of society is based primarily on groups, such as social classes, rather than individuals and the goal of fraternity emphasizes co-operation, duties and the good of the community.

The collectivists' concept of freedom is broad, including both equality and economic security, and it underlies their view of social justice and their rejection of allocation via the market. It is argued that fundamental inequality (of power and wealth) results in the free market inevitably distributing ever more unequally. A more or less continuous redistribution of rights and wealth is required if egalitarian aims (including economic security) are to be met.

Government action (taxation, public production) is, therefore, justified because only it can create the conditions for the full exercise of freedom by all people. The creation of equal opportunities in the distributional process may not be sufficient because substantial inequality in end-states (and freedom) may still persist. Rather, positive equalizing measures are needed - although not necessarily complete equalization. Collectivists disagree about whether such measures should adjust market mechanisms or replace them: socialists believe that the market system can be harnessed to their goals; Marxists argue that the system is inherently in conflict with those goals and so give the state the primary role in production and allocation.

4.4 Underlying behavioural assumptions

These theories of society contain implicit views about human behaviour that influence the just balance between efficiency and equity. Libertarians represent the extreme end of welfare economics, and

assume that individual decision-making is based entirely on self-interest and the pursuit of welfare (efficiency). Liberals and collectivists are concerned, to different degrees, with establishing equity principles for decision-making based on the assumption of more altruistic behavioural patterns. The source of value for efficiency judgements is ultimately the preferences of individuals. The pursuit of equity, in contrast, is often determined independently of the interest of the individual and promotes altruism, caring or other values.

A number of specific views about equity and health are based on this concept of rational altruistic behaviour. For example, altruism has been used as the basis for the defence of the National Health Service (NHS) in the UK and its concern for equity. Equity can also be seen as a caring externality - based on the proposition that individuals care about one another, and in particular care about each others' health status. Their decisions are based, therefore, not just on their own welfare but also on their perceptions of the welfare of others. In practice, the caring externality suggests that as the other person's welfare increases (decreases) so does one's own. More strongly, the concept of 'commitment' suggests that an individual may choose an act that will yield less personal welfare than an alternative act also available for selection. Such a decision runs counter to the view of behaviour based on self-interest in the pursuit of personal welfare.

5. Is Equity Achieved Through the Market?

From this discussion it can be seen that, from an equity perspective, there are two theoretical arguments against the market-based allocation of goods.

First, while the market does not require any particular pattern of income distribution to function, it will reinforce the pattern in existence. This pattern is generally inequitable and is based on accidents of birth, property and education. The market exacerbates income inequalities by allocating goods according to the individual's ability to pay for them, and so, in theory, allocation via the market will undermine the health of those who are both least able to afford health care and are most vulnerable to ill-health (the low-income groups). The pursuit of equity involves allocating goods such as health care in a way that counterbalances income inequalities (i.e. favouring the low-income groups), together with a pattern of taxation that falls more heavily on high- than low-income groups. In practice, health care must be allocated in accordance with the objective of achieving a net subsidy for low-income groups and a net contribution from high-income groups.

Second, the market assumes that individuals are basically self-interested in their behaviour and seek to maximize their own utility. A concern for equity, however, reflects less selfish behaviour; it represents a concern for the well-being of others.

These theoretical arguments are supported by the practical realities that there is often an inverse relationship between income and ill-health, that ill-health may further limit income earning capacity and that the cost of health care is often high in relation to income. Consequently, at a minimum, intervention in the market is required to protect those who are both most likely to be sick and least able to afford health care.

6. Health Care: The Role of Government Intervention

Government intervention in the health care market may be promoted on either efficiency or equity grounds. It can attempt to restore the conditions necessary for the market to work or to limit the undesirable effects of markets and market failure. For example, allocative measures are designed to restore the conditions of perfect competition; and distributive measures are designed to correct an undesirable distribution of economic resources, and then to allow markets to work. Distributive measures may work through the tax system via transfer payments to the poor or negative income taxes; or through price regulation via price support (e.g. minimum wages) or tariffs.

Moreover, government intervention can be minimal, limited to that of umpire, referee, information provider or regulator of market fixing by vested interests. Or government can have a more extensive role ranging from the regulation and control of the health care system to the provision of the finance for services to the direct provision of services for all.

A fully socialized health care system - meaning, in this instance, a system in which health care is provided without charge to the consumer and financed from general revenue - carries the remedial measures to their logical conclusion. For example, with free service (equivalent to a 100% subsidy) there is little distortion of demand through differences in income. Recognition of such advantages of government production is not denied in orthodox Western economics; in the applied field of public finance it has always been recognized that there might be merit goods which the state might choose to provide independently of the market, with education and health care as conspicuous examples.

However, like the competitive market, the socialized health care system does not exist in its ideal form in the real world. Among the qualifications are:

- even with zero price and wide geographical dispersion of services, there will persist some non-price impediments to access that make complete equality hard to achieve
- a socialized system of health care in a non-socialized economy will be exposed to some residual market pressures, for example in the supply and price of inputs it purchases from the rest of the economy

- the degree of equity attained in a socialized health care system depends partly on the nature of the tax system from which it is financed (and on the nature of other sources of finance such as foreign aid). If it is regressive (bearing more heavily on the poor than the rich) a zero price is not enough to achieve overall equality in the health care system
- as a socialized system does not rely on market forces for resource allocation, it must be planned. Its properties therefore depend on the efficiency and cost of the planning system
- at a lower level, the allocation of services is likely to be left largely in the hands of health professionals. In any conflict between consumer needs and demands the result is likely to be biased towards professionally assessed needs, especially when the consumer does not pay and therefore may have less influence
- a socialized health care system, like any other large organisation, is always in danger of bureaucratization, with all this implies in terms of inertia and emphasis of the interest of the producer above those of the consumer.

In practice, neither ideal form of health care organization, the market or a socialized health care system, functions according to theory. Mixed systems exist in reality, involving different degrees of government intervention and reflecting different balances between the two central aims of society, equity and efficiency. Practice tempers theory in policy-making and the realities both of scarce resources and of unequal distribution (of income, of ill-health) are recognized. There is a great need for examination of the specific form that is appropriate for a particular country and period, in order to minimize its side-effects.

Chapter 5

DEMAND, SUPPLY AND THE PRICE SYSTEM

1. Introduction

In order to use scarce resources as efficiently as possible, an efficient information system is required - to bring together the values of the alternative uses of resources (products) with their production cost, and to co-ordinate the many decisions of consumers and producers. The price system, although imperfect (see Chapter 4), attempts this difficult task and plays a central role in economic analysis. It influences many health and health care programmes either directly or indirectly (e.g. through markets for personnel, drugs etc.). It is thus important to consider the workings of the price system, through an examination of demand and supply. An idealised model of demand, supply and the price system is presented, before considering its relevance to health care.

2. Demand

In economics, the quantity of a particular good demanded is seen as a variable determined by a range of factors.

The demand function summarizes this relationship:

$$Q_d = f(P, RP, Y, T, \dots)$$

The quantity demanded (Q_d) is some function of the price of the good (P); the prices of other related goods (RP); income (Y); and the organizational and institutional structure of society and preferences and tastes of individuals (T) which will depend on many socioeconomic and cultural factors. Of these, prices and incomes are the most useful for theoretical analysis. The demand function can represent the demand of an individual, or of all individuals demanding a particular good - in which case it is simply the sum of individual demands.

The relationship between demand and price is often portrayed as a demand curve (Figure 2). It illustrates how much would be demanded at each price. For most goods, more is bought as the price falls and so the demand curve will slope downwards.

In using a two-dimensional drawing of the demand curve it is assumed that the other factors (RP , Y , T) remain constant. This helps to clarify different aspects of demand, but modern econometric techniques allow more sophisticated estimation of the impact of the different variables on demand.

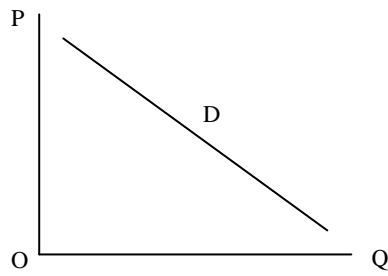


Figure 2

It is also important to know about the responsiveness of demand to changes in any of the variables. Responsiveness is measured by economists and referred to as the elasticity of demand; it is reflected in the shape of the demand curve.

Price elasticity of demand is measured by expressing the percentage change in quantity demanded (Qd) as a proportion of the percentage change in price (P).

$$\frac{\% \text{ change in Qd}}{\% \text{ change in P}} = E_d$$

If the percentage change in Qd is greater than the percentage change in price the demand for a good is elastic, $E_d > 1$. If the percentage change in quantity demanded changes by less than the percentage change in price ($E_d < 1$) the demand is inelastic. If the percentage change in price evokes a similar change in percentage quantity demanded then the demand is unitary elastic, ($E_d = 1$). Elasticity will depend upon the ease with which goods can be substituted for one another.

Income elasticity of demand measures the response in quantity demanded (Qd) which arises from changes in income (Y).

$$\frac{\% \text{ change in Qd}}{\% \text{ change in Y}}$$

Cross-elasticity of demand measures the response in quantity demanded of good A (Qda) which arises from changes in the prices of other goods or services (Pb).

$$\frac{\% \text{ change in Qda}}{\% \text{ change in Pb}}$$

It measures the degree of association amongst goods, i.e. whether they substitute or complement (need to be used with) each other.

3. Supply

In economics, supply is used in a sense symmetrical with demand. The quantity of a good that producers offer for consumption is seen as determined by certain economic variables, of which the most important is the price at which the good can be offered.

The supply function relates the quantity supplied to the variables likely to affect it.

$$Q_s = (P, RP, C, RC, T, \dots)$$

The quantity supplied (Q_s) is some function of the price of the good (P), relative prices (RP), costs (C), relative costs (RC), and tastes (T) which will include socioeconomic and cultural factors.

The supply curve (Figure 3) shows the relationship between price and the quantity supplied. It will normally slope upwards indicating that more will be supplied if prices rise.

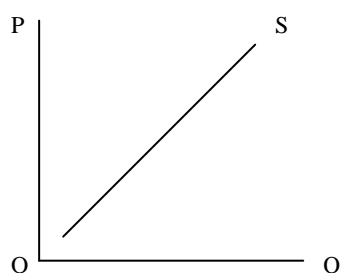


Figure 3

The ease with which quantity supplied will respond to changes in price is measured by the elasticity of supply, and it is reflected in the shape of the supply curve.

The elasticity of supply is roughly measured by expressing the percentage change in quantity supplied (Q_s) as a proportion of the percentage change in price (P).

$$\frac{\% \text{ change in } Q_s}{\% \text{ change in } P} = E_s$$

Like the elasticity of demand, it is described as elastic where ($E_s > 1$) and inelastic where $E_s < 1$. Elasticity will depend upon the ease with which quantities of goods or services can be produced. In some cases over an appreciable time span it may be difficult to obtain any more goods, irrespective of the price offered, e.g. agricultural crops or production where the scale of activities is fixed in the short term.

4. Equilibrium

Equilibrium in a market is achieved when supply equals demand (Figure 4). Price and quantity will adjust until the point is reached where buyers and sellers are content to exchange a given quantity (q_1) at a given price (p_1).

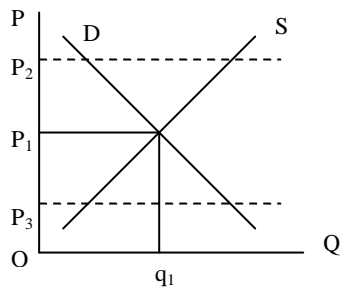


Figure 4

At p_2 suppliers will be willing to supply more than buyers are willing to purchase at that price. There will be excess supply and pressure to reduce prices. At p_3 buyers will be willing to buy more than suppliers are willing to offer and there will be pressure on prices to rise - excess demand.

If factors in the equation other than price change, the supply/demand curve will shift. Assume a change in fashion (T) for a good, leading many people to want more of it at any price (Figure 5).

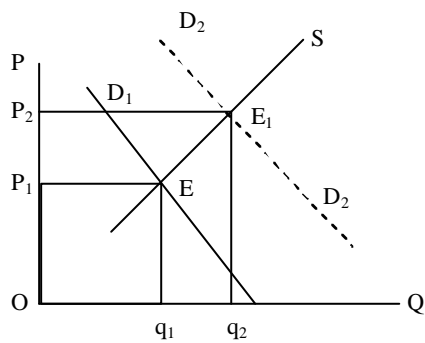


Figure 5

A new equilibrium (E_1) is reached after the demand curve has shifted outwards, and price and quantity have risen, the extent depending on the size of the shift and the elasticity of supply and demand. An example of this is the impact of the 1966 relaxation of the papal instruction to 'eat no meat on Friday', which resulted in a reduction in fish consumption and a 20% price fall for some fish (e.g. haddock & whiting). Another example might be a change in supply conditions. Perhaps a new technique is available which enables more to be supplied at any price, i.e. the supply curve slips downwards (Figure 6).

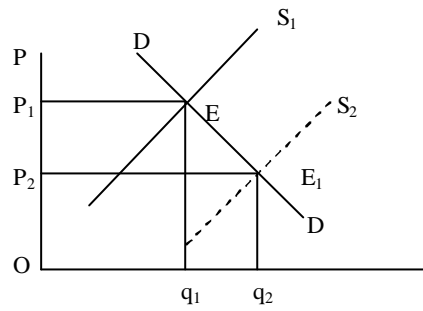


Figure 6

In this case the supply curve has moved down and demand has remained unchanged. The new equilibrium (E_1) will be arrived at with prices lower and quantity purchased greater than before. The extent of the change will again depend upon the elasticity of supply and demand.

Some additional examples (Figures 7 and 8) show the effects of different elasticities. At any particular point on the curve the responsiveness is reflected by the slope of the curve, i.e. the demand and supply curves become steeper as they represent less response of quantity demanded/supplied to price changes.

Figure 7

Inelastic demand
showing no change
in quantity purchased
as price falls.

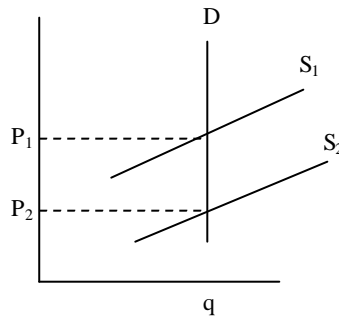
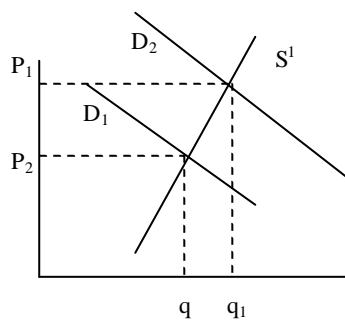


Figure 8

Inelastic supply
showing not much change
in quantity supplied as
price rises.



5. Supply, Demand and Policy

Supply and demand analysis can be used to consider policy changes such as the introduction of a tax on production and consumption of a good or subsidy on price or a policy of price controls.

Figure 9 shows the effect of a tax. Price has risen and quantity purchased has fallen. The extent of the rise or fall will depend upon the tax change and the relative elasticities of the supply and demand curves.

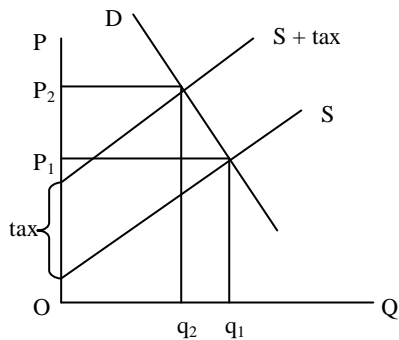


Figure 9

Figure 10 shows the effect of a subsidy. Price has fallen and quantity purchased has risen. Again the extent to which changes occur depend on the responsiveness of demand and supply.

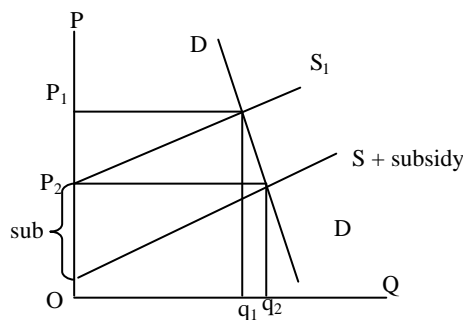


Figure 10

Figure 11 shows the effects of the imposition of a price control (p_2) above equilibrium price (p_1). Price has risen and quantity demanded has fallen but excess supply ($q_1 - q_2$) is available; unless this is managed there will be pressure on the market to bring down prices.

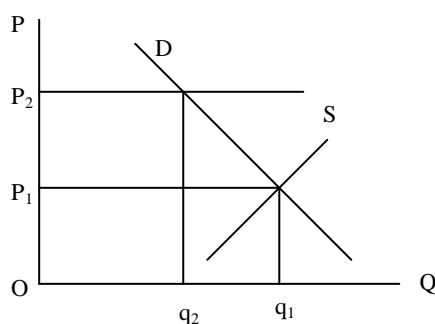


Figure 11

Figure 12 shows the effects of the imposition of a price control (p_2) below equilibrium price (p_1). Price has fallen and quantity has fallen. At this price there will be excess demand ($q_1 - q_2$) and pressure to bid up the price unless some form of rationing is introduced.

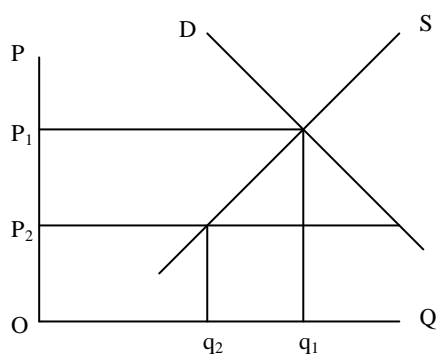


Figure 12

Taxes and subsidies can be used to promote health. For example, taxes can be imposed on health-damaging substances (e.g. cigarettes) or on polluting industries, thus raising their price and reducing consumption/production. The extent of the reduction will depend on the relevant elasticities. Similarly, subsidies can be used to promote consumption or production by in effect lowering the price of health-promoting activities (e.g. zero-priced ante-natal screening or dental and eye checks).

6. Externalities and Demand/Supply Analysis

Values, as expressed by the amount people are willing to pay, reflect only the private assessment of values and no account is taken of society's values. If there is a difference between the individual's and society's valuation of a good or service, externalities exist, i.e. factors are important that are external to those taking part in the transaction.

6.1 Demand or consumption externalities

If there are positive external benefits from the consumption of a good or service, society's demand curve will be to the right of the individual demand curve e.g. immunization, preventive medicine (figure 13); if there are negative external benefits from the consumption of a good or a service, society's demand curve will be to the left of the individual demand curve e.g. smoking and consumption of fat (figure 14).

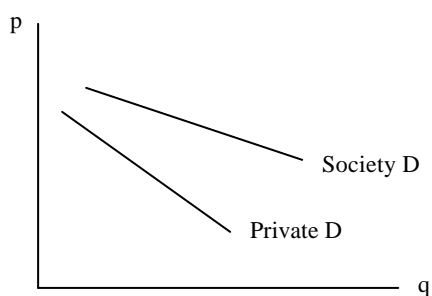


Figure 13

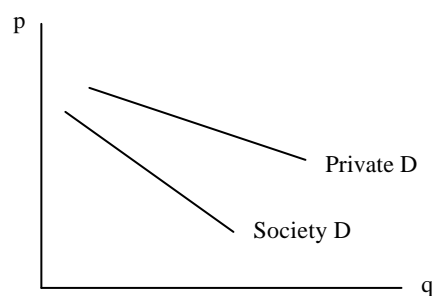


Figure 14

In either case, without intervention, over- or under-consumption would occur.

6.2 Supply and production externalities

A similar situation might occur on the supply side (Figure 15). The costs which underlie the supply schedule only represent the private costs borne by the supplier. There may be costs borne by others as a result of the activity undertaken when producing goods e.g. river pollution, noise pollution or the use of scarce resources.

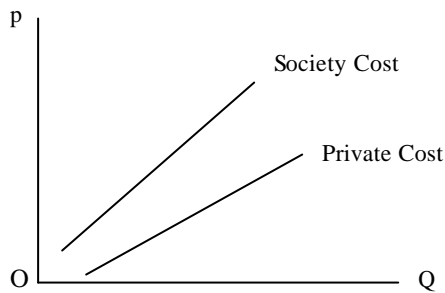


Figure 15

Alternatively, the process may provide benefits in terms of amenities that can be used for other purposes and society's costs would then be below private costs (Figure 16).

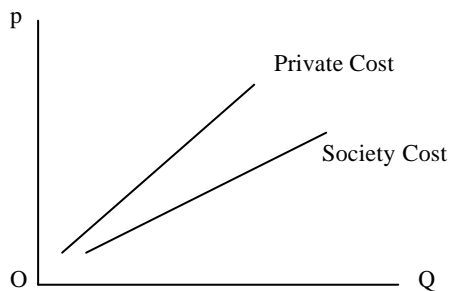


Figure 16

Again, over- and under-consumption would occur if no intervention takes place. Intervention may take the form of taxes or subsidies of services, legislation to control output or price, or direct intervention in production and distribution of services.

7. Is There a Demand for Health Care?

7.1 Need, utilization and demand

It can be argued that the demand for health care is fundamentally different from the demand for 'supermarket' type goods. As discussed in Chapter 4 there are a variety of characteristics that distinguish health care from other goods and there are a number of failures specific to the health care

market. One of the most significant of these is that supply and demand do not interact in the conventional manner. For example, because of uncertainty and information gaps, the supplier is directly involved in the decision whether or not to consume health care. Traditional demand analysis has, thus, been amended to allow for this difference by considering, for example, the existence of an agency relationship between doctor and patient and of supplier-induced demand. In particular, the concept of need has been used to gain a better understanding of decisions in the health care sector, and it is introduced here in order to show some of the complexities of the theoretical debate about whether or not there is a demand for health care.

Need can be defined in a variety of ways. A common definition is that of normative need, which is assessed by an expert on the basis of a comparison of the actual situation with a technically determined standard. Alternatively, need can be defined as felt need i.e. assessed by the individual. These needs can conflict if the consumer is not the best judge of her/his own needs, as may be the case with respect to health.

Normative need and demand differ in principle. Demand reflects individuals' wants, backed by a willingness to pay for them, and so health care may be (normatively) 'needed' but not 'demanded' (e.g. early treatment for hypertension); or it may be 'demanded' but not (normatively) 'needed' (e.g. cosmetic surgery). In contrast, it can be argued that demands are the expression of felt needs - expressed through a willingness and ability to pay.

The decision to use health care reflects a combination of normative and felt need, because for consumption decisions in the health sector consumers often rely on information provided by the supplier in addition to/supplemented by their own preferences. Relatively few studies of demand for health care have been done but it is also possible to glimpse the demand curve for health care through a study of utilization data. From the limited evidence on the demand for and utilization of health care in developing countries, a number of tentative conclusions can be drawn which are not surprising given the economic theory of demand. The main findings are that:

- higher income status is associated with greater demand for health care and, in particular, for modern health services
- price helps to determine health care demand and price rises may reduce the demand of lower income groups more than that of upper income groups
- poor physical access (increasing the time costs of seeking care) reduces demand
- the perceived efficacy and quality of care are important influences over the decision to demand care from any particular provider.

The relative importance of some of the factors influencing demand may differ from theory. For example, where beliefs about illness causation imply certain treatment options, price and income are much less important in health care utilization decisions.

These beliefs might be incorporated in item T of the demand function ('the organizational and institutional structure of society, preferences and tastes'). In practice, however, it is difficult to account for them and adequately to include them in health care demand analysis. The links between the variables of the health care demand function also make it difficult to isolate the impact of any one. Illness causation beliefs may well be linked to income level (a proxy for socioeconomic status and education); income level is itself linked to ill-health and so both directly and indirectly influences the demand for health care.

7.2 Health care planning and the demand for health care

Despite the theoretical debate about the existence and nature of the demand for health care and practical difficulties of estimating demand functions, it is possible to build up a picture of demand by analyzing information about both utilization and need. The quantity of services required to meet demand can also be determined from such an analysis. An assessment of normative need requires consideration of the likely levels of need within the target population (e.g. based on risk factors). Assessment of felt need is based on direct questionnaire surveys. Utilization data can be drawn from service-based statistics or from household surveys.

Such information about demand is especially relevant to broad planning decisions, for example, concerning health financing options (e.g. whether or not to introduce user fees). However, a more realistic view of health care must be taken for detailed planning. This view should take account of the fact that the consumer's demand for health care is a demand not only for treatment but also for information and support; and that it is partly a reflection of consumer desires, and partly a result of the actions of the consumer's agents (relatives, friends and health professionals).

8. The Supply of Health Care

Three peculiarities of the supply of health care complicate its analysis:

- complexity of product: health care is not a simple product, but a related bundle of goods and services directed towards the same objective of improving health. In many cases these goods and services are joint products and the supply and costs of one are not independent of the others. They also have a quality dimension that is difficult to measure objectively

- complexity of organization: commercial, private non-profit and public institutions are found side by side supplying health care to the same, or overlapping, markets. Little is known of the determinants of the behaviour of non-commercial institutions
- specialized input markets: the key inputs of the health care industry are produced in 'dedicated markets', of which the most important (professional labour) is often tightly controlled by the professionals.

Because of its special features, measuring and modelling supply can be as difficult as measuring and modelling demand. Very little is known about how different types of supplier behave. More attention has been devoted to measuring costs within the health care market and inputs into health care production (see Chapter 7). Even so, only limited cost data is currently available for developing countries, particularly for hospitals and for immunization services.

However, the potential value of cost analyses for planning and management purposes is increasingly recognized. For example, cost data can be used:

- in the assessment of operational efficiency
- to determine the resource requirements for alternative planning options (especially recurrent resource requirements)
- to fuel the debate about alternative financing mechanisms and the necessity of cost-recovery.

Chapter 6

CONCEPTS OF ECONOMIC EFFICIENCY

1. Introduction

A central theme running through much of health sector economic analysis relates to the evaluation of health services. The importance attached to this activity has increased in recent years as economic recession has restricted the level of real resources available to health services. Now more than ever it is necessary to ask such questions as:

- are limited resources used in the best ways possible?
- is value for money achieved in their use?

In order to begin to answer these apparently straightforward questions a number of preliminary issues must first be resolved. In particular, it is necessary to have a clear understanding of:

- what one is trying to achieve i.e. objectives
- what 'the best use of resources' means i.e. evaluative criteria
- how to compare both the amount and the value of resources used in an activity with the amount and value of the output of that process i.e. measurement and valuation techniques.

In studying these issues a number of conceptual and methodological problems emerge. By highlighting them and by clarifying a number of associated concepts, definitions and techniques, the application of economic principles and methods to the planning, implementation and evaluation of health services is discussed.

This chapter covers the basic concepts of economic efficiency. Subsequent chapters consider the measurement and valuation of inputs, the measurement and valuation of outputs, and, finally, the way that these are combined in techniques of economic evaluation.

2. The Production Function and the Health Sector

In order to understand better the process of evaluation, it is important to identify three interrelated, although conceptually distinct, components common to all productive activities. These are : input, process, output.

Inputs are the various resources such as manpower and equipment that are available for use in a productive activity. Process describes the transformation or productive technique which changes inputs into the desired outputs. Output refers to the end result of production. In economics these relationships are usually explored within the framework of the production function.

Production functions are merely a way of representing, both qualitatively and quantitatively, the relationship that exists between inputs and outputs. Within this framework the quantity of a particular type of output (Q) is represented as being determined by the various inputs that go into its production, such as labour (L), and capital (K). The way that these inputs are combined is represented by the factor (T), which assumes a given (fixed) state of technology, or the use of a particular type of production process.

The relationship between input and output is often represented by a production function of the type:

$$Q = f (L,K,T)$$

which is simply mathematical shorthand for saying that output (Q) is a function of (or dependent upon) combinations of labour (L) and capital (K), and the way in which they are combined (T).

Although this conceptual distinction seems clear, in practice, analysis is not quite so straightforward. In the social services in general and in the health sector in particular, a number of factors make the study of productive relationships extremely problematical.

In large part these are due to the complexity of the health sector itself. Health care systems are not simple productive units. They are highly complex, dynamic mechanisms, using a wide range of different resources, in a multiplicity of ways. They produce many different types of output simultaneously, and sometimes even unintentionally. The way in which these outputs are produced is often poorly understood and it is often difficult to relate a particular output to any one intervention or activity.

These are not the only difficulties. A large number of different types of resources are used in health activities. Because of poor information systems it can be difficult to identify, measure or value these inputs. Even greater problems exist on the output side of the equation. For example, a major problem

relates to the confusion surrounding the true output of health services. If it is 'health', then it is necessary to define what one means by that term. Even if this were done satisfactorily, the concept still has to be measured and valued.

These problems relating to the output of the health sector have been explored by many researchers and from many different perspectives. The work remains largely at the conceptual stage and only recently have attempts been made to develop and use measures of health outcome in health services planning and evaluation. There is, however, growing support for the development of such measures and the near future may see their practical use increase. These issues will be explored in Chapter 8.

3. Economic Efficiency and Evaluation

In everyday use, the term efficiency is generally understood as the process of acting with a minimum of expense, waste or effort; or of getting 'value for money'. It certainly involves these things, but for the economist it also has a more precise meaning. In order to appreciate this concept of efficiency it is useful to consider its implications, first within the health sector and then outside it.

Within the health sector, economic efficiency requires that available resources are used in a way that guarantees that the population is as healthy as possible. For this state to be achieved, it is necessary that two separate criteria are simultaneously met. These may be termed effectiveness and efficiency. The first identifies the technical possibilities for transforming inputs into outputs; the second ranks these alternatives by comparing the costs of the inputs with the value of the outputs. In order to avoid the confusion that often surrounds their use in everyday language, it is useful to distinguish between these two terms.

Consider effectiveness first. Usually there will exist many ways of producing a particular type, level or mix of outputs. Some of these, however, will use inappropriate technologies or more resources than are necessary to produce a given level of output. In this sense, effectiveness is concerned with the process of minimizing waste or effort when achieving a given objective. Applying the technical consideration of effectiveness allows one to restrict the set of options that one needs to judge by the second criterion, that of efficiency.

Although a necessary first step, effectiveness is concerned only with the technical relationship that exists between input and output, it is not concerned with issues of value or of cost. Efficiency on the other hand, additionally requires that one takes into account the costs of the resources used in production and then compares them with the value of the output produced.

Exclusive concentration on what goes on within the health sector overlooks two important points. First, health and health services comprise only a tiny part of all those things which are valued by society; and they are not so highly valued that society is prepared to forego all those other things in order to obtain perfect health. Full efficiency requires that society compares the value of health and health services (and their resource costs), with the value (and resource costs) of all those other things. Efficiency in this respect requires that resources are allocated to productive activities in accordance with their value relative to their resource costs. In other words, society must choose how much health will be produced.

Second, health services are only one of the many determinants of health. Again, efficiency requires that one looks wider than health services and considers both the relative effectiveness and efficiency of other ways of bringing about changes in health.

In short, economic efficiency is a way of assessing choices with regard to the value that chosen activities yield against the resource costs they imply. In practice such a comparison may prove difficult and a number of techniques have been developed that attempt to make this process easier, such as cost-effectiveness analysis, cost-benefit analysis and cost-utility analysis.

These techniques of economic evaluation differ considerably in their degree of sophistication and information requirements. Which one is appropriate in a particular situation will depend upon the question being asked and the degree of specificity required in the answer. These techniques are considered at greater length in Chapter 9.

In order to implement one of these evaluative techniques, it is necessary that inputs to and outputs from a health sector activity are qualitatively and/or quantitatively assessed. Although the economist is interested in how the production process effects the change from input to output, production techniques lie within the specialist expertise of the medical profession. Economists therefore concentrate on defining, measuring and valuing both health inputs and health outcomes.

Chapter 7

INPUTS, RESOURCES AND COSTS

1. Introduction

The reasons for studying resources and costs have been explained in Chapter 6. It is also worth noting that most health sectors consume between 3 and 10% of GNP. This represents an enormous resource commitment each year which, under the combined pressures of economic, technological and demographic changes, has, until recently, been rising rapidly in many countries.

The usual reasons for studying costs, such as the need to know the resources necessary for existing and future activities and the need to monitor costs in order to achieve efficiency have, therefore, recently been supplemented by the growing need to control resource use and costs.

This chapter provides a brief overview of the economist's view of costs, how they are used and how they are calculated.

2. Resources

This term refers to the various inputs or factors of production which comprise the raw ingredients of any productive activity. A traditional distinction is drawn between three types of factor: land, labour and capital.

Interest lies not so much in the resources themselves, but in their potential contribution to the production of socially valued outputs (such as goods and services) which satisfy human wants and desires. For all intents and purposes the amount of resources available to satisfy these wants is finite, whereas many of the wants are infinite. It is this problem of scarcity that requires consideration of the costs associated with the use of resources for a particular purpose.

Economists are concerned, therefore, with the way that productive activities can be undertaken with different combinations of resource inputs. Typical problems studied are:

- the effect on output of employing different proportions of human and capital inputs
- the possibility for increasing output through specialization and the division of labour
- the effect on output of increasing the input of some resources while holding other inputs constant.

3. Costs

Most people are familiar with the concept of cost as the sacrifice necessary in order to obtain a good or a service. This is also the way that economists understand the term. The difference between economists and others lies in their understanding of the nature of the concept of sacrifice.

In general, people encounter costs in terms of the monetary prices placed upon goods and services, and so they equate sacrifice with price. This is a valid if rather narrow approach. Economists argue that the sacrifice to be considered is the alternative opportunities that are given up when a choice is made to use a resource or resources in a particular way. Choice is exclusive by nature and the opportunity cost represents the value of the sacrifice involved in it. For example, if a budget of say, US\$1m could be used in order to build and run either a rural health service or a car plant, then the economic (opportunity) cost of the health service is the car plant that has to be given up.

In some cases the money price of a good or service may be a good indicator of this opportunity cost, but there are likely to be instances where it is not. For example:

- in certain situations resources will be used that do not have a money price. Goods and services given freely do, however, have alternative uses to which they can be put, and it is the value of these alternative uses that is the true measure of sacrifice or costs
- certain resources do have a monetary price but, for various reasons, this price does not accurately reflect the value of their alternative productive uses
- certain activities may have 'spillover-effects' on other producers or consumers. These effects, called externalities, may be positive (ie. beneficial), or negative (ie. harmful). Money prices may reflect only the private costs (or benefits for producers and consumers, and not allow for the possible existence of social costs (or benefits)
- money prices are often distorted by transfer payments, the redistribution of income from one group to another. Transfers such as taxation or subsidies do not in themselves constitute an opportunity cost and so should be excluded from its calculation.

4. Cost Concepts and Cost Analysis

Cost analysis can focus on many different issues, two of which are explored here to explain cost concepts and to illustrate the measures used in cost analysis. These issues are the relationship between cost and output and the between cost and the time horizon.

4.1 Cost-output relationships

In studying the relationship that exists between cost and output it is useful to separate the analysis into two component parts. The first looks at the physical relationship that exists between inputs and outputs or what is termed the production function. The second looks at the costs attached to those inputs.

Consider the physical production function first. In its simplest form, production consists of using two inputs to produce some good or service. As one input varies, three alternatives are possible within the quantitative relationship between inputs and outputs. Output can rise proportionately faster than the input, proportionately slower, or at the same rate.

This relationship has been much studied within economics. It is the subject of a famous economic hypothesis (the hypothesis of diminishing marginal returns) which postulates that:

"If increasing amounts of a variable factor are applied to a fixed quantity of another factor, the output per unit of the variable factor will eventually decrease".

Empirical evidence shows that this hypothesis holds true in a wide range of circumstances. The hypothesis, however, only suggests that output per unit of the variable factor will eventually fall. At any particular point, total output may be rising due to the effects of specialization or the division of labour.

The second component part of the relationship between output and cost is the relationship between the resource inputs and price. Although it is frequently assumed that the price of resources will remain constant as output rises, it is possible that they will not. They may fall if it is possible to negotiate discounts through bulk purchases; or they may rise if more needs to be paid to attract scarce resources as the demand for them increases.

The overall relationship between output and cost is determined by the interaction of these two components. The cost of a particular level of output is the result of the quantities of resources used to create it multiplied by their price. This overall cost will vary as either the productivity of the inputs varies, or as their price varies, or as both vary.

A variety of problems are typically studied by economists: How do costs vary with the level of output? What level of resources is required to achieve different levels of output? How do the costs of each unit of output vary with the scale of activity? How many resources are associated with a small incremental change in the level of activity? In order to facilitate such analysis, three measures of cost are useful:

- total cost (TC) is a measure of all the costs entailed in producing a given level of output. It is a measure of the aggregate resource requirements of a particular scale of activity and is derived by summing all the costs incurred during production. An example would be the cost of providing an immunization service.
- average cost (AC) is a measure of the total costs of production associated with each unit of output. Average costs indicate the resource requirements for each unit of output and are calculated by dividing total costs (TC) by the number of units of output (Q) ($AC = TC/Q$). An example would be the cost per immunization provided.
- marginal cost (MC) is a measure of the resources associated with a small incremental change in output. Most economic decisions are not about whether to produce all or nothing, but concern small changes in the existing scale of activity. Marginal cost is a measure of the change in costs associated with increasing or decreasing output by one, and is derived by calculating the change in total costs for that one unit. In effect, the marginal cost concept is an attempt to measure the rate of change of costs as output changes. Without a large number of observations and the use of statistical techniques it is difficult to calculate the change in costs for one unit. Therefore, the change in costs for several units is often calculated, and a measure of the incremental unit cost of that larger variation (obtained by dividing incremental costs by incremental units of output) is used instead of marginal cost. An example would be the incremental cost of adding an additional vaccine to the existing service, or of extending the service to an additional village.

The relationship between these three measures can cause some confusion, although in practice it is quite straightforward. If output is increased, say by one unit, total costs will rise by the marginal cost of that unit. How will this affect average costs? If the marginal cost of this extra unit is the same as the previous average cost then the new average cost will be unchanged. However, if the marginal cost of the extra unit is smaller (larger) than the average cost, the new average cost will fall (rise).

These relationships between the different cost measurements can be visualized best in average and marginal cost curves which show how average and marginal costs change as output changes (Figure 17). The average cost curve is hypothesized to be U-shaped, being dominated at low outputs by the

effect of the spreading of fixed costs, and at high outputs by the effect of diminishing returns. The marginal cost curve passes as a matter of mathematical necessity through the minimum point of the AC curve, with MC less than AC at lower outputs and greater than AC at higher outputs.

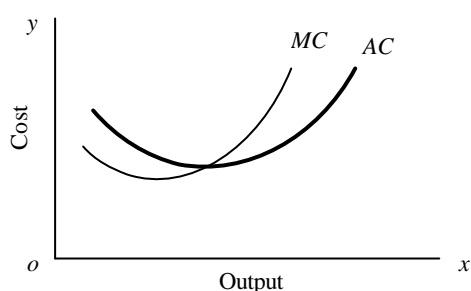


Figure 17

4.2 Costs and the time horizon - fixed and variable costs

One of the purposes of studying costs is to assist decision-making and so it is important to identify those areas over which choice can be exercised. In practice, some resources will have been committed to particular activities by past decisions, and likewise, present decisions will have resource implications well into the future. It is often useful, therefore, to draw a distinction between those resources that are easily varied, and those that are relatively fixed. This classification is crucially dependent upon the time period that is adopted, because in the very long-run, most things can be varied.

Although in practice decisions will be made over a time continuum, economists find it useful to distinguish between a number of discrete decision periods. The two most commonly used periods are the short-run and the long-run.

In the short-run it is assumed that certain resources are fixed and cannot be transferred from their present use. The only way that output can be varied during this period is by changing the amount of variable factors. This is different from the long-run decision period in which all factors are assumed to be variable. The actual time period corresponding to these periods will vary according to how quickly the quantity of fixed factors can be changed. This 'fixedness' can be demonstrated to have important consequences for the output/cost relationship.

In the short-run, with a given level of fixed resource inputs, there is a level of variable resources for which cost per unit of output produced is minimized. Because some inputs are fixed, variations in short-run output can only be achieved by changing the variable resources and thereby altering the proportions of those factors. In doing so there is a move away from the short-run optimum. Short-run variations in output can only be achieved at the expense of increases in unit costs (i.e. the law of diminishing returns).

In the long-run, however, all inputs can be varied. For each level of variable inputs it is also possible to use a level of fixed resources which represents the optimal combination necessary to achieve the lowest unit cost for that level of output. Since in the long-run and the short-run there will be points at which the level of fixed resource inputs coincide, the long-run cost curve can be thought of as the sum of optimal points for all short-run cost curves.

It is possible to study the relationship between input and output as the scale of production varies. When inputs are increased in the same proportions it is possible for output to increase at the same rate (constant returns to scale), at a lower rate (diseconomies of scale), or at a higher rate (economies of scale). Factors such as specialization and the division of labour may account for increasing returns at lower levels of activity. Thereafter, as scale increases, managerial inefficiency and the problems of large scale production are seen as eventually leading to diseconomies of scale.

5. Components of Cost

Some terms have been used above which it is important to define clearly.

5.1 Fixed and variable costs

Fixed costs refer to the costs that are necessarily incurred in setting up a productive activity. Often known as 'overheads', fixed costs are born irrespective of how much output is produced. Variable costs, conversely, are those costs that are directly related to how much output is produced. As most factors of production can be moved to another activity in the long term, fixed costs are usually defined in relation to the time period under consideration (e.g. the financial year).

5.2 Capital and recurrent costs

Capital costs are usually defined as costs for items with a life of more than one year. Typical examples include costs for the:

- construction of buildings
- purchase of equipment
- basic manpower training.

Recurrent costs are the costs that are necessarily incurred each year. These include items such as:

- salaries and wages

- supplies (e.g. drugs, petrol)
- electricity, water
- in-service training.

Most capital items will fall into the category of fixed costs.

6. Comparing Costs

In order to compare costs between institutions, programmes or countries, various adjustments must be made to cost measures. Two adjustments considered here are making capital costs comparable with recurrent costs, and enabling comparisons to be made over time and between countries.

6.1 Comparing capital costs with recurrent costs

The distinction between capital and recurrent costs is important. Capital costs are typically borne in the year the items are acquired and yet those items will provide a service after that time. In order to provide a comprehensive picture of resource use (as opposed to the picture of the timing of expenditures required for budgets), capital costs need to be made comparable with recurrent costs. This can be done by translating the cost of land, buildings and equipment into an equivalent annual cost that takes into account both the expected life of the capital item and a rate of interest (to reflect the opportunity cost of money used for capital purchases that could be used for other purposes).

The simplest way of calculating the equivalent cost of a capital item that yields services over more than one year is to use published tables of present worth factors which provide ready-made calculations for different rates of interest. The tables consist of a matrix with years of life on one axis and interest rates on the other. The initial capital cost is divided by the present worth factor for the number of years the item will be used and the chosen interest rate.

If the cost of tying up funds in capital goods is ignored then the method of straight-line depreciation could be employed to obtain an annual value for capital cost. This simply involves dividing the initial cost of the item by the number of years of expected life.

6.2 Longitudinal and cross-sectional cost comparisons

In many instances it is useful to make a comparison of costs at two different times, or even between two different countries. This is difficult for a number of reasons. Comparing costs over time has to allow for the effects of inflation and changes in relative prices; whereas comparisons across countries

have to adjust for different currencies. In order to understand better the issues involved in such comparisons it is useful to repeat the economist's understanding of cost.

Used by economists, the term 'cost' refers to the sacrifice made in order to obtain something else; that is, its opportunity cost. This is an attempt to measure the 'real' rather than the 'nominal' (or monetary) cost attached to particular activities. In comparing costs over time, inflation disguises the real changes in costs and purchasing power by changing the real value of money. It is possible to compensate for these inflationary changes through the use of price indices or deflators. These attempt to reveal true changes in costs over time by controlling for changes in overall or relative prices, or both. Most governments calculate a number of price indices which can be used to determine changes in real resource use rather than merely changes in price.

For example, the Gross Domestic Product Deflator provides an index of the change in price of the resources produced in an economy each year. A more familiar price index is the Retail Price Index which identifies (for a representative sample of goods and services consumed by a household) the annual changes in nominal prices, and indirectly, the real resources that it is possible to obtain with a given budget.

A persistent problem in calculating changes in real resources within the health sector is that of devising a price index that is representative of the diverse range of goods and services that it purchases.

Similar considerations apply to making comparisons of costs across geographical regions. One problem is that the currency differs, making direct comparison of resource use impossible. Exchange rates may provide a highly imperfect means of comparison because of under-/over-valuation of the domestic currency.

A related problem is that the internal price structure of countries may differ in such a way as to change the relative values of different resources. This particularly affects the price of inputs such as manpower which are not easily traded internationally. For example, a given expenditure in dollar terms on physicians' services may purchase a very different quantity of services in different countries, whereas a given expenditure on drugs may purchase similar quantities.

Even determining the overall level of resources available across countries has proved difficult. In developing countries a high proportion of economic activity may not be adequately reflected in figures for national income or expenditure because it occurs within the informal sector of the economy (e.g. subsistence farming). This will distort the calculation of real resource availability and render international comparison difficult.

Economists have attempted to overcome these difficulties surrounding the valuation of resources and international comparisons by calculating purchasing power parities which reflect the real resources available to countries and to sectors - such as the health sector - within them. Considerable conceptual and methodological problems exist in developing and using these however, and problems of currency conversion remain a persistent problem in international cost comparisons (see Chapter 9 for other approaches to obtaining 'true' measures of opportunity cost).

7. Costing Methods

There are two general approaches to collecting and using cost data in the health sector.

7.1 The direct accounting approach

This approach focuses on the costs directly associated with a particular activity. It requires the identification and measurement of the costs incurred in providing a particular intervention or in treating a particular patient. In other words, it attempts to calculate the costs of the resources that are actually used in an activity or those that are consumed by a particular patient. In practice this often proves difficult because it is difficult to identify the resources used in many activities and to calculate what proportion of the costs of shared items or facilities should be apportioned to individual activities.

Typical examples of the direct accounting approach include:

- activity costing - detailed costing of the activities of different cost centres, such as whole hospitals, speciality departments, or health centres
- disease costing - attempts to identify the cost incurred in treating particular diseases, or particular types of patient, or even in carrying out certain types of procedure.

The direct accounting approach is usually the basis for calculating costs in economic evaluations (e.g. cost-effectiveness and cost-benefit studies).

7.2 The statistical approach

This approach focuses on the costs associated with types of activity or types of patient and not on the costs associated with individual activities or individual patients. Rather than addressing the issue of how much something costs and how these costs are incurred, the statistical approach addresses the issues of why costs differ, and by how much.

The technique used within the statistical approach is to estimate the qualitative and quantitative impact on costs of differences between health units or facilities. Multiple regression analysis is frequently used in order to compare cost profiles across a large range of different units. The aim is to eliminate the possible effects on costs of all factors except the one under study. This allows a simulated comparison of like with like, where it would otherwise be impossible. The remaining cost difference between health units/facilities is arguably a measure of the cost implication of the variable under study.

The statistical approach can be illustrated by the following example. In studying, say, the costs associated with providing hospital care, it is very likely that there will be wide variations in cost between different hospitals. Why should this be the case? Some hospitals may be more effective or more efficient than others, or there may be other reasons why these costs vary. Hospitals may differ with respect to:

- size (number of beds)
- throughput (the number of cases treated in a given time period)
- case-mix (the combination of diagnoses admitted)
- case-severity (more/less severe cases admitted)
- quality of care;
- type of treatments offered
- teaching or research activities
- age and/or location of hospital facilities
- occupancy levels
- manpower availability
- length of stay.

One or more of these factors may account for differences in measured costs between hospitals. In studying these cost variations it is necessary to ask what factors are likely to influence costs, and then try to isolate the effect of those factors using statistical techniques such as multiple regression analysis. For instance, it might be interesting to know whether the cost per patient differs according to whether the hospital is large or small. By controlling in the analysis for other factors that affect cost, the effect on cost of size alone can be examined.

Chapter 8

OUTPUTS, HEALTH AND HEALTH INDICATORS

1. Introduction: The Value of Health Indicators

In Chapter 6, the need for measures of health outcome both in physical and in value terms was noted. These are required because it is necessary to compare the magnitude and value of inputs and outputs in order to evaluate economic efficiency. Why though should this present a problem in the health sector, when many other fields of human activity seem to require no such explicit measures?

The answer lies partly in the fact that in most other areas of production or service provision, it is possible to measure and value output directly in money terms by market prices. It is also possible to infer some measure of efficiency through profits earned and the behaviour of the industry. However, in many of the 'social services', and in the health sector in particular, prices have often been reduced or eliminated, so that profit has ceased to be a useful measure of efficiency or of effectiveness. Furthermore, certain characteristics of the market for health and health services and the institutional characteristics of providers may reduce pressures for efficiency.

Perhaps a more obvious problem is the relative difficulty of identifying just what the output of a health system is. By comparison it is easy to say what the output of, say, a car factory is. These difficulties have meant that historically, the evaluation of health systems has focused on aspects of performance that are relatively easy to measure and define. Measures of service provision (e.g. beds), or activity (e.g. number of patients), or other inputs have frequently been used as if they were measures of outcome.

If it is accepted that the primary purpose of health services is to produce health, it is necessary that one clearly defines what one means by 'health'. This question has been receiving a great deal of research attention in recent years. Amidst the wide variety of approaches, it is possible to discern a number of central themes. Most approaches to the definition of health recognize that it is not merely the absence of disease or infirmity but a generic term seeking to encompass all dimensions of an individual's, group's or population's level of well-being. Likewise most definitions accept that health is not an either/or state, but rather a continuum: at the lower boundary one might place the state of death; and at the upper boundary a state of perfect health, or, perhaps more realistically, some statistically derived norm. All definitions accept that health is a multi-dimensional concept and not merely some state of biological dysfunction, although this acceptance is not necessarily reflected in measures of health used in practice.

However, from this broad level of consensus, opinions differ on a number of crucial points. In particular, they differ over what elements of well-being should be included in the definition, how they should be measured, and how and whether or not they should be valued and incorporated into a single unitary index of health status.

A number of techniques have been developed that attempt to use some of these conceptual approaches to health status measurement. A necessarily brief overview of these is given below.

2. Measures of Mortality and Morbidity

These have been the most frequently used measures of health. The first, mortality, is simply an attempt to measure the quantity of life; whereas morbidity is an attempt to measure the quality or healthfulness of life. Both indicators suffer from serious deficiencies.

Mortality, despite its relative lack of ambiguity, is a highly insensitive measure of health and gives little indication of health dimensions (such as disability) that do not result in death. Morbidity, although a potentially more sensitive measure also has its deficiencies. Whilst it is a measure (if very imperfect) of the quality of life, morbidity studies have often relied on individuals recognizing a health problem and seeking treatment. In practice much illness is not reported and is not therefore recorded in morbidity statistics. In addition it is difficult to make comparisons between a day of illness for different types of morbidity.

3. Measures of Function and Dysfunction

This approach to health definition and measurement recognizes that there is a social role to health and illness as well as a biological one. Health is therefore measured indirectly by its effects on the ability of individuals to perform their normal social roles. For example, different health problems have been studied with regard to the number of days of work that are lost because of them. This provides a useful way of measuring actual health impact, and is a considerable advance on simple measures of resource or service provision. Alternative measures look at the ability to perform everyday activities such as climbing stairs, dressing oneself or shopping.

4. Health profiles

The general feature of the health profile approach is the identification of a number of dimensions of health with a subsequent scaling of an individual or population along each dimension. As the name

implies, the approach provides a profile of certain selected dimensions of health at one time. Repeated measurements make it possible to compare health profiles over time; and arguably, to assess the relative efficacy of particular interventions. A characteristic of the health profile is that it is possible to tailor the dimensions of health measured to suit the particular health problem under study.

5. Global Health Indicators

A common characteristic of all the approaches previously discussed is that they attempt to describe and measure health along one dimension of health at a time without providing any basis for the comparison or ranking of one dimension of health with another. This has been a conscious and intended decision of some writers who argue that any relative valuation or ranking of different dimensions of health is necessarily highly subjective and best avoided.

This view has been challenged by those who argue that it is necessary to rank different dimensions of health relative to each other, and thus to be able to value different levels of health in a single index. They argue that the health profile approach is useful when studying changes in one dimension of health at a time; or to a lesser degree, where more than one dimension changes at the same time, but where the changes are all in the same direction. It is of no use they argue, where one is required to assess changes where one dimension of health is traded-off at the expense of another.

A number of different approaches to the global health indicator have been developed. Although they differ in some respects the common features can be identified as:

- the selection of the dimensions of health to be included in the index
- a classification of health states within these descriptive categories
- the derivation of a scale allowing the relative valuation of the above states.

Typically, different health states will be described in terms of such dimensions of health as pain and distress, or discomfort and disability. These dimensions will then be incorporated together to provide one index along which all feasible health outcomes can be valued relative to each other.

The value of the global indicator approach is that, in theory at least, it allows direct comparison and ranking of options that produce different types of health improvement, or involve different types of interventions. It is the ideal type of indicator for cost-utility or cost-benefit evaluation (see Chapter 9).

6. The Quality Adjusted Life Year (QALY)

A problem common to all of these approaches is that they focus on a measure of health status at a particular time. Health is not a static concept and in practical terms one is also concerned with how health changes over time and with life expectancy. One method that attempts to include the duration as well as the quality of health output is the QALY, or Quality Adjusted Life Year, where the number of years of life gained from an intervention is adjusted by a measure of their quality. This is particularly relevant to the evaluation of treatment for chronic diseases (such as kidney dialysis) where treatment enables survival for a certain period of time at a less than perfect state of health.

Central to this approach is the assumption that it is possible, even necessary, to make comparisons and 'trade-offs' between the quality and the quantity of life. The first part of the QALY approach is essentially the same as that for the global health indicator approach. However, having derived a single index, it is then necessary to provide a weighting for the health index relative to the duration of the health state. This allows assessment of the relative desirability of choices between being relatively healthy for a short period of time, or being in poorer health for a longer period. As this is the sort of choice that needs to be made in the real world the value of the approach is clear. At present, studies treat each additional year of life gained as being of equal value regardless of the age at which it is gained, although this weighting could of course be changed.

Recently the QALY has been used to compare the desirability of different health intervention strategies. The resource costs of different types of intervention are calculated and are then expressed in terms of how many quality adjusted life years they would produce. This makes it possible to identify the relative outputs of different strategies and to determine which procedures represent good value for money in terms of cost per QALY.

One rough form of QALY is the measure Healthy Days of Life Lost, which has been used with Ghanaian data to compare the relative impact of different diseases. The measure combines information on the incidence rate, case- fatality rate and the extent and duration of disability produced by a disease. It has also been modified to allow for time preference (discounted years of healthy life) and for age preference (discounted productive years of healthy life).

Chapter 9

THE TECHNIQUES OF ECONOMIC EVALUATION

1. The Definition and Forms of Economic Evaluation

Economic evaluation studies generally seek to address the two questions posed in Chapter 6: Are limited resources used in the best ways possible? Is value for money achieved in their use? The techniques of economic evaluation are, therefore, an important contribution to the methods of health service evaluation and are based on the economist's concern with economic efficiency and opportunity costs.

Economic evaluation has been defined as:

"the quantitative analysis of the relative desirability to the whole community of investing in alternative projects or programmes",

where desirability is assessed in terms of both costs and consequences. 'Consequences' is used here as the generic term for the beneficial results of a programme (often termed effects or benefits, depending on the techniques of analysis being used). Within this broad definition there are many forms of economic evaluation, as shown in Figure 18. Only those forms which examine both costs and consequences for two or more alternatives fit the above definition and can be described as full economic evaluation studies. In practice, one of the two alternatives examined may be an existing project/programme (the 'do-nothing' alternative).

Cost-minimisation analysis (CMA) is based on prior epidemiological findings which show that the outcome of interest (e.g. reduction of disability) is achieved to the same degree by two (or more) interventions. The technique is used to identify the least cost intervention.

Cost-effectiveness analysis (CEA) investigates the best way of achieving a single objective by comparing effects and costs. It evaluates either:

- which of a number of possible interventions will achieve a given health objective at least cost,
- or
- given a fixed budget, the intervention that maximises the effectiveness of the expenditure.

Its results are expressed either as costs per unit of output (total costs of the intervention divided by total health effect) or as effect per monetary unit (total health effect divided by total available resources).

Cost-benefit analysis (CBA) values both costs and benefits in monetary terms, and compares them, assessing whether the project/programme is desirable through the use of decision criteria (e.g. if the benefit cost ratio (benefits divided by costs) is greater than one, the project/programme is worthwhile).

Cost-utility analysis (CUA) is a form of CEA but it measures the effects of a project/programme in terms of utilities (the quality-adjusted health outcome caused or averted). Like CEA it can focus on either minimizing cost or maximizing effect; and its results are expressed, for example, in terms of costs per Quality Adjusted Life Year (QALY) or QALYs per monetary unit.

Figure 18: Forms of Economic Evaluation

		Are both costs and consequences examined?		
Is there a comparison of 2 or more alternatives?	No	Examines consequences only	Examines costs only	Cost-outcome description
		Outcome description	Cost description	
	Yes	Effectiveness evaluation	Cost analysis	<ol style="list-style-type: none"> 1. Cost minimization analysis 2. Cost effectiveness analysis 3. Cost benefit analysis 4. Cost utility analysis

Adapted from: McMaster University Health Sciences Centre (1984) How to read clinical journals: VII. To understand an economic evaluation (part A) Canadian Medical Association Journal Vol 130 June 1 p.1430.

The root differences between these techniques concern their evaluation of health outcome (consequence) and their breadth of analysis. CMA and CEA tacitly assume that the health objectives which the projects/programmes serve are worthwhile. CUA permits choice between a much wider range of interventions but still ultimately assumes that at some cut-off point of cost per QALY, a programme is worthwhile. CBA in theory permits assessment of whether the health objectives are worth achieving in the first place.

2. The Steps of Economic Evaluation

All economic evaluation techniques involve three basic steps:

- identification of costs and consequences
- measurement of costs and consequences
- valuation of costs or of costs and consequences.

In addition, all economic evaluation studies should consider adjusting costs and consequences for differential timing, and should incorporate an incremental and a sensitivity analysis.

2.1 Identification of costs and consequences

Table 2 establishes the types of costs and consequences relevant to economic evaluation in the health sector. Although it may be impossible to measure and value all of these items it is important that they are, at least, identified, in order to clarify any possible analytical bias in favour of the most easily measurable items.

It may appear obvious that costs and consequences for both the health sector and for patients should be considered, but many studies confine their attention to the health sector only. This approach limits analysis and is contrary to the definition of economic evaluation, which stresses that the relative desirability of alternatives to the whole community should be the focus of evaluation. It can bias study results against certain population groups or interventions.

Table 2: Types of costs and consequences relevant to the economic evaluation of health care projects/programmes

<u>Costs</u>		
1.	Organizing and operating costs within the health care sector (e.g. health care professionals' time, supplies, equipment, power and capital costs)	} } } } direct
2.	Costs borne by patients and their families: out-of-pocket expenses patient and family inputs into treatment time lost from work psychic costs	} costs } } } indirect } costs
3.	Costs borne externally to the health care sector, patients and their families	

Consequences

1. Changes in physical, social or emotional functioning (*effects*)
2. Changes in resources use (*benefits*)
for organizing and operating services within the health care sector:
 - for the original condition } direct
 - for unrelated conditions } benefits
relating to activities of patients and their families:
 - savings in expenditure or leisure time } direct benefits
 - savings in lost work time } indirect benefits
3. Changes in the quality of life of patients and their families (*utility*)

Source: Drummond MF and Stoddart GL (1985) Principles of Economic Evaluation of Health Programmes World Health Statistics Quarterly 38(4): p 360

For example, consider an evaluation of immunization services in urban and rural areas. By focusing on health sector costs and consequences and including 'cost savings to the health sector', the evaluation would probably favour the concentration of immunization services in towns. Health services are more developed in urban than in rural areas and so the potential savings to the health sector from preventing illness are greater. However, in under-provided rural areas, the local provision of immunization services would generate savings for the patients in the form of lower transport costs in obtaining treatment and less private treatment and self-care. These should be considered within the evaluation and may lessen its bias towards urban populations. It is the concern to minimize government (rather than social) costs that often determines the conduct of economic evaluation studies and may be used (wrongly) to justify an exclusive focus on health sector costs/consequences.

The third cost item, costs outside the health sector, is much less obvious and may not be relevant to all studies. Its inclusion stresses the need to adopt a community-wide perspective to clarify the full implication of any project/programme. For example, an industrial safety measure may alter the cost or availability of products if it results in more costly production processes.

It is sometimes difficult to distinguish between the costs of an intervention and the changes in resource use that result from it: Which is the cost and which the consequence? The clearest way of seeing the difference is to remember that the consequences in question are often termed 'benefits', that is they reflect some positive change as a result of the intervention. If a result of an intervention is time saved (less time lost from work) it is, therefore, a benefit; but if time costs are increased, perhaps because of lengthy treatment (more time lost from work), the increase is a cost.

The differences between the consequences identified in the table help to distinguish the differences between CEA, focusing on effects, CBA, focusing on benefits, and CUA, focusing on utilities.

2.2 Measurement of costs and consequences

The measurement of costs and consequences in developing countries is undermined by the lack of routine statistical data. Expenditure data may be considerably out of date and even when it is available is unlikely to be fully accurate or broken down into appropriate expenditure categories (e.g. disaggregated by institution).

It is better to base project/programme costing on detailed consideration of the actual physical inputs used. This would also allow appropriate division of the costs shared by many services; for example, the World Health Organization Expanded Programme of Immunization costing manual recommends that the immunization share of such joint costs (for instance supervision and administration) be determined on the basis of staff time allocations to immunization relative to the other services sharing the resources. However, identification of all the physical inputs used by a programme could be a costly process in itself. It may be most appropriate to focus such detailed consideration on those inputs likely to represent the major portion of costs and/or which can be used to split joint costs (e.g. staff time).

Measurement of consequences in terms of health status effects is even more difficult, given the lack of routine epidemiological data in developing countries and the weaknesses of the epidemiological evidence concerning the links between health care inputs and health status outcomes. As a result of the many influences on health it is extremely difficult to disentangle the health impact of a health intervention from the influence of other factors (such as improved education, better water supplies etc.). Moreover, a service delivery package may be too complex to allow appropriate measurement of effects.

For CEA, the problem is made worse by the need to choose an appropriate single unit of effect. Chapter 8 outlined the measures of health effect that may be used, and stressed that because health is multi-dimensional it is difficult to measure it along one dimension only. For example, the unit of 'death prevented' cannot measure the impact on disability that may be an important health consequence. It is also only projects/programmes that have the same objective in terms of health status improvement whose consequences can be measured using the same unit of effect. 'Deaths prevented' cannot be used to compare, for example, a malaria control programme which prevents many infant deaths and many episodes of acute illness, and a schistosomiasis control programme which prevents widespread chronic disability and relatively few adult deaths.

Using a single unit of effect necessarily ignores the non health status consequences of programmes, such as the time-saving effects of water supplies. Ignoring these wider consequences undervalues those interventions such as water supply programmes which have these broader effects. In general,

primary health care programmes have broad objectives and potential effects that cannot be measured by a single unit of effect, including self-reliance and increased community participation.

Two approaches have been developed to address some aspects of this problem within CEA. The first involves the use of global health indicators and, in particular, quality adjusted life years and healthy days of life lost (see Chapter 8). These measures focus on changes in the quality of life of patients and their families, permitting consideration of a spectrum of health effects, and are used in cost utility analysis. Data collection for both is very complex and, in practice, QALYs have never been used in an economic evaluation study based in developing countries. They require sophisticated techniques for assessing and measuring community preferences for the different qualities of life associated with different states of illness, and seem less relevant to the largely acute diseases generally found in developing countries. Measurement of 'healthy days of life lost' requires relevant epidemiological data and most developing country health information systems cannot yet provide them in sufficient detail. This measure also introduces a bias towards interventions that tackle child health problems as children have most healthy days of life to lose, unless some weighting system is introduced.

The other approach that has been adopted within CEA to overcome the problem of choosing an appropriate unit of effect, is the use of units of intermediate output (e.g. numbers of immunized children, proportion of pregnant women attending an antenatal clinic). These are measures of access, coverage or compliance. Where such units are used it is clearly not changes in physical, social or emotional functioning that are evaluated but rather the success of projects/programmes in extending service provision. The evaluation does not assess whether the specific service is worthwhile but rather seeks to identify the best way of providing a service that has previously been identified as worthwhile. The approach may seem like a weak version of economic evaluation as earlier defined, but it is technically more powerful. It is much less open to criticism about the assumptions embodied in the evaluation (related to identification of the project/programme's priority health effect and to the nature of health effects). The link between the input of a health service and the output of increased coverage may be much stronger than the link between a health programme and health status changes. It also provides a useful focus on the means of service delivery and on operational efficiency. However, the appropriateness of the approach clearly depends on the study objectives.

2.3 Valuation of costs and consequences

Chapter 7 introduced the concept of opportunity cost as the economic definition of the sacrifice necessary to obtain a good or service. Although monetary prices may reflect this opportunity cost, there are instances where there is no monetary price or where this price does not accurately reflect the value of the alternative productive use of resources (the economic price). The prices may also not reflect the social costs of activities (the social price). To allow for these problems economists use the techniques of shadow pricing in order to value costs appropriately.

For example, many health projects/programmes use imported goods (e.g. drugs) that are purchased in international markets using foreign exchange. However, the purchase price of these goods will not reflect the opportunity cost of the resources (foreign exchange) used for their purchase where countries have over-valued exchange rates and foreign exchange shortages. Many developing countries suffer this problem and their purchase of drugs with scarce foreign exchange will prevent their purchase of other products that must be bought internationally. Drug costs must, therefore, be valued using a 'shadow foreign exchange rate' to reflect the true opportunity cost of the purchase - raising their value above the actual purchase price.

Alternatively, distortions within the domestic economy, such as tariffs, subsidies, import and export taxes, may shelter it from international competition and exaggerate the opportunity cost of goods produced domestically. It might be possible to purchase the same goods from abroad at lower real cost by using domestic resources to produce exports and exchanging them for the foreign products. World prices may be a better reflection of opportunity costs than domestic prices and can be used to shadow price the domestic production. Non-traded goods can often be valued roughly at equivalent world prices using a standard conversion factor.

Shadow pricing can also be used to establish social prices, allowing for a country's preference for savings versus consumption and/or for income re-distribution. In either case, weights are used to reflect the national preference and, in effect, to bias the choice of projects towards it (e.g. towards the reduction of income inequality, by over-valuing the benefits gained by low income groups, and under-valuing those of high income groups).

Valuing the consequences of health projects/programmes is even more complex than valuing costs. Although CBA does not suffer from the problems of CEA in choosing an appropriate unit of effect, it does suffer from the problems of valuing a diverse range of consequences that includes changes in health status, savings in public and private expenditure on treatment, and economic returns from the exploitation of natural resources (especially land) permitted by programmes such as malaria control. The greatest problem is the valuation of health status improvements in monetary terms. Two main approaches are generally suggested for these valuations: the human capital approach and the willingness to pay approach.

The first is based on the assumption that health projects/programmes represent an investment in people which enables them to be more productive and to increase their material well-being. The benefits to be considered, therefore, include an increase in the availability of labour for productive work and an increase in the productivity of labour whilst at work. The value of any increase in the supply of labour should be based on its marginal product, for which the wage rate is often used as a proxy. However, if the health improvements occur in the subsistence agriculture sector where

under-employment may exist, the wage rate may overstate the value of additional units of labour. A 'shadow wage' is, therefore, often used. (The same argument applies to valuing unskilled labour employed in the project/programme).

The willingness to pay approach starts from the assumption that the price consumers pay for a good represents its value to them. Therefore, the amount they would be willing to pay for improved health or reduced likelihood of death represents an appropriate valuation of it.

Both approaches are controversial but they are the first steps in the full valuation of health benefits. The first has been the most extensively used in developing countries, mainly because estimates are feasible. In contrast, it is unclear how willingness-to-pay valuations can be obtained in the absence of a market price (as with health care).

2.4 Adjusting costs and consequences for differential timing

Costs and consequences often occur at different times and health projects/programmes may also differ with respect to the point at which consequences occur. Preventive programmes, for example, may have a delayed impact, unlike curative programmes. Allowance for these differences can be made through the technique of discounting which assumes (on a variety of grounds) that people prefer benefits which occur sooner rather than later. Costs and consequences are discounted to present values by use of a discount rate and greater weight is then given to costs and consequences that arise sooner rather than later. In effect, the procedure is the reverse of applying a rate of interest.

2.5 Incremental analysis

Whatever technique of economic evaluation is used it is important to undertake an incremental analysis i.e. to consider marginal costs. As noted in Chapter 7, most economic decisions are not about whether to produce all or nothing, but rather concern changes in the scale of activities. Should an additional vaccine be added to an existing immunization programme? Should the health infrastructure be expanded? In order to make these decisions it is essential that the additional costs be evaluated separately from the total and average costs of the existing programme, and not be assumed to be identical.

An evaluation of the cost-effectiveness of screening for cancer of the colon in the USA showed the importance of incremental analysis. The average cost per case of cancer detected for the advised protocol of six tests was US\$2,500 (the total costs of six tests divided by the total number of cases detected). However, the incremental cost per case detected of the sixth test (having done five) was US\$47 million (the cost of the sixth test divided by the number of cases it detected), since the cost of the test remained the same but by the sixth test almost all cases of cancer had already been detected.

2.6 Sensitivity analysis

Given the range of assumptions inherent in most economic evaluation studies it is essential to test the sensitivity of the study results to changes in these assumptions. What impact do different valuation of costs have on the result? or different rates of discount? or different benefit valuations? or different timing?

Sensitivity analysis also allows clarification of the level of detail required in a study. Some studies may generate fairly reliable results using rough data (i.e. they are insensitive to changes in the assumptions); others may require more accurate estimation of the variables.

3. The Problems of Economic Evaluation

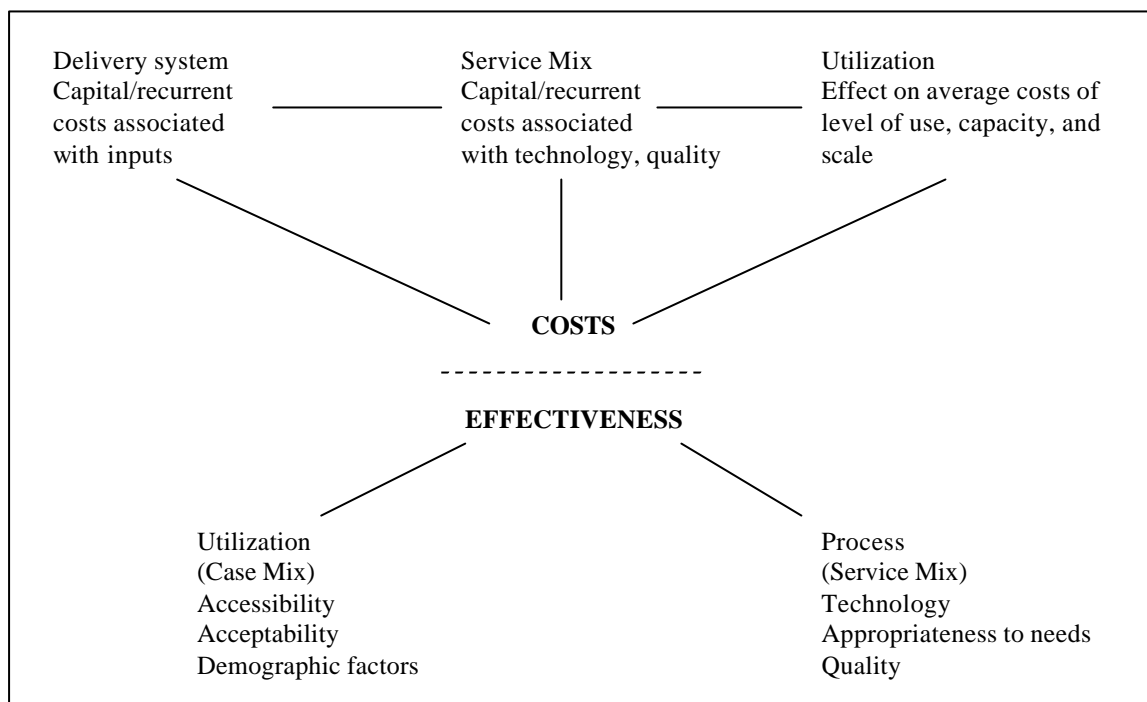
In addition to the technical problems of economic evaluation, concerning the lack of information and the shortcomings of units of effect, other criticisms of the techniques have arisen from the current debate about the nature of primary health care. Selective primary health care strategies provide those services which have been identified as the most cost-effective using crude CEA techniques (e.g. UNICEF's GOBI-FFF). However, the use of CEA in this way has been criticized on the grounds that 'efficient is not sufficient' as the sole criterion for decision-making; and that economic evaluation techniques have been used inappropriately to impose the value judgements of specialists on the community as a whole.

The argument that 'efficient is not sufficient' is based on the problems of choosing a single unit of effect to measure health status and to evaluate complex medical/health care programmes. In order to compare alternative health interventions using cost effectiveness ratios it is important that all the sources of variation in the indicator are considered. Figure 19 sets out the factors that affect this ratio. To assume, for example, on the basis of an experimental study, that the intervention with the lowest cost per death prevented will permit more deaths prevented for the available resources than other interventions is also to assume that the intervention will actually be used by the population. However, utilization is dependent both on acceptability and on accessibility, i.e. on the nature of the delivery strategy, which is often not considered in an experimental situation (for similar reasons it is difficult to transfer CEA findings from one country to another).

For CEA it is also important that the alternatives under examination are similar in all respects. For example, the cost-effectiveness ratio is influenced by the number of people covered by a project/programme and a lower ratio may simply reflect a larger catchment population rather than

effectiveness. Thus, where two interventions have the same cost- effectiveness ratio but serve different populations it is not clear which is the most desirable.

Figure 19: Factors affecting the cost-effectiveness ratio



Source: Berman P (1982) Selective Primary Health Care: Is Efficient Sufficient? Social Science and Medicine 16, p1056.

The steps of economic evaluation discussed in section 2.2 indicate that the generation of even a simple cost effectiveness ratio requires many value judgements - concerning the appropriate objectives of health programmes, the appropriate valuation of costs and consequences, the time preferences of the community etc. Often these judgements are made by specialists with no consideration of community preferences and they are rarely identified clearly in the studies undertaken. In practice, economic evaluation techniques may be used as a black-box into which assumptions are fed and from which the basis of resource use decisions are produced. The resulting decisions may be made without consideration of community views and so in contradiction of one of the basic tenets of primary health care. It is vital that assumptions in studies be clarified, and that study results are not quoted in isolation from their context.

4. The Strengths of Economic Evaluation

Against these criticisms can be set the reality of scarce resources, the need to assess whether value for money is being obtained through their use and the exclusive nature of choice. No person can be in two

places at the same time, and any use of resources denies their use elsewhere. It is, therefore, important to obtain the best value for money from the use of scarce resources.

Economic evaluation techniques provide a framework within which to consider these issues. They are adaptable and can begin to account for many concerns - the nature, magnitude and timing of cost and consequence, consideration of community preferences etc.

In particular, they have an important role to play in addressing 'management issues' where the emphasis is on assessing how best to provide services, rather than identification of priorities for investment. For example: Should services be provided in health facilities or within homes? Should they be provided during the daytime or in the evening? Such issues have been under-evaluated but better strategies will promote more efficient resource use.

The choice of technique obviously depends on the objective of the study. Broad evaluations of health consequences or of programmes in different sectors require the use of CBA or CUA. More constrained evaluations, for which a single unit of effect can validly summarize project/programme impact, can be undertaken with CEA. CEA is especially useful in considering strategies for service provision. However, the techniques used must also reflect an assessment of the methodological difficulties of applying them to the particular study and in the particular context.

A decision on when to use the techniques must rest on consideration of whether:

- the volume of resources at stake justifies the study
- there are clear alternative uses for the resources to be evaluated
- the technology underlying each alternative is sufficiently well understood
- a reasonable length of time is available for the study
- decision-makers are receptive to the results of the study and have not already made up their minds.

5. Ten Questions to Ask of any Study

Decision-makers in the health sector may often find themselves in the position of receiving an economic evaluation study on the basis of which they may be expected to take some action. The following questions provide a framework for assessing the results of any evaluation study. They summarize the various issues discussed in this chapter.

1. Was a well-defined question posed in answerable form?
 - (a) Did the study examine both costs and effects of the service(s) or programme(s)?

- (b) Did the study involve a comparison of alternatives?
 - (c) Was a viewpoint for the analysis stated or was the study placed in a particular decision-making context?
2. Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?
 - (a) Were any important alternatives omitted?
 - (b) Was (should) a 'do-nothing' alternative (have been) considered?
 3. Was there evidence that the programmes' effectiveness had been established?

Was this done through a randomized, controlled clinical trial? If not, how strong was the evidence of effectiveness?

4. Were all important and relevant costs and consequences for each alternative identified?
 - (a) Was the range wide enough for the research question at hand?
 - (b) Did it cover all relevant viewpoints (e.g. those of the community or society, patients and third-party payers)?
 - (c) Were capital costs as well as operating costs included?
5. Were costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits, days lost from work or years of life gained) prior to valuation?
 - (a) Were any identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?
 - (b) Were there any special circumstances (e.g. joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?
6. Were costs and consequences valued credibly?
 - (a) Were the sources of all values (e.g. market values, patient or client preferences and views, policy makers' views and health care professionals' judgements) clearly identified?
 - (b) Were market values used for changes involving resources gained or used?
 - (c) When market values were absent (e.g. when volunteers were used) or did not reflect actual values (e.g. clinic space was donated at a reduced rate) were adjustments made to approximate market values?
 - (d) Was the valuation of consequences appropriate for the question posed (i.e. was the appropriate type, or types, of analysis - cost-effectiveness, cost-benefit or cost-utility - selected)?

7. Were costs and consequences adjusted for differential timing?
 - (a) Were costs and consequences that occurred in the future 'discounted' to their present values?
 - (b) Was any justification given for the discount rate used?

8. Was an incremental analysis of costs and consequences of alternatives performed?

Were the additional (incremental) costs generated by the use of one alternative over another compared with the additional effects, benefits or utilities generated?

9. Was a sensitivity analysis performed?
 - (a) Was justification provided for the ranges of values (for key parameters) used in the sensitivity analysis?
 - (b) Were the study results sensitive to changes in the values (within the assumed range)?

10. Did the presentation and discussion of the results of the study include all issues of concern to users?
 - (a) Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?
 - (b) Were the results compared with those of other studies that had investigated the same questions?
 - (c) Did the study discuss the generalizability of the results to other settings and patient/clinic groups?
 - (d) Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences or relevant ethical issues)?
 - (e) Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme, given existing financial or other constraints, and whether any freed resources could be used for other worthwhile programmes?

Source: Drummond MF and Stoddart GL (1985) Principles of Economic Evaluation of Health Programmes World Health Statistics Quarterly 38(4): p 365.

Chapter 10

NATIONAL ACCOUNTS AND THE HEALTH SECTOR

1. Purpose and History

National accounts provide statistical estimates of the value of national income, product and expenditure, and of their main components such as consumption, investment and government expenditure, on a standard basis which is consistent over time and between countries. They have a central place in economic planning; for example, in the UK various economic models based on the national accounts are used to forecast the effects of changes in government policy. Early forms of national accounts are found from the 18th century onward in Europe. Modern Western national accounting started from about 1900. There has been a rapid growth since then, with technical advances in national accounting methods more or less keeping pace with increasing government responsibility for economic affairs, and the increasing capacity and availability of computers. National income estimation in developing countries began in about 1950 (e.g. Jamaica) and estimates are now available for almost all countries.

National income/product per capita is the most widely used general-purpose indicator of the level of development. Hence it is important to understand how it is calculated and what are its limitations. National income must not be confused with government or public income, which is the share of national income which comes under government control.

2. Methods of Estimation

Modern national accounting started from an interest in welfare problems and concentrated on the income side. Such problems include: How numerous are the poor? How are incomes distributed? What is the effect of taxes?

At the simplest level, given a record of the income for a period - say, a year - of every individual in a country, it might be possible to add these together to give total national income. This approach is basically right if certain complications are considered:

- there are corporate incomes which do not accrue to individuals (the undistributed surpluses of businesses)
- there are similar surpluses (or deficits) related to government income
- part of the income of individuals and governments is passed on as a gift to others through transfer payments.

Provided double counting of transfer payments is eliminated, the national income can be calculated, in theory, as the total of individual, corporate and government incomes. Unfortunately, complete records of individual incomes are not available even in developed countries - much less in developing ones with less complete bureaucracies and tax systems, and many non-cash transactions.

How can the gaps be filled? Part of the answer is to take advantage of an important principle:

- all incomes result ultimately from production
- all incomes are spent, resulting in expenditure on consumption or investment.

Therefore the production and expenditure approaches can be used to supplement the income approach: if, for a part of the economy, there is no information on incomes, it might be possible to replace it by information on production or expenditure. For example, if the incomes of doctors are not known, they might be estimated from the public's expenditure on doctors' services. Information from more than one approach can also be used as a check on accuracy.

In practice, estimation in developed countries is more likely to start with production than incomes. Much of the basic information is in the accounts of businesses (including government enterprises). The basic items in their current accounts are:

<u>Receipts</u>	<u>Expenditure</u>
Sale of products	Purchase of goods and services from other enterprises
	Wages and salaries)
	Rent (land, buildings)) 'factor
	Interest) incomes'
	Profits)
	Depreciation (value of capital used in production)
	Taxes

If the expenditure side for all businesses were added together, all the incomes generated in industry would have been included; and correspondingly for government enterprises. Certain adjustments have, however, to be made to this crude total. The value of goods and services bought from other businesses (intermediate goods) must be left out, because these would involve double counting i.e. because the corresponding incomes will show up somewhere else in the economy, to count both would be to count the same item twice. Factor incomes must be included but depreciation and taxes can be included and/or excluded (see next section).

In some industries, even in developed countries, the basic accounts are poor (for example, self-employed lawyers and doctors, or small-scale farmers). The gaps have to be filled by estimates of what the public spends on the output of these industries (for example, through surveys of household budgets or retail sales).

In developing countries the expenditure approach plays a bigger role in national accounting than the income approach, mainly because of poor information. Another reason is that in some developing countries a large part of the national income is spent on imports, which are relatively accurately recorded for tax purposes. Tracing the money spent on importing and distributing foreign goods provides a framework for the rest of the national accounts.

When national accounts estimates are published, the totals can be broken down in any degree of detail, according to their purpose. But it is universal to distinguish between sectors that have different economic roles - the basic ones are industry, government, households (representing individual consumers and workers) and rest of the world (to take care of imports and exports). All these represent current payments. A capital sector is added to cover payments from the current to the capital account (savings, depreciation, investment). The basic national accounts record the flows between these five sectors.

3. Alternative Forms of the Totals

There are some descriptive words attached to totals like 'income' and 'product' to reflect alternative ways of treating marginal items:

- domestic v. national - all incomes generated within the country v. all incomes accruing to nationals of the country (e.g. if on balance there is an outflow of profits to foreign industrial areas, as is not uncommon in developing countries, the country's domestic product will be greater than its national product)
- gross v. net - including or excluding the value of capital assets used in production (depreciation)
- at market price v. at factor cost - including or excluding indirect taxes.

Ideally, like with like should be compared; but in comparing developing countries this is not always possible.

It is important to notice whether a total is given at current prices or constant prices. In the latter case it will have been adjusted using some index to reflect the prices of some base year.

4. National Income as an Indicator of Welfare

One of the main uses of national income figures is to serve as an indicator of changes in the welfare of the population over time (or welfare differences between countries). For this purpose the estimates have well-known limitations:

- the quality of data is uneven between industries and between countries; for example the economic importance of small farming in developing countries may be underestimated
- omitted items - no money value is attached to many non-market activities, such as domestic work in one's own household
- inclusion of activities which do not contribute to welfare - defence? advertising? alcohol? pornography?
- silence on income distribution - the same per capita average could represent different patterns of income distribution (and so of welfare).

Discussion of these problems has a long history in the developed world, but has recently been emphasized much more for developing countries. Suggestions for improvement are along two main lines:

- to elaborate existing systems of national accounting (for example incorporate the costs of pollution; show incomes of different groups separately in the accounts)
- to use direct welfare measures such as life expectancy, infant mortality, literacy, nutrition, employment, separately or incorporated into a single 'quality of life' index. (Note that between countries, all these measures are in fact highly correlated with per capita incomes.)

At present, the most common approach appears to be to use a statistic such as gross national product per capita as the best single general indicator of welfare and development, supplemented with special indicators for different purposes.

5. The Health Sector and National Accounts

Although the national accounts might appear to provide the information required to assess health sector expenditure, in practice health sector information is generated through special surveys (see Chapter 11). This is because the national accounts do not allow easy or adequate estimation of health expenditure. Among the problems are:

- the full range of health expenditure is broken up in various ways
 - * expenditure on health-related items such as domestic water supplies is excluded from health expenditure (it clearly is difficult to draw the line between such non health items which affect health and more direct health items)
 - * expenditure on the education of health personnel is classified under 'education'
 - * premiums paid for health insurance services are classified under 'financial services' (and are often excluded from national accounts calculations)
- national accounts data are too aggregated to allow easy identification of expenditure such as household payments for the care provided by non-government, non-profit hospitals (which may be large in countries with mission hospitals)
- treatment of money flows in national accounts does not adequately represent the health sector pattern e.g. user fees are seen as transfers to government in national accounts, but for the health sector they are direct payments and part of consumer expenditure.

More generally, there are two central national accounting questions that are especially difficult to answer for the health sector:

- are all relevant activities represented by monetary expenditure?
- are monetary measures appropriate measures of the relative importance of activities (their value)?

As 'health' is a broad-ranging concept it is difficult to identify all relevant activities. However, there are many individual actions (e.g. smoking or not smoking) and services rendered by one household member to another (e.g. discouraging smoking) that are important to health but that cannot be measured in monetary terms. These are missed both by national accounts data and by special health sector expenditure surveys.

Monetary measures have a variety of weaknesses in assessing the relative importance of activities within the health sector. For example, there are fundamental accounting differences between institutions in the government and private sectors. The value of the output of a private hospital is based on the price clients are willing to pay for its services but in government hospitals the value is based on the cost of inputs. In other words, the valuation of a private hospital's output includes an element of profit but the valuation of a government hospital's output does not, and so the public sector's output is underestimated relative to that of the private sector.

In general, the appropriate valuation of activities within the health sector is a difficult issue, as discussed in Chapter 9. National accounts data cannot fully assess the contribution of health or health-related expenditure to welfare - stressing the importance of using complementary indicators in welfare assessment (e.g. access to safe water supplies, percentage literate). Health sector expenditure surveys should also include such assessments.

Chapter 11

HEALTH SECTOR FINANCE AND EXPENDITURE

1. Health Sector Financing Problems

The urgency of the need for information about health sector financing and expenditure in developing countries arises out of the current problems of the health sector. Despite the optimism of the late 1970s and the Alma Ata declaration on Primary Health Care, the limited improvements in health status that had been achieved have now been undermined by the world economic crisis of the late 1970s and 1980s. The health needs of many within developing countries and, particularly, the rural and urban poor populations, remain unmet. Infant mortality rates remain high in most developing countries and 80% of the world's population has no access to any permanent form of health care.

The inadequacy of funds in any country is a mix of absolute deficiencies (resource shortages) and relative deficiencies (inefficient use of resources). Resource use inefficiencies may take the form of inefficient manpower use (because functions are not clearly defined) or inappropriate use of technologies (such as capital-intensive methods of health facility construction). The maldistribution of available resources is also an aspect of inefficiency. Health systems in developing countries are generally dominated by urban (hospital) care, leaving only limited funding for rural facilities and primary health care. In many systems, especially those with social security schemes, resource distribution favours the more wealthy (employed and insured); and in most systems it is likely that the poorest have worst access to, and lower utilization of, health care. Finally, problems are caused by the lack of co-ordination between the sources of health finance and health care (such as government, private, foreign), and between the various providers of care (public, private, different levels of the system). The result may be inadequate service provision or wasteful duplication. Management weaknesses underlie these problems.

Resource shortages have been exacerbated by rising costs within the health sector due to increased utilization rates (in some countries), rising expectations - especially for the adoption of new and expensive technology, and an increasing proportion of elderly within the population (particularly in the more developed countries of Latin America and the Caribbean). Often incentives for cost-containment do not exist to temper the impact of these changes on resource use.

The lack of information about health sector financing and expenditure in many countries has undermined appropriate decision-making. The problems of studying financing and expenditure patterns in the health sector include:

- a tendency to regard the health sector as synonymous with health services and to ignore expenditures outside main government health agencies (i.e. a conceptual/definitional problem)
- bureaucratic reluctance to encroach on the organizational territory of co-providers of finance and services (unless formally instructed to)
- the independence of decision making enjoyed by many providers, which takes priority over the need to coordinate and integrate health policy.

The problems, however, are now sufficiently great to encourage more appropriate information collection, and a common methodology for obtaining the information is emerging.

2. Why Study Health Finance and Expenditure?

Studies of health sector finance and expenditure generally have two main purposes, related to the health sector's problems: to investigate the efficiency of the health sector and to provide information for financial planning.

In practical terms, the information can be valuable for:

- identifying who benefits from services and so clarifying the impact of present policies on equity
- identifying who gets what (i.e. the geographical, age, gender, ethnic and socio-economic distribution) and determining whether this is in line with policies
- identifying current patterns of financing and expenditure and so encouraging discussion of alternative sources of finance
- identifying and quantifying resource deficiencies through clarification of the type of services produced and the resources employed in their production
- facilitating the co-ordination of various funding agencies and spending bodies
- analysing resource deployment and identifying the possibilities for re-deployment
- permitting the comparison of financing and expenditure patterns between countries.

3. Steps in a Health Sector Finance and Expenditure Survey

There are four basic steps in assessing health sector finance and expenditure:

- define the health sector
- describe the scope and objectives of the survey
- collect and tabulate the data
- present the results.

3.1 Definition of the health sector

The precise definition of what services and activities comprise the health sector is necessary to guide data collection and, especially, to make comparisons of health systems across countries or at different times.

The following pairs of items show the difficulty of drawing a line between aspects of the health sector/non-health sector. Which should be included within the definition of the health sector?

- health services; environmental services (e.g. water, sanitation, environmental pollution control, occupational safety etc.)
- hospitals; social welfare institutions
- education and training; pure medical research
- medical social work; social work
- formally trained medical practitioners; traditional medical practitioners.

In practice, the limit of the health sector varies considerably between countries and different definitions have been developed for different purposes. In developing countries, the definition tends to be broader than in developed countries because, for example, of greater deficiencies in certain areas (e.g. environmental health) and extensive use of the traditional health sector. A useful rule of thumb is to include all finance/expenditure whose primary intention (regardless of effect) is to improve health.

Surveys often specifically seek to identify primary health care expenditures, and so require a definition of 'primary health care' (PHC). This encounters the same problems as those of defining the health sector (i.e. where to draw the line between health and non-health activities). In addition, it may be inappropriate to include all of the commonly accepted health expenditures as being PHC: To what extent is hospital care primary health care? To what extent do support services such as laboratories undertake PHC? Of course, each level of the health system supports the others but there are

distinctions between types and levels of care. The common practice is to identify a 'PHC factor' which can be applied to the total within each category of expenditure. Thus, 0.2 might be the factor for secondary services which provide clear support to the primary level, but 0 might be the factor for tertiary facilities. To apply such factors it is important to identify the range of services provided within facilities or at different levels of the health system. The factors, clearly, represent only an approximation of PHC expenditure; it is difficult to capture the expenditure implication of facets of PHC such as decentralization and community participation.

As it is unlikely that uniform definitions of the health sector or of primary health care will ever emerge, it is necessary to describe clearly what has been included and excluded, and to justify the definition adopted.

3.2 Scope and objectives of survey

The scope and objectives of the survey clearly have to be set within the context in which it is undertaken - reflecting the needs and priorities of the specific country, and the feasibility of data collection. The common questions tackled by such surveys include: What is the total expenditure on the health sector? What are its sources? How much does each source provide? On what/who is it spent?

It is important first to clarify whether the emphasis is on financing or expenditure or both, and whether it is to be a full or partial review of sources of finance/items of expenditure.

The range of sources that might be investigated includes: ministry of health, other ministries, local government, other state bodies, missions, industry, local voluntary bodies, direct household payments, insurance, self-help, foreign aid (official and private).

Expenditure categories might include: national hospitals, general hospitals, specialized hospitals, health centres, community-based care, private practitioners, drug sellers, sanitation, nutrition, training, research.

3.3 Data collection and tabulation

A variety of data collection methods will be necessary - both because of the different sources of information and to allow for some cross-checking of results. They include routine accounts/budget data, questionnaires, and interviews with health care providers. It can be especially difficult to obtain information about expenditure on private health care, including traditional care, and information should be sought from many sources including providers, users and tax records.

The data should be tabulated as they are collected; particular care should be taken to guard against double-counting and to record income and expenditure against the appropriate source. For example, government subventions to mission facilities could be recorded against government expenditure or against mission expenditure - but the original source is government and the subvention should be recorded on its account when the survey's purpose is to identify the sources of finance. At the same time it is useful to know who controls how much money, and so in some cases it may be appropriate to record the subvention on the mission account.

Normally, sources of finance are defined as above to be government ministries, missions, insurance agencies etc. In some surveys it might be possible, or relevant to identify the original sources of finance e.g. for government expenditure, the amount originating from taxes (on households, businesses etc.), from borrowing (foreign governments) etc. This is particularly important when there is concern over the equity of financing and expenditure patterns (i.e. who pays, who gets care).

3.4 Presentation of data

The data should be presented clearly, in a form that reflects the survey's objectives and permits relevant decisions to be made. A number of different breakdowns are possible:

- recurrent, capital and total expenditure
- expenditure by line item
- expenditure by source of finance
- primary health care expenditure by source and by line item.

Information about population groups served can be disaggregated by: age, geographical distribution (i.e. urban/rural), insured/uninsured, socioeconomic (income) levels. Similarly, information about the types of health service provided can be disaggregated by: preventive/curative, ambulatory /hospital, primary/secondary/tertiary.

The data can also be used as the basis of other calculations which might add to the survey's usefulness e.g. unit costs (per inpatient day, per outpatient visit), proportion of budget actually spent. Expenditure on programmes or disease categories prevented/treated can be calculated in order to compare expenditure patterns with health objectives and plans, and to monitor the achievement of plans.

4. Results of Health Sector Financing and Expenditure Surveys

The results of surveys that have been undertaken make it more possible to compare health services at different times and across international boundaries. They have provided an impetus to further studies and to refinement of the methods, have contributed towards developing an overall system of national health accounts and have shown that surveys can be undertaken relatively quickly and cheaply.

Studies show that sector finance is contributed by more sources than previously imagined and comprises a higher percentage of national income than hitherto believed (e.g. 9.8% of GDP in Swaziland). The role of government financing within the sector as a whole, however, remains significant. In Sri Lanka 86.0% of sector recurrent expenditure was provided from government revenues, in Togo 52.4% and in Swaziland 30.2%.

Household expenditure on health care and health-related activities may be a high proportion of total recurrent expenditure within the health sector: in Swaziland it represented 32.4% of the total, in Malawi 31.5% and in Togo 27.8%. In Sri Lanka the wide availability of government health care and limited use of private care resulted in household health expenditures representing only about 3% of sector recurrent expenditure. Foreign aid may also be important: in Swaziland it financed 20.8% of the sector's total recurrent expenditure and in Togo 13.2 %.

Other information obtained through such surveys includes: the importance of foreign aid to primary health care in Swaziland (and the lack of government financial backing for its stated policy of PHC), the unequal rural/urban allocation of resources in Malawi and Sri Lanka, the unequal racial allocation of resources in Zimbabwe, the dominance of curative and hospital care within health expenditure and the significance of private care within the sector (e.g. traditional healers in Swaziland).

Chapter 12

SOURCES OF FINANCE FOR THE HEALTH SECTOR

1. Criteria for Assessing Financing Mechanisms

The problems of the health sector that are discussed in Chapter 11 have fuelled the debate about how health care is financed. There have been calls for 'alternative financing' in order to address these problems. However, before considering different financing options it is important to identify criteria for their evaluation. The criteria should be based on the overall objectives of health planners and policy-makers and should also reflect concern with the feasibility of implementing new financing mechanisms.

1.1 Efficiency

Given the limited resources available for health in developing countries, it is essential to raise and use resources as efficiently as possible. Four aspects of the overall efficiency with which financing mechanisms raise and use finance should be considered.

Raising resources: it is clearly important that any mechanism raises adequate resources - channelling sufficient new funds into the health system to support existing or expanded services, and having the potential to raise additional revenue to meet the growing needs of health programmes. The stability or reliability of financing sources are also important considerations. Uncertainty or cyclical fluctuations in the economy and/or political allocation process can undermine the level of revenue raised. Finally, the net yield of any source is a reflection of efficiency. The net resources available for the purposes of health care delivery consist of total revenue minus the administrative and other costs of collection. In addition administrative corruption or evasion on the part of those liable to taxation or fees may reduce the actual yield below its hypothetical yield.

Using resources: freedom and flexibility in the management of funds are important, as excessively stringent reporting requirements can increase administrative costs, and restrictions on the allocation of funds between expenditure categories can cause inefficiencies in the delivery of services. Public sector sources and external sources are typically less flexible than private sector sources in this respect.

1.2 Equity

Most governments explicitly identify equity as a policy objective - and, in particular, as one of the objectives of their health care system. While the precise definition of equity that is adopted is often not clear it generally reflects a concern to distribute health care fairly, in recognition of differences in health need. The question "who gains?" must, therefore, be asked in assessing the equity impact of health systems.

In considering the impact on equity of health care financing options it is equally important to ask "who pays?". One standard against which to assess the answer might be that only those who benefit from/use health care (i.e. the sick or potentially sick) should pay for it - horizontal equity. However, this standard is likely to penalize those who are least able to pay and most likely to be sick (low-income groups). Their in-ability to pay not only reflects the largely chance initial allocations of income, but also is likely to underlie their sickness/vulnerability. An alternative standard that is more fair, therefore, is that the distribution of the burden of paying for health care should reflect differences in ability to pay - vertical equity. Some would argue that health care payments (with taxation) should contribute to a more equal distribution of income.

The impact of the payment burden may influence the distribution of health service benefits, and so the interaction of the payment method with the demand and supply of health care must also be considered (see below).

The first step in evaluating equity, however, must be to define the equity goal of health care and so clarify the nature of the gains sought (e.g. more equal access to health care or more equal utilization). In practice, evaluation should also focus on the socioeconomic status of those who pay for and use health care, as this status underlies both health need and ability to pay.

1.3 Demand/utilization and consumer behaviour

Different financing mechanisms have very different effects on the level and type of service use. Some methods of payment influence consumer behaviour by the incentives given to providers to withhold or provide services; while some may directly stimulate or restrain the utilization of services. Many existing financing policies have paid little attention to the incentives they create or reinforce, or to their ensuing impact upon service providers, households, and government agencies. Some financing mechanisms may encourage undesirable practices such as the in-appropriate utilization of services.

Identifying the effect of financing mechanisms on consumer behaviour requires an understanding of its determinants. The demand for health services can be defined in terms of the coincidence in one individual of both the willingness and ability to pay. These can be related to a set of

sociodemographic factors such as age, education, gender and health status; and a set of economic factors such as the monetary (e.g fees, drug costs and travel costs) and non-monetary (time) costs of seeking care, income levels in relation to the magnitude of the costs of the care, and the degree of access to cash or other accepted forms of pay-ment.

The evidence concerning the demand for health care in developing countries is mixed but suggests that, especially for low income groups, demand may be elastic with respect to price, falling as price increases and resulting in significant shifts in the use of alternative providers. Even where only the time price of health care (resulting from travel and waiting times) has been considered, the evidence supports this finding; and other factors, such as poor access to facilities, is also recognized to undermine utilization. Clearly planners must assess these other in-fluences, as well as monetary prices, in order to evaluate the effects on utilization of financing mechanisms such as user fees.

1.4 Supply/provision and provider behaviour

Methods of financing affect the supply or provision of services. Some financing sources are biased with respect to the types of expenditure that they favour: favouring curative rather than preventive, or capital rather than recurrent, expenditures. The sources may also influence the nature of the production technology and the type of health personnel employed. Public health authorities are more likely to experiment with the use of paramedical personnel, especially in outlying rural areas, and private providers are more likely to favour the use of profes-sionals. Finally, certain payment and reimbursement mechanisms have un-desirable effects on specific aspects of the behaviour of service providers. For example, they may encourage surgical intervention in labour where it is not medically required.

Like demand incentives, these supply-side spin-offs of financing policies are often ignored. Yet their impact on the cost of provision and their encouragement of inappropriate service provision contribute to the inefficiency of resource use.

1.5 Displacement effects

Rather than generating additional resources for the health sector, new or expanded financing mechanisms may merely displace funding from other sources. Displacement is not necessarily an undesirable consequence if the new or expanded source of finance is more efficient or more equi-table than the one it partially displaces.

Examples of displacement effects include foreign assistance which may displace government support for health care; counter-funding, often a precondition for foreign assistance, which may divert funds away from existing priority projects; health insurance schemes, which may in some instances displace

rather than add to the total of resources being allocated to health care (e.g. displacing direct payments); charitable contributions which may be withdrawn when other sources are developed; and government allocations which may be reduced when other sources of finance (such as user fees) are developed.

1.6 Wider effects of the health sector

Health sectors may account for a sizeable share of national resources and are often major employers. Consequently, the activities of the health sector may have spill-over effects on the economy as a whole. These include external effects on costs (e.g. inflation through the repercussions of high increases in staff pay); foreign exchange problems through heavy foreign borrowing for development projects or for payments for imports such as pharmaceuticals or equipment; opportunity costs such as the attraction of scarce manpower into the health sector at the expense of other professions; and disincentives to investment and employment (e.g. as a result of financing health services through high taxes on certain economic activities, enterprises or sectors). These external effects may also be positive as in the case of improved productivity resulting from reduced death and disability in the work force.

2. Public and Quasi-public Sources of Finance

2.1 General tax revenues

General tax revenue is used in almost every country of the world to finance certain components of health care and, in developing countries, it is often the most important source of financing. However, low tax ratios (the proportion of national income collected as tax) in these countries mean that it is often insufficient by itself to support health care. Although tax ratios tend to increase in line with development, this depends in large part on a country's political will to increase the tax burden. In developing countries general tax revenue is composed largely of duties on imports and exports, and sales taxes. Taxes on business transactions, profits and incomes are all of lesser importance.

General tax revenue is currently not the most reliable source of finance for the health sector in developing countries. This results from factors such as the low political priority frequently given to the health sector in national budget decisions; the instability of government finance in countries heavily dependent upon taxes on imports and exports; the frequent use of public expenditure as a tool of macro-economic policy; and frequent disparities between budgeted funds and their actual availability or disbursement. The net yield is usually high, unless bureaucratic overheads are high.

The equity impact of tax systems is dependent on both the proportional burden of taxation and on the use which is made of the revenue raised. Tax systems can be progressive, falling more heavily on the rich than the poor and, therefore, equitable; but they may also be regressive, falling more heavily on the poor than the rich, and inequitable. Developing countries are assumed to have regressive financing systems because they tend to rely on indirect taxation, but in practice their tax systems may be progressive because the poorest sections of society fall outside the formal economy and indirect taxes may be levied primarily on luxury items consumed predominantly by the wealthier population groups. Available evidence on the burden of taxation is in-adequate to permit firm conclusions about its incidence to be drawn. However, it is clear that tax revenue is often used inequitably in health systems: health systems are dominated by high-technology urban- based care and so the rural populations (and the urban poor) have in-adequate access to any form of care.

There is a limit to what can be collected in tax revenue and how much can be allocated to the health sector without conflict with wider primary health care objectives. Taxes that make the poor poorer could seriously damage their health status and undermine their productivity; there are also many other fields of socioeconomic development that com-pete with the health sector for funds and yet give substantial support to primary health care (e.g. agriculture).

2.2 Deficit financing

General tax revenue may be supplemented by deficit financing, that is the decision to borrow and spend funds in the present and repay them over some period of time. Deficit finance may be raised nationally or internationally, through mechanisms such as the issuing of bonds or cer-tificates or long-term low-interest loans. The cost of enjoying the use of these funds in the present rather than the future is the interest that needs to be paid on the loan. In developing countries high infla-tion rates (affecting the real rate of interest on loans) and lack of confidence in the government's abilities to honour eventual redemption of the bond may make it difficult to use deficit financing as a source of support for health systems.

When it is used, deficit financing is typically for specific construc-tion projects (e.g. hospitals and water and sewerage systems). Unless such projects sell their services or contribute directly to increased output that can be taxed to service the debt, the deficit must be repaid from general tax revenue. Thus the agency doing the deficit financing must be endowed with the authority to impose additional taxes or fees, or be given a claim on general tax revenue in order to service the debt.

Deficit finance may also be raised from abroad in the form of bilateral or multilateral aid loans, typically given for a project life of between three and five years, and thereby constituting only a short-term source of support. Although useful for many developing countries in helping to develop and expand health care infrastructure, foreign aid is often limited to supporting import components.

Past reliance on deficit financing in the economy as a whole is now burdening many countries with excessive debt repayment problems.

2.3 Earmarked taxes

Most tax revenues are paid into a national pool and then shared out between different areas of government expenditure. Some governments, however, may ' earmark' a particular tax for a particular purpose. For example, taxes on the sale of particular products may be earmarked for health services at either national or local level. The problem with such taxes is that they are often difficult to administer, may be politically unpopular, and are also often unpopular with tax administrators because they limit their freedom of action. They can be regressive if, as is often the case, taxes are levied on items such as beer, cigarettes, recreational events, or foodstuffs; but they can be progressive if they are imposed on luxury products purchased primarily by the more affluent sections of society.

A clear advantage of this source of finance is that a tax is visibly assigned to priority funding of certain activities or programmes. Although not a major source of health sector finance in most countries, they may constitute an important source of finance for specific projects or programmes.

2.4 Social insurance

Social insurance can finance health care, as well as other needs such as invalidity and old age support, for either the whole population or a part of it. It is conventionally financed by imposing mandatory insurance payments on employed workers as a percentage of their wages, and by imposing a similar or somewhat higher payroll tax on their employers. Governments may in some instances also contribute to the schemes. Beneficiaries (workers and their dependents) may have to pay a user fee (termed copayment) in addition to their wage deduction. In order to include those workers outside the modern employment sector insurance payments may also be calculated on measures of income or wealth other than wages, such as the value of crops produced. Allowance will then have to be made for the fact that cash income is only available seasonally, when crops are sold. In their capacity as employers, governments may either run their own social insurance scheme or contract such schemes to private insurance companies.

The total financial contribution to social insurance schemes is (in theory) determined actuarially on the basis of the incidence of illness, the conditions of eligibility for benefit, and the value of those benefits. Individual contributions are not determined, however, on the basis of expected risks or claims, but in some proportion to income. As risks are pooled, there is an unequal benefit distribution in favour of high-risk (high-need) workers.

The main problems of social insurance are related to issues of equity. It is easiest to cover those in regular employment, who may be as little as 5 to 15% of the population in developing countries; and there are often marked inequalities in the quantity and quality of services available to those covered by insurance relative to those who are not. Overall, it is argued that social insurance reinforces the maldistribution of resources between rural and urban areas in developing countries. It provides extra funds for largely urban, employed workers and leaves the large rural populations and the informally employed urban population even further handicapped than before its introduction. Critics of social insurance also argue that it undermines both public and private health care by competing with these sectors for limited supplies of real medical resources (e.g. personnel). Finally, it tends to promote or reinforce high-cost, hospital-based, doctor-centred, curative health care.

More positively, social insurance can be the means of channelling extra funds into organized health services. By relieving the pressure on ministries of health to devote resources to urban health services, it may even, indirectly, make more resources available to those in rural areas. Governments have in many instances an increasingly favourable attitude to the development of social insurance.

2.5 Lotteries and betting

These may be used as sources of earmarked income for health and social services in developing countries. Often administered by quasi-public bodies under national or local government regulation, these typically non-profit schemes rarely constitute an important component of overall health sector finance. Largely supported by the incomes of the poor and thereby constituting a form of regressive taxation, they typically have low net yields because of the payment of prizes and high administrative costs. The typical net yield from lotteries is between 10-30% of gross receipts.

3. Private Sources of Finance

3.1 Private health insurance

Private health insurance differs from social insurance in two main ways. First, private health insurance typically does not include pensions for invalidity or old age. Second, the price (or 'premium') charged for private health insurance is not based on the pooled risks of a large population, but on personal risk characteristics and the likelihood of illness in the individual or group covered. As a result, premiums are likely to vary for different individuals or groups.

Schemes may be profit or non-profit making, and may be organized for individuals or groups, the latter often benefiting from lower premiums (resulting from lower per capita administration costs as well as a degree of risk-sharing). In many countries the larger employers act as an organizing body for

health insurance, and may pay part of the premium as a fringe benefit. However, in order to control the level of utilization of services, individuals are often required to pay for part of the cost of medical care on a direct fee-for-service basis. In countries where demand is sufficiently high, commercial insurance companies may be active.

In the past few years there has been increasing interest in some developing countries (especially in Latin America) in health maintenance organizations (HMO), an innovative pattern of health insurance and health care organization developed in the USA. Individuals or groups contract with a particular HMO to cover all their health care needs (either in the HMO's own facilities or in facilities under contract to the HMO) in return for an annual payment. The integration of the insurance and provider functions provides an incentive for cost containment, in contrast to a third-party payment system where providers and consumers agree on the quantity of care to be supplied and a third party (the insurance agency) merely reimburses the cost. In addition, competing HMOs enable consumers to choose the one that best suits them and encourages efficient health care provision.

Private insurance is not subject to the political allocation process and may channel extra funds into the health sector. However, it suffers from the problems of low coverage because of its cost and the exclusion of bad risks, of enhancing inequity and promoting the growth of high-technology health care, inappropriate to developing countries.

3.2 Employer-financed schemes

In some instances employers may directly finance health care for their employees. They may, for instance, pay for private sector health services, employ medical personnel directly, or provide necessary facilities and equipment. Oil companies, mining and mineral industries, and large-scale export-centred agricultural enterprises usually provide for the health needs of their workforce. Benefits are seldom extended to families as employers are primarily concerned with maintaining the productivity of the workforce. In developed countries the primary focus is on accident prevention and occupational health, and in developing countries also, employers may have a legal obligation to provide first aid or occupational health services (e.g. sugar and coffee plantations in Latin America, tea and rubber estates in Asia and cocoa farms and mines in Africa).

Problems with employer-financed schemes relate to the quality of care provided, the possible fragmentation of services, difficulties enforcing employer liabilities, and the fact that viability depends upon the performance of the employing agency. Nowhere is employer finance a predominant source of support for health, although employer schemes are often a precursor to national social insurance schemes.

3.3 Charity and voluntary contributions

Charitable or voluntary contributions can take the form of financial support or in-kind donations (e.g. personal services, physical facilities, equipment and supplies), and may originate from business enterprises, wealthy families, religious organizations or private individuals. Often these resources are channelled through foundations or religious bodies.

The problems with this source of finance are often indirect. For example, donors may have different priorities from the recipient nation and may not recognize their most urgent health needs. Donors may prefer to finance visible evidence of their support such as physical facilities and equipment, and thereby commit the recipient country or organization to the recurrent costs of those facilities in the future. Charitable contributions may also take the place of, or reduce, other sources of finance. For example, contributions may be eligible for tax relief, reducing general tax revenues for use elsewhere (although the effects in this instance are likely to be minor).

Charitable contributions have played an important role in health service provision in the past, and in some African countries are still a major source of health care finance, channelled through religious agencies. The general trend, however, is for governments to support or take over mission health services. Thus the role of charitable and voluntary contributions is decreasing, although it may still be important in times of emergency or disaster and can be a useful supplement to other forms of health finance.

3.4 Community financing and self-help

Current primary health care initiatives in developing countries stress the importance of national self-reliance and community participation in health care delivery. By mobilizing under-utilized national and local resources (e.g. organisational skills, manpower and cash) and by developing affordable and culturally appropriate delivery systems, it is hoped that basic health care will become universally accessible. Consequently some governments and many non-governmental agencies are turning to communities for organization, participation and financial support, and communal self-help is increasingly thought of as an important source of financial support for health services in developing countries. The challenge is to develop new types of local institution that can coordinate and systematically utilize the community resources. Self-help can take many forms such as labour, local insurance, support for volunteer health workers, and drug cooperatives.

Opponents of community financing mechanisms argue that it puts the burden of financing on those least able to afford it (often the poorer rural communities). It is also seen as a diversion for governments lacking the political will to generate new sources of revenue, or to re-allocate existing ones. Although in some instances it can make a substantial contribution, community finance is

unlikely to generate sufficient resources by itself to meet country health needs, and should be seen as complementary to, rather than as a substitute for, other sources of finance.

3.5 Direct household expenditure

Household income is ultimately the source of most health care finance, but direct expenditure constitutes a specific category of financing that should be considered separately. Included in this category are any payments a consumer may make directly to health care providers such as fees for services, or prices paid for goods and supplies.

Direct household expenditure is not independent of other sources of finance. Government services may charge user fees (often nominal) for certain services. Even with insurance coverage, there is often a requirement for some degree of copayment, which tends to increase the amount that would otherwise have been spent on health. Health insurance benefits, moreover, may have an upper ceiling, with households required to pay directly for their health care requirements in excess of this level.

Until recently, comparatively little was known about the extent or the characteristics of direct household expenditure on health care, but a range of recent studies have shown that this form of financing is far more common and considerably more important than was hitherto thought.

The extent to which these payments represent a real ability and willingness to pay for health care is, however, unclear. Willingness to pay does not necessarily reflect ability to pay. Current levels of household expenditure partly result from the existing pattern of government health care provision, and the limited access to free/cheap government health care (particularly in rural areas). People may use and buy non-government (e.g. mission, private, traditional) health care partly because they have no cheap or good quality government alternative. Low-income groups tend to delay use of health services until illness is severe, presumably in part to avoid payment, but such delay generally only increases the necessary expenditure. High health care bills may sufficiently undermine their economic position to push them further into poverty. Health care payments also sometimes displace expenditure for other basic necessities of life (e.g. food), because there is only limited ability to pay for the range of household needs.

Utilization of, and payment for, health services is, moreover, likely to depend heavily on the perception of their relevance to a specific health need and the extent to which they provide a service that people value. Use of traditional healers, for example, may reflect a belief in the relevance of their treatments for certain diseases rather than a general willingness to pay for any type of health care. Perceptions of poor quality in government services certainly undermine their use and, therefore, willingness to pay for them. Private services may be more oriented to the preferences and

circumstances of households, for in-stance providing for payment in kind or payment related to ability to pay.

Raising the level of direct household expenditure for health care, for example by user fees, will clearly have a negative impact on equity (by influencing both the distribution of the payment burden and the benefits gained). It may be mitigated by the introduction of an exemption mechanism for the poor, although such a mechanism may itself reduce the demand for health care made by low-income groups because they may not wish to be identified as 'poor'. Moreover, such willingness to pay as exists is attached primarily to curative services, and so can only ex-tend the provision of preventive care if it is possible to re-allocate resources within the health sector. Finally, the potential yield from user fees is unclear. It is dependent on the level and type of fees, the bureaucratic structure required to implement them, the existence of ex-emption mechanisms, the impact of fee systems on the demand for care and the rates of collection. The administrative difficulties of implementing a fee system (e.g. how is ability to pay assessed? who assesses it? who collects the fees? how is abuse of the system restricted?) may cause less revenue to be collected than theoretically is possible.

4. Approaches to Improved Financing of Health Activities

Resource shortages in developing country health systems clearly must be addressed, but the introduction of new financing systems is not an ap-propriate initial response to the problem. Shortages result both from inefficiencies in resource use and from absolute deficiencies, and until the first are adequately addressed any additional resources will also be used inefficiently.

It is also important to recognize that health financing problems are not simply health sector problems, but often reflect economy-wide dif-ficulties. They certainly require national strategies to address them, even where additional resources are to be recruited by actions within the health sector. For example, there must be national agreement that extra finance will be retained for use within the sector (rather than being matched by budget cuts or transferred to other sectors) and that resources can be re-allocated within the sector to meet priority health needs, in order to justify alternative financing strategies.

Within the health sector the first priority must be to improve ef-ficiency, making better use of available resources and enhancing the standing of the sector nationwide. Management can be strengthened through staff training and the development of appropriate tools (including incentives), efforts can be made to understand the community's needs in order that health care better meets them, and resources can be allocated more appropriately (e.g. to preventive rather than curative care). At the same time, the options for increasing funds can be considered - using appropriate evaluation criteria.

If efficiency improvements together with the possibility of additional resources still do not bridge the gap between resource requirements and resource availability, then health sector goals must be reconsidered. Not everything that may have a positive impact on health can be afforded and health plans must be based on a realistic view of resource availability.

Chapter 13

BUDGETARY PROCEDURES: BUDGETARY REFORM AND PROGRAMME BUDGETING

1. Reform of Government Budgeting Procedures

There is a long history of attempts to improve traditional government budgeting procedures. For some years, there has been concern that government budgetary procedures neither serve management efficiency nor provide the information necessary for policy making and planning. The main criticisms of budgetary procedures concern both the actual procedures and the budget structure.

Criticisms of procedures include:

- it is difficult to formulate budgets that will help to achieve a government's objectives because these objectives are often not clearly specified
- the budgeting system is often ignored by planners, although budgets have a crucial influence on existing and future resource allocation patterns
- budgeting is often a completely separate activity from planning: for instance there may be no procedures to ensure that the recurrent costs of new capital developments can be afforded, or that the recurrent budget is increased when new buildings are completed
- budgeting has basically been established as a form of expenditure control and for accounting purposes, not in support of planning activities
- budgets tend to change from year to year on an incremental basis, with changes based on past budgets or, more appropriately, past expenditure; they favour existing facilities and respond to current patterns of utilization, rather than pushing ahead in new directions and ensuring that resource allocation reflects planning priorities
- incremental budgeting brings a little bit more or a little bit less each year regardless of the efficiency or effectiveness of expenditure
- expenditure records are often unavailable for use in budgeting, because accounting systems are themselves inadequate

- shortfalls in the anticipated budget generally result in indiscriminate cuts which do not protect priority budget areas (e.g. a 5% cut across all budget items)
- where budgets are based on development expenditure and its recurrent requirements, planning and budgeting are more integrated but budgets are biased in favour of new facilities rather than supporting primary health care and existing services
- budgeting is often a very centralized procedure (e.g. budgets are planned centrally and central permission is required to reallocate budget items) and so there is little scope for budgetary initiatives or for making lower level managers feel responsibility for the management of budgets; flexibility is reduced as systems grow and the required paperwork becomes uncontrollable.

These problems are reflected and exacerbated by budget structures:

- even if objectives are specified, it is often impossible to associate them with the use of resources because traditional budgets are structured by 'object account' or 'budget code' such as salaries and drugs, and it is difficult to tell what health objectives this expenditure is serving (e.g. primary health care or tertiary level care)
- centralized procedures are reflected in centralized structures, in which individual facilities (even hospitals) may not themselves be 'cost centres' with their own budgets
- expenditure records are often only available by 'sub-vote' (e.g. administration, medical services, preventive services) rather than by 'cost centre' (e.g. hospitals, health centres), and are not helpful for future budgeting activities.

For these and other reasons there have been many proposals for the reform of government budgeting procedures. Of particular interest is programme budgeting.

2. What is Programme Budgeting?

It is difficult to define programme budgeting because:

- programme budgeting means different things to different people: two countries may say they have a programme budgeting system, but the two systems may be very different

- programme budgeting has a number of different names (performance budgeting, output budgeting, PPBS - planning, programming, budgeting system). To some people these all mean the same thing; to others they are different.

The essential feature of programme budgeting is that it focuses on the general character and relative importance of the work to be done, or upon the service to be provided, rather than upon the things to be acquired such as manpower, supplies. These latter items are only the means to an end. Some authors equate programme and performance budgeting. Others distinguish them in the following way:

- programme budgeting is concerned primarily with 'programmes' or the accomplishment of objectives specified in terms of health outputs
- performance budgeting is concerned primarily with 'performance' or accomplishments specified in terms of physical outputs and unit costs.

The basic elements of programme budgeting are:

- classifying expenditures by output-oriented programmes
- using this as a framework for planning and monitoring progress towards objectives
- (possibly) re-structuring budgets.

3. The History of Programme Budgeting

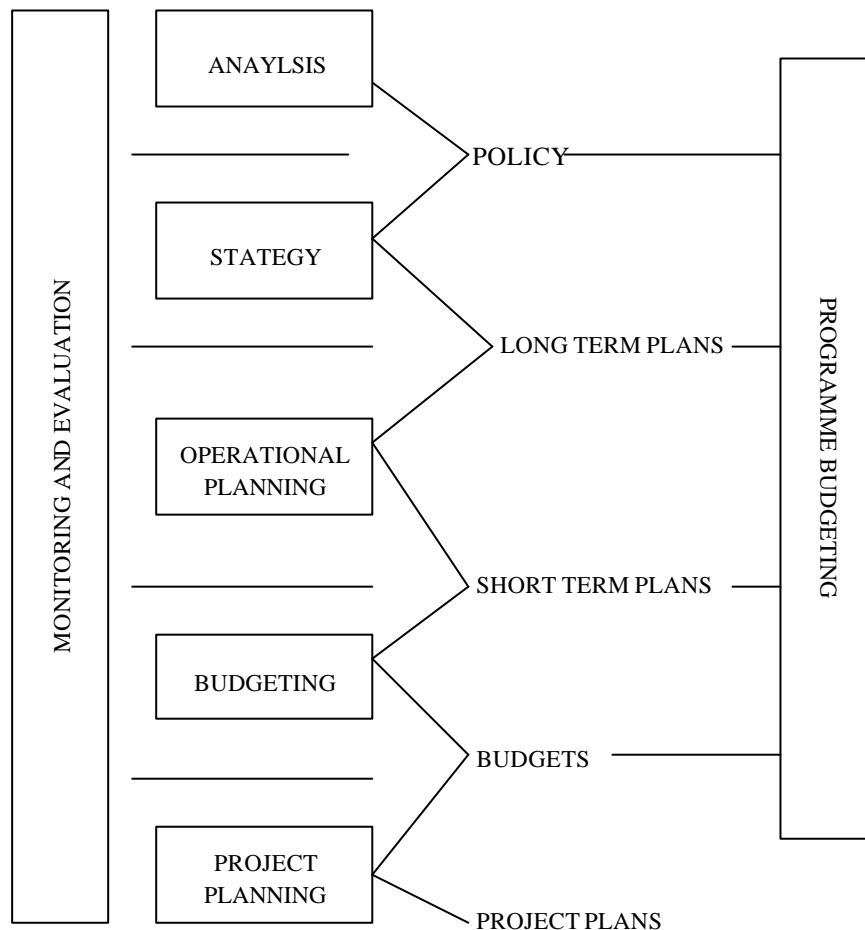
Experiments with programme budgeting began in the 1950s and 1960s. The most publicized experiment was with PPBS in the federal government of the United States, but because of problems the experiment was terminated in 1971. However, the system has since been exported all over the world. It has been adapted and diversified, and does have some approaches that Ministries of Health may find useful.

Nonetheless, the history of programme budgeting provides lessons that are important to keep in mind, particularly that changes in procedures do not necessarily produce changes in behaviour; and that new procedures (like programme budgeting) can easily generate large quantities of information that are then ignored by decision-makers.

4. The Main Issues in Programme Budgeting

Where does programme budgeting fit in? programme budgeting is not a planning procedure or a budgeting procedure but rather provides a framework that links planning and budgeting (see Figure 20).

Figure 20: A Map of Planning Activities



Programme structures: a programme is a group of activities with an objective in common. For instance, one objective might be to improve the health of children and a child health programme could be defined. Under this programme heading, sub-objectives could be specified such as reducing the infant mortality rate; or improving child nutritional status. In theory, a hierarchy of objectives and sub-objectives can be established for programmes, although in practice this can quickly become unnecessarily elaborate.

What programmes should there be within a programme budget for 'health'? A useful guide is that:

- where decisions are primarily a matter of political or moral judgement, the activities to be compared should be in different programmes (e.g. improvement of child care facilities versus psychiatric services)

- where the decision is a technical one of how best to achieve a particular objective, the activities to be compared should be in the same programme (e.g. improvement of curative or preventive services for children).

A programme structure can be based on one or more of the following classification systems. Those higher up the list are more closely related to health objectives; those lower down the list are more feasible given existing information systems.

- target groups (e.g. lower income groups, the unemployed)
- client groups (e.g. physically disabled, mentally ill)
- health service function (e.g. prevention, cure)
- types of activity (e.g. water, sanitation, immunization)
- disease categories (e.g. vector-borne and diarrhoeal diseases)
- levels of care (e.g. primary care, secondary care)
- geographical areas (e.g. districts)
- institutions (e.g. hospitals, health centres).

A government's policy objectives and the availability of information will influence the choice of programmes.

Allocating activities to programmes: each programme should in theory contain all activities (services) that contribute to the programme, regardless of the agency that is responsible for them. For instance a child health programme should contain under-fives clinics, paediatric wards, school health etc.

Costing the programmes: the resources used by each programme should be costed. In theory, all costs should be included whoever pays them (Ministry of Health, community, individuals etc.). Existing accounting systems may make it difficult to identify expenditure with programmes.

Cost information should ideally be presented for the last few years and can be projected into the future.

Adding output measures to programmes: within each programme, the cost of each activity should be matched with information on outputs in order to assess the efficiency of resource use and the value of the activity. Measures of health output should ideally be used (e.g. lives saved), but in practice measures of 'intermediate' output are often used (e.g. number of children immunized).

Evaluation: programme budgeting information can be used in a variety of ways

- to evaluate existing patterns of resource allocation and whether they match government policies
- to evaluate whether a new plan will move resource allocation patterns in the desired direction and whether it has done
- to investigate the implications for health expenditure of changes in the sociotechnical environment: e.g. change in the composition of the population, or the introduction of new technologies
- as the starting point for cost-effectiveness or cost-benefit analysis
- to review the claims of budget-holders for additional resources in the annual budgeting cycle.

The relationship of programme budgeting to the management and accounting system: a crucial question is whether programme budgeting should be used primarily as a tool for planning, or whether the budgeting and accounting system should be changed to incorporate programme budgeting principles.

There is no easy solution. If programmes are defined in terms of client or target groups, these cut across existing institutional structures and budget-holding by programme would complicate accounting systems. If programmes are defined in terms of existing institutional structures (e.g. a vertical malaria control programme, or district hospitals) then this may strengthen their isolation and independence.

However, this problem may be outweighed by the advantages gained in being able to channel resources to a priority programme. For instance, if primary health care is part of a district health care budget, funds may be retained in the hospital. If primary health care has its own programme and budget, then managers can make sure that primary health care obtains an adequate share of total resources.

5. Alternative Budgetary Reforms

Some of the other reforms of existing budgetary procedures that could strengthen the budget process are:

- the decentralization of budgetary control (for example, to the districts or even cost centres at the district level)
- regular consultation with, and involvement of, all levels of the health system and of appropriate non-accounting officers (e.g. medical officer, matron, administrator) during the budget development period

- the provision of budget limits which decentralized budget holders must not exceed but within which they are allowed some flexibility (e.g. for the reallocation of budget items)
- establishing budgets on a 'basic needs basis' ('zero-based' budgeting) to clarify the resource requirements of existing services and to help in identifying budget priorities
- the re-structuring of budgets to reflect the decentralized approach (eg. main division by cost centre, sub-divisions by budget code)
- more timely publication of expenditure data and more appropriate disaggregation of the data (particularly by cost centre, supplemented by summary totals by sub-vote, budget code etc.)
- the adoption of new financial planning approaches (see Chapter 14) together with related budget procedures and structures.

However, any budgeting reforms must also be associated with staff training and skills development. The accountancy cadres in many developing country ministries of health are often under-trained and need support. Budgeting is, moreover, not simply an accountancy exercise but should involve all health managers, who themselves need training in budget development and financial management.

Chapter 14

APPROACHES TO FINANCIAL PLANNING: RESOURCE ALLOCATION PLANNING AND THE FINANCIAL MASTER PLAN

1. The Main Issues

Financial planning takes place within the overall context of health planning, and assists with the following elements of the planning process:

- reviewing present and past expenditure patterns and financing mechanisms of the health sector
- estimating the costs and effects of proposed changes to the health sector to determine whether they are worthwhile
- determining the preferred pattern of resource allocation among population groups, geographical areas and services
- reviewing ways of raising money to obtain the required resources
- translating health priorities and programmes into the budget allocations of the agencies and institutions delivering health and health-related services
- monitoring budgets and controlling expenditure
- producing financial information for decision makers.

Financial planning has thus both a long-term and a short-term component. Long-term financial planning is most concerned with the next 5-15 years, to ensure that current policies concerning the development of the health sector are likely to be affordable and that they do not lead to a situation (such as an increase in hospital capacity or manpower) that cannot be afforded within that period. Short-term financial planning is most concerned with the next 1-3 years, to ensure that plans are translated into action, that costs are estimated, budgets drawn up and expenditure monitored.

2. Resource Allocation Planning

Resource allocation planning has been developed to support primary health care, and is an indicative planning exercise in which priorities are clearly established and resource targets are set for the various elements of the health service. This planning approach seeks to highlight the existing inequalities and imbalances in health care (e.g. between curative and preventive services, between levels of the health

system, between geographical areas, between population groups), to support the process of planning their reduction and to allow progress to be monitored by comparing actual and planned expenditure. It, therefore, evaluates the implementation of primary health care plans and assesses their impact on the structure of health services, with specific consideration of the equity of resource allocations.

Its basic assumption is that only by the appropriate allocation of new resources can the pattern of resource allocation be re-shaped. The existing health infrastructure cannot be substantially altered or reduced and so it is only by allocating new resources to primary health care activities that they will receive funding.

The steps of resource allocation planning are:

- assessment of the development expenditure that may be allocated to the health services in the plan period
- assessment of the likely increase in the Treasury allocation to the health ministry's recurrent budget in the plan period
- the division of these resources within the health service both to meet necessary budget demands and, especially, to address primary health care priorities

In order to improve the equity of resource allocation, the division of resources between geographical areas should be based on population totals; weighted, if possible, by various 'need' factors (such as age and sex) and with respect to the differential costs of care in different areas, patient flows across administrative boundaries etc.

Inevitably the development of plans and budgets in this process involves much negotiation - both among government departments and within the health ministry. It is both a political and a technical exercise. It also requires reforms of budgeting procedures such as: re-structuring to reflect plans and priorities, more flexibility and decentralization of budgetary authority (see Chapter 13).

3. The World Health Organization Approach to Financial Planning

The WHO approach to financial planning falls within the Managerial Process for National Health Development (MPNHD). The MPNHD proposes a systematic way of formulating and implementing a strategy for achieving 'health for all by the year 2000'(HFA2000). The stages of the MPNHD comprise:

- the formulation of national health policies - comprising goals, priorities, and main directions towards priority goals, that are suited to the social needs and economic conditions of the country and form part of national social and economic development policies
- broad programming - the translation of these policies, through various stages of planning, into strategies to achieve clearly stated objectives and wherever possible, specific targets
- programme budgeting - the preferential allocation of health resources for the implementation of these strategies
- the master plan of action - resulting from broad programming and programme budgeting and indicating the strategies to be followed and the main lines of action to be taken in the health and other sectors to implement these strategies
- detailed programming - the conversion of strategies and plans of action into detailed programmes that specify objectives and targets and the technology, manpower, infrastructure, financial resources and time required for their implementation
- implementation - the translation of detailed programmes into action so that they come into operation as integral parts of the health system; the day-to-day management of programmes and the services and institutions for delivering them and the continuing follow-up of activities to ensure that they are proceeding as planned and are on schedule
- evaluation of developmental health strategies and operational programmes for their implementation, in order progressively to improve their effectiveness and impact and increase their efficiency
- reprogramming, as necessary, with a view to improving the master plan of action or some of its components, or preparing new ones as required, as part of a continuous managerial process for national health development
- support, in the form of relevant and sensitive information, for all these components at all stages.

The process is shown in Figure 21.

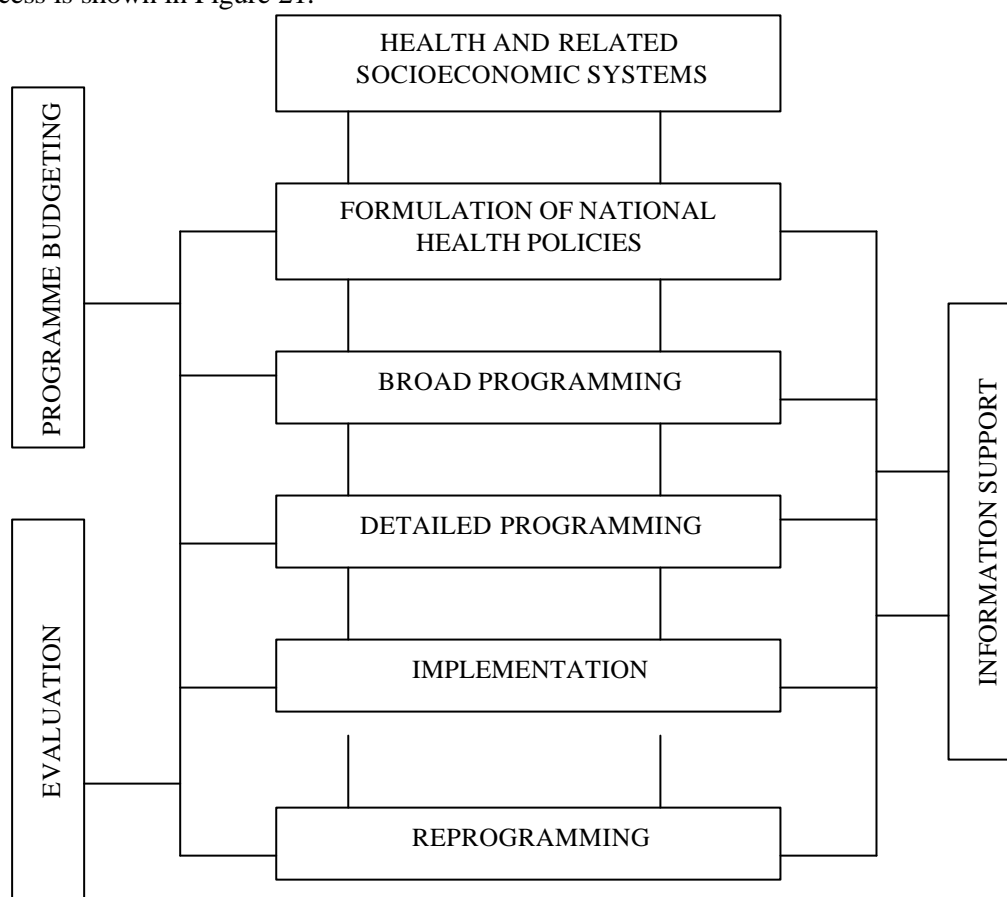


Figure 21: Managerial Process for National Health Development

Financial planning is part of the stage of 'programme budgeting' which here has a much broader definition than that commonly used (see Chapter 13). Programme budgeting, as indicated in the Figure, contributes to all stages of the MPNHD. Long-term financial planning forms part of 'broad programming', and shorter-term financial planning of detailed programming, implementation and reprogramming.

Shorter-term financial planning is closely tied up with project planning and with the annual budgeting cycle. It is thus frequently undertaken by many countries. Long-term financial planning is much more rare, and for this and other reasons, WHO has placed particular emphasis on it.

Long-term financial planning provides a framework for planning strategies to achieve HFA2000 and to assess the financial feasibility of proposed health programmes. The end result of long-term financial planning is a financial master plan which forms part of the master plan of action produced during the broad programming stage of the MPNHD.

4. The Financial Master Plan (FMP)

Four main steps are involved in the development of an FMP:

- the base line: establish what is spent on the health sector and how it was financed for a recent year or years
- the expenditure estimates: make broad estimates of the cost of the HFA plan plus any other commitments
- the income estimates: project existing sources of finance
- reconciling income and expenditure: identify the gap between expenditure and income estimates, and see if it can be filled by new or existing sources of finance; reduce any remaining gap by adapting the costed plan to the finances likely to be available.

The fourth and final step is the most crucial, requiring often not only the mobilization of extra resources but also painful decisions on what can and cannot be afforded.

It is important to note two features that WHO considers as essential to the financial master plan:

- it is a plan for the use of all health sector resources - not merely the new or additional resources that are expected to become available
- it is a plan for the entire health sector - not merely the government part of it.

5. The Experience of Long-term Financial Planning

The approach recommended by WHO for the financial master plan is very new and there are thus very few examples yet available of financial master plans. However, some countries have undertaken long-term financial planning, even if not incorporating all the elements of a financial master plan. In particular, this financial planning may differ from a financial master plan in one or more of the following ways:

- it may not be explicitly oriented to achieving HFA2000
- it may take a shorter time span (e.g. ten years, for instance to 1995)

- it may restrict itself to only one element of the health sector, such as the Ministry of Health
- it may concentrate only on one element of the financial master plan, such as costing the strategy for HFA2000.

Nonetheless, this experience of long-term financial planning is of considerable use because it illustrates the methods required to make a financial master plan and demonstrates many of the practical problems. Five examples of long-term financial planning, from Jamaica, Malawi, Ethiopia, Burkina Faso (Upper Volta) and Sweden, are described briefly below.

5.1 Jamaica

An analysis of health sector expenditure in Jamaica, designed to obtain useful information for analyzing policy, planning and management problems, led on to building up estimates of the future level of spending required to achieve a given level of health care and comparing this with projections of future resource availability. The objective of this analysis was to throw light on the financial constraints that might need to be taken into account in planning.

The method adopted was to use the estimates of health sector expenditure in 1980 and to project the level of expenditure to be expected over the period 1980-2000, given certain assumptions chosen to reflect current plans and policies and foreseeable changes in the economic, demographic and technical environment. Available revenue was also projected, based on assumptions on total population growth, growth of GDP per capita and the relationship of government health expenditure and total health expenditure to GDP.

This paper on Jamaica discusses a most important point: the extent to which the assumptions turned out to be valid and the extent to which the analysis was useful in the light of Jamaica's changing economic circumstances in the early 1980s.

See, Cumper G (1986) Health sector financing: estimating health expenditure in developing countries EPC Publication No. 9, London School of Hygiene and Tropical Medicine.

5.2 Malawi

In 1984, as part of the process of developing a new national health plan, the Ministry of Health (MOH) of Malawi undertook a resource analysis for the period 1984-1995. The objective of the analysis was to provide a framework that would assist the MOH to make decisions on the future distribution of available resources across the different services it aimed to provide by 1995. The

analysis related only to the MOH, but reflects an approach to long-term financial planning for the MOH that is useful for a financial master plan.

The resource analysis was built up through the following steps:

- Step 1: analysis of the current (1984) service provision and of the resources expended on these services in terms of manpower, facilities and finance
- Step 2: identification of all objectives and targets of the MOH for the period up to 1995 covering both health status and the health delivery system
- Step 3: forecast of financial resources likely to be available to the MOH by the year 1995 based upon assumptions about Government allocations to the MOH, revenues to be made from fees, efficiency savings and donor contributions
- Step 4: comparison of total 1995 financial requirements to achieve targets with forecasted available funds. Identification of the resource gap between the funds needed and those likely to be available
- Step 5: examination of options for reducing the identified resource gap including adjustments to target manpower standards and to bed utilization levels. Identification of the reduced resource gap
- Step 6: analysis of different patterns of service delivery which could be considered within the resources forecast to be available.

The gap identified between expenditure and revenue availability clarified the choices facing the MOH on where to place additional resources.

See, Ministry of Health, Malawi (1984) Resource Analysis 1984-1995 Unpublished report.

5.3 Ethiopia

The final three examples, from Ethiopia, Burkina Faso and Sweden, concentrate on the cost side of a financial master plan, aiming to establish the resources required to implement current policies to provide a guide to whether the policies are financially feasible.

The objective of long-term forward planning in Ethiopia is to allocate available resources in ways that most effectively serve policy goals of expanded and equitably distributed health services. The

financial planning component identified resource requirements and sources of finance for the provision of health services. The cost analysis was done by calculating the recurrent and capital cost of different levels of facility, assessing the degree and nature of under-provision or inadequacy of current services and then calculating the cost of the improved level of service required. Analysis of sources of finance consisted of distributing the resources required between different sources - community, foreign assistance and government - assuming that the level of finance required would be forthcoming. The plan implied an (unrealistic) 10% average annual growth rate in government recurrent expenditure and an 11% increase in annual capital expenditure.

See, Ministry of Health, Ethiopia (1984) Financial Planning for Health Care Development in Ethiopia World Health Statistics Quarterly 37 (4).

5.4 Burkina Faso

A study in Burkina Faso (Upper Volta) raised the question of whether HFA2000 was feasible there from an operational and financial point of view, in order to encourage the search for alternative and more realistic policies. The study calculated the capital and recurrent costs of providing a health infrastructure in a province (buildings, manpower, drugs, transport, primary health care, vaccination, evaluation) and then projected this for the whole country.

See, Merkle A (1982) The Cost of Health For All. A Feasibility Study from Upper Volta. Eschborn.

5.5 Sweden

The Swedish study calculated the operational costs and manpower requirements for the Swedish health care system, 1985-2000, in order to evaluate future resource requirements under different assumptions concerning changes in the health care structure and economic development. This approach to resource requirements differs from the others discussed above in that it is based on a computer model that starts with change in the population structure, calculates the service level required, taking account of changes in the quality and pattern of care, then allows for changes in wage and productivity levels, ending up with total manpower needs and total recurrent costs.

See, Lagergren M (1986) Sweden's plan for 2000. World Health May 1986.

6. Major Issues Arising

A number of important issues arise from this consideration of long-term financial planning.

First, should the expenditure projections be based on health programmes (for instance for maternal and child health care, the elderly) or on health facilities (for instance hospital beds, health centres)? The former can more easily be associated with the health needs of the population and with plan objectives stated in terms of health status. However, health programmes may cross-cut the basic institutional and accounting elements of the health sector, namely health facilities. Good information on health programme expenditure is rarely available and thus most expenditure projections are based on the development of physical facilities rather than health programmes.

Second, at what stage should resource constraints and evidence of the opportunity cost of proposals be introduced into the analysis? Plans are frequently based on over-ambitious policies and financial planning may reveal an enormous gap between expenditure projections and revenue projections. There is a strong case for acknowledging the importance of resource constraints at the start of any planning process.

Third, to what extent are detailed expenditure and revenue projections worthwhile and to what extent can simplifying assumptions be used? For example, it might be possible to assume that, say, for every 1% increase in GDP, total health expenditure increases by 1.2%; such assumptions might be based on past government/health expenditure trends and sensitivity analysis could be undertaken to establish the variation in projections resulting from differing simplifying assumptions.

And fourth, long-term planning has proved vulnerable to political and economic changes, even in the more stable countries of the world. Between 1985 and 2000, many unexpected events may happen that will overthrow the projections of financial planners. To what extent, and how, can long term financial plans be protected against these shocks? There is a need for:

simplicity: simple methods of financial planning so that planners do not spend many months refining a financial plan that is then overturned by unexpected changes in the economic or political system

flexibility: introducing flexibility into financial plans, so that changes do not require them to be drastically revised. For instance, different scenarios might be used for the expenditure and revenue projections, to test whether health policies are vulnerable to changes in the forecast levels.

SELECTED BIBLIOGRAPHY

* particularly recommended

1. THE CONTRIBUTION OF HEALTH ECONOMICS TO HEALTH PLANNING

* ABEL-SMITH B (1972) Health Priorities in Developing Countries: The Economist's Contribution. International Journal of Health Services vol 2 no 1 pp 5-12

* ABEL-SMITH B and LEISERSON A (1978) Poverty, Development and Health Policy. World Health Organization Geneva

CULLIS JG and WEST PA (1979). The Economics of Health: An Introduction. Martin Robertson

CULYER AJ (1981). Chapter 1: Health, Economics and Health Economics. In: van der Gaag J and Perlman M 'Health, Economics and Health Economics'. North Holland, Amsterdam

LEE K (1983). Health Care in the Developing World: The role of Economists and Economics. Social Science and Medicine vol 17 no 24 pp 2007-2015

* LEE K and MILLS A (1983). The Economics of Health in Developing Countries. Oxford University Press

WELLS S and KLEES W (1980). Health Economics and Development. New York, Praeger Studies

WORLD BANK (1980). Health Sector Policy Paper. World Bank February 1980

2. ECONOMIC DEVELOPMENT AND HEALTH

* BARLOW R (1979). Health and Economic Development: A Theoretical and Empirical Review. In: Sirageldin I (ed) 'Research in Human Capital and Development' . Vol 1 pp 45-75. JAI Press, Connecticut.

CUMPER G (1983). Jamaica: A Case Study in Health Development. Social Science and Medicine vol 17 no 24 pp 1983-1993

CUMPER G (1984). Determinants of Health Levels in Developing Countries. Research Studies Press Ltd. John Wiley & Sons Inc.

* CUMPER G (1983). Chapter 2: Economic Development, Health Services and Health. In: Lee

K and Mills A (eds) 'The Economics of Health in Developing Countries'. Oxford University Press

* FREUND PJ (1986). Health Care in a Declining Economy: The Case of Zambia. Social Science and Medicine vol 23 (9) pp 875-888

HARDIMAN M and MIDGLEY J (1982). Chapter 2: Poverty, Inequality and Development. In: The Social Dimensions of Development. John Wiley and Sons

HERRICK B and KINDLEBERGER (1983). Chapter 6: Measurement and Development. In: Economic Development. Mc Graw Hill Books

HUGHES and HUNTER BM (1970). Disease and 'Development' in Africa. Social Science and Medicine vol 3 pp 443-493

McKEOWN T (1976). The Modern Rise of Population. Edward Arnold

* MUSGROVE P (1987). The Economic Crisis and its Impact on Health and Health Care in Latin America and the Caribbean. International Journal of Health Services vol 17 No 3 pp 411-441

PRESTON S (1980). Causes and Consequences of Mortality Declines in Less Developed Countries during the Twentieth Century. In: Easterlin R (ed) 'Population and Economic Change in Developing Countries'. University of Chicago Press

RAJ KN (1984). The Causes and Consequences of World Recession. World Development vol 12 no 3 pp 177-185

SANDERS D (1985). The Struggle for Health: Medicine and the Politics of Underdevelopment Macmillan

STREETEN P (1972). Chapter 3: How Poor are the Poor Countries and Why?. In: The Frontiers of Development Macmillan

TROWELL HC and BURKITT DP (1981). Western Diseases: Their Emergence and Prevention. London

3. HEALTH CARE: THE STATE VERSUS THE MARKET

BARR N (1987). The Economics of the Welfare State. Wiedenfeld and Nicolson

LEES D (1976). Economics and Non-Economics of Health Services. The Three Banks Review 110 pp 3-20

* Le GRAND J and ROBINSON R (1976)
Chapters 1 and 2 In: The Economics of Social Problems . Macmillan

ROEMER MI (1982). Market Failure and Health Care Policy. Journal of Public Health Policy December 1982

* ROEMER MI and ROEMER JE (1982). The Social Consequences of Free Trade in Health Care: A Public Health Response to Orthodox Economics International Journal of Health Services vol 12 no 1

WORLD BANK (1987). Financing Health Services in Developing Countries: An Agenda for Reform. World Bank, Washington

4. DEMAND, SUPPLY AND THE PRICE SYSTEM

* AKIN JS et al (1981). The Demand for Child Health Services in the Philippines. Social Science and Medicine vol 15c pp 249-257

* AKIN JS et al (1985). The Demand for Primary Health Services in the Third World. Rowman and Allanheld

BAUMOL WJ and BLINDER AS (1985). Chapter 4 . In: Economic Principles and Policy. Harcourt Brace Jovanovich

CHERNICHOVSKY D and MEESOK OA (1985)
Utilisation of Health Services in Indonesia
World Bank PHN Technical Note 85-18

CULYER AJ (1985). Chapters 3 and 4. In: Economics . Basil Blackwell, Oxford

* DOR A, GERTLER P and van der GAAG J (1987). Non-price Rationing and the Choice of Medical Providers in Rural Cote D'Ivoire. Journal of Health Economics vol 6 no 4 pp 291-304

GILSON L (1988). Government Health Care Charges: Is Equity Being Abandoned? EPC Publication No 15, London School of Hygiene and Tropical Medicine

* HELLER PS (1982). A Model of the Demand for Medical and Health Services in Peninsular Malaysia. Social Science and Medicine vol 16 pp 267-284

JANOVITZ B et al (1982). Caesarean Section in Brazil. Social Science and Medicine vol 16 pp 19-25

* Le GRAND J and ROBINSON R (1976)
Chapter 1: Society's objectives and the allocation of resources . In: The Economics of Social Problems Macmillan

* MAYNARD A (1979). Pricing, Demand, and the Supply of Health Care. International Journal of Health Services vol 9 no 1

McGUIRE A, HENDERSON J and MOONEY G (1988). Chapters 7-10. In: The Economics of Health Care. Routledge and Keegan Paul

WILLIAMS A (1978). Need - an Economic Exegesis. In: Culyer AJ and Wright KG (eds) 'Economic Aspects of Health Services'. Martin Robertson

5. ECONOMIC EVALUATION: GENERAL

* BERMAN P (1982). Selective Primary Health Care: Is Efficient Sufficient?. Social Science and Medicine vol 16 pp 1054-1059

CARRIN G (1984). Economic Evaluation of Health Care in Developing Countries. Croom Helm

CARRIN G (1984). Economic Evaluation of Health Care: A Review of Alternative Methods Social Science and Medicine vol 19 no 11

* DRUMMOND M (1980). Principles of Economic Appraisal in Health Care. Oxford University Press

* DRUMMOND MF, STODDART GL and TORRANCE GW (1987). Methods of Evaluating Health Care Programmes. Oxford University Press

GASPARI C (1983). Use and Misuse of Cost-Effectiveness Analysis. Social Science and Medicine vol 17 no 15 pp 1043-1046

* GREEN A and BARKER C (1988). Priority Setting and Economic Appraisal: Whose Priorities - the Community or the Economist?. Social Science and Medicine vol 26 no 9 pp 919-929

* MILLS A and THOMAS M (1984). Economic Evaluation of Health Programmes in Developing Countries. EPC Publication No 3 Winter 1984, London School of Hygiene and Tropical Medicine

* MILLS A and DRUMMOND M (1987). Value for Money in the Health Sector: The Contribution of Primary Health Care. Health Policy and Planning vol 2 no 2 pp 107-128

* WORLD HEALTH STATISTICS QUARTERLY (1985). Economic Evaluation of Health Programmes. World Health Statistics Quarterly vol 38 no 4

6. THE USE OF OUTCOME MEASURES IN ECONOMIC EVALUATION

* BARNUM H (1987). Evaluating Healthy Days of Life Gained from Health Projects. Social Science and Medicine vol 24 no 10 pp 833-842

CENTRE EIGHT JOURNAL (1985). In Search of Efficiency. Health and Social Service Journal July 18 1985

CONLY G N (1976). The Impact of Malaria on Economic Development. WHO Chronicle vol 30 pp 223-228

* DUNLOP D W (1984). Theoretical and Empirical Issues in Benefit Identification, Measurement and Valuation Related to Parasitic Disease Control in Poor Countries. Social Science and Medicine vol 19 no 10 pp 1031-1037

GHANA HEALTH ASSESSMENT PROJECT TEAM (1981). A Quantitative Method of Assessing the Health Impact of Different Diseases in Less Developed Countries. International Journal of Epidemiology vol 10 no 1 pp 73-80

* PRESCOTT N (1979). Schistosomiasis and Development. World Development vol 7 no 1 pp 1-14

* PRESCOTT N (1984). The Economics of Blindness Prevention in Upper Volta under the Onchocerciasis Control Programme. Social Science and Medicine vol 19 no 10 pp 1051-1055

TORRANCE GW (1986). Measurement of Health State Utilities for Economic Appraisal. Journal of Health Economics vol 5 pp 1-30

* WILLIAMS A (1985). Economics of Coronary Artery Bypass Grafting (plus subsequent correspondence). British Medical Journal 3 August 1985 vol 291 pp 326-329

7. COST ANALYSIS

ABEL-SMITH B.(1976). Chapter 7: The Efficient Use of Hospitals. In: Value for Money in Health Services. Heinemann

* ABEL-SMITH B (1978). Chapter 5: The Analysis of Health Service Expenditure. In:

Poverty, Development & Health Policy Public Health Paper 69, WHO

* ALEXANDER CA et al (1972). Cost Accounting of Health Centre Expenditures. Indian Journal of Medical Research December pp 1849-1863

* BERMAN P (1986). Cost Analysis as a Management Tool for Improving the Efficiency of Primary Care: some examples from Java. International Journal of Health Planning and Management vol 1 pp 275-288

CULLIS JG and WEST PA (1979). Chapter 7 In: The Economics of Health: An Introduction. Oxford: Martin Robertson

CUMPER G (1984). The Costs of Primary Health Care. Tropical Doctor January

GRAY CS (1986). State Sponsored Primary Health Care in Africa: The Recurrent Cost of Performing Miracles. Social Science and Medicine vol 22 no 3 pp 361-368

HELLER PS (1975). Issues in the Costing of Public Sector Outputs: The Public Medical Services of Malaysia. World Bank Staff Working Paper No 207

HUSSAIN AM (1983). Cost Analysis of a Primary Health Care Centre in Bangladesh. Bulletin of the World Health Organization vol 61 no 3 pp 477-483

* JACOBS P (1980). Chapter 6: Behaviour of Health Care Costs. In: The Economics of Health and Medical Care: An Introduction. University Park Press, Baltimore

* KING M (1966). Chapter 12: The economy of a district hospital. In: Medical Care in Developing Countries. Oxford University Press

LEVIN HM (1983). Cost-effectiveness: a primer Sage Publications

* MILLS A (1987). The Financing and Economics of Hospitals in Developing Countries: Key Issues and Research Questions. World Bank PHN Technical Note No 87-20

OVER M (1986). The Effect of Scale on Cost Projections for a Primary Health Care Programme in a Developing Country. Social Science and Medicine vol 22 no 3 pp 351-360

* ROBERTSON RL et al (1984). Service Volume and Other Factors Affecting the Costs of Immunisation in the Gambia. Bulletin of the World Health Organization vol 62 no 5 pp 729-736

UGALDE A (1984). Where there is a Doctor: Strategies to Increase Productivity at Lower Costs. The Economics of Rural Health Care in the Dominican Republic. Social Science and Medicine vol 19 no 4 pp 441-450

* WHO (1980). EPI Costing Guidelines. WHO Geneva

* WHO (1984). Programme Budgeting as a part of the Managerial Process for National Health Development (MPNHD). Guiding Principles. WHO Geneva

8. ECONOMIC EVALUATION STUDIES

BALLANE RC and GUNN RA (1984). Drinking Water and Sanitation Projects: Criteria for Resource Allocation.. WHO Chronicle vol 18 no 6 pp 243-248

CREESE A (1982). Cost-Effectiveness Appraisal of Immunisation Programmes. Bulletin of WHO vol 60 no 4 pp 621-632

HORTON S and CLAQUIN P (1983). Cost-Effectiveness and User Characteristics of Clinic Based Services for the Treatment of Diarrhoea: A case Study in Bangladesh. Social Science and Medicine vol 17 no 11 pp 721-729

* LEE K and MILLS A (eds) (1983). Chapters 7,8 and 9. In: The Economics of Health in Developing Countries. Oxford University Press

* MILLS A and THOMAS M (1984). Economic Evaluation of Health Programmes in Developing Countries. EPC Publication No 3, London School of Hygiene and Tropical Medicine

PHILLIPS MA, FEACHEM RG and MILLS A (1987). Options for Diarrhoeal Diseases Control: The Cost Effectiveness of Selected Interventions for the Prevention of Diarrhoea. EPC Publication No 13, London School of Hygiene and Tropical Medicine

PROST A and PRESCOTT N (1984). Cost-Effectiveness of Blindness Prevention by the Onchocerciasis Control Programme in Upper Volta. Bulletin of the WHO vol 62 no 5 pp 795-802

* SIRAGELDIN I et al (1983). Evaluating Population Programmes. Croom Helm

SOCIAL SCIENCE AND MEDICINE (1984). Special Issue on the Economic Aspects of Parasitic Diseases. vol 19 no 10

WALKER G and GISH O (1977). Mobile Health Services:A Study in Cost-Effectiveness. Medical Care vol XV no 4 pp 267-276

* WORLD HEALTH STATISTICS QUARTERLY (1985). Economic Evaluation of Health Programmes. World Health Statistics Quarterly vol 38 no 4

9. NATIONAL ACCOUNTS AND NATIONAL ECONOMIC DATA

CULYER AJ (1985). Chapters 23 and 24
In: Economics . Basil Blackwell, Oxford

* CUMPER G (1981). National Incomes and Health: Implications of Some Recent Additions to the Data Journal of Tropical Medicine and Hygiene vol 84 pp 49-61

HERRICK B and KINDLEBERGER (1983). Chapter 6: Measurement and Development
In: Economic Development . Mc Graw Hill Books

* LIVINGSTONE I and ORD HW (1980). Chapter 1: National Income and the Standard of Living
In: Economics for Eastern Africa. Studies in the Economics of Africa, Heinemann Educational Books

PREST AR (1985). Public Finance in Developing Countries. Weidenfield and Nicholson

STREETEN P (1972). Chapter 3: How Poor are the Poor Countries and Why? In: The Frontiers of Development. Macmillan

10. HEALTH SECTOR FINANCE AND EXPENDITURE: AN OVERVIEW

* ABEL-SMITH B (1985). Global Perspectives on Health Service Financing. Social Science and Medicine vol 21 no 9 pp 957-963

BLOOM G (1985). Two Models For Change in the Health Services of Zimbabwe . International Journal of Health Services vol 15 no 3 pp 451-468

DUNLOP D W (1983). Health Care Financing: Recent Experience in Africa. Social Science and Medicine vol 17 no 24 pp 2017-2025

de FERRANTI D (1985). Paying for Health Services in Developing Countries: A Call For Realism World Health Forum vol 6(2) 99-105

* HOARE G and MILLS A (1986). Paying for the Health Sector. EPC publication No. 12, London School of Hygiene and Tropical Medicine

HOWARD LM (1981). What Are The Financial Resources For 'Health 2000'? . World Health Forum vol 2 no 1 pp 23-29

JEFFERY R (1986). Health Planning in India 1951-84: The Role of the Planning Commission. Health Policy and Planning 1986 vol 1 no 2 pp 127-137

MACH EP (1978). The Financing of Health Systems in Developing Countries: Discussion Paper. Social Science and Medicine vol 12 pp 7-11

SORKIN AL (1986). Financing Health Development Projects: Some Macro-Economic Considerations. Social Science and Medicine vol 22 no 3 pp 345-349

VIVEROS-LONG A (1986). Changes in Health Financing: the Chilean Experience. Social Science and Medicine vol 22 no 3 pp 379-385

* WORLD BANK (1987). Financing Health Services in Developing Countries: An Agenda for Reform. World Bank, Washington

* WHO TECHNICAL REPORT SERIES 625. Financing of Health Services. Report of a WHO Study Group. WHO Geneva

WORLD HEALTH FORUM (1981). Sharing The Costs of Health Care. World Health Forum Vol 2 No 1 pp 85-95

* WORLD HEALTH ORGANIZATION (1987). Economic Support for National Health For All Strategies. Background document, Fortieth World Health Assembly. WHO, Geneva

* WORLD HEALTH STATISTICS QUARTERLY (1984). Health Costs and Financing World Health Statistics Quarterly vol 37 no 4

* ZSCHOCK D K (1979). Health Care Financing in Developing Countries. APHA Monograph No 1

11. THE UK NATIONAL HEALTH SERVICE

HARRISON A and GRETTON J (1985). Health Care UK 1985. An Economic, Social and Policy Audit . CIFPA

NOTES ON THE NATIONAL HEALTH SERVICE (1986). King's Fund Centre Library and Information Services

RIGDEN M S (1983). Health Service Finance and Accounting. Heinemann, London

12. SURVEYS OF HEALTH FINANCE AND EXPENDITURE

ALAILIMA P and MOHIDEEN F (1984). Health Sector Expenditure Flows in Sri Lanka. World Health Statistics Quarterly vol 37 no 4 403-420

BROTOWASISTO et al (1988). Health Care Financing in Indonesia. Health Policy and Planning vol 3 no 2 pp 131-140

CUMPER G (1986). Health Sector Financing: A Discussion Paper. EPC Publication No 9, London School of Hygiene and Tropical Medicine

GILSON L (1987). Swaziland: Health Sector Financing and Expenditure. Health Policy and Planning vol 2 no 1 pp 32-43 and subsequent correspondence Gish/Gilson vol 3 no 1 pp 74-79

* GRIFFITHS A and MILLS M (1983). Health Sector Financing and Expenditure Surveys. In: Lee K & Mills A (eds) 'The Economics of Health in Developing Countries'. Oxford University Press

LAURENT A (1982). Health Financing and Expenditure in Rwanda and Togo. Sandoz Institute for Health and Socio-Economic studies, Geneva

* MACH EP and ABEL-SMITH B (1983). Planning the Finances of the Health Sector. WHO, Geneva

MINISTRY OF HEALTH, REPUBLIC OF MALAWI (1984). A Survey of Health Sector Costs and Financing in Malawi. World Health Statistics Quarterly vol 37 no 4 375-386

WHO (1985). National Study on Resource Allocation for MCH/FP in Sri Lanka. Geneva, WHO MCH/85.7

13. SURVEYS OF HOUSEHOLD INCOME AND EXPENDITURE

BERMAN P, ORMOND BA and GANI A (1987) Treatment Use and Expenditure on Curative Care in Rural Indonesia. Health Policy and Planning vol 2 no 4 pp 289-300

* CARLSON BA (1985). The Potential of National Household Survey Programmes for Monitoring and Evaluating Primary Health Care in Developing Countries. World Health Statistics Quarterly vol 38 no 1 pp 38-64

* DEEBLE JS (1974). The Economic Approach to Family Health Studies. Social Science and Medicine vol 8 pp 529-533. MUSGROVE P (1983)

Family Health Care Spending in Latin America. *Journal of Health Economics* vol 2 no 3 pp 245-258

NAG M (1985). The Impact of Social and Economic Development on Mortality. Comparative Study of Karelia and West Bengal. In: Halstead SB, Walsh JA and Warren KS (eds) 'Good Health at Low Cost'. The Rockefeller Foundation Conference Report October 1985

PARKER RL (1986). Health Care Expenditures in a Rural Indian Community. *Social Science and Medicine* vol 22 no 1 pp 23-27

14. USER CHARGES

AKIN JS (1986). The Demand for Adult Outpatient Services in the Bicol Region of the Philippines. *Social Science and Medicine* vol 22 no 3 pp 321-328

AKIN JS (1986). Fees for Health Services and the Concern for Equity for the Poor. World Bank PHN Technical Note Series No 86-10

* BEKELE A and LEWIS MA (1986). Financing Health Care in the Sudan. *International Journal of Health Planning and Management* vol 1 no 2 pp 111-127

CASSELS A (1983). Drug Supply in Rural Nepal *Tropical Doctor* vol 13 no 1 pp 14-17

CROSS PN et al (1986). Revolving Drug Funds: Conducting Business in the Public Sector. *Social Science and Medicine* vol 22 no 3 pp 335-343

* de FERRANTI D (1985). Paying for Health Services in Developing Countries: An Overview World Bank Staff Working Papers No 721

ELLIS RP (1987). The Revenue Potential of User Fees in Kenyan Government. *Health Facilities* *Social Science and Medicine* vol 25 no 9 pp 995-1002

* GERTLER P, LOYAY L and SANDERSON W (1987). Are User Fees Regressive? The Welfare Implications of Health Care. Financing Proposals in Peru. *Journal of Econometrics* vol 36 pp 67-88

* GILSON L (1988). Government Health Care Charges: Is Equity Being Abandoned?. EPC Publication No 15, Spring 1988

HELLER P S (1982). A Model of the Demand for Medical and Health Services in Peninsular Malaysia. *Social Science and Medicine* vol 16 pp 267-284

KASONGO PROJECT TEAM (1984). Primary Health Care for Less Than a Dollar a Year. *World Health Forum* vol 5 no 3 pp 211-215

MAYNARD A (1979). Pricing, Demanders, and the Supply of Health Care. *International Journal of Health Services* vol 9 no 1 pp 121-133

MUSGROVE P (1986). What Should Consumers in Poor Countries Pay for Publicly-Provided Health Services? *Social Science and Medicine* vol 22 no 3 pp 329-333

MWABU GM and MWANGI WM (1986). Health Care Financing in Kenya: A Simulation of the Welfare Effects of User Fees. *Social Science and Medicine* vol 22 no 7 pp 763-767

* WORLD BANK (1987). Financing Health Services in Developing Countries: An Agenda for Reform. World Bank, Washington

15. PRIVATE SECTOR FINANCE

ABE M A (1985). Japan's Clinic Physicians and their Behaviour. *Social Science and Medicine* vol 20 no 4 pp 335-340

BERMAN P, ORMOND BA and GANI A (1987). Treatment Use and Expenditure on Curative Care in Rural Indonesia. *Health Policy and Planning* vol 2 no 4 pp 289-300

BLOOM G and LAING R (1986). Doctors, Private Practice and Primary Health Care Development. *Health Policy and Planning* vol 1 no 3 pp 267-269

BROTOWASISTO et al (1988). Health Care Financing in Indonesia. *Health Policy and Planning* vol 3 no 2 pp 131-140

LERMAN S J et al (1985). Treatment of Diarrhoea in Indonesian Children: What it Costs and Who Pays it. *The Lancet* September 21 pp 651-654

* LEWIS MA and MILLER TR (1987). Public-private Partnership in Water Supply and Sanitation in sub-Saharan Africa. *Health Policy and Planning* vol 2 no 1 pp 70-79 and response by . Cairncross in vol 2 no 2 pp 180-182

MWABU GM (1986). Health Care Decisions at the Household Level: Results of a Rural Health Survey in Kenya. *Social Science and Medicine* vol 22 no 3 pp 315-319

PARKER R L (1986). Health Care Expenditures in a Rural Indian Community. *Social Science and Medicine* vol 22 no 1 pp 23-27

* ROEMER MI (1984). Private Medical Practice: Obstacle to Health For All. World Health Forum vol 5 pp 195-210

* SEGALL MM (1984). Planning and Politics of Resource Allocation for Primary Health Care: Promotion of Meaningful National Policy. Social Science and Medicine vol 17 no 24 pp 1947-1960

VIVEROS-LONG A (1986). Changes in Health Financing: The Chilean Experience. Social Science and Medicine vol 22 no 3 pp 379-385

WHO (1981). Section on private medical sector in: National Decision-Making for Primary Health Care Unicef/WHO Joint Committee on Health Policy, Geneva

16. COMMUNITY FINANCE

* ABEL-SMITH B and DUA A (1988). Community-financing in Developing Countries: the Potential for the Health Sector. Health Policy and Planning vol 3 no 2 pp 95-108

* AMERICAN PUBLIC HEALTH ASSOCIATION (1982). Community Financing of Primary Health Care. Primary Health Care Issues, American Public Health Association Washington DC

BLOOM AL (1984). Prospects For Primary Health Care in Africa: Another Look at the Sine Saloum Rural Health Project in Senegal. AID Evaluation Special Study No 20 USAID April

CHRISTIAN MEDICAL COMMISSION (1987). Financing Primary Health Care Programmes: Can They Be Self-sufficient? World Council of Churches, Geneva

JANCLOES M (1985). Financing Urban Primary Health Services. Tropical Doctor vol 15 no 2 pp 9-104

LINK (1983). Newsletter of the Asian Community Health Action Network vol III no 1 Feb/March 1983

PARKER BR et al (1984). A Simple Way of Organizing Opinions. World Health Forum vol 5 no 4 pp 354-360

* RAJAGOPOLAN PK and PANICKER KN (1985). Financial Rewards Ensure Community Involvement. World Health Forum vol 6 no 2 pp 174-176

RUSSELL SS and REYNOLDS J (1985). Community Financing. Pricor Monograph Series: Issues Paper 1 May

* STINSON W (1984). Potential and Limitations of Community Financing. World Health Forum vol 5 no 2 pp 123-125

* STINSON W et al (1987). Community Financing of Primary Health Care: the PRICOR Experience. A Comparative Analysis. PRICOR, Centre for Human Services, Chevy Chase, Maryland

17. HEALTH INSURANCE

* ABEL-SMITH B et al (1986). Funding Health For All - Is Insurance the Answer?. World Health Forum vol 7 no 1 pp 3-32

BASTOS MV (1971). Brazil's Multiple Insurance Programs and their Influence on Medical Care. International Journal of Health Services vol 1 no 4 378-39

EVANS RG and MF WILLIAMSON (1978). Chapter 1: Public Intervention: Objectives and Criteria. In: Ontario Economic Council Research Studies 'Extending Canadian Health Insurance: Options for Pharmaceutical and Dental Care'. Ontario Economic Council 1978

HIGUCHI T (1974). Medical Care through Social Insurance in the Japanese Rural Sector. International Labour Review vol 109 no 3 pp 251-274

HU TEH-WEI (1981). Issues of Health Care Financing in the People's Republic of China. Social Science and Medicine vol 15C pp 233-237

INTERNATIONAL LABOUR OFFICE (1986). Primary Health Care Strategies in Latin American Social Security. ILO Geneva

INTERNATIONAL SOCIAL SECURITY ASSOCIATION GENEVA (1982). Medical Care Under Social Security in Developing Countries. Studies and Research no 18 ISSA Geneva

JANOWITZ B et al (1982). Cesarean Section in Brazil. Social Science and Medicine vol 16 pp 19-25

MALLET A (1980). Social Protection of the Rural Population. International Social Security Review, Year XXIII No 3/4: 35-393

MEIER G (1982). Providing Affordable Medical Care in a Third World Country: The Case of Egypt Inquiry vol 19 pp 346-356

* MILLS A (1983). Economic Aspects of Health Insurance. In: Lee K and Mills A (eds) 'The Economics of Health in Developing Countries'. Oxford University Press

NKANAGU T (1985). African Experience in Sickness Insurance and Health Protection under Social Security. International Social Security Review Year XXXVIII pp 119-139

PARK CK and YEON HC (1981). Recent Developments in the Health Care System of Korea. International Social Security Review Year XXXIV pp 151-167

PHUA KH (1986). Singapore's Family Savings Scheme. World Health May pp 11-12

* ROEMER MI (1971). Social Security for Medical Care: Is it Justified in Developing Countries?. International Journal of Health Services vol 1 no 4 pp 354-361

ROEMER MI (1973). Development of Medical Services under Social Security in Latin America. International Labour Review vol 108 no 1 pp 1-23

SEARO TECHNICAL PUBLICATION NO 5 (1984). Financial Planning for HFA 2000. Annex 5 General Health Scheme for Cyprus

SEARO TECHNICAL PUBLICATION NO 5 (1984). Financial Planning for HFA 2000. Chapter 12 Sub-group 3: Development of Health Insurance Schemes

SINGH HM (1982). Methods of Medical Care Delivery: the Experience of India. International Social Security Review, Year XXXV 17-37

SUH SM (1983). Dilemmas in Developing Social Security Programs for Korea. In: Gunatolleke et al 'Ethical Dilemmas of Development'. Lexington Books

WORLD BANK (1984). China. The Health Sector. A World Bank Country Study. The World Bank 1984

* ZSCHOCK D K (1982). General Review of Problems of Medical Care Delivery Under Social Security in Developing Countries. International Social Security Review Year XXXV pp 3-15

18. PROGRAMME BUDGETING

* BANKS GT (1979). Programme Budgeting in the DHSS. In: Booth TA 'Planning for Welfare: Social Policy and the Expenditure Process'. Basil Blackwell and Martin Robertson

BEYER J C (1973). Budget Innovations in Developing Countries: The experiences of Nepal. Praeger Publishers

CAIDEN N and WILDAVSKY A (1974). Planning and Budgeting in Poor Countries. John Wiley and Sons

LEE K and MILLS A (1982). Plans, Programmes and Budgets. In: Policy-making and Planning in the Health Sector. Croom Helm

MAYNARD A (1984). Budgeting in Health Care Systems. Effective Health Care vol 2 no 2

* MOONEY G (1984). Programme Budgeting: An Aid to Planning and Priority Setting in Health Care Effective Health Care vol 2 no 2

POLE J D (1974). Programmes, Priorities and Budgets. British Journal of Preventive and Social Medicine vol 28 pp 191-195

SATIA JK (1983). The Indian Family Planning Programme. In: Sirageldin I et al Evaluating population programmes. Croom Helm

STEELE R and GRAY AM (1980). Beyond the Programme Budget. Economics and Resource Planning in the NHS Hospital and Health Services Review, March

WATERSTON A (1965). The Budget's Role in Planning. In: Development Planning: Lessons and Experience. Johns Hopkins Press

* WHO (1984). Programme Budgeting as a part of the Managerial Process for National Health Development (MPNHD): Guiding Principles

19. FINANCIAL PLANNING

ABEL-SMITH B (1984). Chapter 3: An Overview of Health Care and Financing Studies. In: Financial Planning for HFA 2000. Searo Technical Publications no 5

BLOOM G (1988). Planning Health Sector Expenditure in Support of PHC. Health Policy and Planning vol 3 no 1 pp 59-68

CUMPER G (1986). Health Sector Financing: Estimating Health Expenditure in Developing Countries. EPC Publication No. 9, London School of Hygiene and Tropical Medicine.

LAGERGREN M (1986). Sweden's Plan for 2000 World Health May pp 7-8

* MACH EP and ABEL-SMITH B (1983)
Planning the Finances of the Health Sector
WHO, Geneva

MERKLE A (1982). The Cost of Health For All. A
Feasibility Study from Upper Volta. Eschborn

MINISTRY OF HEALTH, ETHIOPIA (1984).
Financial Planning for Health Care Development in
Ethiopia. World Health Statistics Quarterly 37 vol
4 pp 421-427

* SEGALL MM (1983). Planning and Politics of
Resource Allocation for Primary Health Care:
Promotion of Meaningful National Policy. Social
Science and Medicine vol 17 no 24 pp 1947-60

GLOSSARY

Accounting system The procedures for recording the money value of transactions - purchases, sales, receipts, payments etc - to show their effect on the financial position of the organisation.

Amortisation The process of writing off the value of an asset over its working life. It can be calculated as an annuitised capital cost.

Annuitised capital cost The cost of a capital good translated into an equivalent annual amount in order to make it comparable with operating costs.

Average cost The cost per unit of output (total costs divided by total number of units of output). Also known as unit cost.

Benefit-cost ratio Total discounted benefits divided by total discounted costs. The outcome should be greater than 1 for an investment to be potentially worthwhile.

Budget A formal, written estimate of income and expenditure for a future period, leading to the allocation of funds to budget holders.

Budget holder The officer responsible for the management of a budget and accountable for the use of budget funds.

Budgetary control The process of checking actual income and expenditure against a budget so that expenditure can be controlled, progress monitored and remedial action taken if necessary.

Budgeting The process of drawing up a budget and monitoring and controlling expenditure against the budget.

Capital The stock of goods which are man-made and used in production (as opposed to consumption). Fixed capital (durable goods such as buildings) is usually distinguished from circulating capital (stocks of raw materials and semi-finished goods which are rapidly used up). In accounting conventions, capital goods are usually taken as those with a life of more than one year. Part of the initial cost of new health programmes will be the purchase of capital goods, which will lead to continuing expenditure in the future to operate buildings, equipment, etc. (see operating costs and recurrent costs). See also human capital.

Capital budget The budget which is used to purchase capital goods.

Capital cost The cost of employing capital goods. In an economic sense, it is the rate of return forgone by not using the funds spent on particular capital goods in other ways. In accounting terms, it is the money expenditure required to purchase capital goods.

Capital expenditure Outlay of money on capital goods.

Capital funds Sums of money available for purchasing capital goods.

Capitation payment Payment per person irrespective of the number of items of service provided.

Charge Price or fee.

Community financing Ways of raising money that are organised and controlled by communities themselves. Contributions may also be provided in the form of materials and community or individual labour.

Constant prices Prices of goods and services which have been adjusted to remove the effect of changes in the purchasing power of money (eg inflation).

Consumer surplus The excess of the amount a consumer is prepared to pay for a good (rather than go without it) over the amount actually paid.

Consumption Acquiring goods or services in order to obtain immediate satisfaction (in contrast to investment which permits greater consumption in the future by increasing a country's productive capacity).

Co-payment Requiring the insured user of services to pay part of the cost directly.

Cost What has to be given up to achieve something. Either:

- (a) the value of opportunities which are forgone in order to achieve something (the economic definition); or
- (b) the total money expenditure required to achieve something (the accounting definition).

Cost-benefit analysis A form of economic evaluation where all the costs and benefits are expressed in money terms. In principle, this form of analysis enables one to assess whether a particular objective is worth achieving. However, estimation difficulties often reduce cost-benefit analysis to a consideration of those costs and benefits that are easy to express in money terms.

Cost-effectiveness analysis A form of economic evaluation where the costs are expressed in money terms but some of the effects are expressed in physical units (eg life years gained, cases detected). It is usually used to compare different ways of achieving the same objective (eg. saving lives) and assumes the objective is worth achieving.

Cost-sharing Sharing the costs of providing a particular type of health care between the patient and agencies such as the provider of care and the employer of the patient.

Demand The quantity of goods or services that consumers wish and are able to buy at a given price in a given period.

Depreciation Decrease in value of a capital good because of passage of time, wear and tear etc. An allowance for depreciation may be included as an operating cost in accounts.

Development budget The budget for activities which promote a country's development. Many governments have a development budget to finance (often from external sources) activities which will increase the country's productive capacity. Some countries may have a capital rather than a development budget.

Development cost The cost of setting up new activities (construction, equipping, training, etc.) as opposed to the costs of operating or running them.

Development expenditure Money outlay on setting up new activities.

Development funds Money available for new activities such as constructing buildings, setting up training programmes, etc. The majority of development funds tend to be spent on capital rather than recurrent items.

Discounting The process of calculating the present value of costs and benefits occurring in the future by applying a discount rate. (The procedure is the reverse of calculating the annual increase in a sum invested at a given rate of interest).

Discount rate The annual rate at which the value of a future cost or benefit is reduced to find its present value. The discount rate expresses society's time preference rate. For example, at a discount rate of r , an event occurring in n years' time has a present value of $(1 + r)^{-n}$.

Earmarked tax A tax assigned to a specific purpose (also called hypothecated).

Economics Economics is concerned with those aspects of human behaviour, and those institutions, which affect the use of scarce resources to produce and distribute goods and services to satisfy human wants.

Economic evaluation A process whereby the costs of programmes, alternatives or options are compared with their consequences (in terms of improved health or savings in resources). Also known as the cost-benefit approach or economic appraisal, it embraces a family of techniques including cost-effectiveness analysis and cost-benefit analysis.

Efficiency Relates to output per unit cost of the resources employed. Resources are being used efficiently if a given output is produced at minimum cost, or maximum output is produced at a given cost ('operational' efficiency). Economists also use the term in the wider sense of cost-effectiveness and cost-benefit analysis ('allocative' efficiency).

Elasticity The degree of responsiveness of one variable to changes in another. If responsiveness is high, it is termed elastic; if low, inelastic. The concept is applied widely, but most commonly to the variables affecting demand and supply (eg price elasticity, income elasticity).

Eligibility The right to benefit from a service by virtue of membership of a defined group such as an insurance scheme, community or workplace.

Equity Fairness or justice.

Equivalent annual cost The recurring annual sum or annuity, which over the life of the project has a present value equal to a lump sum payment made now.

Expenditure Outlay of money to purchase goods or services.

Externality Externalities exist when the level of consumption or production of some good or service by a consumer or firm has a direct effect on the level of welfare of another consumer or firm, as opposed to an indirect effect through the price mechanism. These effects may be desirable or undesirable. An often quoted example in the health field is the protection against disease provided to others when an individual is immunised.

Factors of production The inputs required for making a commodity. Typically classified as land, labour and capital.

Fee-for-service Payment of a charge per item of health care received (eg consultation, drug, diagnostic test).

Finance The provision of money when and where required. May be to purchase capital goods, or recurrent goods and services such as drugs and manpower.

Financing methods Ways of raising financial (and sometimes other) resources to provide services. Financing methods include fee-for-service, insurance schemes and payroll taxes.

Financing sources The origin of the financial resources used to provide services. Sources include individuals, commercial and industrial organisations, governments and external donors.

Financing system The set of activities concerned with the provision of money for health services that can be considered as a connected whole.

Fiscal To do with public expenditure and with raising government revenue, for instance through taxation.

Fixed costs Costs which do not vary with the level of output in the time period considered (usually one year).

Foreign exchange The currency of other countries. It is required by individuals and institutions to buy goods and services from, or make gifts or loans to, people in other countries.

Free market A market in which the forces of supply and demand are allowed to operate unhampered by government regulation or other interference (see market).

Funds Sums of money and credit.

Government revenue The money that government receives (from income tax, customs duties, borrowing, etc).

Gross domestic product (GDP) A measure of the total flow of goods and services produced in a country in a year. Outputs of goods for final consumption and investment and services are valued at market prices and added up. Many socialist countries use instead "material balance" which excludes health care and other services to households.

Gross national product (GNP) GNP equals GDP plus the income received by domestic residents from investment abroad less income earned in the country paid to foreigners abroad.

Health sector The part of the economy which is involved in activities intended to improve health. The term may be used to mean health services but it is often used synonymously with the term health system, to cover both health services and health-related activities.

Health System The health system includes health services and all health-related activities. It is composed of a number of levels - the first level (the first point of contact between the system and the people) and intermediate and central levels.

Household expenditure The total outlay of money by a household (usually over one year). Payments made in kind may be valued in terms of their money equivalent.

Household income The money and/or goods received by a household over some time period. It is often calculated net of tax and includes salaries and wages, and also goods produced by the household, such as agricultural products, which are consumed within the household or sold.

Human capital The skills and capabilities stored in an individual and generated by investment in education training and health, and more generally resulting from work experience.

Hypothecated tax Tax reserved for a particular purpose.

Income The money and/or goods received by individuals, households, companies or governments over some time period (usually a year).

Income distribution The way in which total national income is divided among households in the economy.

Incremental cost The additional cost of one programme, alternative or option over and above another.

Inelastic See elasticity.

In-kind payment Payments made not in terms of money but in goods or services.

Inputs Goods and services used in production, such as capital goods (buildings, equipment), labour, raw materials, etc.

Insurance An agreement to pay a premium at regular intervals for which the insurer will cover the cost or pay compensation if the event that is insured against occurs (eg illness, fire, theft). Insurance spreads the risk between all people contributing so that the cost of treating one person's

illness, for instance, can be paid for by everyone's contributions.

Investment Expenditure on capital goods which are then used in production. In a more general sense, it means undertaking any activity which involves a sacrifice (eg payment of money) followed by a benefit (eg enjoyment of a good).

Marginal benefit The change in total benefit in response to a small change in the level of consumption. Marginal private benefit refers to benefits experienced by the individual doing the consuming; marginal social benefit includes also benefits experienced by others in society.

Marginal cost The change in total cost at a given scale of output when a little more or a little less output is produced. Marginal private cost refers to costs internal to the individual or agency, whereas marginal social cost includes also costs external to the individual/agency (ie costs to society as a whole).

Market A market exists when buyers wishing to exchange money for a good or service are in contact with sellers wishing to exchange goods or services for money.

Means test Making the distribution of a benefit (eg subsidised health care) available to individuals only after an investigation of their 'means' (ie income and/or wealth).

Merit good Goods or services where government believes individuals should not be allowed free choice of whether to consume because of lack of information about their effects (good or bad).

Monopoly A monopoly exists when a firm or individual produces or provides and sells the entire output of same commodity or service.

National income The money value of all goods and services earned in a country over a specified time period. It may be calculated as the sum of either incomes or expenditures of all residents, companies and government bodies. GDP and GNP are related measures.

Operating cost The cost of operating an enterprise or service, also called recurrent costs. In general, those costs of providing a service that vary with the level of output (eg drugs) in contrast to those which are fixed over a given time period, usually a year (eg capital costs). Usually calculated on an annual basis.

Opportunity cost The benefits to be derived from using resources in their best alternative use. It is therefore a measure of the sacrifice made by using

resources in a given programme. When economists use the term 'cost', they mean opportunity cost. This may not be the same as health care expenditures.

Outputs The end-result of production, that is what is produced.

Overheads The costs pertaining to general services (eg administration) which do not necessarily arise from the operation of a given programme.

Payroll tax A tax levied on employers' wage bills. It is often used to finance social insurance.

Per capita national income The total national income of a country divided by the total population.

Per diem The daily rate for reimbursement of hospital expenditures, usually based on the hospital average daily cost.

Present values The value now of future costs or benefits discounted at a given rate.

Price discrimination Charging different prices to different consumers, for the same good or service, where the price differences do not reflect differences in cost of supply.

Price index A price index shows how the prices of a set of goods and services have changed over time. A given physical quantity of items is priced at current prices at regular intervals. The resulting total money value of the items is then expressed as a percentage of their value at some base year.

Private health insurance Health insurance that is sold by either commercial firms or non profit-making organisations to individuals or groups. Such insurance is voluntary for the individual or group as a whole (though it may be compulsory for members of the group, eg employees of a firm).

Private sector That part of the economy of a country which is not directly controlled by the public sector.

Production The process of producing goods and services which satisfy human wants.

Productivity The output per unit of input in a given time period, eg number of visits per physician per day.

Public goods Commodities or services that (a) can be used, consumed or enjoyed by an increasing number of people without diminishing the amount available to others (b) are available to everyone in the catchment area independent of the size or

existence of payment and (c) cannot be with-held from non-payers.

Public sector That part of the economy of a country that comes within the scope of central government, including local government authorities and public corporations.

Real terms A variable (such as national income, or health expenditure) is expressed in 'real terms' if its value has been adjusted to remove the effect of changes in price. The resulting value is said to be at a constant price.

Recurrent budget The budget which can be used to purchase items of a recurrent nature, such as salaries and raw materials, as opposed to once-and-for-all payments for capital goods.

Recurrent costs Costs that 'recur' i.e. the costs of running an enterprise, such as salary and raw materials costs. Also known as operating costs.

Recurrent expenditure Outlay of money on items that 'recur' year after year (salaries, raw materials, etc.)

Recurrent funds Money available to spend on items of a recurrent nature.

Resources The inputs that are used to produce and distribute goods and services. They are conventionally classified into land (including natural resources), labour (people) and capital (goods made to produce other goods). In health programmes they include inputs which are not under the control of the health sector, such as a patient's time.

Scarcity The lack of a commodity in relation to the demand for it. Resources are scarce, and thus choices must be made on how to allocate them.

Sector The economy of a country can be divided into 'sectors'. The broadest classifications are between the 'private' and 'public' sectors, and the economic and social sectors. The latter can be divided into sectors with common activities such as manufacturing and agriculture in the economic sector and education and health in the social sector.

Social benefit The benefit from an activity to society and not merely to the individual or agency carrying it out.

Social cost The cost of an activity to society and not merely to the individual or agency carrying out the activity.

Social insurance An insurance scheme set up and controlled by government or public agencies to provide protection against unemployment, old age, sickness etc. Social insurance is usually compulsory for the whole population or for certain groups and is often financed by a payroll tax. Also called social security.

Supply The quantity of goods or services coming on the market at a given price in a given time period.

Time cost The cost individuals incur in being inactive through illness or in travelling to and waiting for health services. Time can be valued in terms of its opportunity cost (i.e. value of lost production/leisure).

Unit cost The total cost of an activity divided by the number of units of output produced. Also known as average cost.

User charge Requiring the users of a service to pay a fee.

Utility The satisfaction/pleasure derived from consuming some quantity of a good or service.